A 5 m dataset of digital terrain model derivatives across mainland France

Léa Panhelleux, Sébastien Rapinel, Blandine Lemercier, Guillaume Gayet, Laurence Hubert-Moy

Data in Brief 49 (2023) 109369

Contents lists available at ScienceDirect

Data in Brief

journal homepage: www.elsevier.com/locate/dib

Data Article

A 5 m dataset of digital terrain model derivatives across mainland France

Léa Panhelleux[a], Sébastien Rapinel[a], Blandine Lemercier[b], Guillaume Gayet[c], Laurence Hubert-Moy[a,*]

[a] LETG UMR 6554, University of Rennes - CNRS, place du recteur Henri Le Moal, Rennes, 35000, France
[b] SAS UMR 1069, Institut Agro Rennes-Angers - INRAE, 65 rue de Saint-Brieuc, Rennes, 35000, France
[c] PatriNat OFB-MNHN-CNRS-IRD, 57 rue Cuvier, Paris, 75231, France

A R T I C L E   I N F O

Article history:
Received 5 May 2023
Revised 26 June 2023
Accepted 28 June 2023
Available online 4 July 2023

Dataset link: A 5 m vertical distance to channel network index (VDCNI) across France (Original data)
Dataset link: A 5 m multi-scale topographic position color composite (MTPCC) across France (Original data)
Dataset link: A 5 m topographic wetness index (TWI) across France (Original data)

Keywords:
DTM
Wetland
Environmental modeling
Hydrogeomorphology
Soil moisture
Big data

A B S T R A C T

A dataset of three digital terrain model (DTM) derivatives was produced at 5 m spatial resolution across mainland France. This dataset includes (i) a topographic wetness index (TWI) that characterizes potential soil wetness as a function of the contributing area and local slope, (ii) a multi-scale topographic position color composite (MTPCC) that describes the position of a pixel relative to its neighborhood at three spatial scales, and (iii) a vertical distance to channel network index (VDCNI) that expresses the vertical height between the elevation of a pixel and the nearest channel. These three raster layers were derived from the French national airborne DTM at 5 m spatial resolution and the vector layer of the channel network of the national hydrological database. This unprecedented fine-scale dataset opens new insights for geomorphological analysis. It can be used for several purposes, such as environmental modeling, risk assessment, or water-resource management.

DOI of original article: 10.1016/j.heliyon.2023.e13482
* Corresponding author.
E-mail address: laurence.moy@univ-rennes2.fr (L. Hubert-Moy).

https://doi.org/10.1016/j.dib.2023.109369

Specifications Table

Subject	Water Science and Technology
Specific subject area	High-resolution DTM-based derivatives for hydrogeomorphological analysis across France
Type of data	Image
How the data were acquired	The dataset was derived from the national airborne DTM (RGE ALTI® DTM, version 2.0) at 5 m spatial resolution and the vector layer at 1:5,000 scale of the channel network of the national hydrological database (BD TOPAGE®, version 1). Geospatial analyses were performed using R software [1] and its terra [2], Rsagacmd [3], and whitebox [4] packages.
Data format	Analyzed (GIS raster layers in GeoTIFF format)
Description of data collection	The 5 m dataset includes: - a topographic wetness index (TWI), calculated using the FD8 algorithm with a 7 × 7 median filter; - a multi-scale topographic position color composite (MTPCC), calculated using the maximum deviation from mean elevation at micro- (5-300 m), meso-(300-1,000 m), and macro- (1,000-5,000 m) scales; - a vertical distance to channel network index (VDCNI), generated by subtracting the channel network base level from the DTM. Chunk-based processing was used to address computing resources, and a buffer was used to prevent edge effects.
Data source location	- Country: Whole mainland France, including a 500 m buffer zone beyond the national borders, and the foreshore up to the edges of topographic acquisitions. - Latitude and longitude: Upper left corner: 51°05'20"N and 5°09'00"W Lower right corner: 41°20'02"N and 9°15'43"E.
Data accessibility	TWI is publicly provided per 50 × 50 km raster tile as follows: Repository name: Zenodo Data identification number: 10.5281/zenodo.7797365 Direct URL to data:https://doi.org/10.5281/zenodo.7797365 MTPCC is publicly provided per 50 × 50 km raster tile as follows: Repository name: Zenodo Data identification number: 10.5281/zenodo.7798631 Direct URL to data: https://doi.org/10.5281/zenodo.7798631 VDCNI is publicly provided per 50 × 50 km raster tile as follows: Repository name: Zenodo Data identification number: 10.5281/zenodo.7795503 Direct URL to data: https://doi.org/10.5281/zenodo.7795503 The DTM primary data (RGE ALTI®, version 2.0) are released in raster format by French department as follows: Repository name: Geoservices (IGN) Data identification number: IGNF_RGEALTIr_2-0 Direct URL to data: https://geoservices.ign.fr/rgealti The channel network primary data (BD TOPAGE®, version 1) are released as one vector file for mainland France as follows: Repository name: SANDRE Data identification number: http://www.sandre.eaufrance.fr/?urn=urn:sandre:donnees:CoursEau:FXX:::ressource:2019:topage:1:html Direct URL to data: https://www.sandre.eaufrance.fr/atlas/srv/fre/catalog.search#/metadata/3b3d3c56-d9b6-4625-a57e-ba054e798274
Related research article	S. Rapinel, L. Panhelleux, G. Gayet, R. Vanacker, B. Lemercier, B. Laroche, F. Chambaud, A. Guelmami, L. Hubert-Moy, National wetland mapping using remote-sensing-derived environmental variables, archive field data, and artificial intelligence, Heliyon. 9 (2023) e13482. https://doi.org/10.1016/j.heliyon.2023.e13482

