Title

Setting the Flow Accumulation Threshold Based on Environmental and Morphologic Features to Extract River Networks from Digital Elevation Models

Journal

ISPRS International Journal of Geo-Information, 10(3)

ISSN

2220-9964

Authors

Zhang, HuiHui

Loáiciga, Hugo A

Feng, LuWei

et al.

Publication Date

2021

DOI

10.3390/ijgi10030186

Peer reviewed

International Journal of Geo-Information

MDPI

Article

Setting the Flow Accumulation Threshold Based on Environmental and Morphologic Features to Extract River Networks from Digital Elevation Models

HuiHui Zhang [1,2], Hugo A. Loáiciga [2], LuWei Feng [1], Jing He [1] and QingYun Du [1,3,4,*]

[1] School of Resources and Environmental Sciences, Wuhan University, 129 Luoyu Road, Wuhan 430079, China; Zhanghuih@whu.edu.cn (H.Z.); lwfeng@whu.edu.cn (L.F.); Jing.he@whu.edu.cn (J.H.)
[2] Department of Geography, University of California, Santa Barbara, CA 93106, USA; hugo@geog.ucsb.edu
[3] Key Laboratory of Geographical Information System, Ministry of Education, Wuhan University, Wuhan 430079, China
[4] Key Laboratory of Digital Mapping and Land Information Application Engineering, Ministry of Natural Resources, Wuhan University, Wuhan 430079, China
* Correspondence: qydu@whu.edu.cn

Abstract: Determining the flow accumulation threshold (FAT) is a key task in the extraction of river networks from digital elevation models (DEMs). Several methods have been developed to extract river networks from Digital Elevation Models. However, few studies have considered the geomorphologic complexity in the FAT estimation and river network extraction. Recent studies estimated influencing factors' impacts on the river length or drainage density without considering anthropogenic impacts and landscape patterns. This study contributes two FAT estimation methods. The first method explores the statistical association between FAT and 47 tentative explanatory factors. Specifically, multi-source data, including meteorologic, vegetation, anthropogenic, landscape, lithology, and topologic characteristics are incorporated into a drainage density-FAT model in basins with complex topographic and environmental characteristics. Non-negative matrix factorization (NMF) was employed to evaluate the factors' predictive performance. The second method exploits fractal geometry theory to estimate the FAT at the regional scale, that is, in basins whose large areal extent precludes the use of basin-wide representative regression predictors. This paper's methodology is applied to data acquired for Hubei and Qinghai Provinces, China, from 2001 through 2018 and systematically tested with visual and statistical criteria. Our results reveal key local features useful for river network extraction within the context of complex geomorphologic characteristics at relatively small spatial scales and establish the importance of properly choosing explanatory geomorphologic characteristics in river network extraction. The multifractal method exhibits more accurate extracting results than the box-counting method at the regional scale.

Keywords: multi-source satellite data; flow accumulation threshold; river networks; geomorphological complexity; correlation model

Citation: Zhang, H.; Loáiciga, H.A.; Feng, L.; He, J.; Du, Q. Setting the Flow Accumulation Threshold Based on Environmental and Morphologic Features to Extract River Networks from Digital Elevation Models. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 186. https://doi.org/10.3390/ijgi10030186

Academic Editor: Wolfgang Kainz

Received: 19 January 2021
Accepted: 19 March 2021
Published: 21 March 2021

1. Introduction

River networks constitute topographic features that are widely used in the analysis of regional geomorphic features and hydrogeological environments [1,2]. The accurate extraction of river networks is necessary for delineating water pollution sources, monitoring disturbances on rivers caused by human activities, determining flood levels in periods of heavy rain, assessing soil and water conservation measures, and enacting comprehensive management practices [3–6].

Much effort has been devoted to extracting river networks [7–9]. GeoNet [10] is a well-developed and representative open-source software for extracting channel networks based on geodesic minimization principles. However, the extraction efficiency is not high due to its complex geometric-based channel head recognition algorithm and the loosely

coupled program design. The flow accumulation method still dominates in large-scale drainage network extraction from digital elevation model (DEM) data because of its simple form and efficient computational design [11–16]. The flow accumulation threshold (FAT) is a user-defined parameter that directly affects the structure and density of extracted river networks from DEMs [17]. Too large a FAT would omit useful details of river networks, and too small a FAT would extract pseudo rivers. This work focuses on the determination of the FAT seeking to maximize river extraction accuracy with computational efficiency.

Several methods have been proposed to determine the FAT. Tarboton et al. (1991) reported a quantitative method to determine the FAT through power functions involving slope and drainage area [18]. However, power functions must be validated further to evaluate the influences of land morphology, soil, and climate on channel initiation [19]. Widely used GIS tools apply 1% of the maximum flow accumulation value as a default FAT [20]. The FAT proposed by Tang (2000) and Olivera et al. (2002) is easily determined from flow accumulation statistics, but the accuracy is relatively low [20,21]. Jones (2002) applied a trial-and-error search to determine the FAT [22]. Trial-and-error is subjective and time-consuming for determining a suitable FAT. Lin et al. (2006) determined FAT with the headwater-tracing approach with a fitness index in Taiwan's upstream watersheds [19]. However, this method requires comparison with known stream data that must be obtained from aerial photography or a topographic map and it is time-consuming when applying it to multiple sub-basins. Tantasirin et al. (2016) determined the optimal FAT by considering the rate of change of the trend line of the "error stream cells" obtained with multiple thresholds [23]. FAT was obtained at the point where the difference of variation decreases, and the trend line approaches zero. This method is mainly used to produce a drainage network consisting of the shortest length of stream links (stream tributaries) and to delineate a watershed area into smaller-sized hillslopes. Gökgöz et al. (2006) assessed the effect of the FAT selection on the extraction of drainage lines and watershed delineation by considering the adjacency and direction relationships between the cells including the beginning points of the lines in the drainage network to be derived [24]. Ibrahim and Gökgöz (2018) improved the method proposed by Gökgöz et al. (2006) [17,24]. FAT was determined by considering the adjacency and direction functions between all cells including the drainage lines. The latter two algorithms are based on the assumption that among the cells containing parallel lines there must be other cells in the model to describe a ridge, which may or may not be the case.

Initiation of natural river networks depends on the climate, slope, lithology, vegetation cover, soil properties, and drainage density (δ), and is spatially variable [10,25,26]. The cumulative action of erosional processes over long timescales drives the formation of river networks commensurate with the evolution of large-scale topography [27]. Drainage (or, perhaps more accurately, river) density is defined as the total length of river reaches (l) within a basin divided by the basin area. It is a macroscale measure of the amount of area (say, in km^2) encompassing one unit length (in km) of a river reach [28]. Several studies have analyzed the spatial dependence of δ on environmental parameters at the continental scale. Luo et al. (2016) extracted the land dissection density (considered equivalent to the drainage density) within the United States with a geomorphological detection method and a 30 m resolution DEM and discerned a spatial dependence on climate, lithology, and several terrain-based attributes [29]. Schneider et al. (2017) proposed a new method to extract river networks based on the spatial variable δ as a function of slope, lithology, and climate with the 15" HydroSHEDS DEM in France and Australia [30]. In the Holocene, or perhaps more accurately, in the Anthropocene, human activities have gradually become the main driving force of global ecosystem degradation [31,32]. Song et al. (2019) concluded that the anthropogenic influence on river network structure is more significant than those of morphology and lithology in the urbanized regions by evaluating the river network change and its multifractal dynamics in the Yangtze River Delta [33]. River networks have become increasingly degraded, altered, or modified by anthropogenic actions such as road and reservoir construction [34]. Urban expansion has exerted dramatic changes

in structural or developmental processes in approximately 60% of rivers worldwide [35]. Thus, multiple geomorphologic factors and river characteristics must be considered when estimating the FAT.

Key obstacles to the mapping of river networks are the existence of intermittent and perennial streams, insufficient resolution, vegetation masking, and scarce field data [36,37]. Previous research on river network extraction has focused on basins that are not intensely urbanized, thus neglecting anthropogenic impacts [10]. Remote sensing provides a complementary source of river network analysis, yielding data on surface water extent, surface runoff, and evapotranspiration, which creates opportunities to derive river-related information where and when in-situ data are not available [38]. Global, high-resolution, DEMs have become available. There were 30-m resolution DEMs like SRTM digital elevation data [39] and ASTER GDEM ten years ago. There are now 12.5-m resolution ALOS global DEM data. The use of high-resolution DEMs provide accurate drainage networks and minimizes the extraction error, especially when the drainage network is dense [40]. Significant improvements in the resolution and accuracy of DEMs lead to more reliable DEM-derived drainage networks [41]. The extracted river networks are validated against delineated drainage networks from the Google imagery to assess the accuracy of extracted river networks [42]. Improving drainage density estimates requires consideration of suitable control factors such as land-use information, detailed geomorphologic information, and vegetation indicators [29,30,43]. Evidently, there are gaps of knowledge in the understanding of the statistical association between the drainage density, river length, and the FAT.

This paper's objective is the development and testing of a proposed methodology to calculate the FAT accurately. One method herein proposed applies an empirical function relating the drainage density and the FAT: it is a quantitative method for determining the FAT that evaluates multi-source parameters in basins with complex topographic and environmental characteristics that may control channel initiation. This method considers meteorological conditions and other factors such as lithology, landscape patterns, and anthropogenic impacts (especially notable in urbanized basins) in the determination of FAT, herein named the drainage density-FAT method. The second method proposed in this work relies on fractal geometry that accounts for morphological characteristics at the regional scale, in which case the large areal extent precludes the use of basin-wide representative regression predictors as done in the first method. Fractal algorithms provide an alternative approach to estimate the FAT at the regional scale; they constitute a useful alternative compared with the methodologies that are best suited to capture complex local features in relatively small basins. Fractal geometry [44] has been implemented in the study of river network morphology with successful results [40,45–47]. The most widely used fractal methods are the box-counting and multifractal methods. Yet, the estimation of the FAT based on fractal algorithms has not been fully addressed, a gap this paper intends to narrow down.

A case study relying on data from Hubei Province and Qinghai Province, China, a region with significantly variable landscape properties demonstrates the application of this paper's methods and their validation based on visual, statistical criteria, the comparison with several classical FAT methods, and robustness tests. This study contributes to the field of river network extraction in basins with complex topographic and environmental characteristics in which basin-wide regression factors are suitable for the determination of the FAT, and at the regional scale based on the morphological characteristics of river networks. The methods applied in this study are helpful to improve the accuracy of river network extractions in areas without LIDAR data.

2. Methodology

The common procedure of extracting a river network involves filling sinks, calculating the flow direction, calculating the flow accumulation, and extracting the river network [48]. A deterministic eight-node (D8) algorithm is herein applied to obtain the flow direction with ArcGIS 10.6 [49]. The D8 algorithm determines the flow direction of each grid cell by

choosing the steepest among a set of slopes oriented towards the neighboring eight cells. The FAT must be specified prior to extracting a river network. The determination of the optimal FAT is herein made with two methods, as shown in the methodology flowchart displayed in Figure 1. The first method calculates the FAT at the small basin scale by means of an empirical scaling equation (see Section 2.1). The second method calculates the FAT at the regional scale with the fractal method (see Section 2.2). The second method is intended for FAT determination when the large areal extent of a river basin precludes the use of basin-wide representative regression predictors. The application of the fractal method at the small basin scale is not addressed in this study. The accuracies of the river-network extracting methods are assessed based on the similarity score evaluation index (SSEI), river length error (err(l)), root mean square error (RMSE), and qualitative visual inspection (see Section 2.6). The first method is also compared with several classical FAT determination methods and is evaluated with a robustness test.

