



IMPORTÂNCIA DA MORFOLOGIA NA DETERMINAÇÃO DO REGIME DE FLUXO EM NASCENTES

RESUMO

Determinar a localização das nascentes e seu regime de fluxo é essencial para a proteção dos recursos hídricos, garantia da qualidade da água, do ciclo hidrológico e preservação de processos biológicos associados. O regime de fluxo das nascentes (perenes, intermitente ou efêmero) resulta da complexa interação entre diferentes fatores exógenos e endógenos, naturais ou antrópicos, que atuam na resistência à erosão, permeabilidade, capacidade de retenção de água e controle dos fluxos sub e superficiais. A legislação brasileira define APPs (Áreas de Proteção Permanente) como áreas destinadas à proteção dos recursos hídricos e, neste sentido, a revegetação e a preservação da vegetação são, em geral, os fatores apontados como mais favoráveis a proteção e recuperação de nascentes. Contudo, características morfológicas também atuam no regime de vazão. A morfologia, pouco associada a formação das nascentes, atua na captação, concentração do fluxo superficial, controle da velocidade do escoamento e no favorecimento ou restrição dos processos de infiltração e recarga. Conhecer a importância da morfologia no regime e capacidade de geração de fluxo em cabeceiras de drenagem é fundamental para promover e gerir ações de preservação e recuperação das nascentes. Esta pesquisa objetiva quantificar a importância da morfologia na determinação do regime de fluxo das nascentes (perenes vs intermitentes), para tanto emprega dados de campo de 164 bacias de drenagem (com e sem nascentes perenes) distribuídas pelos quatro macro compartimentos geomorfológicos do estado do Paraná e técnicas de aprendizado de máquina (RFR - Random Forest Regressor) com objetivo de quantificar a importância da morfologia, frente a variáveis climáticas e índices de vegetação na determinação do regime de fluxo de nascentes em bacias de primeira ordem. A análise da importância foi feita no RFR a partir do uso do valor de Gini o qual avalia a importância das variáveis a partir da redução da impureza (impurity reduction) no processo de classificação das nascentes. Dois tipos de variáveis morfométricas foram testadas, o RUNOFF e o ICR (Índice de Concentração da Rugosidade), junto a dados climáticos e de vegetação disponibilizados pelo projeto ERA-Land (European Centre for Medium-Range Weather Forecasts). A pesquisa demonstrou que a morfologia regional (ICR) e o RUNOFF como variáveis que apresentam importância superior à precipitação e a vegetação na determinação do tipo de fluxo observado nas nascentes pesquisadas (perene, intermitente). Os resultados encontrados acrescentam dados relevantes para medidas de proteção e recuperação de nascentes e apresentam novas informações no âmbito das pesquisas sobre geomorfologia fluvial e sobre o papel da morfologia na dinâmica de encostas.

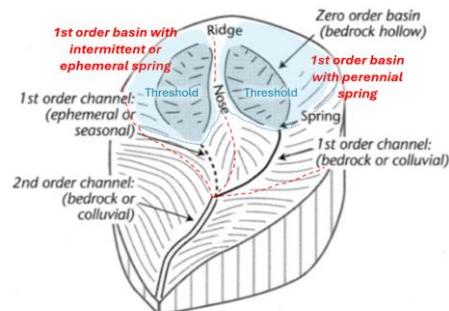
Palavras-chave: Morfologia, Nascentes, Aprendizado de Máquina, Random Forest.

Introdução

O regime de fluxo em bacias de drenagem de primeira ordem resulta da complexa interação entre diferentes elementos e mecanismos que atuam na resistência à erosão (Garrett; Wohl, 2017), permeabilidade (Nag; Chakraborty, 2003; Rushton, 2003; Soulsby et al., 2007), capacidade de retenção de água (Stanislawski et al., 2018) e aceleração/desaceleração dos fluxos sub e superficiais (Sangireddy et al., 2016a).

Nesse sentido, a morfologia atua na captação, concentração do fluxo superficial, controle da velocidade do escoamento (Jaeger; Montgomery; Bolton, 2007) e no favorecimento ou restrição dos processos de infiltração e recarga (Dunne, 1980).

Parte essencial no processo de geração de fluxo, a cabeceira de drenagem corresponde a área na qual o escoamento superficial é concentrado e direcionado (Benda et al., 2005a) (Marcus, 1980) até que dê início ao canal de drenagem (*headchannel*) e as nascentes. Essa área de captação é também denominada de “área limite” (LAMBERT, 1975) (Chen et al., 2018; Wu et al., 2017) (Figura 1).