Value of the Data

- These three DTM derivatives have a very high spatial resolution (5 m) across France.
- These raster layers are ready to use in GIS and avoid time-consuming processing.
- The dataset could be useful for scientists in many fields of research (e.g. environmental sciences, hydrology, geomorphology, pedology, archeology, risk assessment), as well as for stakeholders and engineering agencies working on land-use planning, conservation, or water-resource management.

- The three layers can be used as predictive variables in spatial modeling and as a visual aid for field collection.

1. Objective

The three topographic layers were initially produced to model the spatial distribution of wetlands at 5 m resolution across mainland France, which is unprecedented at the national scale, for which coarser-resolution (30-50 m) variables are usually used [5]. TWI, MTPCC, and VDCNI were selected since they are relevant for characterizing slope, riverine, depressional, coastal, and lacustrine-fringe wetlands [6–8]. The objective of this data paper is to describe the method used to produce these layers, which are publicly available for other users and uses.

2. Data Description

The dataset consists of three DTM derivatives, produced at 5 m spatial resolution in GIS raster format across mainland France (upper left corner: 51°05'20"N and 5°09'00"W // lower right corner: 41°20'02"N and 9°15'43"E), which are illustrated in Fig. 1 and described in Table 1. They were derived from the national airborne DTM at 5 m spatial resolution (RGE ALTI®, version 2.0) and the vector layer of the channel network of the national hydrological database at 1:5,000 scale (BD TOPAGE®, version 1). The first DTM derivative is the topographic wetness index (TWI), which characterizes potential soil wetness as a function of the contributing area and local slope [9]. The index ranges from 0-30, with a larger value indicating a higher probability of wet soil.

The second DTM derivative is the multi-scale topographic position color composite (MTPCC), which uses 3 bands to describe the position of a pixel relative to its neighborhood at micro- (5-300 m, band 3), meso- (300-1,000 m, band 2), and macro- (> 1,000 m, band 1) scales. As a result, default viewing parameters for bands 1-3 in the widely used freeware QGIS (red, green,

