

XXIV SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS

SENSORIAMENTO REMOTO E REDES NEURAIS ARTIFICIAIS PARA ESTIMATIVA DA CONCENTRAÇÃO DE SEDIMENTOS EM SUSPENSÃO

*Renata Barão Rossoni¹; Olavo Corrêa Pedrollo², Juliana Andrade Campos³, Leonardo Laipelt⁴
& Fernando Mainardi Fan⁵*

Palavras-Chave – Modelos empíricos, modelagem hidrossedimentológica, transporte de sedimentos, reflectância

RESUMO

A concentração de sedimentos em suspensão (CSS) é uma importante componente da qualidade da água que influencia diretamente os processos biológicos, físicos e químicos nos corpos d'água. Contudo, nem sempre é de fácil obtenção dados observados de CSS. Por isso, técnicas como modelagem e uso de sensoriamento remoto vêm sendo utilizadas para estimativa da CSS. Neste trabalho, utilizamos dados de fácil obtenção (níveis de água e reflectância da banda vermelha e infravermelha) em conjunto com redes neurais artificiais (RNA) para estimar as CSS da bacia do Rio Santo Antônio, localizada na bacia do Rio Doce, em Minas Gerais. Essa bacia possui uma área de 6500 km² e características mais homogêneas, permitindo a aplicação de poucos dados para obtenção das estimativas. Foram encontrados resultados satisfatórios (NSE > 0,85 para período de treinamento e NSE > 0,6 para período de verificação) para as estimativas de CSS. Assim, devido à boa capacidade de ajuste de relações não-lineares das RNA, foi possível encontrar boas estimativas, ainda que o modelo tenha utilizado dados simples para variáveis de entrada.

INTRODUÇÃO

A concentração de sedimentos em suspensão (CSS) é uma importante componente da qualidade da água que influencia diretamente os processos biológicos, físicos e químicos nos corpos d'água. O monitoramento e o conhecimento dessa variável são importantes para a gestão dos recursos hídricos nas bacias hidrográficas, auxiliando em estudos de processos erosivos, em planejamento e operação de reservatórios ou na recuperação de ambientes aquáticos degradados. O monitoramento da CSS geralmente consiste em coletas de amostras em campo, em seções previamente determinadas, com posterior análise laboratorial. Devido à necessidade de elevados recursos humanos e financeiros, a frequência espacial e temporal do monitoramento da CSS é insuficiente na maior parte das bacias

1) Universidade Federal do Rio Grande do Sul: 51 998872363, renata.rossoni@ufrgs.br

2) Universidade Federal do Rio Grande do Sul:

3) Universidade Federal do Rio Grande do Sul:

4) Universidade Federal do Rio Grande do Sul:

5) Universidade Federal do Rio Grande do Sul:

hidrográficas ao redor do mundo (Campos e Pedrollo, 2021; Heng e Suetsugi, 2014; Sari et al., 2017). Consequentemente, diversas pesquisas vêm sendo desenvolvidas para melhorar o conhecimento do transporte de sedimentos nas bacias hidrográficas por meio de técnicas de estimativas indiretas, tais como a modelagem. Existem diversos tipos de modelos matemáticos para estimar a concentração de sedimentos em suspensão. Modelos conceituais são normalmente aplicados em grandes bacias, apresentando bons resultados (da Silva et al., 2016; Fagundes et al., 2019, 2020; Vigiak et al., 2015). Contudo, em bacias menores, esses modelos são limitados e não conseguem representar muito bem a dinâmica dos sedimentos em suspensão.

Modelos físicos são normalmente aplicados em pequenas bacias, e necessitam de uma grande quantidade de dados de entrada e dados para validação do modelo. Em algumas localidades, com poucos dados, a aplicação desses modelos pode gerar resultados inferiores a modelos mais simples, como os modelos empíricos (Sadeghi et al., 2014). Esses modelos são construídos a partir de dados observados e podem apresentar boas estimativas, mesmo que com uma quantidade limitada de dados (Millares et al., 2020; Sadeghi et al., 2014). Este é o caso das redes neurais artificiais (RNA), que são modelos baseados em inteligência artificial e que vem sendo amplamente utilizados no campo dos recursos hídricos devido à sua elevada capacidade de modelar relações não-lineares, mesmo sem física explícita, o que limita a aplicação dos modelos físicos. As RNA, quando aplicadas corretamente, tendem a gerar modelos com estimativas satisfatórias para CSS, tanto para bacias homogêneas quanto heterogêneas (Campos e Pedrollo, 2021), de grande ou pequena escala (Teixeira et al., 2020).