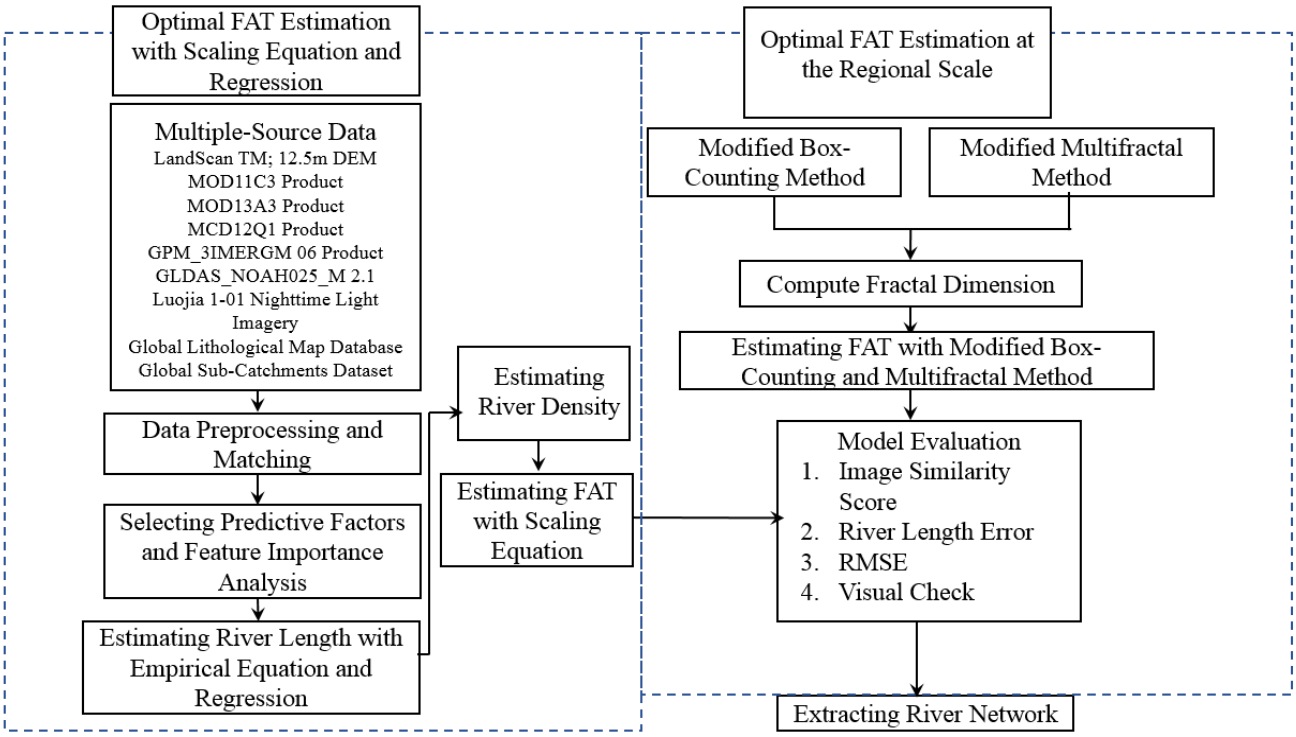


Figure 1. Flowchart of this paper's methodology.

2.1. The Estimation of the FAT with Empirical Scaling Equation and Statistical Regression at the Basin Scale

2.1.1. Empirical Scaling Equation Relating the Drainage Density and the FAT

The drainage density of basin i is given by Equation (1):

$$\delta_i = \frac{l_i}{A_i} \tag{1}$$

in which δ_i, l_i, and A_i denote the drainage area, the length of river reaches, and the drainage area of basin i, respectively. Tarboton et al. (1991) [18] proposed an empirical function expressing the drainage density in terms of the flow accumulation threshold of basin i (FAT_i) and drainage length is given by Equation (2):

$$\delta_i \propto \frac{1}{\sqrt{FAT_i}} = K\frac{1}{\sqrt{FAT_i}} \tag{2}$$

K denotes a proportionality constant that is empirically determined. Equation (2) may be interpreted as an empirical "scaling law" for the FAT given that it can be solved as a function of the drainage density.

2.1.2. Estimation of the River Length

The river length is estimated by multilinear regression involving multiple factors or regressor variables. A list of considered factors is presented in Table 1. The dependent variable of the regression equation is the total length of the river network within each basin. Predictive factors exhibiting strong correlations (determined with the Pearson's correlation coefficient) are filtered for the purpose of reducing data redundancy and model complexity. The reference river networks are obtained from the National Earth System Science Data Center, National Science and Technology Infrastructure of China, after being corrected by remote sensing images (0.4 m spatial resolution from GeoEye) accessed from Google Earth. The total length of the reference river network within each basin is calculated with the ArcGIS software.

Table 1. List of possible predictor factors.

Category	Brief Description	Year	Symbol	Unit
Meteorological factors	Average monthly precipitation	2001–2018 (5–9)	PRE	Mm
	Average monthly surface runoff	2001–2018 (5–9)	ROF	Mm
	Average monthly evapotranspiration	2001–2018 (5–9)	EVP	Mm
	Average surface temperature	2001–2018 (5–9)	LTD	°C
Topographical factors	Maximum terrain altitude	2018	ALT_{max}	M
	Mean terrain altitude	2018	$\overline{ALT}$	M
	Maximum slope	2018	SLP_{max}	°
	Mean slope	2018	$\overline{SLP}$	°
Vegetation factors	Maximum surface roughness	2018	SR_{max}	/
	Mean surface roughness	2018	$\overline{SR}$	/
	Normalized Difference Vegetation Index	2001–2018 (5–9)	NDVI	/
Anthropogenic factors	Population ratio	2018	Q	%
	Basin light area	2018	L_1	/
	Basin light intensity	2018	L_2	/
Landscape factors	Landscape percentage	2018	PLAND	%
	Shannon's diversity index	2018	SHDI	%
	Species evenness	2018	SEI	%
	Patch density	2018	PD	number/km^2
Lithology factors	Lithology type	2012	LT	/

2.1.3. NMF

The degree of association between selected regression factors was further verified by nonnegative matrix factorization (NMF). NMF has previously been shown to be a useful decomposition for multivariate data belonging to unsupervised learning [50]. This study introduces the NMF to reflect the effect of several environmental factors on the estimation of river length [51]. A detailed description of NMF is presented in Section 3.1.3. NMF captures the data traits by identifying the correlation between data parts and finding internal patterns of association among the data. The final model was constructed by the multilinear stepwise regression method with the software SPSS 22.0, as described in the Results section.

2.2. The Estimation of the FAT at the Regional Scale

The modified box-counting algorithm and the multifractal algorithm for FAT calculation applied in this study were proposed by De Bartolo et al. (2004) and Ariza-Villaverde et al.

(2015) [40,52]. The determination of the FAT by box-counting and the multifractal method are compared in this study.

2.2.1. The Box-Counting Method

A mesh of d-dimensional cubes of size r^d is used to discretize the region within which a river network is defined [53]. Let $N(r)$ denote the smallest number of cubes of size r needed to cover a region. This work extracts river networks relying on thirteen FATs which range from 1000 to 25,000. The corresponding fractal dimensions of the thirteen extracted river networks were calculated.

The formula for computing the fractal dimension D_f is as follows [54]:

$$N(r) = r^{-D_f} \tag{3}$$

Taking the logarithm (log) on both sides of Equation (3) yields:

$$\log N(r) = -D_f log r \tag{4}$$

D_f equals the slope of a plot of $N(r)$ vs. r in logarithmic scales. The rate of change d_i of the fractal dimension is given by Equation (5):

$$d_i = \left[\frac{|D_{i+1} - D_i|}{V_{i+1} - V_i} \right] \times 100 \tag{5}$$

where V_i, V_{i+1} denote the ith and (i+1)th thresholds, respectively (the FAT = 1000 corresponds to V_1). D_i, D_{i+1} denote respectively the box-counting fractal dimensions corresponding to the ith and (i+1)th flow accumulation thresholds. $D_{i+1} - D_i$ equals the fractal dimension difference of the V_i, V_{i+1} thresholds, and $V_{i+1} - V_i$ denotes the difference of the thresholds.

There is a fractal dimension corresponding to a river network extracted with a FAT. That is, there is an optimal FAT associated with the shape of the function of the rate of change d_i. The FAT associated with the optimal estimate of the river network is determined from the shape of the function of the rate of change d_i. Specifically, the FAT for river network extraction occurs at that point defining where the change rate of the fractal dimension exhibits a rapidly declining trend. This method is illustrated in the Results section (Section 3.2.1, specifically).

2.2.2. The Multifractal Method

The length of a river network is herein employed as the performance variable with which to assess the performance of river extraction methods. The box-counting method is implemented to calculate the probability distribution of the river network's length [55]. Specifically, $N(\varepsilon)$ squares of size $\varepsilon \cdot \varepsilon$ are used to superimpose a grid over the river network. The sum of a river network's length within a box of size $\varepsilon \cdot \varepsilon$ is calculated and divided by the sum of the lengths of all rivers within the entire basin. The river distribution ratio, $Q_i(\varepsilon)$, with box size $\varepsilon \cdot \varepsilon$ is given by the following expression:

$$Q_i(\varepsilon) = \frac{M_i(\varepsilon)}{M} \tag{6}$$

where $M_i(\varepsilon)$ equals the summed lengths of the rivers in a box of size $\varepsilon \cdot \varepsilon$, and M denotes the total lengths of the rivers within the river network spanning the entire study area.

The partition function $X_q(\varepsilon)$ is defined as follows [56]:

$$X_q(\varepsilon) = \sum_{i=1}^{N(\varepsilon)} [Q_i]^q \tag{7}$$

$$X_q(\varepsilon) \propto \varepsilon^{\tau(q)} \tag{8}$$

The partition function $X_q(\varepsilon)$ denotes the q-th quadratic weighted summation of the probability, where $N(\varepsilon)$ represents the number of grid cells with box size ε, and $q \in (-\infty, +\infty)$ denotes the moment order. $\tau(q)$ denotes the sequence of mass exponents estimated by the slope of the fitted function $\ln(X_q(\varepsilon))vs\ln(\varepsilon)$. The value of q reflects the importance of the subset of the distribution ratio $Q_i(\varepsilon)$ in the partition function [57]. When $q \to +\infty$ the maximum probability plays a significant role. $X_q(\varepsilon)$ reflects the nature of the subset with the largest probability. On the other hand, when $q \to -\infty$ the minimum probability is dominant, representing the subset with the smallest probability. The multifractal dimension $D(q)$ is calculated as follows:

$$D(q) = \frac{1}{q-1}\lim_{\varepsilon \to 0}\frac{lnX_q(\varepsilon)}{ln(\varepsilon)}, (q \neq 1) \tag{9}$$

$$D(q) = \lim_{\varepsilon \to 0}\frac{\sum_{i=1}^{N(\varepsilon)} Q_i(\varepsilon) lnQ_i(\varepsilon)}{ln(\varepsilon)}, (q = 1) \tag{10}$$

$\tau(q)$ is zero when q = 1. A larger weight is assigned to cells having a larger amount of information (i.e., a greater mass) when q > 1; smaller weights are assigned to cells where the quantity to be measured is less concentrated (when q < 1). The singular exponent $\alpha(q)$ and the multifractal spectrum $f(\alpha(q))$ are deduced from the Legendre transform as follows:

$$\alpha(q) = -\frac{d(\tau(q))}{dq} \tag{11}$$

$$f(\alpha(q)) = q \cdot \alpha(q) + \tau(q) \tag{12}$$

where the smallest and largest values of the singular exponent α (i.e., α_{min} and α_{max}, respectively) constitute the minimum and maximum Lipschitz-Holder exponents, respectively, and $\Delta\alpha = \alpha_{max} - \alpha_{min}$ denotes the range of the multifractal exponent that quantitatively characterizes the degree of difference and the range of variation within the river network. A large value of $\Delta\alpha$ indicates the river network exhibits greater heterogeneity and complexity, and vice versa. The singular exponent α reflects the structural characteristics of the river network. The singular exponent changes with changes in the FAT. When the FAT is too large the extracted river network omits details in the river network; when it is too small a virtual "pseudo-river network" is extracted. The river network extracted with an optimal FAT captures the watershed-scale structure and does not include small, artificial, river networks. The optimal FAT corresponds to an extremal point of the line of the maximum singular exponent expressed as a function of the FAT, as shown in the Results section. $f(\alpha_{max})$ and $f(\alpha_{min})$ represent the maximum and minimum spectral values, respectively. $\Delta f = f(\alpha_{max}) - f(\alpha_{min})$ denotes the multifractal spectral elevation difference, which indicates the structural changes between subsets with different probabilities associated with the fractal dimension. $\Delta f > 0$ when the number of maximum probability subsets exceeds that of the minimum probability subsets. Similarly, $\Delta f < 0$ when the number of minimum probability subsets exceeds that of the maximum probability subsets. Furthermore, $\Delta f = 0$ when the number of maximum probability subsets is equal to that of the minimum probability subsets.

