Article

Scale Optimization in Topographic and Hydrographic Feature Mapping Using Fractal Analysis

Christos Karydas ^{1,*} and Bin Jiang ²

¹ Ecodevelopment S.A., Filyro P.O. Box 2420, 57010 Thessaloniki, Greece

² Division of GIScience, Faculty of Engineering and Sustainable Development, University of Gävle, SE-801 76 Gävle, Sweden; bin.jiang@hig.se

* Correspondence: karydas@ecodev.gr

Received: 26 August 2020; Accepted: 23 October 2020; Published: 26 October 2020

Abstract: A new method for selecting optimal scales when mapping topographic or hydrographic features is introduced. The method employs rank-size partition of heavy-tailed distributions to detect nodes of rescaling invariance in the underlying hierarchy of the dataset. These nodes, known as head/tail breaks, can be used to indicate optimal scales. Then, the Fractal Net Evolution Assessment (FNEA) segmentation algorithm is applied with the topographic or hydrographic surfaces to produce optimally scaled objects. A topological transformation allows linking the two processes (partition and segmentation), while fractal dimension of the rescaling process is employed as an optimality metric. The new method is experimented with the two biggest river basins in Greece, namely Pinios and Acheloos river basins, using a digital elevation model as the only input dataset. The method proved successful in identifying a set of optimal scales for mapping elevation, slope, and flow accumulation. Deviation from the ideal conditions for implementing head/tail breaks are discussed. Implementation of the method requires an object-based analysis program and few common geospatial functions embedded in most GIS programs. The new method will assist in revealing underlying environmental processes in a variety of earth science fields and, thus, assist in land management decision-making and mapping generalization.

Keywords: head/tail breaks; fractal dimension; FNEA; Pinios; Acheloos

1. Introduction

Most landscape features or processes, either natural or anthropogenic (e.g., a forest, a rail network, an urban area, a flood, or erosion), are scale-dependent [1]. Scale dependency is due to the fact that every feature or process has its own functional scale, i.e., a specific spatiotemporal interval within which it integrates or smooths to provide a message. The functional scale, in turn, will determine an appropriate mapping scale, i.e., the spatiotemporal unit required to visualize these properties for map readers.

Landscapes are composed of loosely coupled levels which function at distinct spatial and temporal scales and interact as nodes of an inherent hierarchy. In order to interpret possible interactions between these levels, it is necessary to decompose them into their fundamental parts [2].

Scale dependency and necessity of a multi-scale hierarchical approach has been indicated as an important issue in hydrological research for a long time now. Selection of an appropriate scale when mapping topographic or hydrographic features has been one of the trickiest problems, especially in land use and management projects [3].

The major concern about scale is how it may affect geospatial data collection and analysis results with respect to accuracy and reliability [1]. The scale problem has also been expressed by Zhang et al. (2018) [4] through the question: how many scales are necessary (authors' note: to describe a specific

feature or process) and what is the ‘interval’ between these scales? The scales which would correspond to this double inquiry (i.e., ‘how many’ and ‘at what interval’) can be called ‘optimal’.

Conventionally, ‘optimal’ scales in topographic and hydrographic mapping tasks are selected by intuition or experience. For example, Karydas and Panagos (2016) [5] combined a coarse and a fine scale of application within the framework of mapping sediment yield in Cyprus, in terms of an erosion model calibration. If determination of optimal scales were possible in advance of studying an environmental or land management problem, then the appropriate features or processes could be modelled in a rapid, correct, effective, and robust manner. As a paradigm, the introduction of functional scale by Jiang (2013) [6] in the problem of city mapping has led to the conceptualization of ‘natural cities’, as a bottom up, data-driven process, so that all mapped cities are comparable across different countries [7].

For many years, scale in raster-type datasets of topographic or hydrographic features was associated only with the cell size [8]. Cells, however, have a rectangular shape by default, and thus, cell resizing as a rescaling process has a serious limitation in capturing the natural shape and the spatial distribution of the properties of interest. Alternative to cell resizing, conversion of cells into objects using image segmentation techniques has been shown to be a more appropriate methodology for studying landscape on multiple scales (provided the spatial resolution of the original dataset is high enough to capture the studied features or processes) [9].

In this direction, much research has been conducted over time within the framework of geographic object-based image analysis (GEOBIA). Selection of optimal scales with trial-and-error techniques has been, so far, the most common practice, although at the expense of time and labor [10]. In the most known attempt so far, Drăguț et al. (2010) [11] developed ‘ESP’ (estimation of scale parameter), a software embedded in Definiens program (the most widely used OBIA software). ESP allows for a fast estimation of scale parameters in ‘multiresolution segmentation’ (an equivalent term for multi-scale segmentation), either for images or thematic layers. ESP segments the user-defined data automatically after insertion of six different parameters by the user (including hierarchical schemes if available) and then calculates local variance as standard deviation within the created objects for each segmentation level. The results are illustrated in a user-friendly graphical environment.

Other less generic attempts have also taken place. Romstad and Etzelmüller (2012) [12] introduced a method for geomorphic segmentation, which creates objects based on their mean curvature values. It can be said that this approach ensures geometric simplicity, irrespective of possible topographic variation within each created object. More recently, Newton et al. (2019) [13] compared three segmentation methods (morphological segmentation on edges, contour stability analysis, and active contours) for feature identification and performance in feature boundary determination. The results indicate that no segmentation method is consistently superior for all test cases, but the comparison approach was useful to explore and optimize segmentation alternatives for feature-based characterization scenarios.

However, none of the methods developed so far have achieved a fully automated bottom-up and data-driven segmentation approach. ESP may have facilitated the segmentation process in a generic multi-scale framework but it has not truly overcome the trial-and-error approach. Other methods either lack generality or are confined in a single-scale framework. For a critique to pre-existing methods in image segmentation and analysis, see Karydas (2020) [14].

According to Lam et al. (2002) [15], “an appropriate tool for detecting changes through scales in complex landscape process hierarchies” is fractal analysis. Jiang and Brandt (2016) [1] argue that fractal nature is inherent in all geographic phenomena and, moreover, it is the main source of their scale-dependency. Most topographic properties, such as slope and aspect, or hydrological features, such as stream network and watersheds, all demonstrate a fractal nature.