Contudo, mesmo para modelos empíricos em pequena escala, utilizar pelo menos duas variáveis com forte correlação com as CSS podem gerar estimativas mais robustas (Sari et al., 2017). Contudo, nem sempre temos dados observados, como a turbidez, para utilizar nesses modelos. Por isso, utilizar informações de sensoriamento remoto, que possuem correlação com CSS (Ritchie et al., 1987), pode ser uma alternativa para substituir dados observados. Alguns estudos vêm utilizando esse tipo de informação para calibração de modelos conceituais (Fagundes et al., 2019; Rossoni et al., 2021, 2020), para calibração e validação de modelos bidimensionais (Yang et al., 2014), ou ainda para compreensão dos processos hidrossedimentológicos (Fassoni-Andrade e Paiva, 2019; Martinez et al., 2009).

Assim, este trabalho visa utilizar dados de sensoriamento remoto em conjunto com níveis de água observados, para desenvolver um modelo de redes neurais artificiais para estimativa da concentração de sedimentos em suspensão. O uso de modelos de RNA com dados de entrada simples (níveis de água e reflectância na banda vermelha e infravermelha de sensoriamento remoto) é uma alternativa para estimar as concentrações de sedimentos em suspensão em localidades sem dados de

sedimentos. Assim, o modelo apresenta a simplicidade de utilizar RNA, necessitando de apenas três dados de entrada, que podem ser facilmente obtidos.

MATERIAIS E MÉTODOS

Foi utilizada a bacia do Rio Santo Antônio, na bacia do Rio Doce, para o desenvolvimento da nossa pesquisa. A estação fluviométrica Fazenda Ouro Fino, com dados disponibilizados pela CEMIG (Companhia de Energia de Minas Gerais), possui uma área de drenagem de aproximadamente 6500 km² (Figura 1). Foram utilizados dados observados de nível de água e concentração de sedimentos em suspensão. Os dados foram medidos uma vez ao dia no período chuvoso e a cada dez dias no período seco. Escolhemos essa localidade por causa da disponibilidade de dados, ao todo 4771 amostras entre 1984 e 2013. Os dados de sensoriamento remoto utilizados foram a reflectância nas bandas vermelha e infravermelha, por causa da forte correlação entre esses dados com as concentrações de sedimentos em suspensão (Ritchie et al., 1987). O satélite utilizado foi o Landsat 5, que possui dados abrangendo todo período de simulação. Esse satélite possui uma frequência temporal de 16 dias de revisita. Esses dados foram obtidos através da metodologia desenvolvida por (Rossoni et al., 2021, 2020), obtidas através da plataforma em nuvem Google Earth Engine (GEE) (Gorelick et al., 2017), tornando o processamento mais fácil.

