

A Importância da Ciência para o Futuro do Rio São Francisco

ESTIMATIVA DA QUALIDADE DA ÁGUA DO RESERVATÓRIO DE TRÊS MARIAS A PARTIR DOS SENSORES SENTINEL-2 MSI E LANDSAT-8 OLI

Pizani, F. M. C.¹; Maillard, P.¹; Ferreira, A. F. F.²; Amorim, C. C.²

¹ Universidade Federal de Minas Gerais – UFMG/IGC
Campus Pampulha, Belo Horizonte/MG, CEP 31270.901
fm.coelho@yahoo.com.br – philippermaillard@yahoo.com.br

² Universidade Federal de Minas Gerais – UFMG/DESA
Campus Pampulha, Belo Horizonte/MG, CEP 31270.901
adrielly.fonseca@gmail.com – camila@desa.ufmg.br

Abstract. The low operational cost of using freely available remote sensing data is a strong incentive for water agencies to complement their field campaigns and produce spatially distributed maps of some water quality parameters. The objective of this study is to compare the performance of Sentinel-2 MSI and Landsat-8 OLI sensors to produce multiple regression models of water quality parameters in the Três Marias hydroelectric reservoir. Water quality parameters were measured in loco using sensors and also analyzed in laboratory from water samples collected simultaneously. The date of sampling corresponded to the almost simultaneous overflight of Sentinel-2B and Landsat-8 satellites. Two optically active parameters were considered in this study case: chlorophyll-a (chl-a) and turbidity. The multiple regression models used the spectral reflectance bands from both sensors (separately) as predictors. The reflectance values were based on an averaging window of 90x90 m instead of single pixel vectors. Stepwise variable selection combined with a priori knowledge were used to optimized the choice of predictors. The optically active parameters yielded strong regression models and significant results with both the Sentinel and Landsat sensors. All results with $r^2 > 0.75$ and p -value < 0.01 . The spectral bands resulting from the selection methods were good predictors.

Palavras-Chave – Qualidade da água; Sensoriamento remoto; Modelos estatísticos; Análise de regressão; Reflectância.

INTRODUÇÃO

Aplicações de técnicas de sensoriamento remoto e de processamento de imagens em ambientes aquáticos são possíveis devido à relação entre o comportamento espectral e os componentes opticamente ativos (COAs) da água, que respondem de forma diferente quando comparados aos constituintes terrestres. Assim, a reflectância e absorção dos COAs da água podem ser estudados em determinados comprimentos de onda. A água contém uma assinatura espectral única que pode ser distinguida pela alta absorção na faixa do infravermelho próximo, o que faz com que os corpos d'água possam ser detectados e delimitados facilmente (Teodoro, 2016, p.167). As alterações observadas nas propriedades ópticas da água vêm dos efeitos da interação com a radiação solar que permite que os COAs sejam usados como indicadores de qualidade da água (Barbosa *et al.*, 2019, p.25).

Atualmente, o monitoramento dos parâmetros de qualidade da água demanda custos altos de coletas e de análise das amostras (Karaoui *et al.*, 2019, p.34; Yepez *et al.*, 2018, p.21), além dos custos referentes ao tempo despendido e à mão-de-obra especializada Giardino *et al.*, 2010, p.3887). O monitoramento a distância promove a diminuição destes custos e possibilita a estimativa de alguns parâmetros de qualidade da água com alta frequência temporal e de forma distribuída espacialmente. Diversos satélites apresentam aquisição sistemática de imagens com resoluções variadas da ordem dos centímetros a quilômetros. Dentre os satélites ópticos mais comuns estão as

A Importância da Ciência para o Futuro do Rio São Francisco

séries Landsat (como o Landsat-8, portando o sensor OLI) e Sentinel (como o Sentinel-2, A e B, portador do sensor MSI) que fornecem imagens com frequência temporal variando de 16 dias (Landsat) a 5 dias (Sentinel). As missões Landsat-8 e Sentinel-2 fornecem gratuitamente imagens com alta resolução espacial além de facilidade de aquisição dos dados.