2.3. Study Area and Original Data

Two regions in China were selected as study areas to implement this paper's methods, namely the middle Yangzi River delta and the northeastern Tibetan Plateau. The Yangzi River delta lies in eastern China and encompasses 14 municipalities, covering an area equal to 95,400 km^2. It produces 37.6% of exports and attracts 41.8% of transnational investments, contributing 17.6% to the GNP of China, and harboring 7.6% of the Chinese population in 2010 [58]. Hubei Province (108°21′42″ to 116°07′50″ E, 29°01′53″ to 33°06′47″ N) is located in the middle part of the Yangzi River delta. It features numerous rivers and lakes, and thus has been called the "land of a thousand lakes". Most rivers in Hubei Province are part of the

Yangtze River fluvial network. In addition, this province is host to a variety of landforms, including mountains (about 56%), plains (about 20%), and hills (about 24%) [59]. Hubei Province has a sub-tropical, humid, and monsoonal climate. Its annual mean temperature ranges from 13 °C to 18 °C, and the average annual precipitation ranges between 800 and 1609 mm depending on location, most of which falls in the wet season [60]. The wet season ranges from May through September. The geographical distribution of vegetation in the Yangtze River Basin is heterogeneous. The climate, topography, and vegetation of Hubei Province interact to form the Yangtze River network [61,62].

The Tibetan Plateau is the source of several major rivers such as the Yangtze, Yellow, and LanCang rivers (the three rivers). Qinghai province (89°35′ to 103°04′ E, 31°9′ to 39°19′ N) lies in the northeastern part of the Qinghai-Tibet Plateau, with average elevation ranging from 3000 to 5000 m. The climate is dry and cold. There are many rivers and lakes in Qinghai Province, and its topography and climate patterns are complex [63]. The speed of industrialization and urbanization is accelerating. The locations of the study areas are shown in Figure 2.

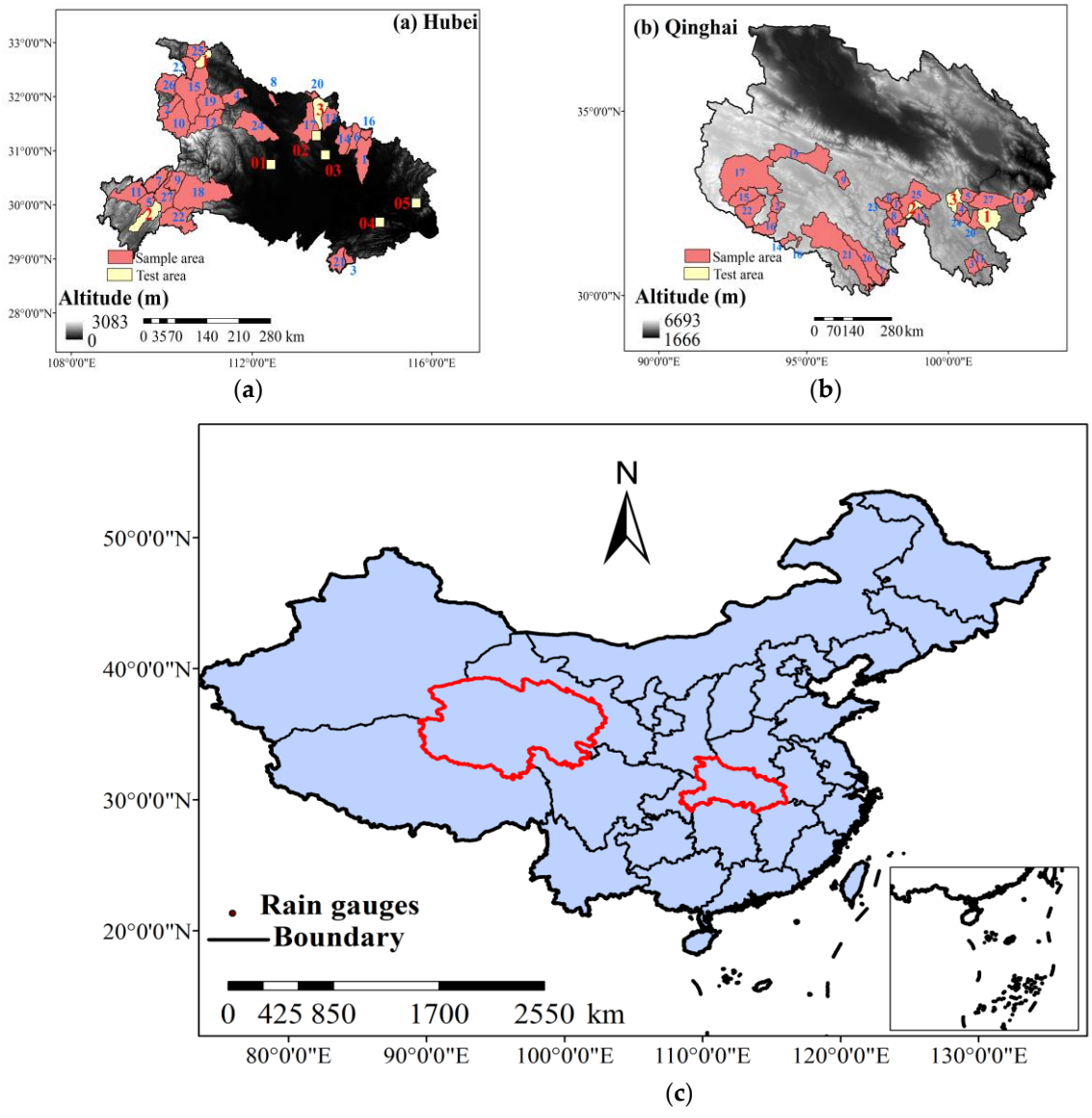


Figure 2. (**a**) Map of Hubei Province depicting 27 sample areas used for model development; (**b**) Map of Qinghai Province depicting 27 sample areas used for model development. (**c**) Map of China with the study areas.

The data used for model building span from 2001 to 2018 (May to September), and include ten main components: (1) Actual river network dataset: The 1:25,0000 reference river networks were obtained from the National Earth System Science Data Center, National Science and Technology Infrastructure of China (http://www.geodata.cn, accessed on 1 August 2020). The data were further corrected by remote sensing images (0.4-m spatial resolution from GeoEye) accessed from Google Earth. Read Yan et al. (2019) for the data-correction procedure [64]. (2) The normalized difference vegetation index (NDVI): The NDVI data used in this study were obtained from the MOD13A3 version 6 products (https://search.earthdata.nasa.gov/search, accessed on 1 August 2020), which provide monthly measurements at a 1 km spatial resolution. (3) Topographic data: The ALOS DEM with 12.5 m resolution, downloaded from the Alaska Satellite Facility (ASF) Distributed Active Archive Center (DAAC), was used to identify and extract drainage networks in the study area. The ALOS DEM from which maximum altitude, average altitude, maximum slope, average slope, maximum surface roughness, and average surface roughness were extracted has a resolution of 12.5 m at the equator. (4) Precipitation data: the data issued from the GPM_3IMERGM 06 product (https://disc.gsfc.nasa.gov/, accessed on 1 August 2020) with a resolution of 0.1°. (5) Surface runoff and evapotranspiration data: the data were provided by GLDAS_NOAH025_M 2.1 (https://disc.gsfc.nasa.gov/, accessed on 1 August 2020). Surface runoff and evapotranspiration data were extracted from the datasets with a resolution of 0.25°. (6) Surface temperature: the data acquired from the MOD11C3 product (https://search.earthdata.nasa.gov/search, accessed on 1 August 2020) had a spatial resolution of 0.05°. (7) Population data: The population distribution was provided by the LandScan^{TM} (2018) High-Resolution Global Population Data Set (https://landscan.ornl.gov/, accessed on 1 August 2020), which has a spatial resolution of approximately 1 km (30″ × 30″). (8) Luojia 1-01 Night-time Light Imagery: the images were jointly provided by Wuhan University and the Gaofen Hubei Center (http://www.hbeos.org.cn/, accessed on 1 June 2020) with a resolution of 130 m. Night-time lights mainly reflect human activities at night and include residential lighting, traffic lights, commercial lighting, and factory lighting. Therefore, night-time lighting area reflects the spatial distribution of residential, transportation, commercial, and factory lighting [65]. (9) Land use data: these data were acquired from the MCD12Q1 product (https://search.earthdata.nasa.gov/search, accessed on 1 June 2020) with a 500 m resolution and comprise seventeen types (see Table S1 in the online Supplemental Materials). The last five land-use types of Table S1 are not found in Hubei province. The evergreen broadleaf forests closed shrublands, and cropland/natural vegetation mosaics are not found in Qinghai province. Land pattern indicators were computed from the land use data. (10) Lithology describes the geochemical, mineralogical, and physical properties of rocks [66]. Global lithology data acquired from the Global Lithological Map Database v1.0 (http://dx.doi.org/10.1594/PANGAEA.788537, accessed on 1 July 2020) with a 0.5° resolution and comprised sixteen types. The Hubei study area includes seven types of soil and rock (su, ss, sm, sc, mt, pa , and py). The Qinghai study area includes twelve types of soil and rock (su, ss, sm, sc, mt, pa, pb, va, vi, wb, pi and ig). The full name of lithology is listed in Table S2 in the online Supplemental Materials. The data description with 47 features is listed in Table 1. The detailed description of each feature is displayed in Table S3.

2.4. Data Pre-Processing and Normalization

The topographic indexes were extracted with the spatial analyst toolbox of ArcGIS 10.6, namely, maximum terrain altitude, mean terrain altitude, maximum slope, mean slope, maximum surface roughness, and mean surface roughness. The surface area ratio grids are applied to measure the surface roughness [67]. The steps to calculate the surface roughness are as follows: (1) calculate the slope of DEM; (2) calculate the surface roughness = 1 /Cos ([angle of slope of DEM] * π/180) [68–70].

The population ratio (Q) is calculated with Equation (13). Two light indicators were constructed from Night-time Light Imagery: basin light area (L_1) and basin light intensity

(L_2), as shown in Equation (14) and Equation (15), respectively. Equation (16) shows how the digital numbers (DN) are converted to the radiance for accurate analysis of lighting brightness and discrepancy. Thirdly, landscape pattern indexes including landscape percentage, patch density, Shannon's diversity index, and Species evenness were computed by FRAGSTATS 4.2 [71]. Finally, all data were projected to the same coordinate system.

$$Q_i = \frac{U_i}{A} * 100 \tag{13}$$

where U_i denotes the population size of basin i. A denotes the population of the province.

$$L_I = \frac{S_N}{S} \tag{14}$$

where L_1, S_N, and S denote the urbanization sprawl, the lighting area of a basin, that is, the total area of lights in a basin, and the total area of a basin, respectively [65].

$$L_2 = n \cdot \frac{Sum(G_i)}{G_{max}} \tag{15}$$

where L_2, G_i, G_{max}, and n denote the urbanization intensity, the radiance of the i-th pixel in a basin (indicating light intensity of the i-th pixel), the maximum radiance of the basin, and the total number of lighted pixels in a basin, respectively [65].

$$G = 10^{-10} \cdot D^{\frac{3}{2}} \tag{16}$$

where G, and D denote the radiance of a pixel in the Night-time Light Image ($W/(m^2 \cdot sr \cdot \mu m)$) and the Digital Number (DN) of a pixel [72].

