ESTIMATIVA DA CONCENTRAÇÃO DE SÓLIDOS SUSPENSOS EM RIOS SOBRE O GRUPO SERRA GERAL - REGIÃO SUL

Otávio Cristiano Montanher [1], Waterloo Pereira Filho [2]

[1]UEM – Universidade Estadual de Maringá, Av. Colombo, 5790, Maringá, PR, otaviocmontanher@yahoo.com.br;
[2]Universidade Federal de Santa Maria, Av. Roraima, 1000 – Bairro Camobi – Santa Maria – RS, waterloo@ufsm.br;

RESUMO

Este artigo tem como objetivo apresentar resultados de uma modelagem para estimativa da concentração de sólidos em suspensão (CSS) para parte dos rios que ocorrem na Região Sul do Brasil. O modelo foi elaborado para rios sobre afloramentos das formações geológicas do Grupo Serra Geral, utilizou uma abordagem empírica e dados do programa Landsat (5, 7, 8 e 9). Os dados de campo são provenientes da Agência Nacional de Águas, de 16 estações situadas em seis rios. Foram utilizadas 388 amostras de campo com uma concentração máxima de 929 mg/l. O modelo utiliza as bandas do verde, vermelho, infravermelho próximo e infravermelho de ondas curtas, e teve R^2 ajustado de 0,547. Este modelo tem como finalidade dar suporte para uma avaliação espaço-temporal abrangente da CSS na região, como uma variável associada aos processos erosivos em escala de bacia hidrográfica.

***Palavras-chave* —** Sensoriamento Remoto, Hidrologia, Programa Landsat, erosão.

ABSTRACT

This paper aims to present the results of a modelling approach to estimate suspended solids concentration (SSC) for some rivers located in the Southern Region of Brazil. The model was developed for rivers overlying outcrops of the Serra Geral Group geological formations, using an empirical approach and data from the Landsat programme (5, 7, 8, and 9). Field data were obtained from the Brazilian National Water Agency, from 16 stations located across six rivers. A total of 388 field samples were used, with a maximum concentration of 929 mg/l. The model utilises the green, red, near-infrared, and shortwave-infrared bands, yielding the following results: adjusted R^2: 0.547. The purpose of this model is to support a comprehensive spatio-temporal assessment of SSC in the region, serving as a proxy variable for erosive processes at the watershed scale.

***Key words* —** *Remote Sensing, Hydrology, Landsat Programme, erosion.*

1. INTRODUÇÃO

A disponibilização de dados orbitais gratuitos em massa, bem como a possibilidade de processamento de grandes volumes de dados, em plataformas como o *Google Earth Engine,* têm revolucionado diversas áreas do Sensoriamento Remoto e, em específico, a estimativa de séries históricas de parâmetros de qualidade d'água.

Embora preconizado por [1], a estimativa generalizada da concentração de sólidos suspensos (CSS) com o uso de Sensoriamento Remoto permaneceu como um desafio por décadas, mas pesquisas recentes têm demonstrado avanços importantes. Modelos de ampla abrangência espacial tem sido desenvolvidos para estimativa da CSS [2 - 4], bem como de outros parâmetros de qualidade de água [5 - 7].

Após a etapa de modelagem, as equações ajustadas são aplicadas para recuperação de séries históricas de CSS, o que possibilita a análise evolutiva de fatores físicos, ambientais e socioeconômicos que envolvem a erosão do solo nas bacias hidrográficas e as consequentes variações de CSS nos sistemas fluviais [3, 8 - 9].

Este artigo tem como objetivo apresentar resultados de uma modelagem empírica para estimativa da CSS para uma parcela dos rios que ocorrem na Região Sul do Brasil, a partir de dados do programa Landsat. Este modelo tem como finalidade maior dar suporte para uma avaliação espaço-temporal abrangente da CSS nesta região, como uma variável relacionada aos processos erosivos em escala de bacia hidrográfica.

2. MATERIAL E MÉTODOS

A primeira etapa metodológica foi a definição da área de estudo. Buscou-se definir rios importantes da Região Sul que atendessem, em conjunto, quatro parâmetros: i) estivessem em contextos geológicos semelhantes; ii) abrangessem importantes áreas agrícolas; iii) abrangessem trechos onde houve represamentos de rios durante o período de 1984-2024 e iv) possuíssem largura superior a 90 m e não tivessem interferência de fundo.