Em paralelo ao uso de sensores remotos, a modelagem espacial associada à estatística torna-se uma ferramenta capaz de gerar modelos distribuídos de parâmetros bio-ópticos de qualidade da água. As abordagens que utilizam métodos estatísticos são comuns em estudos de diagnósticos de qualidade da água, sobretudo, aquelas que utilizam técnicas de regressão, que estão entre as mais frequentemente aplicadas (Mainali *et al.*, 2019, p.802).

Os grandes reservatórios brasileiros demandam monitoramento extensivo de um elevado número de parâmetros ambientais, uma vez que muitos deles fornecem subsídio à população que deles se beneficia, seja pelo lazer ou pela economia, como é o caso do Reservatório de Três Marias na bacia do Rio São Francisco. Atualmente, os monitoramentos são realizados de forma tradicional e *in situ*. Este fato ocasiona a alocação significativa de recursos humanos e financeiros, o que dificulta as ações de preservação e de controle ambiental em tempo satisfatório. Outros fatores também podem ser limitantes neste monitoramento como a dificuldade de se obter dados tendo em vista as particularidades e a dimensão do corpo hídrico (Karaoui *et al.*, 2019, p.31).

Este estudo buscou avaliar e comparar o desempenho dos sensores a bordo dos satélites Landsat-8 e Sentinel-2 para estimar parâmetros de COA. Foram realizadas análises de regressão, de maneira a correlacionar parâmetros de qualidade da água coletados *in situ* com valores de reflectância e gerar modelos distribuídos de toda área da represa.

METODOLOGIA

Área de Estudo e Aquisição de Dados Limnológicos

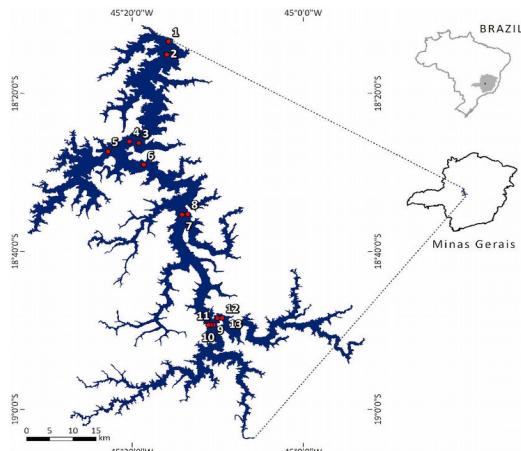
O Reservatório de Três Marias está localizado no contexto hidrográfico da Bacia do Alto Rio São Francisco (Figura 1). O reservatório apresenta área útil de inundação de 1090 km² e cerca de 20 Gm³ de volume máximo. O reservatório se insere entre os dez maiores reservatórios brasileiros quanto à área e ao volume (von Sperling, 1999, p.117–119). A região de inserção do reservatório se caracteriza por duas estações climáticas bem definidas, sendo uma chuvosa durante o período do verão e outra seca, no período do inverno.

A coleta de dados *in situ* ocorreu no dia 4 de outubro de 2019 durante o período da manhã e utilizou um barco a motor que percorreu 85 km entre a barragem e a confluência dos rios Paraopeba e São Francisco. A data foi escolhida de forma a sincronizar a coleta com a passagem dos satélites Landsat-8 e Sentinel-2, que coincidem a cada 80 dias. As condições climáticas previstas foi outro fator que motivou para a escolha da data, uma vez que havia indicativo de céu claro e limpo.

Foram levantados 13 pontos amostrais ao longo do reservatório. Mediú-se *in loco* o parâmetro turbidez com o apoio de um turbidímetro enquanto o parâmetro clorofila-a (chl-a) foi determinado em laboratório com amostras. Outros parâmetros não opticamente ativos também foram mensurados mas, não serão considerados para efeitos deste estudo. Apenas os parâmetros chl-a e turbidez foram considerados, ambos passíveis de monitoramento por sensoriamento remoto.

A Importância da Ciência para o Futuro do Rio São Francisco

FIGURA 1. Localização dos pontos de coleta de amostras no Reservatório de Três Marias.