Anatomy of headwater drainage basins (from Benda et al 2005).

Figura 1. Bacias de zero, primeira e segunda ordem (hollow ou área limite) adaptado de (Benda et al., 2005b).

Apesar de interagir com a definição do regime de fluxo, a “área limite” não atua de forma isolada e padronizada. Ariza-Villaverde et al (Ariza-Villaverde; Jiménez-Hornero; Gutiérrez de Ravé, 2013) apontam para a necessidade de uso de diferentes valores de área limite para modelar a ocorrência de nascentes e representação da rede de drenagem. Isso porque, fatores como clima e o uso e cobertura da terra também interferem a partir de diferentes processos como a precipitação, temperatura e evapotranspiração (Sangireddy et al., 2016b; Tucker; Slingerland, 2010) .

A exemplo do observado no mapeamento de processos complexos que demandam o uso de um grande número de dados (Kavzoglu Taskinand Colkesen, 2019; Khaledian; Miller, 2020), a identificação do regime de fluxo pode ser favorecido pelo uso de algoritmos de aprendizado de máquina (Machine Learning – ML) (Kulkarni; Lowe, 2016). Isto porque, os algoritmos de aprendizado possibilitam a identificação de padrões e correlações em conjuntos de dados complexos, o que pode permitir uma compreensão mais acurada dos fenômenos relacionados à formação de nascentes (Biau; Scornet, 2016).

Esta pesquisa avalia o potencial e a relevância dos dados morfológicos, climáticos e de uso e cobertura, associados a técnicas de aprendizado de máquina (Random Forest Regressor - RFR), para classificação do regime de fluxo das nascentes em bacias de primeira ordem (perenes e intermitentes). Para tanto, utiliza dados climáticos fornecidos pelo ERA-Land, dados morfológicos extraídos de modelos digitais de elevação e dados de campo sobre o regime de fluxo de nascentes em bacias de primeira ordem no estado do Paraná (Brasil).

Materiais e Métodos

Área de estudo:

O Paraná localiza-se na região sul do Brasil e conforme classificação Climática de Köppen, apresenta dois tipos climáticos: Cfa, subtropical com temperaturas superiores a 22°C no verão e mais de 30mm de chuva no mês mais seco e Cfb, temperado com verão ameno e chuvas uniformemente distribuídas sem estação seca e temperatura média inferior a 22°C (Aparecido et al., 2016).

No âmbito geomorfológico o Paraná pode ser dividido em diferentes unidades geomorfológicas, das quais destaca-se: Planície Litorânea, Serra do Mar, Primeiro e Segundo e Terceiro Planaltos (Santos et al., 2006). (Figura 2)