The fundamental characteristic of fractal features is that length, area, or volume are a function of the scale of measurement [16]. Fractals were introduced mainly because of the difficulty in analyzing spatial forms and processes by the (integer) topological dimension of the Euclidean geometry. The most

intrinsic parameter in fractal geometry is fractal dimension, defined as the ratio of the logarithmic change in detail (n) to the logarithmic change in scale (s) of geographical features [17]:

$$D = \frac{\log\left(\frac{n_{v+1}}{n_v}\right)}{\log\left(\frac{s_{v+1}}{s_v}\right)} \quad (1)$$

From Equation (1), it is derived that fractal dimension may take decimal values, for example 1.26 or 2.65.

Scale in fractal geometry is primarily defined in a manner in which a series of scales are related to each other in a scaling hierarchy. For example, a coastline is a set of recursively defined bends, forming the scaling hierarchy of far more small bends than large ones. Fractal analysis should not be considered only as a method but mainly as a new paradigm involving the universal scaling pattern across all scales from smallest to largest [1].

In this direction, Jiang and Yin (2014) [18] introduced the ‘ht-index’ as a metric of complexity in the dataset’s hierarchical structure, which fractal dimension alone could not express effectively. The authors experimented successfully with a grey-scale image, a digital elevation model, and road networks of very different cities. More recently, Karydas (2020) [14] employed head/tail breaks to develop a new method for selecting optimal scales for multi-scale segmentation of satellite imagery as a tool for facilitating a bottom-up approach in object-based image classification. The method employs fractal dimension as a metric of scale invariance in segmentation outputs; a topological transformation allows linking the partition with the segmentation process.

The main aim of this research was to develop a new, rapid, easy, objective, and robust methodology for selecting optimal scales when mapping topographic or hydrographic features. According to Meentemeyer (1989) [19], optimal scales are those where the most influencing variables of the mapped features remain invariant with rescaling; Sun et al. (2006) [16] argue that invariance over rescaling can be indicated using fractal dimension. Thus, the basic hypothesis of this research was that selection of optimal scales when mapping topographic or hydrographic features could be guided using fractal dimension as a judging metric, which, if it remains constant over rescaling, is expected to indicate optimal scales. The new method expands the method of Karydas (2020) [14], from object-based image analysis to multi-scale segmentation of digital elevation models (DEM) and their derivatives.

2. Methodology

The new method detects scale optimality in the structure of topographic or hydrographic datasets and then defines optimally scaled patterns in the spatial context by combining three distinct techniques:

- Partition of topographic or hydrographic data with the rank-size rule (also known as head/tail breaks), thus determining the hierarchical data structure [20].
- Segmentation of the topographic or hydrographic surfaces using the Fractal Net Evolution Assessment (FNEA), thus converting thematic cells into meaningful objects [21].
- Linkage of the above two processes through a topological transformation, which projects the partitioning results to the segmented topographic or hydrographic data [14].

Partition with the rank-size rule is a process through which a dataset is divided recursively into two subsets: one containing all the values above the mean, called the ‘head group’, and one containing all the values below the mean, called the ‘tail group’ [22]. Partitioning allows determination of the inherent hierarchical structure by the data itself. The possibility of partitioning a dataset with the rank-size rule is associated with heavy-tailed distributions, i.e., datasets where far more small values can be found than large ones [6,23]. Types of heavy-tailed distributions include power-law, exponential, or lognormal distributions.

Segmentation is an appropriate rescaling methodology of a raster dataset, as the local heterogeneity of the property of interest is minimized, thus indicating those scales at which meaningful objects

(i.e., naturally delineated surfaces) can be detected. One of the most common segmentation algorithms is the Fractal Net Evolution Assessment (FNEA), also known as ‘multiresolution segmentation’. With FNEA, objects are created by minimizing the internal heterogeneity of contained cell values through a stepwise region-growing procedure until a user-defined threshold (called the ‘scale factor’) is reached [21]. The scale factor influences the desired object size and shape, according to a spectral and/or two shape criteria [24]. The spectral criterion applied with image datasets is defined as the standard deviation of the spectral values contained in a candidate object [25]. In this research, however, the so-called ‘spectral criterion’ is used to measure standard deviation of the relevant thematic values (e.g., elevation, slope, aspect, etc.) instead of spectral values.

FNEA has proven to be one of the most successful image segmentation algorithms in a GEOBIA framework [10]. Bobick and Bolles (1992) [26] argue that the information stored in a layer processed with FNEA is considered as fractal, in terms of holding the same degree of non-regularity at all scales or (inversely) self-similarity across every scale. Previous research findings suggest that when FNEA is applied with satellite images, a power-law equation of scale factor vs. mean object size is derived [14,27,28]:

$$f = a s_f^b \quad (2)$$

where f : scale factor; s_f : the mean object size (at scale f (in m^2); a , b : constants (a in m^{-2} and b dimensionless). Newman (2005) [29] argues that “power-law behavior may be identified empirically using already seen results”. Furthermore, Vuduc (1997) [30] states that fractal objects follow, by default, power-law statistical distributions of scaling.

Following the paradigm of Karydas (2020) [14] for satellite images, a power-law equation between scale factor and mean object size for each of the topographic or hydrographic features of interest can be extracted for a study area, using a series of data segmentation—for example, the first twelve powers of 2, i.e., {1, 2, 4, 8, ..., 2048}; the experimentation with imagery shows that such a series is enough to indicate a quite accurate function (with a high coefficient of determination, $R^2 > 0.9$).