Aquisição de Dados de Sensoriamento Remoto e Pré-processamento

Este trabalho utilizou os produtos gratuitos de nível 2A provenientes dos satélites Sentinel-2 (Agência Espacial Europeia - ESA) e Landsat-8 (Serviço Geológico dos Estados Unidos - USGS). Imagens de nível 2A são processadas e tem corrigidas a radiometria e geometria, dispondo de produtos reflectância, reparados de acordo com parâmetros próprios. A passagem do Sentinel-2 ocorreu às 10:12:49 AM e do Landsat-8 às 10:03 AM (horário local). Ambas imagens apresentaram excelente condição climática e ausência total de nuvens. Os dados foram convertidos de radiância para reflectância na superfície. As bandas de 20 m do satélite Sentinel-2 e a banda 1, de 60 m de resolução espacial, foram reamostradas para 10 m pelo método bilinear. As bandas do satélite Landsat-8 foram mantidas na resolução espacial de 30 m.

Os pixels de reflectância dos pontos amostrados foram extraídos em janelas de 90x90 m em torno da coordenada dos pontos para ambas as imagens. Isto permitiu desconsiderar o erro inerente ao GNSS portátil utilizado e pela deriva natural do barco causada pela ação dos ventos. As estatísticas de cada janela foram computadas sobre as imagens e a média do valor espectral foi utilizado como valor de referência para a análise.

Modelos Estatísticos e Análise de Regressão

O software estatístico RStudio (v. 1.5.5019) foi utilizado para as análises de regressão lineares múltiplas a fim de correlacionar os valores dos parâmetros coletados *in situ* e as médias espetrais das bandas de cada imagem. Técnicas de regressão estatística podem fornecer resultados robustos quando aplicados à áreas de estudo específicas (Matthews, 2011, p.6872).

Foram aplicadas regressões lineares múltiplas aos dois parâmetros considerando todas as bandas disponíveis estipuladas. Dois métodos de seleção de variáveis foram utilizados:

- *Stepwise*: consiste na seleção do melhor conjunto de preditores por iteração, de forma a indicar as variáveis que apresentam o melhor desempenho para o modelo (James *et al.*, 2013, p.207). O método adotado utilizou a abordagem híbrida ou bidirecional que considera

A Importância da Ciência para o Futuro do Rio São Francisco

a adição de variáveis uma a uma (forward) e, ao final, remove aquelas que não são importantes para o modelo (backward);

- Referência bibliográfica: o segundo método combinou preditores amplamente utilizados, com base em levantamento bibliográfico de estudos de sensoriamento remoto aplicados à chl-a e turbidez. Esta escolha objetivou refinar os resultados gerados pela regressão *Stepwise* garantindo a presença de importantes bandas na análise (Babyak, 2004, p.419).

A opção por estratégias combinadas teve por finalidade reduzir o número de variáveis preditoras, tornando o modelo mais objetivo, mais significativo e evitando um possível *overfitting*.

Validação dos Modelos e Avaliação da Acurácia

Foram aplicadas métricas estatísticas com a finalidade de descrever a diferença entre os dados previstos e coletados, além de avaliar a performance dos modelos. As métricas foram: Coeficiente de Determinação (r^2), Erro Médio Absoluto (MAE), Raiz do Erro Médio Quadrático (RMSE) e Raiz do Erro Médio Quadrático Normalizado (NRMSE). Após a aplicação das métricas estatísticas, os modelos de regressão foram aplicados às bandas dos satélites correspondentes por meio de álgebra de mapas para gerar mapas de chl-a e de turbidez.