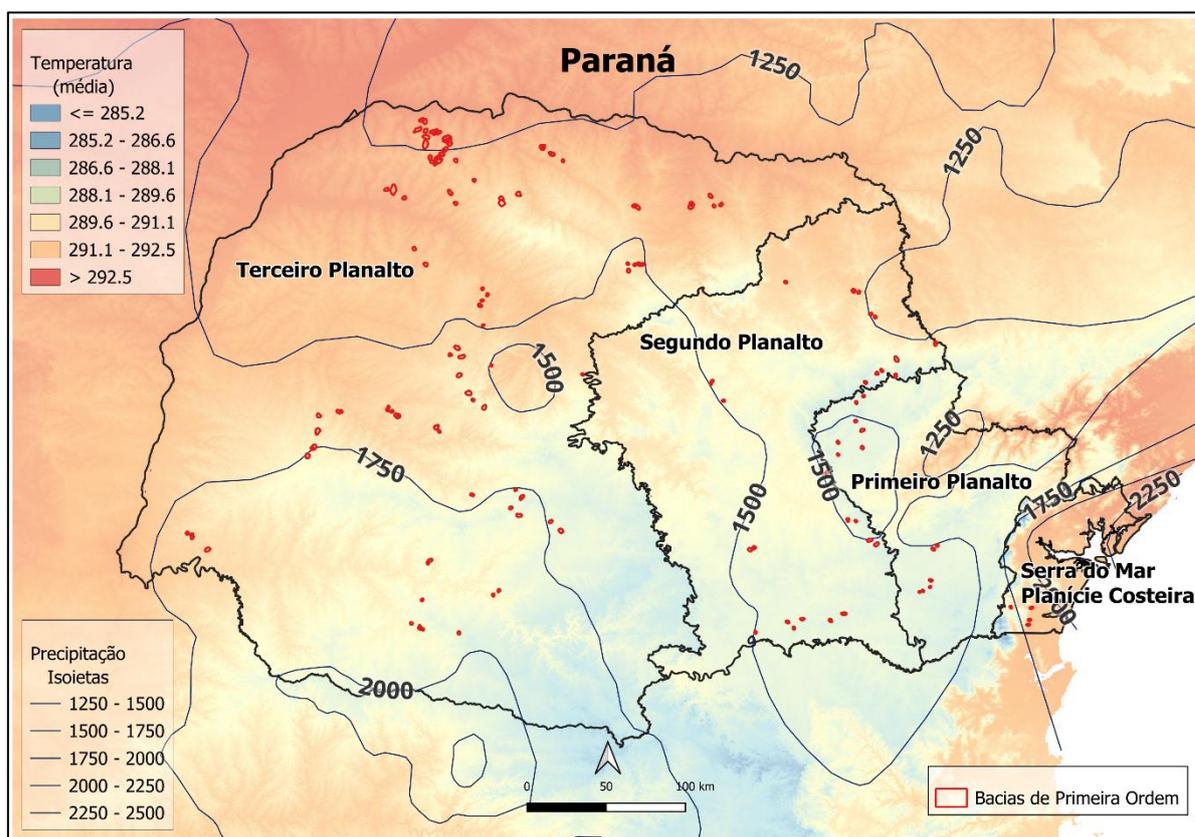


Figure 2. Área de estudo e bacias de primeira ordem amostradas.

Bacias amostrais

Foram selecionadas 164 bacias de drenagem distribuídas pelas maiores unidades geomorfológicas do Paraná (Figura 2). O regime de fluxo das nascentes (perene ou intermitente) presentes nestas bacias foi verificado diretamente em campo. Das bacias selecionadas 84 bacias apresentam regime de vazão perene (tipo 1) e 80 regime de vazão intermitente (tipo 0).



Variáveis utilizadas

As bacias foram delimitadas manualmente com apoio de modelos digitais de elevação e interpretação de curvas de nível e as áreas calculadas utilizando a Projeção Equivalente de Albers, para evitar distorções promovidas por sistemas projetados, como a exemplo o UTM (Universo Transverso de Mercator).

Para caracterização do relevo foi utilizado o ICR (Índice de Concentração da Rugosidade) (Sampaio; Augustin, 2014). O ICR representa para cada pixel, a média dos valores de declividade (calculados em porcentagem) observados em aproximadamente 1km². Para o cálculo do ICR foi utilizado o modelo ALOS Palsar com resolução espacial de 30m.

Os dados climáticos e de uso e cobertura utilizados são do ERA5-Land, disponibilizados pelo Programa Copernicus (European Centre for Medium-Range Weather Forecasts – ECMWF). O ERA5-Land combina dados de modelos, observações globais e leis físicas. Utiliza variáveis atmosféricas do ERA5 para controlar os campos simulados e correção para as diferenças de altitude. O ERA-Land fornece 69 temas, disponibilizados na forma de 150 variáveis.

A descrição completa e detalhada das variáveis fornecidas pelo ERA-Land pode ser obtida em (https://developers.google.com/earth-engine/datasets/catalog/ECMWF_ERA5_LAND_HOURLY#bands).

Os dados para a pesquisa foram obtidos via plataforma Google Earth Engine. Foi gerada uma imagem com 150 bandas (variáveis), com valor de média para o intervalo de tempo entre 01/01/1950 e 31/12/2023 (73 anos).

Os valores dos dados do ERA5-Land e do ICR foram adicionados ao arquivo contendo o identificador de cada bacia (ID), a área e o regime de fluxo (0 – intermitente ou 1 - perene) das bacias de drenagem (camada de dados vetorial) por meio do QGIS com uso do plugin desenvolvido pela Dymaxion Labs: Zonal statistics (multiband).

Procedimentos metodológicos (Testes)

Como a área de captação é o principal parâmetro empregado no mapeamento por extração automática (Roelens et al., 2018; Rueda; Noguera; Martínez-Cruz, 2013) e localmente, bacias de drenagem maiores tendem a apresentar nascentes perenes enquanto bacias menores, tendem a apresentar fluxo intermitente ou efêmero, a seleção de bacias com fluxo perene e intermitente foi feito de modo a evitar o enviesamento da pesquisa a partir da área.