A topological transformation was introduced to transfer optimal scales detected in the frequency distribution with the head/tail breaks to the physical space. This transformation is justified by the fact that the surface corresponding to a head group (produced with the rank-size rule), which in reality is dispersed into smaller pieces throughout the entire studied surface, could form a single object if these small pieces were contiguous (a kind of virtual aggregation); this single object would have the optimal size, provided that it is the result of an optimal scale segmentation. Hofmann et al. (1998) [31] argue that optimal information for further processing and analysis is expected to be provided by the best segmentation results. According to Karydas (2020) [14], the size which an object is supposed to have after the above transformation can be called the ‘simulated mean object size’ (SMOS), and the number of objects corresponding to SMOS can be called the ‘simulated object number’ (SON). Practically, SMOS is computed for every partition level by dividing the total extent of the study area over the extent of the head group (corresponding to that partition level), while SON is equal to the ratio of the head group extent over SMOS:

$$s_n = \frac{i}{h_n} \quad (3)$$

$$N_n = \frac{h_n}{s_n} \quad (4)$$

where, s_n : simulated mean object size (SMOS) at n -th partition (dimensionless); i : extent of the entire study area (m^2); h_n : extent of the head group at n -th partition (m^2); N_n : simulated object number (SON). This transformation allows for the projection of the spatial outputs of the partitioning process (after the recursive means) to the extent of the entire study area (where segmentation is applied). In other words, the physical space has been projected to a simulated one and this kind of projection can be considered as a mapping tool, transferring scale optimality from a (global) hierarchical structure to (local) spatial patterns (Figure 1).

Figure 1. An illustration of the introduced topological transformation of the new method: from n -th partition after the mean (M_n) and creation of head groups (h_n), to their (virtual) aggregation in the physical space and their projection to a simulated space.

Finally, the computed SMOSs resulting from the topological transformation are inserted in Equation (2) to derive the optimal scale factors for the segmentation. To verify that the computed scales are the optimal ones, the fractal dimensions of all the successive scale factor values are computed and must be found to remain constant.

The whole computational procedure of the new method can be accomplished in three distinct steps (see also Figure 2):

1. Systematic segmentation of the surface of interest and extraction of a trend line (corresponding to Equation (2)) using the Fractal Net Evolution Assessment (FNEA) algorithm (usually embedded in an OBIA program).
2. Implementation of partition with the rank-size rule and computation of the head/tail breaks in a common spreadsheet; simple geospatial functions embedded in a geographic information system (GIS) program can be used to compute the means of the consecutive partitions.
3. Computation of the optimal scale factors using the trend equation on the head/tail breaks that resulted from the partition; the simulated mean object size (SMOS) is computed as an intermediate output; a common spreadsheet is enough to carry out these simple computations.

As it has become clear, partition process by the rank-size rule employs the mean, whereas segmentation process employs standard deviation. The values of the entire dataset under partitioning follow a heavy-tailed distribution, whereas the values contained by every object resulting from segmentation follows a normal distribution. With regard to the spatial context of the two processes, head/tail breaks is by default a non-spatial process applied on a global scale, whereas segmentation is a geospatial process applied on a local scale. Here is where the introduced topological transformation links the two procedures in order to transfer the partition results from the frequency plot to the segmentation outputs, thus converting a hierarchical structure into a spatial one (Table 1).

Table 1. Overview of the main processes in the new methodology.

Properties Process	Head/Tail Breaks	Segmentation (Fractal Net Evolution Assessment, FNEA)
Subject of application	Hierarchical structure (frequency plot)	Spatial structure (cells)
Spatial context	Global (entire dataset)	Local (neighboring cells/objects)
Frequency distribution type	Power-law (heavy-tailed)	Topological transformation
Statistic employed	Mean (global)	Normal (Gauss) Standard deviation (local)

Figure 2. An overview of the computational procedure of the new method.

3. Experimentation

3.1. Study Area and Data Preparation

The two largest river basins in Greece, namely Pinios river basin and Acheloos river basin, located in the central-east and the central-west part of Greece, respectively, were selected to experiment with the new methodology. The Pinios river basin has a projected surface of about 10,701 km² with a perimeter of about 888 km and pours out to the Aegean Sea (eastern coast of mainland Greece); the area is mostly plain covered by agricultural crops. The Acheloos river basin has a projected surface of about 5686 km² with a perimeter of about 823 km and pours out to the Ionian Sea (western coast of mainland Greece); the area is covered mainly by forested mountains (Figure 3).

Figure 3. The study areas of the Pinios and Acheloos river basins in Greece and their main river network (source: Catchment Characterisation and Modelling, CCM—River and Catchment Database for Europe).

An ASTER GDEM V2 (Advanced Spaceborne Thermal Emission and Reflection Radiometer Global Digital Elevation Model, Version 2) product was used as the only input dataset (freely downloaded from: <https://earthdata.nasa.gov/>). Digital elevation models (DEMs) are used widely for river basins' mapping and analysis [8,32].

The ASTER GDEM product was generated with stereo-pair images acquired with the ASTER instrument onboard of Terra satellite (EOS AM-1) and was released in October 2011 by the Ministry of Economy, Trade, and Industry (METI) of Japan and the National Aeronautics and Space Administration (NASA) of the USA. The ASTER GDEM products are provided in GeoTIFF format in a 30-m pixel size and 1×1 degree tiles.

It must be noted, however, that the original values in ASTER GDEM are estimated through optical instruments, i.e., those recording reflectance measurements and thus are less precise compared to LIDAR products (light detection and ranging, or laser imaging, detection, and ranging). LIDAR uses laser remotely sensed data in regular point grids and can be used to create 3D surface representations too [33,34].

The experimentation focused on four common topographic or hydrographic features, namely elevation, slope, aspect, and flow accumulation. The required tiles of the original ASTER GDEM products were mosaicked and clipped to the outlines of the two studied river basins. Elevation values were taken identical to the respective original cell values of the DEMs (in meters of altitude), while slope, aspect, and flow accumulation were computed from the DEM values in 3×3 windows using Spatial Analyst in ArcGIS program; the D8 algorithm was applied for flow accumulation estimation [32,35]). Ariza-Villaverde et al. (2015) [36] argue that the D8 algorithm is a reasonable method when convergent flow conditions predominate and has shown to be consistent among different flow patterns.