RESULTADOS E DISCUSSÃO

Qualidade da Água no Reservatório de Três Marias

Os resultados obtidos em laboratório coletados durante a campanha de campo no Reservatório de Três Marias estão apresentados na Tabela 1. Analisando os dados obtidos nos pontos amostrais de uma maneira geral, o que se percebe claramente é uma certa (ainda que pequena) deteriorização da qualidade da água no reservatório a medida em que se avança no sentido norte-sul. Particularmente, nos pontos próximos aos braços dos principais afluentes (Rio São Francisco e Rio Paraopebas) há os maiores índices de turbidez se comparados aos demais pontos da represa. Os dois primeiros pontos, localizados próximo à barragem no vetor norte do reservatório, apresentam os melhores resultados para o parâmetro relativo à transparência, indicando turbidez extremamente baixa (0,2 NTU).

Os pontos 3, 4, 5 e 6 estão inseridos na porção centro-norte do reservatório próximo a áreas caracterizadas por atividades agrícolas e piscicultura, com forte presença de pivôs centrais e tanques de pesca. Entre estes pontos estão os maiores índices de chl-a (8,10 $\mu\text{g/L}$). Para o Instituto de Gestão das Águas do Estado de Minas Gerais (IGAM), considerando apenas a chl-a identificada na data da coleta, estes valores não indicam, neste momento, estado de eutrofização do reservatório. Entretanto, cabe ressaltar que esta classificação não é simplória e ainda depende de uma série de fatores que não serão destrinchados nesta pesquisa. É igualmente importante considerar o período seco em que a coleta foi realizada, onde a qualidade da água é habitualmente mais alta.

Os pontos 7 e 8 localizam-se na extremidade norte de uma pequena ilha, denominada Ilha do Mangabal, e apresentam altos índices de turbidez e chl-a, em comparação aos demais. Próximo a estes pontos, foram observadas possíveis colônias de algas e cianobactérias. Este fato reafirma a necessidade do monitoramento contínuo, especialmente para parâmetros como chl-a, fósforo e

A Importância da Ciência para o Futuro do Rio São Francisco

nitrogênio (carga de nutrientes), que podem contribuir para uma futura eutrofização do reservatório além de alteração do ecossistema local de modo significativo (Barbosa *et al.*, 2019, p.49).

TABELA 1. Mensuração dos parâmetros opticamente ativos coletados *in situ*.

| PONTO | CHL-a (µg/L) | TURBIDEZ (NTU) |
|-------|--------------|----------------|
| 1 | 1.94 | 0.24 |
| 2 | 1.78 | 0.20 |
| 3 | 2.08 | 0.65 |
| 4 | 2.15 | 0.32 |
| 5 | 1.63 | 0.53 |
| 6 | 8.10 | 1.06 |
| 7 | 5.68 | 0.53 |
| 8 | 7.76 | 1.11 |
| 9 | 4.40 | 1.04 |
| 10 | 4.38 | 1.10 |
| 11 | 5.35 | 1.08 |
| 12 | 4.30 | 1.08 |
| 13 | 3.82 | 1.11 |

Os pontos de 9 a 13 estão localizados na bifurcação entre os rios São Francisco (pontos 9, 10 e 11) e Paraopebas (12 e 13) e representam um conjunto de valores que indicam menor qualidade da água para o reservatório. Além da baixa transparência, observam-se os maiores índices de turbidez presentes no reservatório (variando entre 1,04 e 1,11 NTU).

Análise de regressão e avaliação da acurácia

A construção de modelos relacionando reflectância espectral da água e a sua qualidade no Reservatório de Três Marias foi derivada, primeiramente, pelo método *Stepwise*. O método consiste na seleção das variáveis independentes mais impactantes para explicar a variância de cada parâmetro. A combinação entre o método *Stepwise* e seleção a priori de preditores baseado em bibliografias publicadas levou em consideração as aplicações recentes para chl-a (Yadav *et al.*, 2019; Dutra *et al.*, 2019; Watanabe *et al.*, 2018) e turbidez (Liu e Wang, 2019; Gonzalez-Marquez *et al.*, 2018). Os resultados obtidos estão apresentados na Tabela 2.