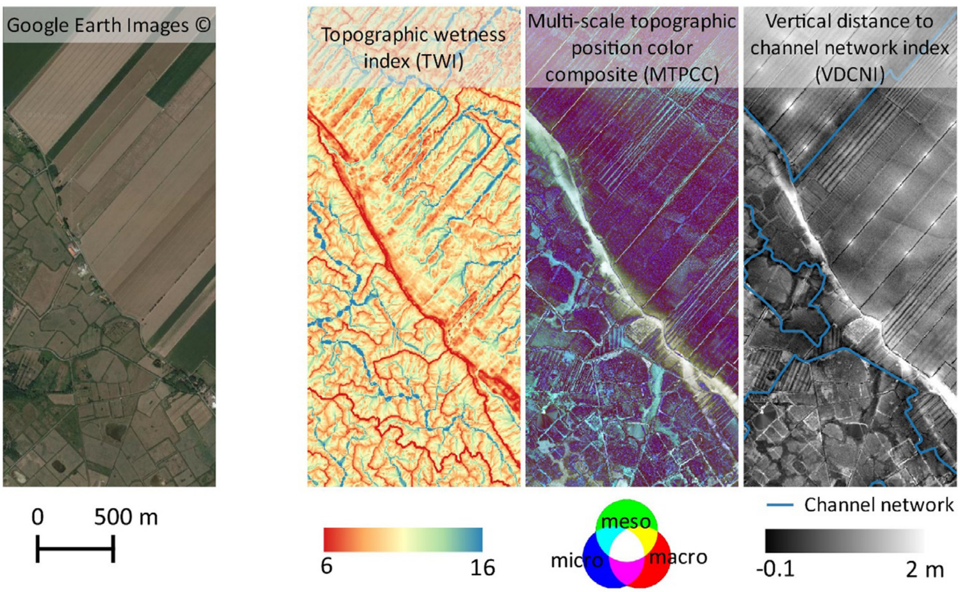

Fig. 1. The dataset that includes the three digital terrain model derivatives produced at 5 m spatial resolution. This subset focuses on wet grasslands and crops along the Gironde Estuary (45.47°N, 1.03°W).

Table 1
Description of the three 5 m digital terrain model derivatives included in the dataset. The ranges of values are the 1st–99th percentiles.

Layer name	Acronym	Data format	Unit	Range of values	Band name
Topographic wetness index	TWI	32-bit float	Unitless	4.1–16.5	Band 1: TWI
Multi-scale topographic position color composite	MTPCC	8-bit integer	Unitless	0–221	Band 1: Macro-scale
				0–199	Band 2: Meso-scale
				0–189	Band 3: Micro-scale
Vertical distance to channel net work index	VDCNI	32-bit float	Meter	-0.1–450.5	Band 1: VDCNI

and blue, respectively) [10] generate a 3-color composite (Fig. 1). The largest values for a given scale indicate the highest topographic positions.

The third DTM derivative is the vertical distance to channel network index (VDCNI), which expresses the vertical height (in m) between the elevation of a pixel and that of the nearest channel. The elevation of the channel represents that of the water's surface on the date of airborne acquisition. The largest values indicate the largest difference in elevation from the channel. Negative and positive values indicate that a pixel is located below or above the nearest channel, respectively.

The three DTM derivatives are publicity released in 280 tiles of 50 × 50 km in GeoTIFF raster format and projected in the French Lambert-93 system (EPSG code 2154), which is the original coordinate system of the DTM data.

Each of the three DTM derivatives is available in a specific Zenodo repository that includes:

- 280 GeoTIFF raster files (<XXX_000>.tif), with each file corresponding to a 50 × 50 km tile. The letters "XXX" correspond to the acronym of the DTM derivative, and the number indicates the tile of interest
- 1 vector tile index in Google Earth format (tile_index.kmz) showing the location of each tile. This file was created to facilitate downloading a layer only for an area of interest. All three DTM derivatives have the same tiling.

To decrease storage space and download time, each raster file was packed in 7-Zip freeware format (.7z or .tar.gz) [11].

3. Experimental Design, Materials and Methods

The materials and methods are summarized in Fig. 2. All geospatial analyses were performed using R software [1] and its terra [2], Rsagacmd [3], and whitebox [4] packages.

The 5 m French airborne DTM raster (RGE ALTI®, IGN) and channel network vector (BD TOPAGE®, SANDRE) used as primary data were downloaded in 2020. The vertical accuracy of the DTM ranges from 30-70 cm [12], depending on the type of land use and the acquisition technology. For water bodies, the DTM considers the elevation of the water's surface on the date of airborne acquisition. This DTM has a buffer zone of 500 m beyond the national borders and partially covers the foreshore up to the edges of topographic acquisitions. The DTM raw data were converted from ASCII to GeoTiff format and merged into a virtual raster image across France. In the areas not covered by this 5 m airborne DTM product (9% of mainland France), the 25 m contour DTM product was resampled at 5 m resolution using cubic interpolation. The channel network vector layer (BD TOPAGE®, SANDRE) was produced at 1:5,000 scale by visual interpretation of aerial photographs and verified using field surveys. It includes most of rivers, large ditches, and canals. It was released in 2019 in ESRI vector shapefile format (.shp). Most of artificial canals were removed from the vector layer because they do not follow natural flow