Stream and watershed networks in the two river basins were mapped by thresholding the flow accumulation (raster) layers (a process called 'flow accumulation thresholding' or FAT) [37]. The output layers can be as many as the number of the different thresholds selected and usually are converted and stored in vector format (polylines and polygons for streams and watershed, respectively). Here, $FAT = 10,000$ was applied to extract the river network in both river basins for geovisualization purposes only; selection of any FAT threshold for the display would not affect the analysis (neither the partition nor the segmentation processes), which are applied on the whole of the flow accumulation cells. The Pinios river basin was found to have an average elevation of 421.5 m, a maximum value of 2810 m, and a coefficient of variation of 86.5%; the mean slope of the basin is 16.7%, with a coefficient of variation of 102%; the mean aspect of the basin is 167.5, with a coefficient of variation of 63.2%. The slightly eastward aspect, compared to the expected mean of 180° , means that the mountain ranges lying on the west side are predominant in the basins' delineation. The shape index of the basin was computed to 2.147. Shape index is a metric of geometric complexity of an object and is defined as the ratio of the perimeter over four times the square root of the object's area; higher values correspond to higher complexity; rectangular is the least complex (or most regular) shape, with a shape index value of 1 [25] (Figure 4).

The Acheloos river basin was found to have an average elevation of 811.8 m, a maximum value of 2398 m, and a coefficient of variation of 63.2%; the mean slope of the basin is 34.8%, with a coefficient of variation of 69.5%; the mean aspect of the basin is 168.9, with a coefficient of variation of 62.5%. The slightly eastward aspect, compared to the expected mean of 180° , also means that the mountains on the west side are predominant in the basins' delineation. The shape index of the basin was computed to 2.729 (Figure 5). Some descriptive information and metrics of the two studied river basins are summarized in Table 2.

Table 2. Statistics, properties, and metrics of the main topographic and hydrographic features computed for the two experimental river basins.

Parameter		Pinios River Basin	Acheloos River Basin
Elevation (m)	Mean	421.5	811.8
	CV (%)	86.5	63.2
Slope (%)	Mean	16.7	34.8
	CV (%)	102	69.5
Aspect	Mean	167.5	168.9
	CV (%)	63.2	62.5
Watershed	Surface (km ²)	10,701	5686
	Perimeter (km)	888.5	823.2
	Shape index	2.147	2.729

Figure 4. Topographic and hydrographic maps of the Pinios river basin: (a) elevation; (b) slope; (c) aspect; (d) drainage network (derived with flow accumulation thresholding (FAT) = 10,000).

Figure 5. Topographic and hydrographic maps of the Acheloos river basin: (a) elevation; (b) slope; (c) aspect; (d) drainage network (derived with FAT = 10,000).

3.2. Preconditions and Method Implementation

The new methodology was applied only with elevation, slope, and flow accumulation layers, excluding the aspect layers. The reason is that the aspect values were distributed almost constantly (or only with slight differentiations) throughout the entire range of possible values (i.e., 0–360°), thus demonstrating an almost straight and horizontal frequency distribution. Therefore, aspect did not meet the fundamental condition for applying the rank-size rule, i.e., that the data should follow a heavy-tailed distribution. This limitation is expected also to exist for a single mountain, and only in case of a small hillslope could the partition be implemented (Figure 6).

For each of the elevation, slope, and flow accumulation layers, a series of twelve segmentations with the FNEA algorithm was carried out using eCognition Developer program. Plotting scale factor vs. mean object size allowed for the extraction of a predictive equation, in which the computation of the optimal scale factors (dependent variable, y) will be based by inputting SMOS (as the independent variable, x). The plotted segmentation results and the predictive equations (numerically and as trend lines) for the six layers are shown in Figure 7.

Figure 6. Frequency distribution plots of aspect values for the two studied river basins: (a) Pinios and (b) Acheloos.

Figure 7. Log-log plots of scale factor vs. mean object size and the extracted predictive equations for: (a) Pinios elevation; (b) Acheloos elevation; (c) Pinios slope; (d) Acheloos slope; (e) Pinios flow accumulation; (f) Acheloos flow accumulation.

Then, the successive partition of all the layers according to the rank-size rule was implemented. The extents resulting from the head/tail breaks of the partition were recorded and the SMOS and SON were computed according to Equations (3) and (4), respectively. Three obvious prerequisites were taken into account for the determination of the optimal scale factors: (a) SMOS had to be larger than the cell size (here, 900 m²), (b) SON had to be equal or larger than 1, and (c) the computed scale factors had to be larger than zero.

3.3. Results

The outputs of the new method implementation for each of the studied properties (elevation, slope, and flow accumulation) for the studied river basins (Pinios and Acheloos) are presented in Tables 3–8.

Table 3. Partial and final outputs for elevation in the Pinios river basin; 0-partition level corresponds to the original elevation surface.

Partition Level	River Basin Extent (m ²)	Mean value of Elevation (m)	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s_n) (m ²)	Simulated Object Number (N_n)	Computed Scale Factor (f_n) *
0	10,696,217,517	556	29.2	70.8	1	10,696,217,517	0
1	3,125,455,714	902	38.8	61.2	3	913,264,283	0
2	1,211,422,430	1186	36.1	63.9	9	137,202,175	0
3	437,142,350	1452	34.4	65.6	24	17,865,515	0.1
4	150,553,077	1738	32.4	67.6	71	2,119,088	0.1
5	48,712,500	2100	48.7	51.3	220	221,845	0.3
6	23,737,284	2334	48.3	51.7	451	52,678	0.4
7	11,470,495	2456	39.0	61.0	932	12,301	0.6
8	4,474,933	2556	39.0	61.0	2390	1,872	1.2
9	1,747,422	2638	44.3	55.7	6121	285	2.2
10	774,734	2702	48.7	51.3	13,806	56	3.7
11	377,131	2737	43.5	56.5	28,362	13	5.9
12	164,120	2762	45.5	54.5	65,173	3	10.1
13	74,700	2781	45.8	54.2	143,189	1	16.9

* optimal scale factors in bold.

Table 4. Partial and final outputs for elevation in the Acheloos river basin; 0-partition level corresponds to the original elevation surface.