Neste sentido, o primeiro procedimento metodológico objetiva avaliar se as bacias de drenagem escolhidas para a pesquisa (tipo 1 e tipo 0), apresentam dimensões similares. Para tanto será aplicado o teste T para as áreas das bacias.

O script utilizado para o Random Forest foi escrito em Python e executado no ambiente do Google colab. Foi utilizado o módulo `sklearn.ensemble`, componente da biblioteca `scikit-learn`. O módulo possibilita o uso do Random Forest Regressor, bem como a análise da importância das variáveis testadas.

Para cada teste, o nível de importância das variáveis foi avaliado via índice de Gini e, as variáveis não utilizadas, bem como as de menor importância eram descartadas. A redução do número de variáveis se deu por processos sucessivos enquanto a acurácia do processo de classificação se manteve constante.



Resultados

O Teste t

O resultado do teste t (alfa = 0.05) apresentou valor de $p = 0.102$ (maior que 0.05), o que indica que a dimensão média das bacias de drenagem utilizadas na pesquisa com nascentes perenes é similar a das bacias com drenagem intermitente. Isso reforça a importância das demais variáveis na análise, uma vez que as bacias de drenagem com drenagem são similares (em relação a dimensão).

A similaridade na dimensão das bacias de drenagem, já havia sido apontada por (Roelens et al., 2018; Rueda; Noguera; Martínez-Cruz, 2013) como complicador ao processo de mapeamento por extração automática e por (Lopez García; Camarasa, 1999) como dificultador ao processo de mapeamento por fotointerpretação.

Random Forest

A segunda análise consistiu em avaliar o desempenho do Random Forest e, para tanto, foram realizadas 200 simulações com uso do K-Fold (com split = 10), ou seja, para cada simulação o grupo amostral (164 bacias) foi dividido em 10 partes, sendo 03 para teste e 7 para treinamento.

Foram utilizadas todas as variáveis (152), nomeadas de b0 a b151, sendo: b0 o valor da área das bacias, b1 a b150 variáveis climáticas disponibilizadas pelo ERA-Land e b151 o valor de ICR (Índice de Concentração da Rugosidade).

O resultado indicou capacidade de identificação do regime de fluxo das nascentes superior ao observado por Sampaio e Rocha (Sampaio; Rocha, 2022) para o estado do Paraná, no qual o índice de concordância total (TCI) foi de 0,528. Neste caso, a técnica de mapeamento empregada foi a de fotointerpretação.

Das variáveis (152) utilizadas, a análise da importância das variáveis indicou o RUNOFF e o ICR (variáveis morfológicas) como variáveis mais relevantes, seguida das variáveis precipitação, índice de área foliar (vegetação alta), evapotranspiração, volume de água no solo e fluxo de calor sensível (superficial).

Como as bacias pesquisadas apresentam dimensões similares, a variável área (b0) resultou na variável de menor importância, dentre as variáveis empregadas.

Tabela 1. Importância média das 8 variáveis mais importantes.

Variável	Nome	Importância média
b40	RUNOFF	0.281284
b151	ICR	0.257864
b48	Precipitação total	0.123506
b32	Fluxo de calor sensível (superficial)	0.088625
b49	Índice de área foliar (vegetação alta)	0.081891
b38	Evapotranspiração	0.076181
b24	Volume de água no solo	0.054425
b0	Área	0.036225

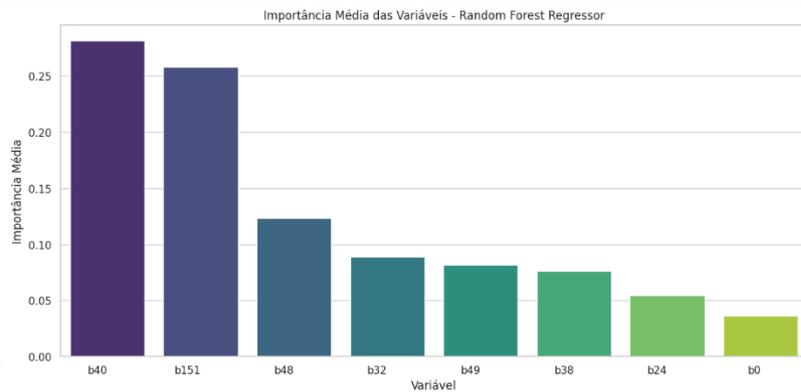


Figura 3. Importância média das 8 variáveis mais relevantes, segundo RFR.