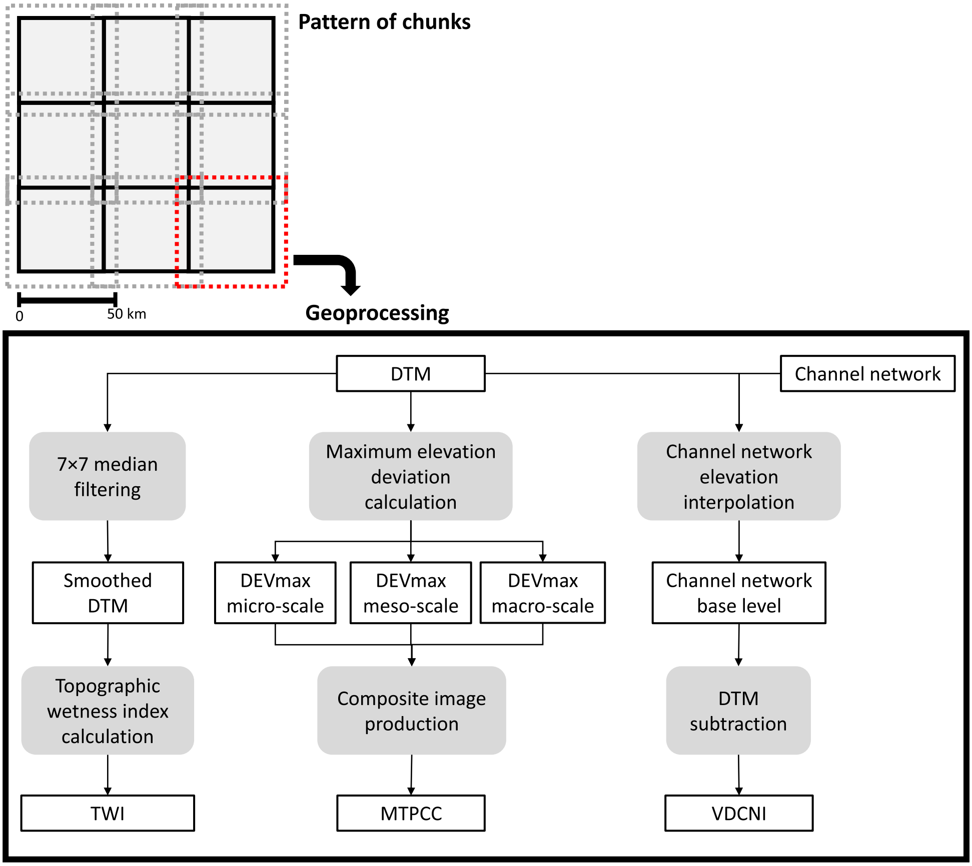

Fig. 2. Flowchart of the method used to produce the three digital terrain model (DTM) derivatives at 5 m resolution across mainland France: (top) the pattern of chunks (black lines) and their buffer zones (gray dotted lines), which were iteratively geoprocessed (bottom) to derive the topographic wetness index (TWI), the multi-scale topographic position color composite (MTPCC), and the vertical distance to channel network index (VDCNI).

paths and are usually parallel to rivers. The channel network was then rasterized in GeoTIFF format at 5 m spatial resolution according to the DTM grid.

Given the huge size of the DTM raster grid over France (228,640 × 212,167 pixels), it was iteratively geoprocessed in individual chunks. Each chunk was buffered to consider the elevation pattern beyond its actual area. At the end of geoprocessing, all chunks were merged to produce a seamless raster. This approach managed computing resources, parallelized the geoprocessing, and prevented edge effects.

TWI was derived from a smoothed DTM to eliminate non-natural linear patterns that could influence water flow, such as transport networks or ditches. For this, a 7 × 7 median filter was applied to the original DTM using the *focal* function of the terra package [2]. TWI was then calculated using the SAGA GIS *Topographic Wetness Index (One Step)* tool chain [13,14] called in R software using the Rsagacmd package [3]. Specifically, it calculated TWI using the standard Beven and Kirkby equation [9] in a single command line that specified the DTM and selected flow-accumulation algorithm. The Freeman FD8 multiple-flow accumulation algorithm [15] was selected since it is the most appropriate for predicting soil moisture in natural environments [16]. All other parameters were set to default values [16]. TWI was calculated per 30 × 30 km