Partition Level	River Basin Extent (m ²)	Mean Value of Elevation (m)	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s_n) (m ²)	Simulated Object Number (N_n)	Computed Scale Factor (f_n) *
0	5,685,890,000	812	49.9	50.1	1	5,685,890,000	0
1	2,836,198,998	1251	46.4	53.6	2	1,414,734,502	0
2	1,315,364,537	1501	41.9	58.1	4	304,294,291	0
3	551,598,126	1687	41.5	58.5	10	53,511,498	0
4	228,827,778	1832	40.8	59.2	25	9,209,139	0
5	93,361,845	1948	40.2	59.8	61	1,532,994	0
6	37,567,139	2045	39.4	60.6	151	248,209	0
7	14,803,403	2130	40.2	59.8	384	38,541	0.6
8	5,951,459	2202	42.4	57.6	955	229	1.0
9	2,524,270	2257	43.4	56.6	2252	1121	1.7
10	1,096,217	2295	40.4	59.6	5187	211	3.0
11	443,175	2326	45.0	55.0	12,830	35	5.3
12	199,342	2346	43.6	56.4	28,523	7	8.8
13	86,899	2362	39.2	60.8	65,431	1	15.0
14	34,058	2375	36.6	63.4	166,947	0	27.3

* optimal scale factors in bold.

The fractal dimension of the consecutive partition pairs for all the properties of interest were derived using the Mandelbrot equation (Mandelbrot 1982) [17], adapted to this method by Karydas 2020 [14], as follows:

$$D_n = \frac{\log\left(\frac{s_n}{s_{n-1}}\right)}{\log\left(\frac{f_n}{f_{n-1}}\right)} \quad (5)$$

where s_n : simulated mean object size at n -th partition (dimensionless); f_n : scale factor at n -th partition (dimensionless). The outputs of Equation (3) were found to be constant for each of the studied features (Tables 3–8), ranging from 1.1587 for Acheloos flow accumulation up to 1.9146 for Acheloos slope. In general, the two river basins demonstrated similar fractal dimension values for elevation (1.5461 vs.

1.5537) and flow accumulation (1.1935 vs. 1.1587), whereas they differed substantially only for flow accumulation (1.3784 vs. 1.9146).

Table 5. Partial and final outputs for slope in the Pinios river basin; 0-partition level corresponds to the original slope surface.

Partition Level	River Basin Extent (m ²)	Mean Value of Slope (%)	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s _n) (m ²)	Simulated Object Number (N _n)	Computed Scale Factor (f _n) *
0	10,696,217,517	16.72	39.9%	60.1%	1	10,696,217,517	0
1	4,268,123,100	34.09	40.9%	59.1%	3	1,703,113,719	0
2	1,745,121,600	47.38	38.2%	61.8%	6	284,722,089	0
3	666,109,877	59.45	37.6%	62.4%	16	41,482,175	0
4	250,299,899	71.06	35.9%	64.1%	43	5,857,214	0
5	89,766,000	83.14	35.4%	64.6%	119	753,344	0
6	31,743,899	96.16	36.0%	64.0%	337	94,209	0
7	11,440,800	109.94	36.6%	63.4%	935	12,237	0
8	4,188,599	123.79	36.3%	63.7%	2554	1640	0.9
9	1,520,100	137.88	37.8%	62.2%	7037	216	2.0
10	574,199	151.62	36.2%	63.8%	18,628	31	4.0
11	207,900	165.17	35.1%	64.9%	51,449	4	8.4
12	72,900	179.58	44.4%	55.6%	146,725	0	17.9
13	32,399	191.17	38.9%	61.1%	330,140	0	32.3
14	12,600	201.14	35.7%	64.3%	848,906	0	64.1

* optimal scale factors in bold.

Table 6. Partial and final outputs for slope in the Acheloos river basin; 0-partition level corresponds to the original slope surface.

Partition Level	River Basin Extent (m ²)	Mean Value of Slope (%)	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s _n) (m ²)	Simulated Object Number (N _n)	Computed Scale Factor (f _n) *
0	5,685,890,000	34.89	45.7	54.3	1	5,685,890,000	0
1	2,600,604,256	56.04	40.8	59.2	2	1,189,460,664	0
2	1,061,724,540	72.78	38.7	61.3	5	198,255,506	0.1
3	410,375,695	87.60	36.5	63.5	14	29,618,619	0.1
4	149,638,380	102.39	35.2	64.8	38	3,938,107	0.2
5	52,730,337	118.37	34.4	65.6	108	489,016	0.4
6	18,156,022	136.50	33.6	66.4	313	57,975	0.6
7	6,106,861	157.85	33.9	66.1	931	6559	1.1
8	2,072,364	183.01	35.0	65.0	2744	755	2.0
9	724,382	211.24	37.8	62.2	7849	92	3.5
10	274,005	238.02	38.5	61.5	20,751	13	5.8
11	105,507	262.60	36.6	63.4	53,891	2	9.5
12	38,581	287.08	44.9	55.1	147,375	0	16.1
13	17,322	305.06	27.3	72.7	328,247	0	24.5
14	4724	332.21	50.0	50.0	1,203,618	0	48.3

* optimal scale factors in bold.

Table 7. Partial and final outputs for flow accumulation in the Pinios river basin; 0-partition level corresponds to the original flow accumulation surface.

Partition level	River Basin Extent (m ²)	Mean Value of Flow Accum.	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s _n) (m ²)	Simulated Object Number (N _n)	Computed Scale Factor (f _n) *
0	10,696,217,517	3177	1.2%	98.8%	82	130,060,800	85,174
1	130,060,800	257,817	10.7%	89.3%	767	169,658	552,894
2	13,952,700	2,141,397	15.4%	84.6%	4991	2795	2,657,081
3	2,142,900	8,300,061	38.4%	61.6%	12,989	165	5,921,296

* optimal scale factors in bold.

Table 8. Partial and final outputs for flow accumulation in the Acheloos river basin; 0-partition level corresponds to the original flow accumulation surface.