Importante observar que o RFR demonstrou desempenho médio na identificação do regime de fluxo acima de 77%, com desvio padrão entre 75,9% e 78,3% (Figura 4). Este resultado indica que tanto o RFR quanto as variáveis utilizadas apresentam potencial de uso em demais estudos que relacionam o regime de fluxo em bacias de primeira ordem com outras variáveis preditivas.

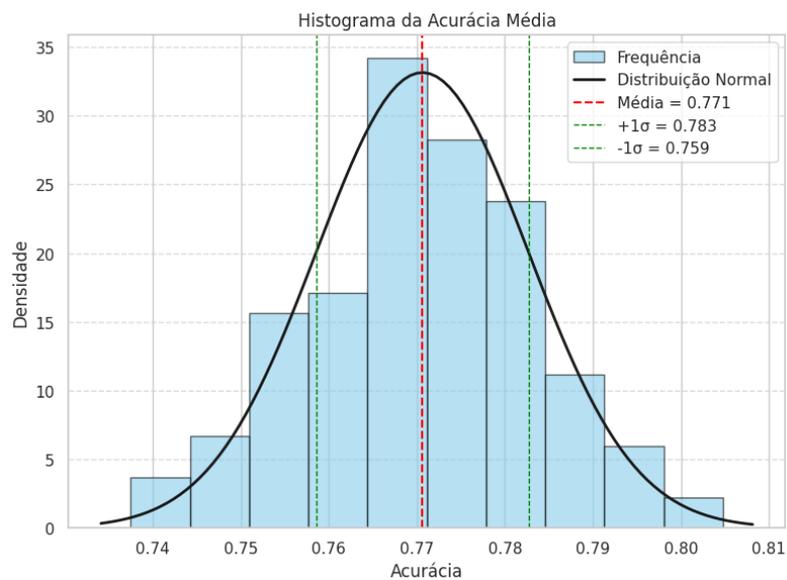


Figura 4. Histograma do desempenho médio do RFR.



Conclusão

O RFR demonstrou potencial de uso para estudos que envolvem a caracterização do regime de fluxo em bacias de primeira ordem, com desempenho superior ao observado em processos de fotointerpretação e em processos de classificação automatizada, via algoritmos de extração da rede de drenagem.

As variáveis morfológicas demonstraram maior importância na determinação do regime de fluxo em bacias de primeira ordem via uso do RFR, sendo os valores observados superiores ao da precipitação e evapotranspiração, usualmente mais associados ao regime de fluxo e, superiores à variáveis como o índice de área foliar.

Referências:

- APARECIDO, Lucas Eduardo de Oliveira *et al.* Köppen, Thornthwaite and Camargo climate classifications for climatic zoning in the State of Paraná, Brazil. **Ciencia e Agrotecnologia**, v. 40, n. 4, p. 405–417, 1 jul. 2016.
- ARIZA-VILLAVERDE, a. B.; JIMÉNEZ-HORNERO, F. J.; GUTIÉRREZ DE RAVÉ, E. Multifractal analysis applied to the study of the accuracy of DEM-based stream derivation. **Geomorphology**, v. 197, p. 85–95, set. 2013.
- BENDA, Lee *et al.* Geomorphology of steepland headwaters: The transition from hillslopes to channels. **Journal of the American Water Resources Association**, v. 41, n. 4, p. 835–851, 2005a.
- BENDA, Lee *et al.* Geomorphology of steepland headwaters: The transition from hillslopes to channels. **Journal of the American Water Resources Association**, v. 41, n. 4, p. 835–851, 2005b.
- BIAU, Gérard; SCORNET, Erwan. A random forest guided tour. **Test**, v. 25, n. 2, p. 197–227, 1 jun. 2016.
- CHEN, Liya *et al.* Threshold Selection of River Network Extraction Based on Different DEM Scales Using ATRIC Algorithm. *In*: IOP Publishing, 2018.
- DUNNE, Thomas. Formation and controls of channel networks. **Progress in Physical Geography**, v. 4, n. 2, p. 211–239, 1980.



GARRETT, Krista K.; WOHL, Ellen E. **Climate-invariant area-slope relations in channel heads initiated by surface runoff.** *Earth Surface Processes and Landforms*, 2017.