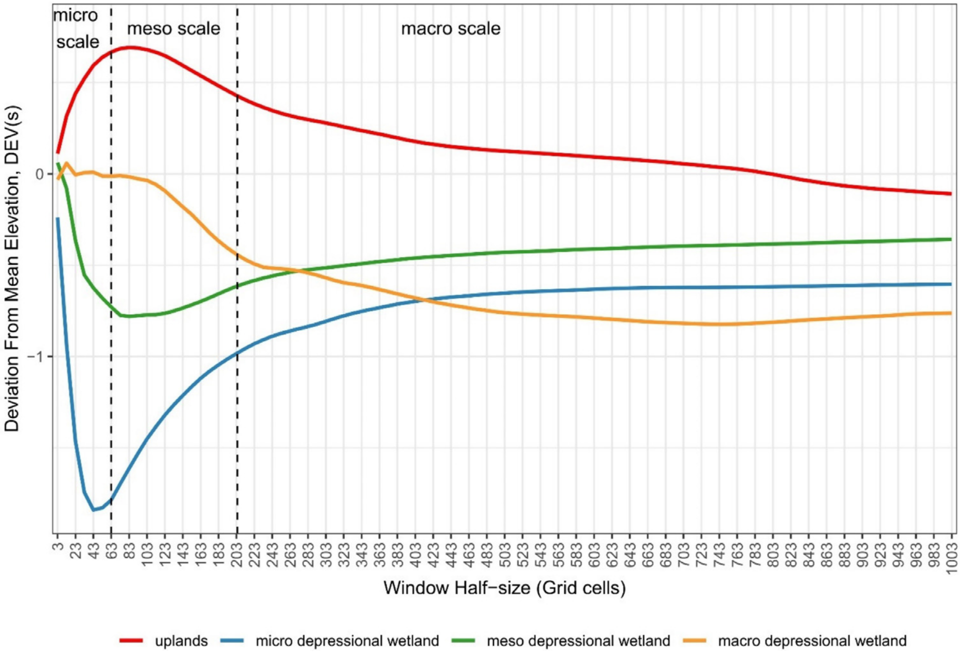

Fig. 3. Scale signature of deviation from mean elevation (DEV) for uplands and depressional wetlands at three spatial scales (i.e. micro-, meso-, and macro-; dashed horizontal lines) used to calculate DEVmax rasters. DEV ranges from −2 to 2, with negative and positive values indicating positions lower or higher than the nearest channel, respectively.

chunk with a 5 km buffer. The value of overlapping pixels equaled the maximum of the two pixels.

MTPCC was based on integral image transformation [4]. The first step entailed calculating the maximum deviation from mean elevation at micro- (5-300 m, with a 50 m step-size), meso- (300-1,000 m, with a 50 m step-size), and macro- (1,000-5,000 m, with a 200 m step-size) scales using the *wbt_max_elevation_deviation* function of the whitebox package [4]. The distance interval for each of the three scales was set using the *wbt_max_elev_dev_signature* function, which displays a graph of scale signatures of maximum deviation from mean elevation, for 92 depressional wetlands, whose area ranged from a few m^2 to several thousand m^2 (Fig. 3). MTPCC was then generated from the three DEVmax rasters at micro-, meso-, and macro-scales using the *wbt_multiscale_topographic_position_image* function. MTPCC was calculated per 100 × 100 km chunk with a 20 km buffer.

VDCNI was calculated using the SAGA GIS *Vertical Distance to Channel Network* tool [13] called in R software using the Rsagacmd package [3]. For this, the DTM raster and the channel network raster were combined. The elevations of pixels located on a channel were interpolated over the raster extent to generate a channel network base level, which was then subtracted from the DTM to generate the VDCNI raster, in which each pixel expresses the elevation relative to the nearest channel. VDCNI was calculated per 40 × 40 km chunk with a 10 km buffer. The value of overlapping pixels equaled the minimum of the two pixels.

Ethics Statements

Ethics statements do not apply to this dataset since it does not involve the use of human subjects, animal subjects, or social media resources.

CRediT Author Statement

Léa Panhelleux: Formal analysis, Investigation, Software; **Sébastien Rapinel:** Methodology, Supervision, Writing - Original Draft; **Blandine Lemercier**: Conceptualization, Validation; **Guillaume Gayet**: Conceptualization, Validation; **Laurence Hubert-Moy**: Conceptualization, Methodology, Supervision, Project administration, Writing - Original Draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

A 5 m vertical distance to channel network index (VDCNI) across France (Original data) (zenodo).