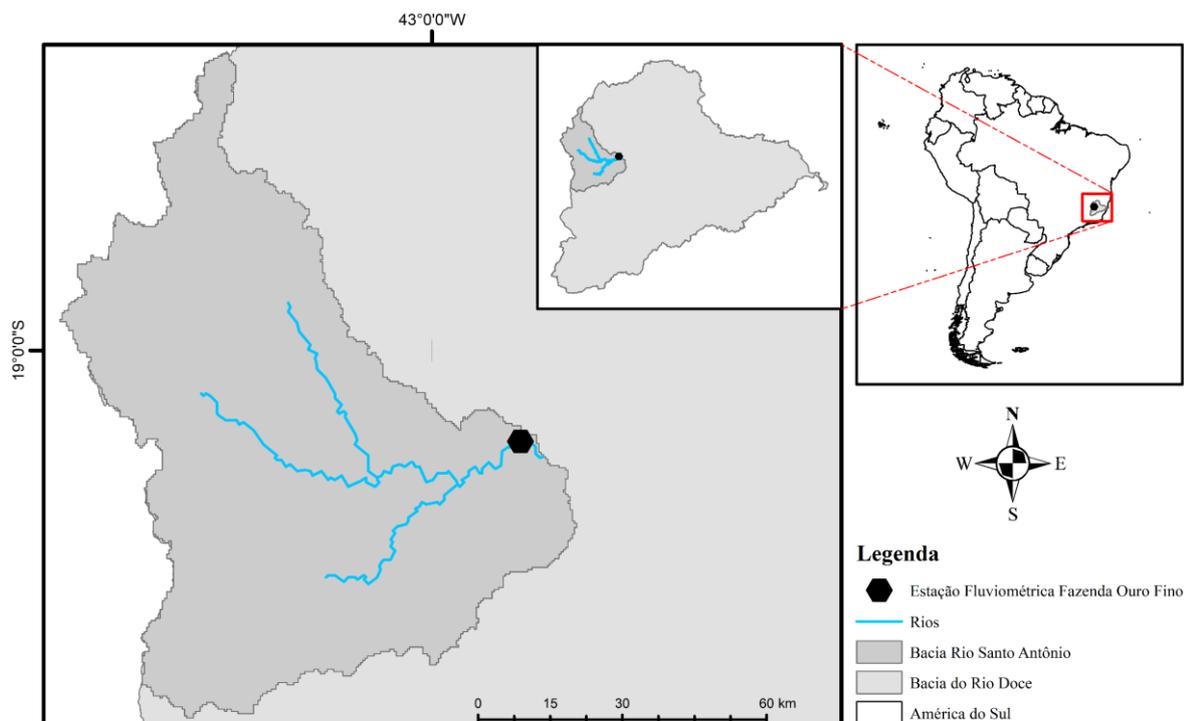


Figura 1. Localização da área de estudo

As variáveis de entrada utilizadas para o desenvolvimento das redes neurais são: níveis de água observados, reflectância da banda vermelha e infravermelha. A correlação linear entre as concentrações de sedimentos em suspensão e níveis de água, reflectância da banda vermelha e infravermelha são 0,71, 0,52 e 0,39, respectivamente. A variável de saída utilizada foi a concentração de sedimentos em suspensão. Escolhemos essas variáveis devido a facilidade de obtenção, que podem ser aplicadas em bacias com poucos dados.

Para a modelagem foram selecionados os dias em que havia dados de reflectância em conjunto com níveis de água e concentração de sedimentos em suspensão. Nos meses secos, nos dias em que havia dados de satélite, mas não havia dados observados, foi feita uma média dos níveis de água e concentração de sedimentos em suspensão dos três dias anteriores e posteriores. Nesses períodos mais secos, o rio possui um comportamento mais constante, não havendo tantas mudanças de descarga sólida e líquida entre os dias. Ao todo, foram utilizados 144 registros na modelagem.

Foram utilizadas RNA do tipo *feedforward* com três camadas. A primeira camada é destinada à inserção dos dados de entrada. Na camada intermediária é onde ocorre a transformação não-linear dos dados de entrada em direção aos dados de saída (neste trabalho nós utilizamos a função de ativação sigmóide unipolar, amplamente aplicada em estudos semelhantes (Campos and Pedrollo, 2021; Sari et al., 2017; Teixeira et al., 2020). Já na terceira camada é destinada aos dados de saída, que neste caso é a concentração de sedimentos suspensos. As camadas da rede neural artificial são interligadas por conexões chamadas de pesos sinápticos, que são ajustados durante o processo de treinamento. O algoritmo retropropagativo foi utilizado para o ajuste dos pesos sinápticos (Rumelhart et al., 1986) e a validação cruzada foi utilizada para evitar o superajustamento da rede (Hecht-Nielsen, 1990). Nós utilizamos 65 amostras para o treinamento (1984 até 1997), 25 amostras para validação (1997 até 2001) e 54 amostras para verificação (2001 até 2013). O número de neurônios utilizados foi 40, o número de ciclos foi 2.000.000 e o número de repetições igual a 50.

Foram feitos três testes de desenvolvimento do modelo com os seguintes dados de entrada: (1) apenas nível de água, (2) apenas reflectância na banda vermelha e infravermelha, e (3) níveis de água, reflectância na banda vermelha e infravermelha. Escolhemos esses testes para verificar a influência de cada uma das variáveis no desenvolvimento das redes. Foram utilizadas as mesmas configurações das RNA para cada um dos modelos, alternando apenas as variáveis de entrada.

Para avaliar o modelo e calcular as métricas de avaliação, comparamos as saídas do modelo de redes neurais com os dados observados. Utilizamos a métrica Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970). A performance foi avaliada de acordo com (Moriassi et al., 2015), sendo os critérios: “muito bom” para $NSE > 0,80$, “bom” para $0,70 \leq NSE \leq 0,80$, “satisfatório” para $0,45 <$

$NSE \leq 0.70$ e “não satisfatório” para $NSE \leq 0.45$. Nós utilizamos esses valores para escala temporal diária, apesar de serem para escala temporal mensal, devido a falta de valores de referência.