TABELA 2. Seleção das bandas espetrais e método utilizado.

| PARÂMETRO | SENTINEL-2/MSI | LANDSAT-8/OLI | MÉTODO |
|-----------|----------------|---------------|------------------------------|
| Chl-a | B2, B3, B4, B5 | B2, B3, B4 | <i>Stepwise</i> + Referência |
| Turbidez | B1, B3 | B1, B3 | <i>Stepwise</i> + Referência |

A partir da seleção de bandas, os modelos puderam ser desenvolvidos por meio de regressão linear múltipla. A Tabela 3 descreve as equações de regressão para estimar cada parâmetro de qualidade da água a partir dos valores de reflectância selecionados para cada modelo.

A Importância da Ciência para o Futuro do Rio São Francisco

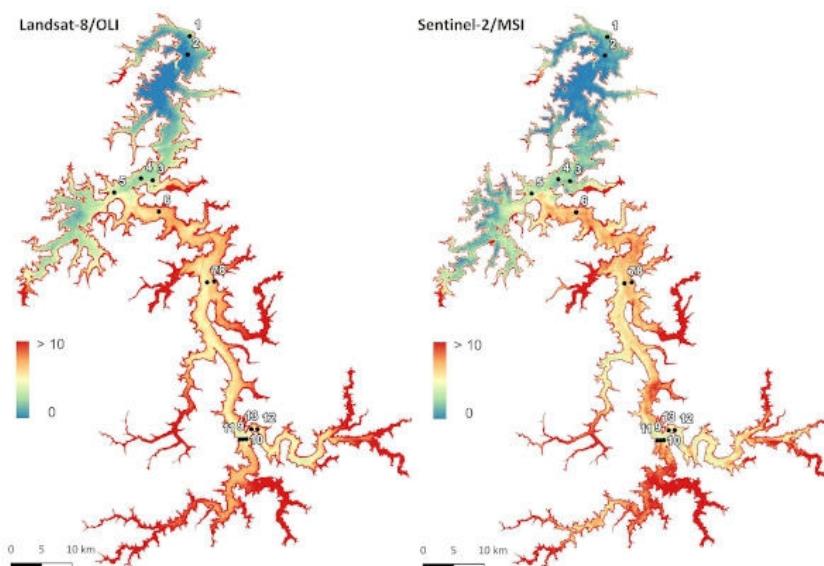
TABELA 3. Equações de regressão dos modelos de qualidade da água.

| DEPENDENTE | Equações SENTINEL-2/MSI | r^2 |
|------------|---|-------|
| Chl-a | $18.420 + (-1263.615 * B2) + (413.934 * B3) + (52.127 * B4) + (710.274 * B5)$ | 0.75 |
| Turbidez | $0.4532 + (-56.94 * B1) + (43.57 * B3)$ | 0.80 |
| DEPENDENTE | Equações LANDSAT-8/OLI | r^2 |
| Chl-a | $15.717 + (-1247.530 * B2) + (192.236 * B3) + (846.290 * B4)$ | 0.67 |
| Turbidez | $0.23 + (-59.28 * B1) + (48.00 * B3)$ | 0.80 |

A estatística descritiva apresentada revela altos índices referentes ao r^2 para os parâmetros opticamente ativos avaliados, variando entre 0.67 e 0.80 para ambas as janelas. Isto significa que as bandas espectrais definidas por meio dos métodos propostos demonstraram ser bons preditores dos parâmetros utilizados, em especial para a janela 90x90 m do sensor MSI.

Entretanto, é necessário considerar, além dos índices apontados pelos coeficientes adotados nesta pesquisa, o quanto os modelos são estatisticamente significantes e os erros associados aos modelos ajustados. Primeiramente, para a validação dos modelos, foi considerada a observação do p -value < 0.05 , ainda que a aplicação do método *Stepwise* permita um valor mais liberal para este índice (Babyak, 2004, p.416). Ainda assim, todos os modelos apresentaram valores abaixo de 0.01. Ao analisar os modelos da janela de 90x90 m do sensor OLI, percebe-se que os ajustes tiveram desempenho levemente inferior àqueles referentes aos modelos do sensor MSI. As Figuras 2 e 3 apresentam os mapas gerados para os modelos de chl-a e turbidez para os sensores OLI e MSI. As métricas utilizadas para a avaliação dos modelos podem ser observadas na Tabela 4.