A 5 m multi-scale topographic position color composite (MTPCC) across France (Original data) (zenodo).

A 5 m topographic wetness index (TWI) across France (Original data) (zenodo).

Acknowledgments

The authors thank François Chambaud (water agency Rhone-Méditerranée-Corse), Joanie Catrin, Sébastien Abric, Ghislaine Ferrere (French Ministry of Ecology), Anis Guelmami (Tour du Valat), Rachel Vanacker (MNHN), and Bertrand Laroche (INRAE) for their productive discussions. The authors also thank Jean Nabucet (CNRS) for maintaining the IT equipment and Alexandre Guyot (Hytech-imaging) for his help with computer programming. This study was supported by the French Ministry of Ecology (grant no. 2103207978).

References

[1] R. Core TeamA Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2012 URL https://www.R-Project.org (2019).
[2] R.J. Hijmans, R. Bivand, J. van Etten, K. Forner, J. Ooms, E. Pebesma, Package ‘terra’, 2022.
[3] S. Pawley, Rsagacmd: Linking R with the Open-Source “SAGA-GIS” Software, (2022). https://CRAN.R-project.org/package=Rsagacmd (Accessed 31 March 2023).
[4] J.B. Lindsay, GAT Whitebox, A case study in geomorphometric analysis, Comput. Geosci. 95 (2016) 75–84, doi:10.1016/j.cageo.2016.07.003.
[5] S. Rapinel, L. Panhelleux, G. Gayet, R. Vanacker, B. Lemercier, B. Laroche, F. Chambaud, A. Guelmami, L. Hubert-Moy, National wetland mapping using remote-sensing-derived environmental variables, archive field data, and artificial intelligence, Heliyon 9 (2023) e13482, doi:10.1016/j.heliyon.2023.e13482.
[6] J.W. Riley, D.L. Calhoun, W.J. Barichivich, S.C. Walls, Identifying small depressional wetlands and using a topographic position index to infer hydroperiod regimes for pond-breeding amphibians, Wetlands 37 (2017) 325–338.
[7] S.M. Chignell, M.W. Luizza, S. Skach, N.E. Young, P.H. Evangelista, An integrative modeling approach to mapping wetlands and riparian areas in a heterogeneous Rocky Mountain watershed, Remote Sens. Ecol. Conserv. 4 (2018) 150–165, doi:10.1002/rse2.63.
[8] S. Rapinel, B. Clément, S. Dufour, L. Hubert-Moy, Fine-Scale Monitoring of Long-term Wetland Loss Using LiDAR Data and Historical Aerial Photographs: The Example of the Couesnon Floodplain, France, Wetlands. 38 (2018) 423–435.
[9] K. Beven, M. Kirkby, A physically based, variable contributing area model of basin hydrology, Hydrol. Sciences Bull. 24 (1979) 43–69.
[10] J.M. Flenniken, S. Stuglik, B.V. Iannone, Quantum GIS (QGIS): An introduction to a free alternative to more costly GIS platforms: FOR359/FR428, 2/2020, EDIS. 2020 (2020) 7–7.
[11] I. Pavlov, 7-Zip, 2022 https://7-zip.org/.
[12] IGNRGE ALTI V2 - Descriptif de contenu, IGN, Saint Mandé, 2018 https://geoservices.ign.fr/ressources_documentaires/Espace_documentaire/MODELES_3D/RGE_ALTI/DC_RGEALTI_2-0.pdf.

[13] O. Conrad, B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, J. Böhner, System for automated geoscientific analyses (SAGA) v. 2.1. 4, Geosci. Model Devel. Discuss. 8 (2015).
[14] P. Mattivi, F. Franci, A. Lambertini, G. Bitelli, TWI computation: a comparison of different open source GISs, Open Geospatial Data, Softw. Stand. 4 (2019) 1–12.
[15] T.G. Freeman, Calculating catchment area with divergent flow based on a regular grid, Comput. Geosci. 17 (1991) 413–422, doi:10.1016/0098-3004(91)90048-I.
[16] M. Kopecký, Š. Čížková, Using topographic wetness index in vegetation ecology: does the algorithm matter? Appl. Veget. Sci. 13 (2010) 450–459.