Z-score normalization was adopted for comparison purposes to render the data dimensionless according to Equation (17):

$$X = \frac{x - \mu}{\sigma} \tag{17}$$

where X, μ, and σ denote the random variable that takes values equal to x, the mean value, and the standard deviation of the independent or dependent variable X, respectively.

2.5. Sample and Test Areas

The sample and test areas feature unique morphological characteristics with variable elevations, slopes, surface roughness, landform types, and network morphology. This study employs 27 sample areas in Hubei and Qinghai Provinces for model building. The sample and test areas are depicted in Figure 2a,b, respectively. These areas were obtained from a global dataset (https://doi.org/10.6084/m9.figshare.8044184.v6, accessed on 1 July 2020) with a 3-arcsecond resolution which has higher precision than the HDMA dataset released by the USGS [64].

The test areas are employed for model testing under a wide range of conditions. There are eight and three test areas in Hubei and Qinghai Provinces, respectively. The NDVI and average elevations of test areas of Hubei Province are significantly lower than those of Qinghai Province. Of the eight test areas of Hubei three serve to test the determination of the FAT based on an empirical scaling equation and statistical regression. These three independent sub-basins are depicted as yellow polygons in Figure 2a. The average elevation of test area 2 is above 1000 m a.s.l., which is significantly higher than those of the other test areas. The average elevations of test area 1 is near 500 m a.s.l., and has fewer mountainous features than test area 2. Test area 3 encompasses plain landform, with average elevation near 200 m a.s.l., and features relatively mild slopes compared with other test areas of Hubei Province. The other five test areas in Hubei Province were obtained with the fishnet tool of ArcGIS 10.6. Each has an area equal to 400 km^2 area. These

five test areas are used for the evaluation of the FAT algorithm at the regional scale and are shown as yellow rectangles in Figure 2a. The test areas in Qinghai Province are depicted as yellow polygons in Figure 2b. The average elevations of test areas (Figure 2b) are over 4400 m a.s.l. The test areas in Qinghai Province are used for evaluating the FAT model based on an empirical scaling equation and statistical regression.

2.6. Validation Method

This paper compares the extracted with the actual river network in each area based on three validation criteria. One is the modified matching algorithm based on the Frobenius norm [73] and SSEI. The steps to implement this criterion are as follows:

(1) Image pre-processing: Carrying out image binarization with 0 and 1 representing background (i.e., non-river) area and the river network, respectively;

(2) The Frobenius norm of the actual river network is calculated;

(3) Difference matrices between the actual and extracted river networks are calculated with several algorithms. This is followed by a calculation of the Frobenius norm of each difference matrix.

(4) The evaluation score is determined by the Frobenius norm calculated in step (3) divided by the Frobenius norm of the actual river network. The SSEI is given by Equation (18):

$$score(i) = \frac{norm(a-i)}{norm(a)} \tag{18}$$

where score(i) denotes the similarity between the actual and extracted river networks in basin i; a represents the binary matrix of the actual river network and i is that of the extracted river network calculated by different algorithms, norm(a) denotes the value of the Frobenius norm of the actual river network, and $norm(a-i)$ denotes the Frobenius norm of the matrix difference between the actual and the extracted network in basin i. The similarity between the actual and the extracted river networks increases with decreasing value of the SSEI.

The second evaluation criterion is the percentage error of the river length, which is given by the following equation:

$$err(l) = \frac{|l_e - l|}{l} \cdot 100\% \tag{19}$$

where $err(l)$, l_e, and l denote the percentage river length error, the total length of rivers within the extracted river network, and the total length of the rivers in the actual river network. The $err(l)$ is a quantitative criterion well suited to assess the accuracy of extracted river networks. It is herein supplemented with visual inspection by comparison of the extracted and actual rivers networks as shown in the Results section.

The third evaluation criterion is the root mean square error (RMSE), which is expressed as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(l_e - l)^2}{N}} \tag{20}$$

where N denotes the total number of rivers. The l_e and l have been defined in Equation (19).

3. Results and Discussion

3.1. Estimation of the FAT by Empirical Scaling Equation and Statistical Regression at the Basin Scale

3.1.1. Estimation of the River Length

Five characteristics were selected from among the 40 characteristics considered as predictive factors to predict the river length of Hubei province (see a complete list in Table S3) by multiple stepwise regression. The original data of Qinghai Province are listed in Table S4 of the online Supplemental Materials. The estimation of river length is based on Equation (21):

$$l_{th} = 456233.597 \times Q - 144667.232 \times \mathrm{LTD} - 483.984 \times \overline{\mathrm{ALT}} + 73147.410 \times \mathrm{SR_{max}} + 3548.086 \times \mathrm{EVP} + 3551837.787 \quad (21)$$

where l_{th} denotes the estimated length of rivers in Hubei province (in m) based on the environmental factors within the basins. The meanings of the symbols of the predictive factors listed on the right-hand side of Equation (21) are found in Table 1. The coefficient of determination (R^2) associated with the regression Equation (21) equals 0.902.

Four characteristics were selected from among the 40 characteristics (Table S3) considered as predictive factors to predict the river length of Qinghai province according to multiple stepwise regression. The estimation of river length is based on Equation (22):

$$\mathrm{l_{tq}} = 5366498.512 \times \mathrm{SHDI} - 45851158 \times \mathrm{SEI} + 2242448.637 \times \overline{\mathrm{SR}} + 0.001 \times \mathrm{L_2} - 2150428.604 \quad (22)$$

where $\mathrm{l_{tq}}$ denotes the estimated length of rivers in Qinghai province (in m). The meaning of the symbols and parameters can reference Equation (21). The coefficient of determination ($\mathrm{R^2}$) associated with the regression Equation (22) equals 0.725.

The formation of the river networks is governed by physical-geographic characteristics. Hubei Province has a sub-tropical and humid climate with relatively high temperatures in the wet season (May through to September). It features a relatively high population density. Evapotranspiration, surface temperature, and population ratio play a significant role in the formation of river networks (see Equation (21)). The climate of Qinghai Province, on the other hand, is dry and cold; its population density is relatively sparse ($319/\mathrm{km^2}$) compared with that of Hubei province ($7.80/\mathrm{km^2}$). Equations (21) and (22) indicate that anthropogenic factors and surface roughness are useful predictors of river length. The regression factors vary from region to region, in general.

3.1.2. Estimating the FAT

The river networks were extracted from DEMs with the ArcHydro tools of ArcGIS 10.6 software within basins with different FATs and calculated the lengths of corresponding extracted river networks. Graphs relating the drainage density to the FAT in Hubei Province are shown in Figure 3a–c. An interval equal to 1% was applied to calculate the percentile of flow accumulation. Notice the fitted equations in Figure 3a–c take the form $\delta = KFAT^{-b}$ with an average of $b = 0.50$ as implied by Equation (2). The optimal FAT was estimated by the equations fitted to each basin shown in Figure 3a–c based on the estimated river length of Equation (21). The optimal FAT is given by Equation (23):

$$FAT_{optimal} = \sqrt[b]{\frac{A_i \times K}{l_i}} \quad (23)$$

Figure 3d–f depicts the captured and missed streams in test areas 1, 2, and 3 in Hubei Province.

Graphs relating the drainage density to the FAT in Qinghai Province are shown in Figure 4a–c. An interval equal to 1% was applied to calculate the percentile of flow accumulation. Notice the fitted equations in Figure 4a–c take the form $\delta = KFAT^{-b}$ with an average of $b = 0.50$ as implied by Equation (2). The optimal FAT was estimated by the equations fitted to each basin shown in Figure 4a–c based on the estimated river length of Equation (22), in a manner analogous to the approach followed with Hubei Province areas. Figure 4d–f depicts the captured and missed streams in Qinghai Province.

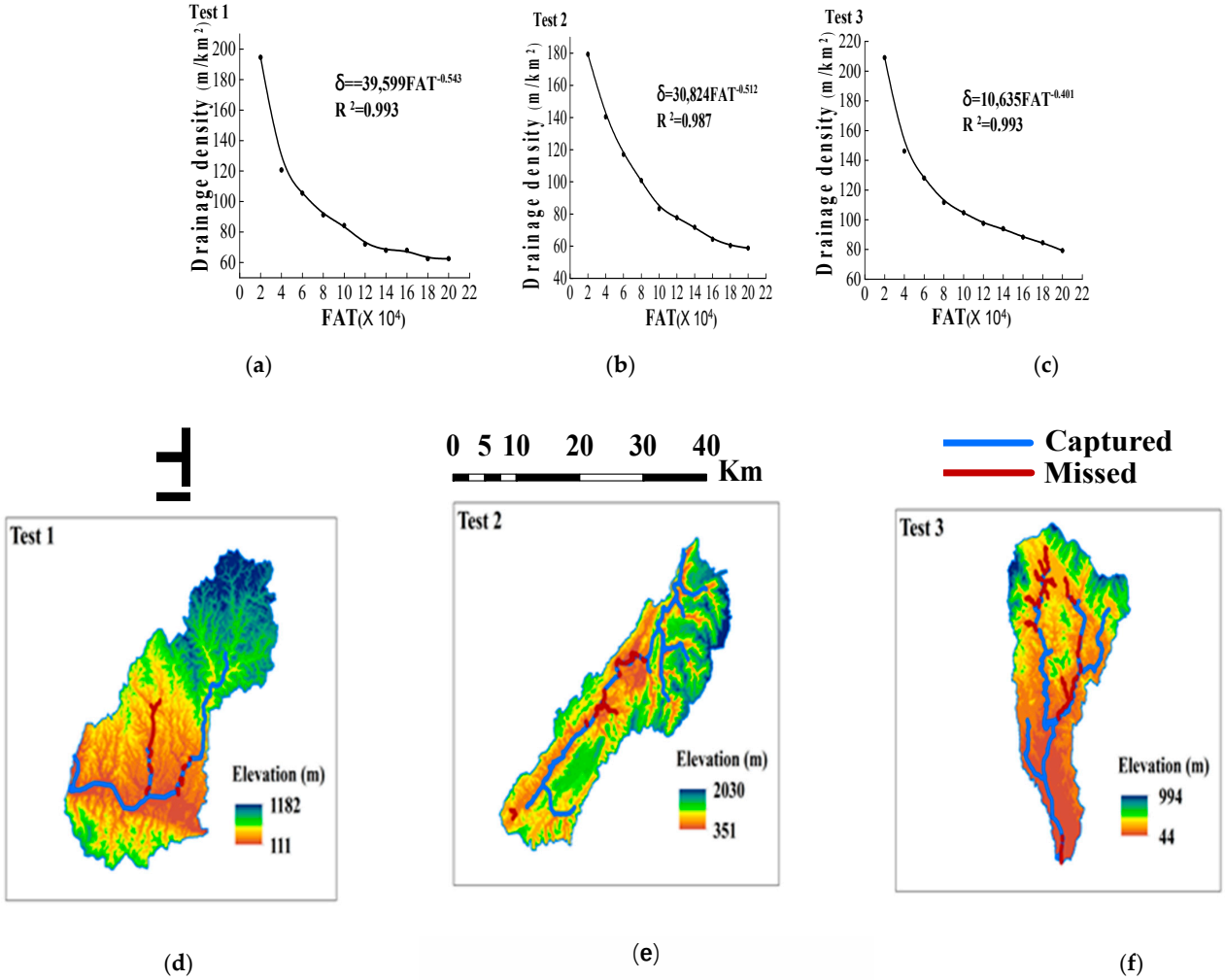


Figure 3. Maps of selected river networks in Hubei Province. (**a**) Scaling equation for test area 1. (**b**) Scaling equation for test area 2. (**c**) Scaling equation for test area 3. (**d**) Map of captured and missed streams in test area 1. (**e**) Map of captured and missed streams in test area 2. (**f**) Map of captured and missed streams in test area 3.