Partition Level	River Basin Extent (m ²)	Mean Value of Flow Accum.	Head Extent (%)	Tail Extent (%)	Simulated Mean Object Size (s _n) (m ²)	Simulated Object Number (N _n)	Computed Scale Factor (f _n) *
0	5,685,890,000	3586	1.1%	99.0%	94.75465	60,006,446	91,275
1	60,006,446	335,833	12.9%	87.1%	732.3881	81,933	533,113
2	7,763,493	2,360,027	26.7%	73.3%	2746.804	2826	1,668,231
3	2,070,002	5,807,163	39.2%	60.8%	7011.016	295	3,745,045

* optimal scale factors in bold.

3.4. Discussion

The computed optimal scale factors for elevation or slope segmentations as well as the corresponding mean object size were found to be quite small for landform units. Indicatively, even for the highest computed scale factor for elevation in the Pinios river basin (i.e., 16.9), the corresponding mean object size was found to be only about 16 ha (Figure 8). Similarly, for the Pinios slope, the highest computed scale factor (i.e., 8.4) resulted in a 7-ha mean object size. In contrast to that, the computed optimal scale factors for flow accumulation are shown to be much larger than those computed for elevation and slope for both river basins. This can be attributed to the fact that flow accumulation layers are ‘empty’ of high values or inversely full of 0 values to a large extent. In other words, high values follow linear patterns, leaving big homogeneous objects of low values to fill the surface everywhere else. As a result of this particular spatial pattern, the mean object size for flow accumulation was found to be much larger than that of the other two feature types, and consequently, the scale factor was found to be similarly large.

Figure 8. An indicative abstract from the segmentation results of Pinios elevation case; 100-m contour lines and the elevation values reclassified with the head/tail breaks are overlaid for comparison.

The total number of rows in the computational tables (Tables 3–8) corresponds to another fractal metric known as the ht-index, introduced by Jiang and Yin (2014) [18]. Ht-index is an integer indicating the number of hierarchical levels of a dataset and, as such, a metric of its hierarchical structure complexity. In their work, the authors estimated the ht-index of the elevation in the United States of America to 16 using a DEM of 1800-m resolution. Similar ht-index values (14 or 15) were computed in the current study for elevation and slope too. The difference between the two resolutions (30 vs. 1800 m) should not be considered as an obstacle for comparison, as the geographic scale between the two case studies also differs greatly. Significantly lower values were found for flow accumulation in the current study (four in both cases). However, no other similar studies have been made available to do a comparison. The overall results from the two river basins indicate high complexity for the elevation and slope surfaces and low complexity for the flow accumulation surfaces (Table 9).

In some cases (for example, in Pinios slope mapping), noticeable divergences were detected between the fractal dimension derived from the Mandelbrot (1982) [17] equation (refer to Equation (3)) and that inferred from the ‘slope’ of the trend line, as defined by the plots of segmentations series (refer to Figure 4). This deviation was expected and could be attributed to the lower (compared to

that of the other features) coefficient of determination (R^2) of the equations derived from the plots. In all cases, the fractal dimension values computed with Equation (3) are the most precise ones as they are derived from the optimal data, whereas the trend lines correspond only to some systematic but not necessarily optimal segmentations. Similar deviations were noticed by Karydas (2020) [14] in the original implementation of the method with a variety of satellite images.

Table 9. A summary of the fractal dimension and ht-index values estimated for the studied topographic and hydrographic features in Pinios and Acheloos river basins.

Feature	Pinios River Basin		Acheloos River Basin	
	Fractal Dimension	Ht-Index	Fractal Dimension	Ht-Index
Elevation	1.5461	14	1.5537	15
Slope	1.3784	15	1.9146	15
Flow	1.1935	4	1.1587	4
Accumulation				

However, the implementation of the new method with the DEM and its derivatives revealed more problems than that with the imagery. As it can be understood from the distribution plots of Acheloos elevation and slope, their frequency distribution does not follow an ideal heavy-tailed shape. More specifically, the Acheloos elevation distribution demonstrates two additional peaks and valleys, while the Acheloos slope distribution imitates a normal distribution skewed to the left. In contrast, Pinios elevation demonstrates only a small extra peak, while both the slope distributions are closer to heavy-tailed shapes than any other else (Figure 9). Thus, only the flow accumulation distributions of both river basins were of an ideal heavy-tailed shape.

Another problem was that the rank-size rule in both river basins resulted in many ups and downs towards reaching the end-up threshold, while some of the head groups were detected close to 50% early in the partitioning process (Figure 10). In such cases, it is debatable if partitioning could be applied or better be avoided. Note that for digital elevation models, the end-up threshold is set usually to less than or equal to 40% (Jiang 2015) [19].

Nonetheless, regardless of the above abnormalities, the current results indicate clearly that the new method was successful in all cases. Moreover, plotting segmentation vs. partitioning data of any of the studied features shows that optimal scale factors could be computed directly from the mean values derived from the partitions. Indicatively, the two studied river basins resulted in the following equations for elevation (see also Figure 11):

$$\text{For Pinios : } f = 4 \cdot 10^{-13} e^{0.0111 M_n} (R^2 = 0.9023) \quad (6)$$

$$\text{For Acheloos : } f = 2 \cdot 10^{-18} e^{0.0184 M_n} (R^2 = 0.8046) \quad (7)$$

where f : scale factor; M_n : mean at partition n . If Equations (4) and (5) were generalized into a single one for any river basin, optimal scale factors for each feature could be computed directly from the successive mean values of the partition process alone and, thus, the segmentation process could be skipped. Considering that scale factor is equivalent to standard deviation within candidate objects, Equations (4) and (5), in reality, link the two main statistics in every dataset. Furthermore, high correlation between other series of data used in this study has also been detected (e.g., head extents with computed scale factors). This could be attributed to the fact that head groups correspond to the successive means of the dataset, while segmented objects at different scales correspond to the successive standard deviations of the created objects.