RESULTADOS E DISCUSSÕES

Os resultados encontrados são apresentados na Tabela 1. As Figuras 2 e 3 apresentam os resultados para os períodos de treinamento e verificação. O que representou melhor os dados observados foi o experimento realizado no Teste 3, utilizando ambos dados de nível e reflectância na banda vermelha e infravermelha. Esse resultado era esperado, já que um modelo com mais variáveis de entrada tende a apresentar melhores resultados.

Tabela 1. NSE para os períodos de modelagem (treinamento e verificação) para cada um dos testes realizados

Período	Teste 1	Teste 2	Teste 3
Treinamento	0,52	0,89	0,86
Verificação	0,41	-0,06	0,60

O resultado do Teste 3 foi considerado “satisfatório” para o período de verificação (Moriassi et al., 2015). Em Campos e Pedrollo (2021), para uma bacia de grande escala e heterogênea, os autores encontraram o melhor modelo utilizando oito variáveis de entrada. Os valores de NSE foram 0,94 e 0,85 para os períodos de treinamento e verificação, respectivamente. Ao utilizarem apenas os dados de turbidez e vazão, dados similares aos utilizados nesse estudo, os valores de NSE encontrados foram de 0,52 e 0,40, para os períodos de treinamento e verificação, respectivamente. Para um modelo de grande escala, em uma bacia heterogênea, um modelo simples com apenas duas variáveis de entrada apresentou um resultado considerado “não satisfatório” pelos autores. Assim, mais variáveis foram necessárias para representar a heterogeneidade da bacia.

No presente estudo, devido a escala da bacia ser menor, e da maior homogeneidade da bacia, com menos variáveis foi possível representar melhor as concentrações de sedimentos em suspensão. Isso foi observado também em Sari et al. (2017), onde os autores utilizaram 59 registros, em uma bacia menor e mais homogênea. Os autores estimaram concentração de sedimentos em suspensão, utilizando as variáveis nível de água e turbidez e encontraram estimativas muito boas, com NSE acima de 0,90, tanto no treinamento quanto na verificação. Nossos resultados não foram tão bons quanto os apresentados por Sari et al. (2017), uma vez que foi utilizado dados de turbidez (com correlação de 0,91), enquanto utilizamos no presente estudo dados de reflectância (com correlações de 0,52 e 0,39).

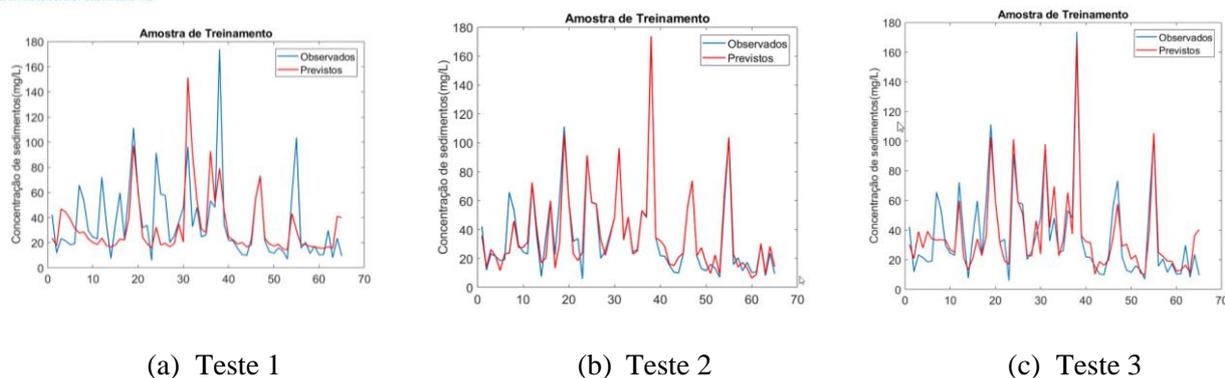


Figura 2. Resultados do período de treinamento para o Teste 1, 2 e 3

Em Fagundes et al. (2019), os autores utilizaram um modelo conceitual de grande escala para estimar a concentração de sedimentos em suspensão na estação fluviométrica Fazenda Ouro Fino. Os autores encontraram um NSE igual a 0,44, após o processo de calibração, sendo que o NSE antes da calibração era igual a -0,44. Isso demonstra o quanto é difícil de encontrar boas métricas de avaliação para modelos de transporte de sedimentos.