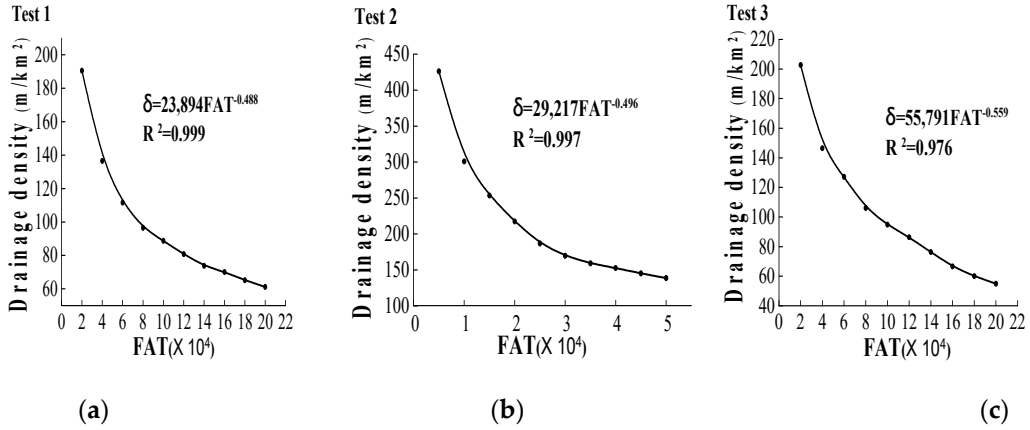
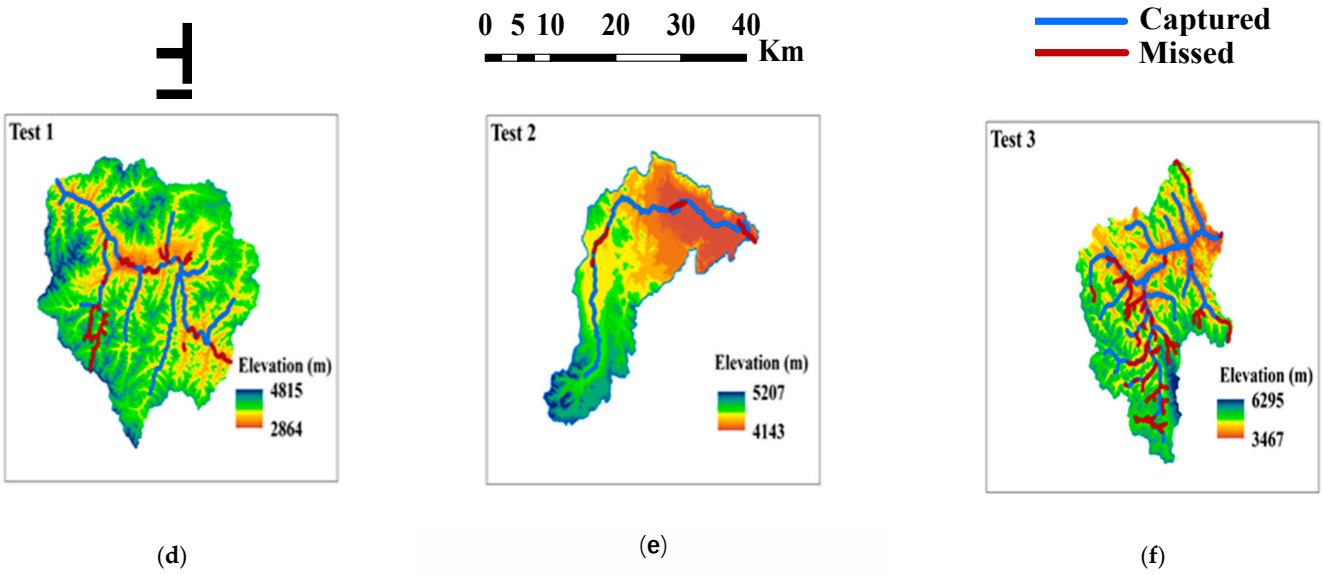


Figure 4. *Cont.*



(**d**) (**e**) (**f**)

Figure 4. Maps of selected river networks in Qinghai Province. (**a**). Scaling equation for test area 1. (**b**). Scaling equation for test area 2. (**c**). Scaling equation for test area 3. (**d**). Map of captured and missed streams in test area 1. (**e**). Map of captured and missed streams in test area 2. (**f**). Map of captured and missed streams in test area 3.

3.1.3. NMF

The NMF method was applied to reflect the importance of different features in Equation (21) (for Hubei Province) and Equation (22) (for Qinghai Province). Taking Equation (21) as an example, the NMF quantifies the strength of association between factors (see a list in Table 1) and river length. Environmental, human, and topographic factors have multiple interactions and influence river length. Matrix decomposition in NMF performs a mapping of the original matrix to a subspace. The mapping is performed by approximating the low rank of the original matrix. $\overline{\text{ALT}}$, SR_{max}, LTD, Q, l, and EVP data from 27 sample sub-basins were normalized by means of Equation (17). M denotes a 27 row $\times$ 6 column matrix:

$$\text{M} = \begin{pmatrix} \overline{ALT_1} & \cdots & EVP_1 \\ \vdots & \ddots & \vdots \\ \overline{ALT_{27}} & \cdots & EVP_{27} \end{pmatrix} \tag{24}$$

The mean value of the NMF was calculated 1000 times and the relations between the factors were expressed by a heat map (Figure 5). The third row of Figure 5a shows the influence of SR_{max} on l is relatively weak in comparison to other factors. Thus, the factor SR_{max} was excluded from the original matrix, and NFM was performed on the remaining 5 factors of the matrix. The last two rows of Figure 5b denote the influences of $\overline{\text{ALT}}$ on l is less significant, and NMF was performed on the remaining 4 factors, excluding $\overline{\text{ALT}}$. Figure 5c shows that EVP is closely related to the LTD (see the first row), and river length exhibits a high correlation with EVP in the last row, which demonstrates the relatively large contribution of factor EVP. Overall, EVP, Q, and LTD ranked among the top three factors in terms of importance. The same procedure was repeated to the factors of Equation (22). L_2 ranked as the top factor in terms of importance. The introduction of the anthropogenic factor significantly enhanced both models' accuracy. The NMF clarifies the identification of key factors employed in Equations (21) and (22).

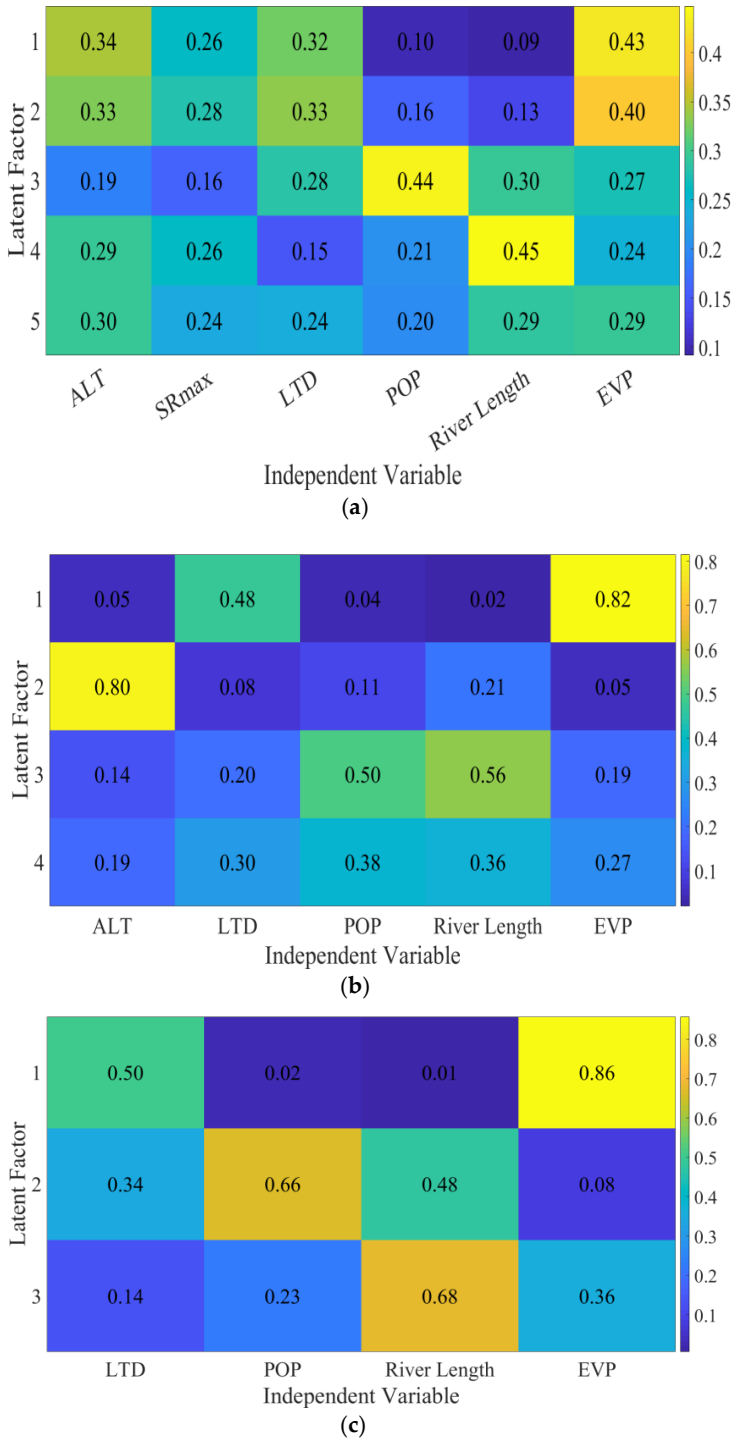


Figure 5. Non-negative matrix factorization (NMF) heat maps. (**a**) Heat map of six NMF factors. (**b**) Heat map of five NMF factors. (**c**) Heat map of four NMF factors.

The correlations between actual drainage density in the sample areas and rock types are shown in Table 2 (taking Hubei province as an example), where the areas are arranged and ranked in descending order by the magnitude of δ. Rock type which covers less than 30% of sub-basin areas was not included as a dominant factor. δ is largest for the carbonate sedimentary rocks and unconsolidated sediments compared to other rock types in Hubei province. δ is the largest for the mixed sedimentary rocks in Qinghai province.

Table 2. Calculated drainage density (δ) and rock type in the sample areas in Hubei Province (see Table 1 for symbology).

Ranking	Drainage Density δ (m/km^2)	Dominant Rock Type	Ranking	Drainage Density δ (m/km^2)	Dominant Rock Type
1	246.922	sc, su	15	119.891	Mt
2	230.128	Su	16	118.144	Sm
3	228.912	Sc	17	115.381	Sm
4	221.086	Sm	18	113.214	Sm
5	191.177	Sc	19	104.364	su, sm
6	158.465	su, ss	20	97.355	Sm
7	145.309	Py	21	95.585	sm, sc
8	145.045	sm, sc	22	93.063	Sm
9	144.632	sm, sc	23	91.383	su, ss, pa
10	144.534	Pa	24	89.913	su, py
11	139.167	Sm	25	83.102	Sm
12	135.286	sm, sc	26	73.790	Py
13	133.594	sm, sc	27	53.544	Sm
14	122.085	Sm			

3.1.4. Validation

The SSEI was applied to measure the extracted river network with the actual river network based on image similarity. The extracted river network was calculated by the ArcHydro Tools of ArcGIS 10.6 based on the estimated optimal FAT. The performance of the drainage density-FAT model was tested statistically, with results for Hubei Province listed in Table 3. The image similarity of test area 2 is the highest with a low SSEI (0.356) while that of test area 1 is the lowest with a relatively high SSEI (0.441). The results evaluated by means of the river length error do not generally coincide with those from the SSEI. The smallest length error of the extracted river network was 1.17% (corresponding to test area 1) and the average length error was 4.89%. All length errors were smaller than 15%. The RMSE of extracted river length equaled 1.109×10^4 m. Overall, the three indexes (SSEI, $\mathrm{err}(l)$, RMSE) demonstrated that the drainage density-FAT model performed well in basins of Hubei Province.