Elevation means

(a)

Slope means

(b)

Elevation means

(c)

Slope means

(d)

Figure 9. The topographic data distributions of the study: (a) Pinios elevation distribution (one small extra peak); (b) Pinios slope distribution; (c) Acheloos elevation distribution (two extra peaks); (d) Acheloos slope distribution (skewed normal distribution).

Figure 10. Head group extent (denoted by the columns; in m^2) and percentage (denoted by the lines) vs. partition means for: (a) Pinios elevation; (b) Acheloos elevation; (c) Pinios slope; (d) Acheloos slope; (e) Pinios flow accumulation; (f) Acheloos flow accumulation.

Figure 11. Indicative log-log plots of optimal scale factor vs. mean elevation for the two studied river basins; (a) Pinios and (b) Acheloos.

4. Conclusions

In this research, a new method for scale optimization when mapping elevation, slope, or flow accumulation was developed and experimented with the two biggest river basins in Greece. Aspect does not meet fundamental conditions for applying the rank-size rule and thus the new method could not be applied with aspect layers.

The new method uses head/tail breaks to indicate the optimal scales in the hierarchical data structure and a topological transformation to project the results to the spatial context. Fractal dimension is employed to confirm optimality in the spatial data. The method expands the work of Karydas (2020) [14], for multi-scale segmentation of image data, to topographic and hydrographic raster datasets.

The experimentation revealed some deviations from the ideal conditions regarding heavy-tailed distributions and partitioning. Regardless of the possibility of not meeting the ideal conditions, however, the method proved successful to identify a set of optimal scales in all cases. Therefore, the basic hypothesis of the research, that selection of optimal scales in topographic and hydrographic mapping could be guided by fractal dimension as a scale optimality metric, was verified in all terms.

Optimal scales known in advance of environmental problem solving and land management tasks would reveal underlying (or even hidden) processes, thus supporting decision making. Furthermore, the new method may define appropriate mapping scales for object generalization within the framework of topographic and terrain analysis, where an appropriate proportion of variability among feature densities and textures has to be retained. For example, the object-scale selection process could be equated to a filtering process for near real-time web-mapping.

Implementation of the new method prerequisites an object-based analysis program (to conduct segmentation tasks), a GIS-program (to prepare the required layers and compute spatial statistics), and a spreadsheet (to carry out the computations of the partial and final outputs).

Regarding the studied river basins, it was shown that the Pinios river basin was more heterogeneous for elevation and less for slope from that of Acheloos, while they were similarly heterogeneous for the aspect. Both river basins were found to be similarly complicated, as they had similar values of ht-index. The fractal dimension differed for slope between the two basins, whereas it was similar for elevation and flow accumulation.

In summary, the new method proved to be rapid (short time for accomplishment), easy (simple processes), objective (fully numerical), and robust (results were derived even with non-ideal preconditions). The method could be further simplified by directly using the mean values of the partition to compute the optimal scale factors, provided that more tests with diverse river basins would take place towards generalization. In such a case, partition would be the only required process to carry out, so that data plotting would result in the necessary non-linear equations. In its general form, the method could help identify the best datasets for extracting features of a particular size.

Finally, new light must be shed on the properties of heavy-tailed distributions with a view to clarify and further specify the required preconditions. Future research also should expand the method to the vector data type of hydrographic features, for example, to stream networks and watersheds.

Author Contributions: Conceptualization, Christos Karydas and Bin Jiang; Methodology, Christos Karydas and Bin Jiang; Software, Christos Karydas; Validation, Christos Karydas; Formal Analysis, Christos Karydas and Bin Jiang; Investigation, Christos Karydas and Bin Jiang; Resources, Christos Karydas and Bin Jiang; Writing—Original Draft Preparation, Christos Karydas; Writing—Review and Editing, Bin Jiang; Supervision, Bin Jiang; Project Administration, Bin Jiang; Funding Acquisition, Bin Jiang. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: Segmentation tasks were conducted with the eCognition Developer program in demonstration mode; geospatial analysis was conducted with ArcGIS Pro; statistics and graphs were produced with Excel spreadsheets. Many thanks to the anonymous reviewers, whose substantial critique helped us improve the paper a lot.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Jiang, B.; Brandt, S.A. A Fractal Perspective on Scale in Geography. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 95. [\[CrossRef\]](#)
2. Hay, G.J.; Blaschke, T.; Marceau, D.J.; Bouchard, A. A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS J. Photogramm. Remote Sens.* **2003**, *57*, 327–345. [\[CrossRef\]](#)
3. Klemeš, V. Conceptualization and scale in hydrology. *J. Hydrol.* **1983**, *65*, 1–23.
4. Zhang, X.; Du, S.; Ming, D. Segmentation Scale Selection in Geographic Object-Based Image Analysis. In *High Spatial Resolution Remote Sensing: Data, Analysis, and Applications*; He, Y., Weng, Q., Eds.; Taylor & Francis Group LLC: Abingdon, UK; CRC Press: Boca Raton, FL, USA, 2018; pp. 201–228.
5. Karydas, C.G.; Panagos, P. Modelling monthly soil losses and sediment yields in Cyprus. *Int. J. Digit. Earth* **2016**, *9*, 766–787. [\[CrossRef\]](#)
6. Jiang, B. Head/tail breaks: A new classification scheme for data with a heavy-tailed distribution. *Prof. Geogr.* **2013**, *65*, 482–494. [\[CrossRef\]](#)
7. Jiang, B. Living structure down to earth and up to heaven: Christopher Alexander. *Urban Sci.* **2019**, *3*, 96. [\[CrossRef\]](#)
8. Wu, S.; Li, J.; Huang, G.H. A Study on DEM-Derived Primary Topographic Attributes for Hydrologic Applications: Sensitivity to Elevation Data Resolution. *Appl. Geogr.* **2008**, *28*, 210–223. [\[CrossRef\]](#)
9. Blaschke, T. Object based image analysis for remote sensing. *ISPRS J. Photogramm. Remote Sens.* **2010**, *65*, 2–16. [\[CrossRef\]](#)
10. Witharana, C.; Civco, D.L. Optimizing multi-resolution segmentation scale using empirical methods: Exploring the sensitivity of the supervised discrepancy measure Euclidean distance 2 (ED2). *ISPRS J. Photogramm. Remote Sens.* **2014**, *87*, 108–121. [\[CrossRef\]](#)
11. Drăguț, L.; Tiede, D.; Levick, S. ESP: A tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 859–871. [\[CrossRef\]](#)
12. Romstad, B.; Etzelmüller, B. Mean-curvature watersheds: A simple method for segmentation of a digital elevation model into terrain units. *Geomorphology* **2012**, *139–140*, 293–302. [\[CrossRef\]](#)
13. Newton, L.; Senin, N.; Smith, B.; Chatzivagiannis, E.; Leach, R. Comparison and validation of surface topography segmentation methods for feature-based characterisation of metal powder bed fusion surfaces. *Surf. Topogr. Metrol. Prop.* **2019**, *7*, 045020. [\[CrossRef\]](#)
14. Karydas, C.G. Optimization of multi-scale segmentation of satellite imagery using fractal geometry. *Int. J. Remote Sens.* **2020**, *41*, 2905–2933. [\[CrossRef\]](#)
15. Lam, N.S.-N.; Qiu, H.L.; Quattrochi, D.A.; Emerson, C.W. An evaluation of fractal methods for characterizing image complexity. *Cartogr. Geogr. Inf. Sci.* **2002**, *29*, 25–35. [\[CrossRef\]](#)
16. Sun, W.; Xu, G.; Gong, P.; Liang, S. Fractal analysis of remotely sensed images: A review of methods and applications. *Int. J. Remote Sens.* **2006**, *27*, 4963–4990.