O primeiro parâmetro refere-se ao fato de que a resposta espectral (magnitude e forma da reflectância) da água para diferentes CSS possui variações conforme o tipo de sólido em suspensão [10 -11]. Assim, uma fonte de erro das estimativas da CSS é a modelagem conjunta de sólidos com diferentes características (composição e tamanho, por exemplo). Dessa forma, adotou-se apenas rios que predominantemente drenam porções da área exposta do Grupo Serra Geral na Região Sul. Esse grupo é composto predominantemente por rochas ígneas extrusivas, principalmente basaltos, que formam solos argilosos e ricos em óxidos de ferro. Claramente, na escala

espacial adotada nesta pesquisa, existem variações internas de litologia das formações geológicas, bem como de mineralogia dos solos, mas trabalhando-se com rios localizados neste contexto geológico buscou-se minimizar tal fonte de variabilidade espectral.

Com relação aos parâmetros ii) e iii) relacionados com a abrangência de áreas agrícolas e reservatórios recentes, estas definições foram adotadas para que posteriormente os modelos desenvolvidos pudessem avaliar os efeitos dessas variáveis sobre os valores de CSS ao longo das séries históricas. O quarto parâmetro refere-se à necessidade de pureza espectral dos pixels que compõem as amostras a serem extraídas das imagens.

Definida a área e os rios de interesse, foi realizada uma busca por dados de estações *in situ* a partir da base de dados da Agência Nacional de Águas (ANA - www.snirh.gov.br/hidroweb). Após a obtenção dos dados de campo, foram extraídas séries históricas de reflectância de superfície e de reflectância aparente (TOA – *top of atmosphere)* com o uso da plataforma GEE [12], para os mesmos locais das estações *in situ*. Foram utilizados dados dos sensores TM (Landsat 5), ETM+ (Landsat 7) e OLI (Landsat 8 e 9). Após a extração dos dados via GEE foi realizada a normalização dos dados em que os valores de reflectância foram compatibilizados entre si, dadas as diferenças que existem principalmente entre os sensores OLI (Landsat 8 e 9) e os sensores mais antigos. Assim, adotou-se os coeficientes expressos em [13] para simulação dos dados OLI em ETM+.

Após a obtenção de séries temporais de reflectância normalizada, foram selecionados pares de dados concomitantes (reflectância – dado *in situ* de CSS), em que o parâmetro máximo de defasagem foi de dois dias [2]. Tendo-se uma base composta por valores de reflectância e CSS de campo, o próximo passo foi a modelagem com o uso de regressão múltipla, com uso de bandas e razões entre bandas [2, 4].

Para avaliação do modelo foram calculados índices gerais como o R^2 ajustado e o RMSE (erro quadrático médio), bem como foram seguidas recomendações de [14] para avaliação da precisão de simulações de modelos hidrológicos. Conforme [14], foram calculados os índices PBIAS (viés percentual), Nash-Sutcliffe efficiency (NSE) e a razão RSR, entre o RMSE e o desvio-padrão dos dados medidos.

3. RESULTADOS E DISCUSSÕES

A busca por dados da ANA resultou em 3477 amostras *in situ*, oriundas de 16 estações distribuídas em seis rios da Região Sul (Figura 1). Observe que os rios e estações ocorrem predominantemnete sobre o Grupo Serra Geral, mas em alguns casos, partes das bacias estão sobre outras formações geológicas. Foi necessário incluir estações nessas condições para aumentar tanto a abrangência espacial quanto a quantidade de dados para modelagem.