Table 3. Statistical criteria values for the comparison of the calculated and actual river length in Hubei province.

Area	Actual River Length (m)	Estimated River Length (m) see Equation (21)	FATs	Extracted River Length (m)	River Length Error (%)	SSEI
1	62,109	61,323	101,237	60,050	1.17	0.441
2	151,113	161,006	80,916	169,379	10.34	0.356
3	176,724	176,287	54,547	171,143	3.16	0.400
Average					4.89%	

Table 4 lists the statistical criteria values for the comparison of the calculated and actual river length in test areas of Qinghai Province (Figure 2b). The smallest length error of the extracted river network was 3.56% (corresponding to test area 1) and the average length error was 11.44% in test areas of Qinghai Province. The RMSE of extracted river length equaled 2.135×10^4 m. The image similarity of test area 3 is the highest with a low SSEI (0.478), while that of test area 1 is the lowest which does not coincide with the results of length error. Thus, a comprehensive evaluation must consider the image similarity and length error of river networks.

Table 4. Statistical criteria values for the comparison of the calculated and actual river lengths in Qinghai Province.

Area	Actual River Length (m)	Estimated River Length (m) see Equation (21)	FATs	Extracted River Length (m)	River Length Error (%)	SSEI
1	302,294	322,167	190,275	291,539	3.56	0.622
2	95,765	127,176	140,641	119,354	24.63	0.583
3	429,522	409,901	41,356	403,155	6.14	0.478
Average					11.44%	

Tables 3 and 4 provide evidence that the river length error in test area 2 is larger than in areas 1 and 3. This may be due to the fact that the average altitude of test area 2 is higher than those of areas 1 and 3 in the Hubei and Qinghai provinces. The accuracy of regression factors significantly influences the performance of the drainage density-FAT method. The satellite products that are bias-corrected are more accurate than the satellite-only products because the meteorological stations are densely distributed in flat areas. The performance of the MODIS satellite products is relatively poor in the high-altitude area with complex terrain [74]. Moreover, previous studies have shown that the ecosystem exhibits greater vulnerability [75] with increasing altitude, which may influence the land pattern indicators.

In addition, the drainage density-FAT method was compared with Lin's method, known as the Fitness Index method hereafter [19], the OnePercent method [20], and the Mean method [21]. Choosing test area 2 of Qinghai Province as an example, the FAT was estimated by the Fitness Index method at 135,000. The FAT was calculated as 1,884,010 × 0.01 = 18840.1 based on the maximum accumulation value according to the OnePercent method. The mean accumulation value for the FAT was 1908 according to the Mean method. Table 5 indicates that a drainage network of acceptable accuracy was obtained by the drainage density-FAT method. In addition, the drainage density-FAT method is less computationally burden than the Fitness Index method.

Table 5. The statistics for four river networks derived with four flow accumulation threshold (FAT) estimates (test area 2 of Qinghai Province).

Method	FAT	Length Error (%)	Accuracy
drainage density-FAT	140,641	24.63	High
Fitness index	135,000	25.44	High
OnePercent	18,840	87.96	Low
Mean	1908	504.06	Very Low

The 10-fold cross-validation [76] was applied to quantify the variation in performance due to the choice of sample and test cases (choosing Qinghai province as an example). The results indicate the average length error estimated by different models ranges from 11.42% to 30.35%. Most river length errors are within 25%. The model accuracy would slightly fluctuate with the change of sample and test datasets, but its accuracy remains acceptable.

Next, test area 2 of Qinghai province was discarded from the dataset. The 10-fold cross-validation was applied to the remaining dataset. The river length of test area 2 (Qinghai Province) was calculated by different models. The results indicate that the regression factors L_2, SHDI, and SEI play a significant role in ten models, which is consistent with the factors selected by Equation (21). The parameters of the objective equation changed with different training groups, yet, the average length error estimated by models only slightly changed from 24.63% (estimated by Equation (21)) to 22.37%. This robust performance confirms the accuracy of the drainage density-FAT model.

3.2. Estimation of the FAT at the Regional Scale (Hubei Province)

Fractal geometry methods are applied in this study to estimate the optimal FAT by considering the morphologic characteristics of river networks (using Hubei Province data), followed by an evaluation of the corresponding results.

3.2.1. Estimation of the Optimal FAT with the Box-Counting Method

Box-counting was applied to ascertain the relation between the fractal dimension and the FAT (see Table 6). The fractal dimension of river networks whose cube sizes r equaled 500, 1000, 2000, 5000, 10,000, 15,000, and 20,000 were calculated. The values of r and FAT used in this study may be not suitable for other regions. The selection of the r and FAT data must be based on the river network data's statistical characteristics (i.e., drainage density, area). The variable difference between chosen FATs (e.g., 500, 1000, 2000, 5000, etc.) does not alter the box-counting method results. The fractal dimension of the river network declined from 1.851 to 1.432 as the FAT increased respectively from 1000 to 25,000, implying the fit between the fractal dimension and the FAT was excellent. It is seen in Table 6 the rate of change of the fractal dimension stabilizes at 10,000, and the change rate diminished more rapidly when the FAT is equal to or larger than 15,000. Thus, the optimal FAT is estimated to be close to 10,000 for the entire study area.

Table 6. Fractal parameters obtained with the box-counting method (Hubei Province).

FAT	Fractal Dimension (D_f)	R^2	Change Rate (d_i)	FAT	Fractal Dimension (D_f)	R^2	Change Rate (d_i)
1000	1.851	0.999	0.0072	8000	1.594	0.994	0.0017
2000	1.779	0.997	0.0049	9000	1.577	0.994	0.0015
3000	1.730	0.996	0.0041	10,000	1.562	0.993	0.0011
4000	1.689	0.995	0.003	15,000	1.507	0.994	0.00084
5000	1.659	0.995	0.0025	20,000	1.465	0.994	0.00066
6000	1.634	0.994	0.0021	25,000	1.432	0.994	
7000	1.613	0.994	0.0019				

According to Figure 6, the fractal dimension decreases with increasing FAT. When the FAT increases, the slope of the function decreases. The fractal dimension of the reference river network (i.e., the river network) was estimated to be 1.543 which corresponds to a FAT equal to 10,000. This is the ascertained optimal FAT of the study area established with the box-counting method.

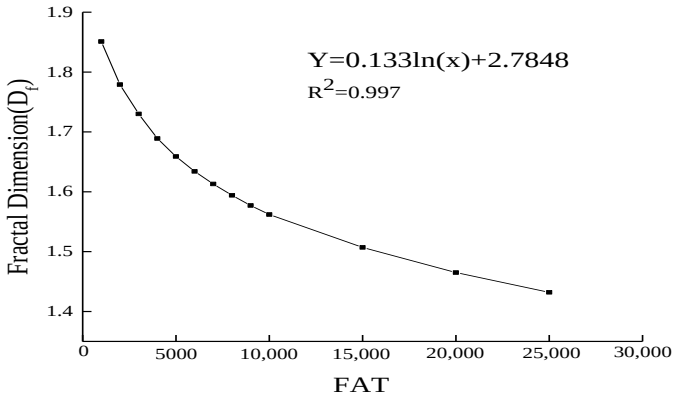


Figure 6. Graph of the fractal dimension vs. the FAT (Hubei Province) where Y and X denote the fractal dimension and the FAT, respectively.

3.2.2. Estimation of the Optimal FAT with Multifractal Analysis

River networks were extracted with FATs set equal to 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000, 15,000, 20,000, 25,000. The distribution index $Q(\varepsilon)$ was calculated based on box sizes 500 × 500, 1000 × 1000, 2000 × 2000, 5000 × 5000, 10,000 × 10,000, 15,000 × 15,000, and 20,000 × 20,000. The q-th order moment was set in the range [−3,3], and the step Δq was 0.5.

Figure 7 displays thirteen sequences of mass exponents $\tau(q)$, thereby illustrating $\tau(q)$ is a convex function, which indicates there is a nonlinear function between $\tau(q)$ and q. Therefore, the river network in Hubei Province exhibits multiple fractal characteristics.

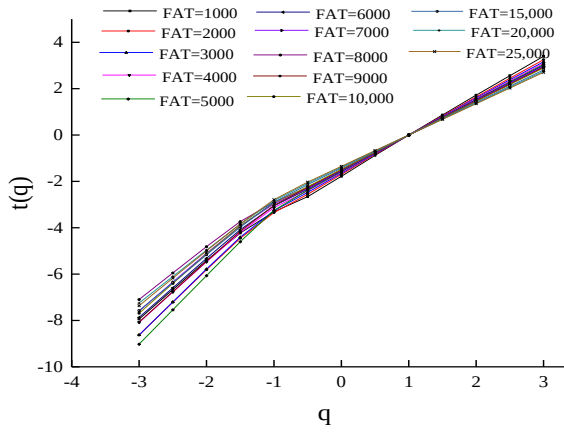


Figure 7. $\tau(q)$ vs. q functions corresponding to FAT values (Hubei Province).

The maximum singular exponent α_{max} exhibited a decreasing trend with increasing FAT. Figure 8 depicts a graph of the change of α_{max} vs. FAT. It is seen in Figure 8 there are three extremal points corresponding to FATs equal to 3000, 5000, and 9000. Thus 3000, 5000, and 9000 are possible optimal FATs of Hubei province. The optimal FAT was further estimated by the range of the singularity exponent $\Delta\alpha$. The results listed in Table 7 indicate $\alpha_{min} \in$ [0.926,1.448], $\alpha_{max} \in$ [2.033,2.937], $f(\alpha_{max}) \in$ [0.209,1.317], and $f(\alpha_{min}) \in$ [−0.239,0.951]. $\Delta\alpha$ describes the degree of heterogeneity of the probability distribution of river networks. It is demonstrated below there is no significant difference between the probability density distributions of the actual river networks in Hubei Province, that is, their distributions are similar. Comparing the $\Delta\alpha$ associated with FATs equal to 3000, 5000, and 9000, shows the $\Delta\alpha$ changes significantly with changing FAT. $\Delta\alpha$ is smaller when the FAT = 9000 than when it equals 3000 or 5000. A FAT = 9000 is estimated as being optimal within the entire study area given the heterogeneity of river networks in Hubei province.

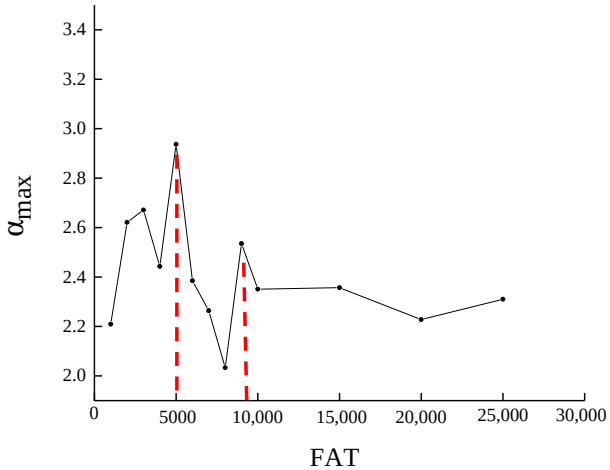


Figure 8. The change of α_{max} vs. FAT (Hubei Province).

Table 7. Calculated parameters of the multifractal method (Hubei Province).

FAT	α_{min} (1)	$f(\alpha_{min})$ (2)	α_{max} (3)	$f(\alpha_{max})$ (4)	$\Delta\alpha$ (5)	$\Delta f(\alpha)$ (6)
1000	1.448	0.951	2.209	1.317	0.761	−0.366
2000	1.195	0.337	2.621	0.761	1.426	−0.424
3000	1.109	0.167	2.671	0.614	1.562	−0.447
4000	1.152	0.362	2.443	0.754	1.291	−0.392
5000	0.935	−0.239	2.937	0.209	2.002	−0.448
6000	1.111	0.333	2.385	0.715	1.274	−0.382
7000	1.125	0.408	2.264	0.807	1.139	−0.399
8000	1.196	0.650	2.033	0.999	0.837	−0.349
9000	0.995	0.071	2.535	0.467	1.540	−0.396
10,000	1.034	0.208	2.351	0.635	1.317	−0.427
15,000	0.989	0.162	2.357	0.506	1.368	−0.344
20,000	0.998	0.250	2.228	0.565	1.230	−0.315
25,000	0.926	0.077	2.310	0.433	1.384	−0.356

(1) minimum singular exponent; (2) fractal spectrum corresponding to αmin; (3) maximum singular exponent; (4) fractal spectrum corresponding to α_{max}; (5) range of the singularity exponent; (6) multifractal spectral elevation difference.