JAEGER, Kristin L.; MONTGOMERY, David R.; BOLTON, Susan M. Channel and perennial flow initiation in headwater streams: Management implications of variability in source-area size. **Environmental Management**, v. 40, n. 5, p. 775–786, 2007.

KAVZOGLU TASKIN AND COLKESEN, Ismail and Sahin Emrehan Kutlug. Machine Learning Techniques in Landslide Susceptibility Mapping: A Survey and a Case Study. *In: PRADHAN S.P. AND VISHAL, V. and Singh T. N. (Org.). Landslides: Theory, Practice and Modelling.* Cham: Springer International Publishing, 2019. p. 283–301.

KHALEDIAN, Yones; MILLER, Bradley A. Selecting appropriate machine learning methods for digital soil mapping. **Applied Mathematical Modelling**, v. 81, p. 401–418, 2020.

KULKARNI, Arun D.; LOWE, Barrett. Random Forest Algorithm for Land Cover Classification. **International Journal on Recent and Innovation Trends in Computing and Communication**, v. 4, n. 3, p. 58–63, 2016.

LOPEZ GARCÍA, Maria Jose; CAMARASA, Ana M. Use of geomorphological units to improve drainage network extraction from a DEM. **International Journal of Applied Earth Observation and Geoinformation**, v. 1999, n. 3–4, p. 187–195, 1999.

MARCUS, Alan. First-Order Drainage Basin Morphology -Definition and Distribution. **Earth Surface Processes & Landforms**, v. 5, p. 389–398, 1980.

NAG, S. K.; CHAKRABORTY, Surajit. Influence of rock types and structures in the development of drainage network in hard rock area. **Journal of the Indian Society of Remote Sensing**, v. 31, n. 1, p. 25–35, 2003.

ROELENS, Jennifer *et al.* Extracting drainage networks and their connectivity using LiDAR data. **Hydrological Processes**, v. 32, n. 8, 2018.

RUEDA, Antonio; NOGUERA, José M.; MARTÍNEZ-CRUZ, Carmen. A flooding algorithm for extracting drainage networks from unprocessed digital elevation models. **Computers & Geosciences**, v. 59, p. 116–123, set. 2013.



RUSHTON, K. R. **Groundwater hydrology: Conceptual and Computational Models**. [S.l.: S.n.].

SAMPAIO, Tony Vinicius Moreira; AUGUSTIN, Cristina Helena Ribeiro Rocha. Índice de concentração da rugosidade: uma nova proposta metodológica para o mapeamento e quantificação da dissecação do relevo como subsídio a cartografia geomorfológica. **Revista Brasileira de Geomorfologia**, v. 15, n. 1, p. 1–14, 2014.

SAMPAIO, Tony Vinicius Moreira; ROCHA, Jorge. On the quality of the drainage network cartographic representation. **Ecological Indicators**, v. 143, p. 109350, 2022.

SANGIREDDY, Harish *et al.* Controls of climate, topography, vegetation, and lithology on drainage density extracted from high resolution topography data. **Journal of Hydrology**, v. 537, 2016a.

SANGIREDDY, Harish *et al.* Controls of climate, topography, vegetation, and lithology on drainage density extracted from high resolution topography data. **Journal of Hydrology**, v. 537, p. 271–282, 2016b.

SANTOS, Leonardo José Cordeiro *et al.* MAPEAMENTO GEOMORFOLÓGICO DO ESTADO DO PARANÁ. **Revista Brasileira de Geomorfologia**, v. 7, n. 2, p. 3–12, 2006.

SOULSBY, C. *et al.* Inferring groundwater influences on surface water in montane catchments from hydrochemical surveys of springs and streamwaters. **Journal of Hydrology**, v. 333, n. 2, p. 199–213, 2007.

STANISLAWSKI, Lawrence V. *et al.* An open source high-performance solution to extract surface water drainage networks from diverse terrain conditions. **Cartography and Geographic Information Science**, v. 45, n. 4, p. 319–328, 2018.

TUCKER, Gregory E.; SLINGERLAND, Rudy. Drainage basin responses to climate change. **Water Resources Research**, v. 33, n. 8, p. 2031–2047, 2010.

WU, Miao *et al.* Impacts of DEM resolution and area threshold value uncertainty on the drainage network derived using SWAT. **Water SA**, v. 43, n. 3, p. 450–462, 2017.