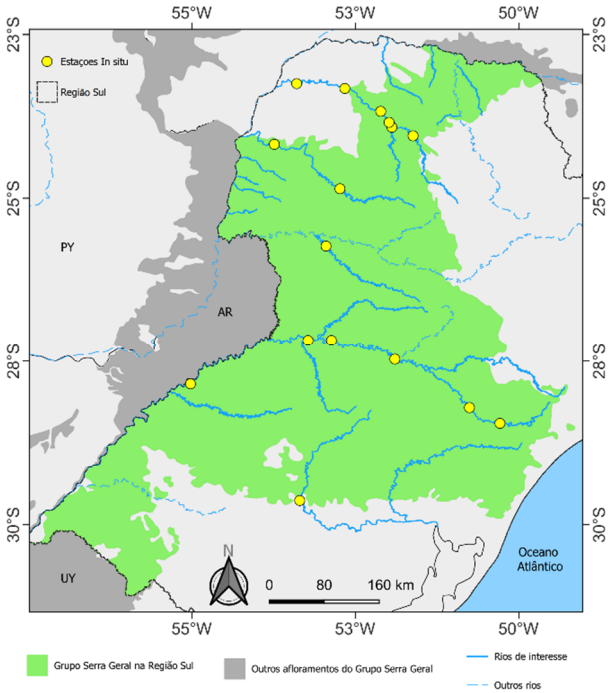


Figura 1. Localização das estações com dados *in situ* da ANA em rios sobre o Grupo Serra Geral na Região Sul.

Desse conjunto de dados *in situ*, 388 amostras foram obtidas em mesma data de passagem de satélite, ou com no máximo dois dias de defasagem, as quais constituíram a base de dados para a modelagem. Os dados *in situ* de CSS possuem as seguintes estatísticas descritivas: mínimo: 1,6, média: 50,87, mediana: 27,95 e máximo: 929 mg/l. É um conjunto assimétrico, em que grande parte dos dados se concentram em valores mais baixos (Figura 2).

Devido aos efeitos de não-linearidade e heterocedasticidade dos resíduos da regressão linear múltipla, a variável dependente, CSS, foi transformada pelo logaritmo natural. Neste artigo são apresentados apenas os resultados obtidos com uso da reflectância TOA, por ser a unidade cujo modelo apresentou maior precisão para estimativas.

A dispersão entre os valores observados e estimados está apresentada na Figura 3, e as variáveis selecionadas na Tabela 1. Basicamente, o modelo utiliza as bandas do verde, vermelho, infravermelho próximo e infravermelho de ondas curtas, além da razão entre as bandas do vermelho e verde. O modelo de regressão apresentou valor-p (estatística F) < 0,0000001 (significativo), um R^2 ajustado de 0,547, um erro médio absoluto de 29,17 mg/l, e o NRMSE (erro médio quadrático normalizado) de 80,1%. Também se ressalta que a distribuição dos resíduos da regressão foi avaliada, sendo próxima de uma normal (teste de Kolmogorov-Smirnov).

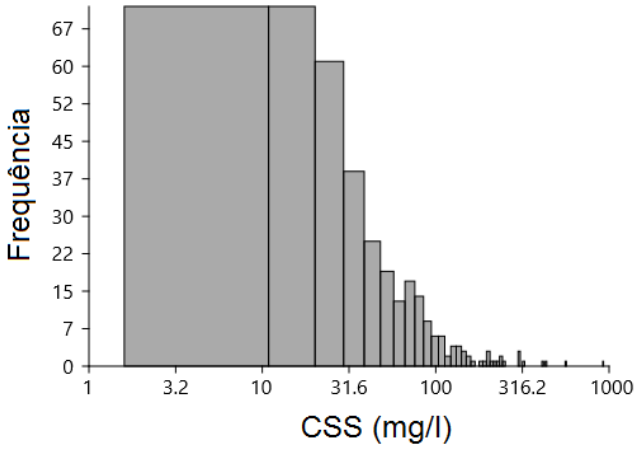


Figura 2. Histograma dos dados *in situ* de concentração de sólidos suspensos utilizados na modelagem (eixo X em logaritmo natural).

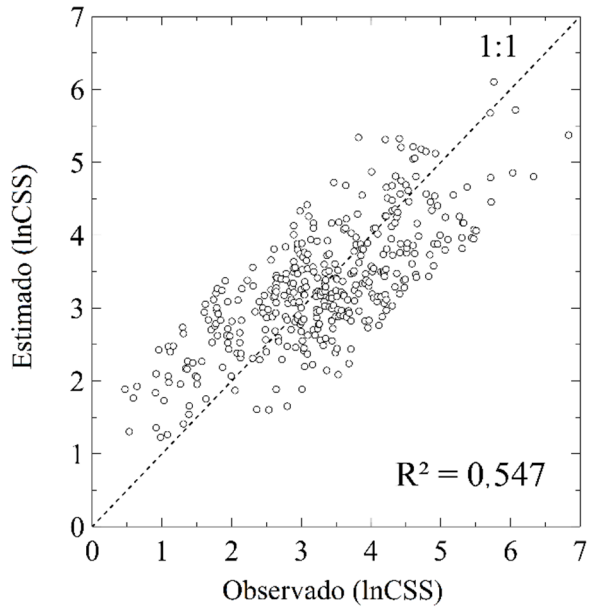


Figura 3. Localização das estações com dados *in situ* da ANA em rios sobre o Grupo Serra Geral na Região Sul.