Figure 9 demonstrates that multifractal spectral curves of the river networks have similar shapes with different thresholds. The multifractal spectrum $f(\alpha)$ first shows an increasing trend with increasing α, followed by a decreasing trend $\Delta f(\alpha) < 0$ (see Table 7). These results reveal the number of minimum probability subsets in the river networks in Hubei Province exceeds that of the maximum probability subsets. Therefore, the minimum probability subsets are dominant, indicating the distribution of differences between the river networks in all directions is small and the river networks are relatively homogeneous in most areas. This result is consistent with the actual conditions in Hubei province.

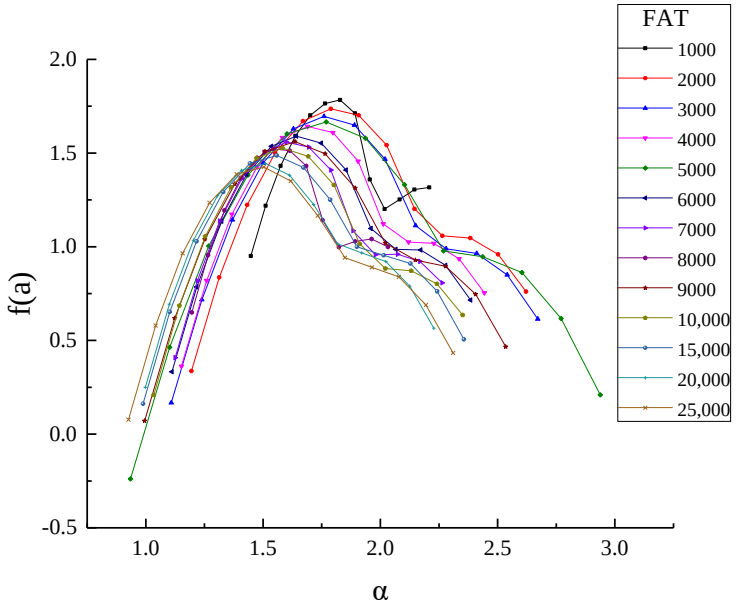


Figure 9. $f(\alpha)$ vs. α functions corresponding to several FAT values (Hubei Province).

3.2.3. Validation

Five test areas were chosen to further verify the accuracy of the optimal FAT estimated by the box-counting and the multifractal method (the test areas are depicted in Figure 2a with the rectangular shape).

(1) Similarity score

Table 8 displays the SSEI calculated by two algorithms which shows the degree of image matching between the actual and the extracted river networks. The river network

extracted with the multifractal method achieved greater similarity with the actual river network than the river networks extracted with the box-counting method, featuring a lower similarity score than box-counting except in test area 02.

Table 8. Similarity score and river network error corresponding to Hubei test areas obtained with the box-counting and multifractal methods (Hubei Province).

Test Area	Similarity Score (SSEI)		River Length Error (err(l)) (%)	
	Box-Counting Method	Multifractal Method	Box-Counting Method	Multifractal Method
01	1.0394	1.0334	21.49	16.49
02	1.0387	1.0395	20.13	18.42
03	1.2207	1.2202	26.53	20.15
04	1.0247	1.0223	36.46	29.26
05	1.3728	1.3674	26.28	22.90

(2) River length error

The percentage errors associated with the river length among the 5 test areas with box-counting and multifractal models are listed in Table 8. In terms of the box-counting model, the err(l) of test areas 03, 04, and 05 were over 25%. Compared with the average err(l) (26.18%) and RMSE (2.728×10^4 m) of box-counting, the average err(l) of the multifractal model's calculation of the river length of the five test areas equaled 21.44%, and the RMSE of the river length equaled 2.222×10^4 m.

The fractal dimension reflects the morphologic characteristics of basins and indicates the degree of land erosion and river network evolution. The larger the fractal dimension is, the more advanced the evolutionary stage of a river network. The fractal dimension calculated in this study was in the range of $1.432 \leq D_f \leq 1.851$. The method for classifying erosional stages proposed by He and Zhao (1996) suggested the basins of Hubei Province are in their early stage of erosion [77]. Lateral and down-cutting erosion played a central role during this stage.

Overall, the image similarity reflected by SSEI and err(l) indicate the modified multifractal performed better in estimating the optimal FAT than the modified box-counting model at the regional scale.

3.3. Discussion

This work's results provide a theoretical basis for the accurate determination of the FAT and river network extraction. The D8 algorithm is relatively simple and convenient and has found wide applicability in hydrologic analysis. Its popularity stems from its simplicity and reasonable representation for convergent flow conditions, and it preserves the consistency between the flow patterns, calculated contributing area, and spatial delineation of sub-catchments [44,78]. Thus, it was applied in this study. Other algorithms (e.g., D∞ and MD∞) may improve the calculation of flow directions on a topographic surface, and their evaluation is the topic of future research. These authors will evaluate the performance of the multifractal method at the basin scale in future work.

4. Conclusions

The hydrologic analysis must produce extracted drainage networks with accuracy for numerous applications. Depending on the intended application the FAT is not unique. For other purposes, drainage networks at a more or less detailed level may be needed. This study extracted river networks with suitable accuracy with the drainage density-FAT and multifractal methods. The drainage density-FAT method is a quantitative procedure that considers the physiographic and geomorphologic features of basins. The fractal method can be seen as complementary to the drainage density-FAT method when the large areal extent precludes the use of basin-wide representative regression predictors. The methods'

performances were evaluated in this work with statistical indexes and visual inspection, and compared with several FAT determination methods with data from Hubei and Qinghai Provinces, China. These authors conclude the density-FAT and the multifractal methods are useful for FAT determination. Based on this study's results it is concluded that:

The drainage density-FAT method performed well by fully considering possible environmental, topographic, and human activity factors, and selected the significant regression factor from the 47 factors in basins where the regression factors are representative of basin-wide characteristics. This method can be used to estimate the optimal FAT in areas where basin-wide regression factors are representative for the purpose of river network extraction. River initiation depends on environmental and topographical characteristics. The best regression factors in a region might not be the best in other regions due to differences in topography, climate, and human activities. Anthropogenic factors improve the extraction accuracy of river networks in urbanized regions. The topographic characteristics, specifically, the surface roughness, play a significant role in the estimation of FAT. The regression factors of the extraction models account for basin-wide features and can be obtained from accurate satellite datasets.

There are also limitations with the proposed drainage density-FAT method that may lead to systematic or random errors in extracting river networks. First, the model based on empirical scaling equation and statistical regression was developed in areas with abundant water resources (Hubei Province) and in headwater areas (Qinghai Province). The statistical association between river length and model parameters applied in this study may not always be represented accurately in dry areas or in non-urbanized regions. Second, the satellite datasets may introduce random errors in building the regression model used to estimate the FAT (errors that may arise from extreme weather and radiation-limited ecosystems). The drainage density-FAT model may not be applicable at large spatial scales when regression factors are not basin-wide representatives.

The fractal geomorphology is efficient in the estimation of the FAT, overcoming the limitation of the drainage density-FAT method at the regional scale. This work investigated the performance of two classical fractal methods on the estimation of the FAT. The multifractal method featured a better performance than the box-counting method. The parameters of the multifractal spectrum were applied to describe the hierarchical structure of river networks and highlight the variability of characteristics across study areas. The accuracy of fractal geomorphology in relatively small basins is not addressed in this study. This constitutes the theme of future work.

The methods proposed in this study are applicable with DEMs of resolutions other than the one used in this work. This work indicates, however, that river networks extracted from very high-resolution DEMs are the most accurate [30,36,79]. Experimental results using data from Hubei and Qinghai Provinces, China, provide us with insights as to how the regression and multifractal methods may be applied and expected to perform in other regions with spatially variable topographic and environmental characteristics. The drainage density-FAT method is applicable to dendritic networks, meanders, braided networks, and trellis networks in urbanized basins where the regression factors are representative of basin-wide characteristics. The drainage density-FAT method presented in this paper may have limitations in the extraction of coastal river networks. Complex environmental phenomena such as hydrologic alteration, landscape fragmentation, coastal cutting and filling, land subsidence, sea-level rise, and hurricane activity may introduce uncertainties in extracting coastal river networks [80]. Nevertheless, the accuracy of extracted river networks may be high in some instances by means of the density-FAT method applied to coastal regions. This paper applied the Frobenius norm matching algorithm, the percentage error of the river length combined with a conventional index (root mean square error) by the mining of the contents of images, and the river length error to evaluate the accuracy of extracted river networks. A thorough procedure for the evaluation of river-network extraction methods must consider image similarity and the length error of river networks, as herein demonstrated.

Supplementary Materials: The following are available online at https://www.mdpi.com/2220-9964/10/3/186/s1, Table S1: The land use types of MCD12Q1 products. Table S2: The full name and corresponding abbreviation of lithology datasets. Table S3: All factors extracted from the original data (/= dimensionless). Table S4: All factors of Qinghai Province for estimating the river length and the FAT.

Author Contributions: Formal analysis, HuiHui Zhang; Software, HuiHui Zhang; LuWei Feng; Validation, HuiHui Zhang; Visualization, HuiHui Zhang; Jing He; Writing—original draft, HuiHui Zhang; Writing—review and editing, Hugo A. Loáiciga and QingYun Du All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key R&D Program of China, grant number 2016YFC0803106".

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data and codes that support the findings of this study are available with the identifier(s) at the private link: (https://figshare.com/s/795ea26a3baa91e8dd69) (accessed on 10 March 2021).