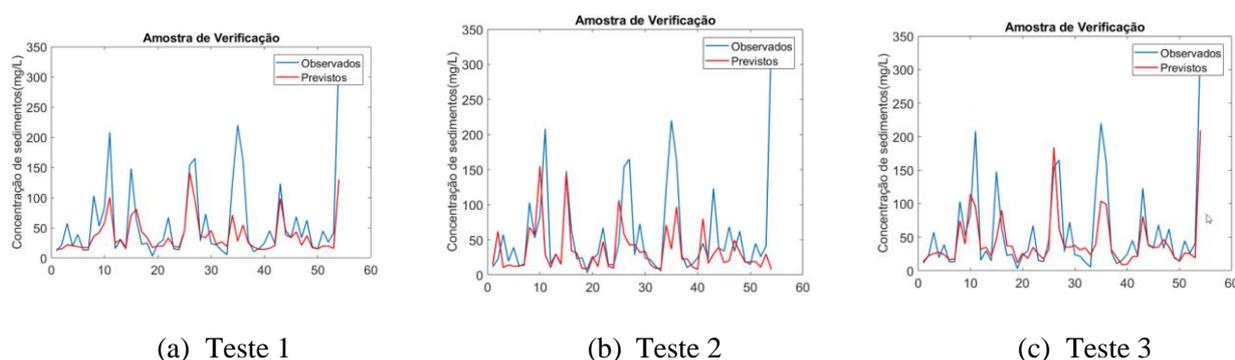


Figura 3. Resultados do período de verificação para o Teste 1, 2 e 3

Uma das limitações do presente estudo é a necessidade de validar esse modelo em outras bacias. Contudo, é uma limitação encontrar localidades que possuam uma grande quantidade de dados observados para validar o modelo. Por este motivo, esse trabalho contribui para um maior entendimento da dinâmica de sedimentos em pequenas bacias, especialmente em localidades com poucos dados, devido a simplicidade dos dados de entrada e aplicação do modelo.

CONCLUSÕES

Nós desenvolvemos um modelo para estimar a concentração de sedimentos em suspensão em pequenas bacias, utilizando redes neurais artificiais. O modelo apresentou resultados satisfatórios, apesar da simplicidade dos dados de entrada (nível de água, reflectância da banda vermelha e infravermelha). Isso foi possível devido à boa capacidade de ajuste de relações não-lineares das RNA, e devido à homogeneidade das características do meio ambiente na bacia, que permitiram estimativas

satisfatórias da CSS apenas a partir de dados de nível d'água e dados de reflectância. Como desenvolvimentos futuros, o trabalho precisa ser testado em outras localidades, para validar a eficiência da técnica aplicada.

AGRADECIMENTOS – Os autores agradecem a CEMIG (Companhia de Energia de Minas Gerais) pela disponibilidade dos dados. Ainda, agradecemos à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela bolsa concedida à primeira e terceira autoras do artigo.