Variável	Coeficiente	Valor-p (teste t)
Intercepto	-5,1334	0,000002
Verde	57,2745	< 0,0000001
Vermelho	-64,2005	
Infravermelho próximo	22,1677	
Infravermelho de ondas curtas	-21,3738	
Vermelho / verde	8,2787	

Tabela 1. Variáveis e coeficientes da regressão múltipla para estimativa da CSS.

Conforme as recomendações de [14] para avaliação da precisão de modelos hidrológicos, obteve-se um PBIAS de -21,7%, que indica uma tendência média subestimada. Isso mostra que o modelo tem dificuldades para estimar altos valores de CSS, o que tem relação com o comportamento físico da reflectância com o aumento da CSS na água, cuja relação tende a uma assíntota horizontal [10, 11]. Esse resultado é comumente registrado na bibliografia em modelos que buscam estimar altos valores de CSS (por exemplo CSS > 100 mg/l) [2, 4]. Mesmo com a tendência média de subestimativa, esta estatística está abaixo do limiar recomendado por [14], de ±55% para séries de CSS.

Com relação ao coeficiente NSE, obteve-se um valor de 0,36, abaixo do recomendado, de no mínimo 0,5 [14]. No entanto, o coeficiente rNSE (NSE relativo) foi de 0,55, acima do limiar mínimo. O rNSE é uma variação do coeficiente NSE tradicional que minimiza a influência de *outliers*. Assim, entende-se que a análise da aderência entre os valores observados e simulados do modelo proposto está sendo bastante influenciada pela presença de *outliers*. Por fim, a razão RSR observada foi de 0,8, acima do limiar máximo recomendado, de 0,7 [14]. Assim, o modelo apresentado atende parcialmente as recomendações sugeridas por [14] para a simulação de dados hidrológicos.

A análise em conjunto de todas as estatísticas obtidas mostrou que, embora o modelo tenha alguns resultados satisfatórios, outros não indicam uma precisão tão alta quanto registrado em modelos similares na bibliografia [2, 4, 15]. Por exemplo, [15] mencionam valores de R^2 ajustado de 0,82 e NRMSE de 13,86% para a estimativa da CSS no Rio Araguaia (resultados para o modelo apenas com variáveis radiométricas). No entanto, há que se considerar que o modelo descrito em [15] lidou com valores de CSS entre 4 e 217 mg/l, enquanto o modelo descrito aqui abrange uma amplitude maior, entre 1,6 e 929 mg/l.

Comparando-se os resultados do modelo descrito neste artigo com os de [4] que trabalham com amplitudes silimares de CSS, observou-se que este modelo teve resultados inferiores em termos de precisão. Este resultado pode ser justificado por uma série de fatores, como pela heterogeneidade interna dos tipos de solos da área de estudo, mesmo que tenha sido adotado o critério de se utilizar estações situadas no mesmo contexto geológico. As variações dos tipos de sólidos [10, 11] e as variações nas concentrações de materiais orgânicos dissolvidos [2, 3] entre as estações de coleta e entre as épocas do ano podem explicar parte dos erros.

Adicionalmente, a área de estudo também possui heterogeneidade de condições climáticas e de uso e cobertura do solo que podem influenciar os tipos de sólidos em suspensão em diferentes bacias hidrográficas. Algumas das estações *in situ*, por exemplo, abrangem extensas áreas de agricultura mecanizada, que podem fornecer maior proporção de sedimentos inorgânicos, enquanto outras regiões possuem maiores taxas de cobertura natural, como florestas e

formações campestres, as quais podem fornecer maior proporção de sólidos orgânicos.