FIGURA 2. Mapa dos modelos de chl-a para os sensores OLI e MSI, respectivamente.



A Importância da Ciência para o Futuro do Rio São Francisco

FIGURA 3. Mapa dos modelos de turbidez para os sensores OLI e MSI, respectivamente.

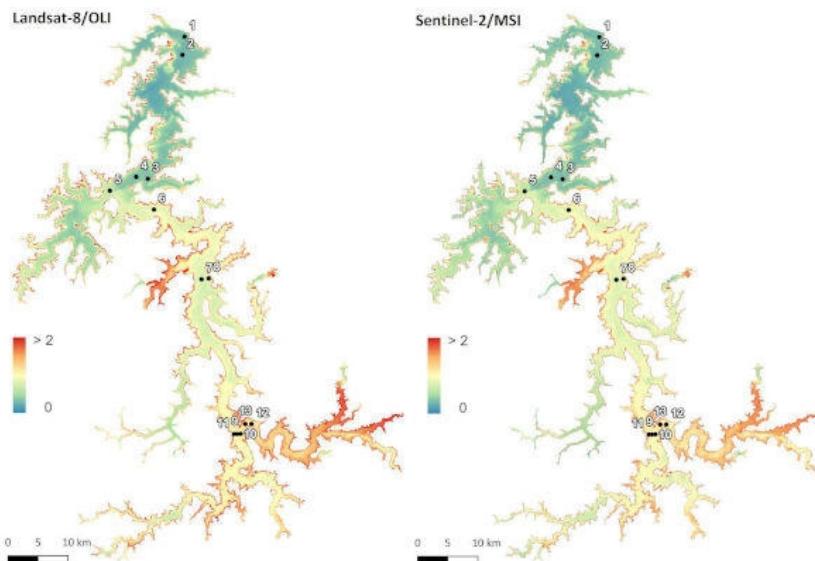


TABELA 4. Validação dos modelos com as métricas estabelecidas para os sensores MSI e OLI.

| PARÂMETRO | SENTINEL-2/MSI | | | | | LANDSAT-8/OLI | | | | |
|-----------|----------------|------|------|-------|-------|---------------|------|------|-------|-------|
| | r^2 | MAE | RMSE | NRMSE | p | r^2 | MAE | RMSE | NRMSE | p |
| Chl-a | 0.75 | 0.65 | 0.84 | 38.6 | <0.01 | 0.67 | 0.84 | 1.03 | 47.3 | <0.01 |
| Turbidez | 0.80 | 0.13 | 0.16 | 42.9 | <0.01 | 0.80 | 0.13 | 0.16 | 42.7 | <0.01 |

CONCLUSÕES

O estudo comparou o desempenho dos sensores OLI e MSI quando aplicados à qualidade da água. Conclui-se que tanto o Landsat-8 quanto o Sentinel-2 são capazes de predizer parâmetros de qualidade da água (chl-a e turbidez) com altos índices de confiabilidade. Os pontos seguintes sintetizam as conclusões do estudo:

- O sensor MSI (Sentinel-2) foi superior ao OLI (Landsat-8) para produzir modelos de chl-a e turbidez;
- A resolução espacial mais fina pode ter contribuído em parte para a melhor performance do sensor MSI. Esta conclusão é apoiada por outros autores com estudos semelhantes (Yadav *et al.*, 2019; Govedarica *et al.*, 2019);
- A resolução espectral mais refinada do MSI também contribui para explicar o resultado, bem como o maior número de bandas (*red-edge*) do sensor;
- Ainda assim, ambos os sensores permitiram produzir resultados muito bons com $r^2 \geq 0.75$ para o MSI e $r^2 \geq 0.67$ para o OLI.

Os resultados preliminares aqui alcançados encorajam a continuidade de novos estudos na área pesquisada a fim de aperfeiçoar a precisão dos valores obtidos por meio de análise estatística.