17. Mandelbrot, B. *The Fractal Geometry of Nature*; W. H. Freeman and Co.: New York, NY, USA, 1982.
18. Jiang, B.; Yin, J. Ht-Index for Quantifying the Fractal or Scaling Structure of Geographic Features. *Ann. Assoc. Am. Geogr.* **2014**, *104*, 2014. [\[CrossRef\]](#)
19. Meentemeyer, V. Geographical perspectives of space, time, and scale. *Landscape Ecol.* **1989**, *3*, 163–173. [\[CrossRef\]](#)
20. Jiang, B. Geospatial analysis requires a different way of thinking: The problem of spatial heterogeneity. *Trends. Spatial. Anal. Modell. GeoJournal* **2018**, *19*, 23–40. [\[CrossRef\]](#)
21. Benz, U.C.; Hofmann, P.; Willhauck, G.; Lingenfelder, I.; Heynen, M. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.* **2004**, *58*, 239–258. [\[CrossRef\]](#)
22. Zipf, G.K. *Human Behavior and the Principle of Least Effort*; Addison-Wesley: Boston, MA, USA, 1949.
23. Chen, Y. The mathematical relationship between Zipf's law and the hierarchical scaling law. *Phys. A Stat. Mech. Appl.* **2012**, *391*, 3285–3299. [\[CrossRef\]](#)
24. Kim, M.; Madden, M.; Warner, T. Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery. In *Object-Based Image Analysis-Spatial Concepts for Knowledge Driven Remote Sensing Applications*; Blaschke, T., Lang, S., Hay, G.J., Eds.; Springer: Berlin, Germany, 2008; pp. 291–307.
25. Baatz, M.; Benz, U.; Dehghani, S.; Heynen, M.; Holtje, A.; Hofmann, P.; Lingenfelder, I.; Mimler, M.; Sohlbach, M.; Weber, M.; et al. *ECognition User's Guide*; Definiens Imaging GmbH, Digital: Munich, Germany, 2002.
26. Bobick, A.; Bolles, R. The representation space paradigm of concurrent evolving object descriptions. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 146–156. [\[CrossRef\]](#)
27. Karydas, C.; Tompoulidou, M.; Kalaitzidis, C.; Gitis, I. Segmentation of a WorldView2 image for feature identification in highly heterogeneous environments. In Proceedings of the 5th GEOBIA Conference, Thessaloniki, Greece, 21–24 May 2014; pp. 717–720.
28. Manakos, I.; Technitou, E.; Petrou, Z.; Karydas, C.; Tomaselli, V.; Veronico, G.; Mountrakis, G. Multi-modal knowledge base generation from very high resolution satellite imagery for habitat mapping. *Eur. J. Remote Sens.* **2016**, *49*, 1033–1060. [\[CrossRef\]](#)
29. Newman, M.E.J. Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **2005**, *46*, 323–351. [\[CrossRef\]](#)
30. Vuduc, R. *Image Segmentation Using Fractal Dimension*; GEOL634, Report; Cornell University: Ithaca, NY, USA, 1997.
31. Hofmann, P.; Blaschke, T.; Strobl, J. Quantifying the robustness of fuzzy rule sets in object-based image analysis. *Int. J. Remote Sens.* **2011**, *32*, 7359–7381. [\[CrossRef\]](#)
32. Zhang, H.; Huang, G. Building Channel Networks for Flat Regions in Digital Elevation Models. *Hydrol. Process.* **2009**, *23*, 2879–2887. [\[CrossRef\]](#)
33. Shan, J.; Toth, C.K. *Topographic Laser Ranging and Scanning: Principles and Processing*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2018.
34. Jenson, S.K.; Domingue, J.O. Extracting Topographic Structure from Digital Elevation Data for Geographic Information System Analysis. *Photogramm. Eng. Remote Sens.* **1988**, *54*, 1593–1600.
35. Burrough, P.A.; McDonell, R.A. *Principles of Geographical Information Systems*; Oxford University Press: New York, NY, USA, 1998; 190p.
36. Ariza-Villaverde, A.B.; Jiménez-Hornero, F.J.; Gutiérrez de Ravé, E. Influence of DEM resolution on drainage network extraction: A multifractal analysis. *Geomorphology* **2015**, *241*, 243–254. [\[CrossRef\]](#)
37. O'Callaghan, J.F.; Mark, D.M. The extraction of drainage networks from digital elevation data, Computer Vision. *Graph. Image Process.* **1984**, *28*, 328–344.

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).