4. CONCLUSÕES

Apresentamos um modelo empírico para a estimativa da CSS para alguns rios da Região Sul do Brasil, especificamente aqueles localizados principalmente sobre o Grupo Serra Geral. O modelo é simples, tendo utilizado bandas e uma razão entre bandas, e gerou resultados parcialmente satisfatórios, quando comparado com outros modelos e critérios recomendados pela bibliografia.

Embora tenha pontos fracos em termos de precisão, a partir da análise em conjunto de todas as estatísticas, concluímos que o modelo tem resultados aceitáveis para sua finalidade, de extração de séries históricas de CSS, em um contexto em que não existem outras fontes de dados.

5. AGRADECIMENTOS

O primeiro autor agradece ao CNPq pela concessão de bolsa de Pós-doutorado Sênior, edital 32/2023, processo n. 171664/2023-0 e o segundo autor pelo auxílio do CNPq processo 311977/2022-7.

6. REFERÊNCIAS

[1] R. J. Holyer. Towards universal suspended sediment algorithms. *Remote Sensing of Environment*, v. 7, p. 323 – 338, 1978.

[2] E. N. Dethier, C. E. Renshaw, and F. J. Magilligan. Toward improved accuracy of remote sensing approaches for quantifying suspended sediment: Implications for suspended-sediment monitoring. *Journal of Geophysical Research: Earth Surface*. v. 125, n. 7, 2020.

[3] E. N. Dethier, C. E. Renshaw, and F. J. Magilligan. Rapid changes to global river suspended sediment flux by humans. *Science*, v. 376, n. 6600, p. 1447–1452, 2022.

[4] O. C. Montanher, E. M. L. M. Novo, C. C. F. Barbosa, C. D. Renno, and T. S. F. Silva. Empirical models for estimating the suspended sediment concentration in Amazonian white water rivers using Landsat 5/TM. *International Journal of Applied Earth Observation*, v. 29, p. 67–77, 2014.

[5] D. A. Maciel, N. Pahlevan, C. C. F. Barbosa, V. S. Martins, B. Smith, R. E. O'Shea, S. V. Balasubramanian, A. M. Saranathan, E. and M. L. M. Novo. Towards global long-term water transparency products from the Landsat archive. *Remote Sensing of Environment*, v. 299, 2023.

[6] J. R. Gardner, X. Yang, S. N. Topp, M. R. V. Ross, E. H. Altenau, and T. M. Pavelsky. The color of rivers. *Geophysical Research letters. v. 48, p. 1-12, 2021*.

[7] P. Ramtel, D. Feng, and J. Gardner, J. Toward large-scale riverine phosphorus estimation using remote sensing and machine learning. *Journal of Geophysical Research: Biogeosciences*, v. *129*, 2024.

[8] F. L. Lobo, M. P. F. Costa, and E. M. L. M. Novo. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. *Remote Sensing of Environment*, v. 157, p. 170-184. 2015.

[9] O. C. Montanher, E. M. L. M. Novo, and E. E. Souza filho. Temporal trend of the suspended sediment transport of the Amazon River (1984–2016). *Hydrological Sciences Journal***.** v. 63, p. 1901-1912. 2018.

[10] E. M. L. M. Novo, J. D. Hanson, and P. J. Curran. The effect of sediment type on the relationship between reflectance and supended sediment concentration. *International Journal of Remote Sensing*, v. 10, n.7, p. 1283-1289, 1989.

[11] M. A. Lodhi, D. C. Rundquist, L. Han, and M. S. Kuzila. The Potential for Remote Sensing of Loess Soils Suspended in Surface Waters. *Journal of the American Water Resources Association*, v. 33, n. 1, p. 111 – 117, 1997.

[12] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, v. 202, p. 18-27, 2017.

[13] D. P. Roy, V. Kovalskyy, H. K. Zhang, E. F. Vermote, L. Yan, S. S. Kumar, and A. Egorov. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sensing of Environment*, v. 185, p. 57-70, 2016.

[14] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, v. 50, n. 3, p. 885-900, 2007.

[15] E. H. Cremon, A. M. S. Silva, and O. C. Montanher. Estimating the suspended sediment concentration from TM/Landsat-5 images for the Araguaia River – Brazil. *Remote Sensing Letters*, v. 11, n. 1, p. 47-56, 2020.