REFERÊNCIAS

- Campos, J.A., Pedrollo, O.C., 2021. A regional ANN-based model to estimate suspended sediment concentrations in ungauged heterogeneous basins. *Hydrological Sciences Journal*. <https://doi.org/10.1080/02626667.2021.1918695>
- da Silva, V.D.P.R., Silva, M.T., de Souza, E.P., 2016. Influence of land use change on sediment yield: A case study of the sub-middle of the São Francisco river basin. *Engenharia Agrícola* 36, 1005–1015. <https://doi.org/10.1590/1809-4430-Eng.Agric.v36n6p1005-1015/2016>
- Fagundes, H. de O., Fan, F.M., de Paiva, R.C.D., 2019. Automatic calibration of a large-scale sediment model using suspended sediment concentration, water quality, and remote sensing data. *Revista Brasileira de Recursos Hídricos* 24. <https://doi.org/10.1590/2318-0331.241920180127>
- Fagundes, H.O., Fan, F.M., Paiva, R.C.D. de, Siqueira, V.A., Buarque, D.C., Kornowski, L.W., Laipelt, L., Collischonn, W., 2020. Sediment flows in South America supported by daily hydrologic-hydrodynamic modeling. *Water Resources Research*. <https://doi.org/10.1002/essoar.10503046.2>
- Fassoni-Andrade, A.C., Paiva, R.C.D. de, 2019. Mapping spatial-temporal sediment dynamics of river-floodplains in the Amazon. *Remote Sensing of Environment* 221, 94–107. <https://doi.org/10.1016/j.rse.2018.10.038>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Hecht-Nielsen, R., 1990. *Neurocomputing*. Addison-Wesley Publishing Company, Massachusetts.
- Heng, S., Suetsugi, T., 2014. Comparison of regionalization approaches in parameterizing sediment rating curve in ungauged catchments for subsequent instantaneous sediment yield prediction. *Journal of Hydrology* 512, 240–253. <https://doi.org/10.1016/j.jhydrol.2014.03.003>
- Martinez, J.M., Guyot, J.L., Filizola, N., Sondag, F., 2009. Increase in suspended sediment discharge of the Amazon River assessed by monitoring network and satellite data. *Catena* 79, 257–264. <https://doi.org/10.1016/j.catena.2009.05.011>
- Millares, A., Chikh, H.A., Habi, M., Morsli, B., Galve, J.P., Perez-Peña, J. v., Martín-Rosales, W., 2020. Seasonal patterns of suspended sediment load and erosion-transport assessment in a Mediterranean basin. *Hydrological Sciences Journal* 65, 969–983. <https://doi.org/10.1080/02626667.2020.1724294>
- Moriasi, D.N., Gitau, M.W., Pai, N., Daggupati, P., 2015. Hydrologic and water quality models: Performance measures and evaluation criteria. *Transactions of the ASABE* 58, 1763–1785. <https://doi.org/10.13031/trans.58.10715>
- Nash, J.E., Sutcliffe, J. v., 1970. River flow forecasting through conceptual models part I - a discussion of principles. *Journal of Hydrology* 10, 282–290.
- Ritchie, J.C., Cooper, C.M., Yongqing, J., 1987. Using Landsat multispectral scanner data to estimate suspended sediments in Moon Lake, Mississippi. *Remote Sensing of Environment* 81, 23–65.
- Rossoni, R., Fan, F., Laipelt, L., 2021. Application of Google Earth Engine (GEE) to obtain calibration data for a large-scale hydrosedimentological model: a test case in South America, in: EGU General Assembly 2021, Online. EGU21, pp. 19–30. <https://doi.org/10.5194/egusphere-egu21-11926>
- Rossoni, R.B., Fernando, J., Fan, M., Laipelt, L., 2020. Aplicação do Google Earth Engine (GEE) para obtenção de dados de calibração de modelos hidrosedimentológicos de grande escala, in: XIV Encontro Nacional de Engenharia de Sedimentos. ABRHidro, Campinas/SP.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323.
- Sadeghi, S.H.R., Gholami, L., Khaledi Darvishan, A., Saeidi, P., 2014. A review of the application of the MUSLE model worldwide. *Hydrological Sciences Journal*. <https://doi.org/10.1080/02626667.2013.866239>

- Sari, V., Castro, N.M. dos R., Pedrollo, O.C., 2017. Estimate of Suspended Sediment Concentration from Monitored Data of Turbidity and Water Level Using Artificial Neural Networks. *Water Resources Management* 31, 4909–4923. <https://doi.org/10.1007/s11269-017-1785-4>
- Teixeira, L.C., Mariani, P.P., Pedrollo, O.C., dos Reis Castro, N.M., Sari, V., 2020. Artificial Neural Network and Fuzzy Inference System Models for Forecasting Suspended Sediment and Turbidity in Basins at Different Scales. *Water Resources Management* 34, 3709–3723. <https://doi.org/10.1007/s11269-020-02647-9>
- Vigiak, O., Malagó, A., Bouraoui, F., Vanmaercke, M., Poesen, J., 2015. Adapting SWAT hillslope erosion model to predict sediment concentrations and yields in large Basins. *Science of the Total Environment* 538, 855–875. <https://doi.org/10.1016/j.scitotenv.2015.08.095>
- Yang, X., Mao, Z., Huang, H., Wang, T., Liu, D., 2014. Numerical simulation of suspended sediment transport merging with satellite derived data in coastal waters, in: *Remote Sensing and Modeling of the Atmosphere, Oceans, and Interactions V*. SPIE, p. 92650E. <https://doi.org/10.1117/12.2068510>