A Importância da Ciência para o Futuro do Rio São Francisco

REFERÊNCIAS

- BABYAK, M. A. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. **Psychosomatic Medicine**, v.66, n.3, p.411–421, 2004.
- BARBOSA, C. C. F., NOVO, E. M. L. M., MARTINS, V. S. **Introdução ao sensoriamento remoto de sistemas aquáticos: princípios e aplicações**. São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2019, 178p.
- DUTRA, V. A. B., TAVARES, P. A., RIBEIRO, H. M. C. Analysis of cloud condition on Sentinel-2 MSI and Landsat-8 OLI images of a public supply lake in Belém-Pará-Brazil. **Ambiente e Água-An Interdisciplinary Journal of Applied Science**, v.14, n.4, p.1–11, 2019.
- GIARDINO, C., BRESCIANI, M., VILLA, P., MARTINELLI, A. Application of remote sensing in water resource management: the case study of Lake Transimeno, Italy. **Water Resour Manage**, v.24, p.3885–3899, 2010.
- GONZÁLEZ-MÁRQUEZ, L. C., TORRES-BEJARANO, F. M., TORREGROZA-ESPINOSA, A. C., HANSEN-RODRIGUEZ, I. R., RODRIGUEZ-GALLEGOS, H. B. Use of Landsat-8 images for depth and water quality assessment of El Guajaro reservoir, Colombia. **Journal of South American Earth Sciences**, v.82, p.231–238, 2018.
- GOVEDARICA, M., JAKOVLJEVIĆ, G. Monitoring spatial and temporal variation of water quality parameters using time series of open multispectral data. **International Conference on Remote Sensing and Geoinformation of the Environment**, v.11174, p.1–11, 2019.
- JAMES, G., WITTEN, D., HASTIE, T., TIBSHIRANI, R. **An introduction to statistical learning**. New York: Springer Science, 2013. 441p.
- KARAOUI, I., ARIOUA, A., BOUDHAR, A., HSSAISOUNE, M., EL MOUATASSIME, S., OUHAMCHICH, K. A., ELHAMDOUNI, D., IDRISI, A. A., NOUAIM, W. Evaluating the potential of Sentinel-2 satellite images for water quality characterization of artificial reservoirs: The Bin El Ouidane Reservoir case study (Morocco). **Meteorology Hydrology and Water Management: Research and Operational Applications**, v.7, n.1, p.31–39, 2019.
- LIU, L., WANG, Y. Modelling Reservoir Turbidity Using Landsat 8 Satellite Imagery by Gene Expression Programming. **Water**, v.11, n.7, p.1479, 2019.
- MAINALI, J., CHANG, H., CHUN, Y. A review of spatial statistical approaches to modeling water quality. **Progress in Physical Geography: Earth and Environment**, v.43, n.6, p.826–841, 2019.
- MATTHEWS, M. W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. **International Journal of Remote Sensing**, v.32, n.21, p.6855–6899, 2011.
- TEODORO, A. C. Optical satellite remote sensing of the coastal zone environment: an overview. **Environment Applications of Remote Sensing**, v.1, p.165–196, 2016.
- VON SPERLING, E. **Morfologia de lagos e represas**. Belo Horizonte: DESA/UFMG, 1999. 137p.
- WATANABE, F., ALCANTARA, E., RODRIGUES, T., ROTTA, L., BERNARDO, N., IMAI, N. Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil). **Anais Da Academia Brasileira de Ciências**, v.90, n.2, p.1987–2000, 2018.
- YADAV, S., YAMASHIKI, Y., SUSAKI, J., YAMASHITA, Y., ISHIKAWA, K. Chlorophyll Estimation of Lake Water and Coastal Water Using Landsat-8 and Sentinel-2A Satellite. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v.42, n.3/W7, p.77–82, 2019.
- YEPEZ, S., LARAQUE, A., MARTINEZ, J., DE SA, J., CARRERA, J. M., CASTELLANOS, B., GALLAY, M., LOPEZ, J. Retrieval of suspended sediment concentrations using Landsat-8 OLI satellite images in the Orinoco River (Venezuela). **Comptes Rendus Geoscience**, v.350, n.1–2, p.20–30, 2018.