Acknowledgments: This work was partly supported by the China Scholarship Council (NO.201806270111).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yang, B.; Ren, L.L. Identification and comparison of critical support area in extracting drainage network from DEM. *Water Resour. Power* **2009**, *27*, 11–14.
2. Lindsay, J.B.; Dhun, K. Modelling surface drainage patterns in altered landscapes using LiDAR. *Int. J. Geogr. Inf. Sci.* **2015**, *29*, 397–411. [CrossRef]
3. Ariza-Villaverde, A.B.; Jimenez-Hornero, F.J.; De Ravé, E.G. Multifractal analysis applied to the study of the accuracy of DEM-based stream derivation. *Geomorphology* **2013**, *197*, 85–95. [CrossRef]
4. Garneau, C.; Sauvage, S.; Sánchez-Pérez, J.M.; Lofts, S.; Brito, D.; Neves, R.; Probst, A. Modelling trace metal transfer in large rivers under dynamic hydrology: A coupled hydrodynamic and chemical equilibrium model. *Environ. Model. Softw.* **2017**, *89*, 77–96. [CrossRef]
5. Obida, C.B.; Blackburn, G.A.; Whyatt, J.D.; Semple, K.T. River network delineation from Sentinel-1 SAR data. *Int. J. Appl. Earth Obs. Geoinform.* **2019**, *83*. [CrossRef]
6. Wu, T.; Li, J.; Li, T.; Sivakumar, B.; Zhang, G.; Wang, G. High-efficient extraction of drainage networks from digital elevation models constrained by enhanced flow enforcement from known river maps. *Geomorphology* **2019**, *340*, 184–201. [CrossRef]
7. Montgomery, D.R.; Dietrich, W.E. Where do channels begin? *Nat. Cell Biol.* **1988**, *336*, 232–234. [CrossRef]
8. Montgomery, D.R.; Dietrich, W.E. Channel Initiation and the Problem of Landscape Scale. *Science* **1992**, *255*, 826–830. [CrossRef]
9. Passalacqua, P.; Trung, T.D.; Foufoula-Georgiou, E.; Sapiro, G.; Dietrich, W.E. A geometric framework for channel network extraction from lidar: Nonlinear diffusion and geodesic paths. *J. Geophys. Res. Space Phys.* **2010**, *115*, 01002. [CrossRef]
10. Sangireddy, H.; Stark, C.P.; Kladzyk, A.; Passalacqua, P. GeoNet: An open source software for the automatic and objective extraction of channel heads, channel network, and channel morphology from high resolution topography data. *Environ. Model. Softw.* **2016**, *83*, 58–73. [CrossRef]
11. Freeman, T. Calculating catchment area with divergent flow based on a regular grid. *Comput. Geosci.* **1991**, *17*, 413–422. [CrossRef]
12. Lea, N.L. An aspect driven kinematic routing algorithm. In *Overland Flow: Hydraulics and Erosion Mechanics*; Chapman and Hall: New York, NY, USA, 1992.
13. Costa-Cabral, M.C.; Burges, S.J. Digital Elevation Model Networks (DEMON): A model of flow over hillslopes for computation of contributing and dispersal areas. *Water Resour. Res.* **1994**, *30*, 1681–1692. [CrossRef]
14. Tarboton, D.G. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resour. Res.* **1997**, *33*, 309–319. [CrossRef]
15. Orlandini, S.; Moretti, G.; Franchini, M.; Aldighieri, B.; Testa, B. Path-based methods for the determination of nondispersive drainage directions in grid-based digital elevation models. *Water Resour. Res.* **2003**, *39*, 1144. [CrossRef]
16. Li, J.; Li, T.; Zhang, L.; Sivakumar, B.; Fu, X.; Huang, Y.; Bai, R. A D8-compatible high-efficient channel head recognition method. *Environ. Model. Softw.* **2020**, *125*, 104624. [CrossRef]
17. Ibrahim, M.O.; Türkay, G. Examining the stream threshold approaches used in hydrologic analysis. *Int. J. Geo.-Inf.* **2018**, *7*, 201.
18. Tarboton, D.G.; Bras, R.L.; Rodriguez-Iturbe, I. On the extraction of channel networks from digital elevation data. *Hydrol. Process.* **1991**, *5*, 81–100. [CrossRef]

19. Lin, W.T.; Chou, W.C.; Lin, C.Y.; Huang, P.H.; Tsai, J.S. Automated suitable drainage network extraction from digital elevation models in Taiwan's up-stream watersheds. *Hydrol. Proc.* **2006**, *20*, 289–306. [CrossRef]
20. Oliveira, F. *Arc Hydro: GIS for Water Resources*; Maidment, D.R., Ed.; ESRI, Inc.: Redlands, CA, USA, 2002; Volume 1, pp. 55–86.
21. Tang, G. *A Research on the Accuracy of Digital Elevation Models*; Science Press: Beijing, China, 2000.
22. Jones, R. Algorithms for using a DEM for mapping catchment areas of stream sediment samples. *Comput. Geosci.* **2002**, *28*, 1051–1060. [CrossRef]
23. Tantasirin, C.; Nagai, M.; Tipdecho, T.; Tripathi, N.K. Reducing hillslope size in digital elevation models at various scales and the effects on slope gradient estimation. *Geocarto Int.* **2016**, *31*, 140–157. [CrossRef]
24. Gökgöz, T.; Ulugtekin, N.; Basaraner, M.; Gulgen, F.; Dogru, A.O.; Bilgi, S.; Yucel, M.A.; Cetinkaya, S.; Selcuk, M.; Ucar, D. Watershed delineation from grid DEMs in GIS: Effects of drainage lines and resolution. In Proceedings of the 10th International Specialised Conference on Diffuse Pollution and Sustainable Basin Management, Istanbul, Turkey, 18–22 September 2006.
25. Vogt, J.V.; Colombo, R.; Bertolo, F. Deriving drainage networks and catchment boundaries: A new methodology combining digital elevation data and environmental characteristics. *Geomophology* **2003**, *53*, 281–298. [CrossRef]
26. Camporeale, C.V.; Perucca, E.; Ridolfi, L.; Gurnell, A.M. Modeling the Interactions between River Morphodynamics and Riparian Vegetation. *Rev. Geophys.* **2013**, *51*, 379–414. [CrossRef]
27. Kirkby, M.J. Long term interactions between networks and hillslopes. In *Channel Network Hydrology*; Beven, K., Kirkby, M.J., Eds.; John Wiley: New York, NY, USA, 1993; pp. 255–393.
28. Horton, R.E. Drainage-basin characteristics. *Trans. Am. Geophys. Union* **1932**, *13*, 350–361. [CrossRef]
29. Luo, W.; Jasiewicz, J.; Stepinski, T.; Wang, J.; Xu, C.; Cang, X. Spatial association between dissection density and environmental factors over the entire contermi-nous United States. *Geophys. Res. Lett.* **2016**, *43*, 692–700. [CrossRef]
30. Schneider, A.; Jost, A.; Coulon, C.; Silvestre, M.; Théry, S.; Ducharne, A. Global-scale river network extraction based on high-resolution topography and constrained by lithology, climate, slope, and observed drainage density. *Geophys. Res. Lett.* **2017**, *44*, 2773–2781. [CrossRef]
31. Strohbach, M.W.; Haase, D. Above-ground carbon storage by urban trees in Leipzig, Germany: Analysis of patterns in a European city. *Landsc. Urban Plan.* **2012**, *104*, 95–104. [CrossRef]
32. Chi, Y.; Shi, H.; Zheng, W.; Sun, J.; Fu, Z. Spatiotemporal characteristics and ecological effects of the human interference index of the Yellow River Delta in the last 30 years. *Ecol. Indic.* **2018**, *89*, 880–892. [CrossRef]
33. Song, S.; Zeng, L.; Wang, Y.; Li, G.; Deng, X. The response of river network structure to urbanization: A multifractal perspective. *J. Clean. Prod.* **2019**, *221*, 377–388. [CrossRef]
34. Shao, X.; Fang, Y.; Cui, B. A model to evaluate spatiotemporal variations of hydrological connectivity on a basin-scale complex river network with intensive human activity. *Sci. Total Environ.* **2020**, *723*, 138051. [CrossRef]
35. Chen, Y.; Xu, Y.; Fu, W. Influences of urbanization on river network in the coastal areas of East Zhejiang province. *Adv. Water Sci.* **2007**, *18*, 73.
36. Benstead, J.P.; Leigh, D.S. An expanded role for river networks. *Nat. Geosci.* **2012**, *5*, 678–679. [CrossRef]
37. Persendt, F.; Gomez, C. Assessment of drainage network extractions in a low-relief area of the Cuvelai Basin (Namibia) from multiple sources: LiDAR, topographic maps, and digital aerial orthophotographs. *Geomorphology* **2016**, *260*, 32–50. [CrossRef]
38. Hou, J.W.; van Dijk, A.I.J.M.; Beck, H.E. Global satellite-based river gauging and the influence of river morphology on its application. *Remote Sens. Environ.* **2020**, *239*, 11629. [CrossRef]
39. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*. [CrossRef]
40. Ariza-Villaverde, A.; Jiménez-Hornero, F.; De Ravé, E.G. Influence of DEM resolution on drainage network extraction: A multifractal analysis. *Geomorphology* **2015**, *241*, 243–254. [CrossRef]
41. Woodrow, K.; Lindsay, J.B.; Berg, A.A. Evaluating DEM conditioning techniques, elevation source data, and grid resolution for field-scale hydrological parameter extraction. *J. Hydrol.* **2016**, *540*, 1022–1029. [CrossRef]
42. Niipele, J.N.; Chen, J.P. The usefulness of also-palsar dem data for drainage extraction in semi-arid environments in The Iishana sub-basin. *J. Hydrol. Reg. Stud.* **2019**, *21*, 57–67. [CrossRef]
43. Colombo, R.; Vogt, J.V.; Soille, P.; Paracchini, M.L.; de Jager, A. Deriving river networks and catchments at the European scale from medium resolution digital elevation data. *Catena* **2007**, *70*, 296–305. [CrossRef]
44. Mandelbrot, B.B. *The Fractal Geometry of Nature/Revised and Enlarged Edition*; WH Freeman and Co.: New York, NY, USA, 1983; 49p.
45. Shen, X.; Zou, L.; Zhang, G.; Su, N.; Wu, W.; Yang, S. Fractal characteristics of the main channel of Yellow River and its relation to regional tectonic evolution. *Geomorphology* **2011**, *127*, 64–70. [CrossRef]
46. Joanna, F.B. Fractal structure of the Kashubian hydrographic system. *J. Hydrol.* **2013**, *488*, 48–54.
47. Zhang, S.; Guo, Y.; Wang, Z. Correlation between flood frequency and geomorphologic complexity of rivers network—A case study of Hangzhou China. *J. Hydrol.* **2015**, *527*, 113–118. [CrossRef]
48. Bai, R.; Li, T.; Huang, Y.; Li, J.; Wang, G. An efficient and comprehensive method for drainage network extraction from DEM with billions of pixels using a size-balanced binary search tree. *Geomorphology* **2015**, *238*, 56–67. [CrossRef]
49. O'Callaghan, J.F.; Mark, D.M. The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process.* **1984**, *28*, 323–344. [CrossRef]

50. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nat. Cell Biol.* **1999**, *401*, 788–791. [CrossRef]
51. Zhang, H.; Loáiciga, H.; Ren, F.; Du, Q.; Ha, D. Semi-empirical prediction method for monthly precipitation prediction based on environmen-tal factors and comparison with stochastic and machine learning models. *Hydrol. Sci. J.* **2020**, *65*, 1–15. [CrossRef]
52. De Bartolo, S.G.; Gaudio, R.; Gabriele, S. Multifractal analysis of river networks: Sandbox approach. *Water Resour. Res.* **2004**, *40*, 02201. [CrossRef]
53. Grassberger, P. On Efficient Box Counting Algorithms. *Int. J. Mod. Phys. C* **1993**, *4*, 515–523. [CrossRef]
54. Ge, M.; Lin, Q. Realizing the box-counting method for calculating fractal dimension of urban form based on remote sensing image. *Geo.-Spatial. Inf. Sci.* **2009**, *12*, 265–270. [CrossRef]
55. Pavón-Domínguez, P.; Rincón-Casado, A.; Ruiz, P.; Camacho-Magriñán, P. Multifractal approach for comparing road transport network geometry: The case of Spain. *Phys. A Stat. Mech. Appl.* **2018**, *510*, 678–690. [CrossRef]
56. Halsey, T.C.; Jensen, M.H.; Kadanoff, L.P.; Procaccia, I.; Shraiman, B.I. Fractal measures and their singularities: The characterization of strange sets. *Phys. Rev. A* **1986**, *33*, 1141–1151. [CrossRef]
57. Chakraborty, B.; Haris, K.; Latha, G.; Maslov, N.; Menezes, A. Multifractal Approach for Seafloor Characterization. *IEEE Geosci. Remote Sens. Lett.* **2013**, *11*, 54–58. [CrossRef]
58. Ge, Y.; Dou, W.; Gu, Z.; Qian, X.; Wang, J.; Xu, W.; Shi, P.; Ming, X.; Zhou, X.; Chen, Y. Assessment of social vulnerability to natural hazards in the Yangtze River Delta, China. *Stoch. Environ. Res. Risk Assess.* **2013**, *27*, 1899–1908. [CrossRef]
59. Zhang, H.; Loáiciga, H.A.; Ha, D.; Du, Q. Spatial and Temporal Downscaling of TRMM Precipitation with Novel Algorithms. *J. Hydrometeorol.* **2020**, *21*, 1259–1278. [CrossRef]
60. Wang, R.; Li, C. Spatiotemporal analysis of precipitation trends during 1961–2010 in Hubei province, central China. *Theor. Appl. Clim.* **2016**, *124*, 385–399. [CrossRef]
61. Gregory, K.J.; Gardiner, V. Drainage density and climate. *Geomorphology* **1975**, *19*, 287–298.
62. Moglen, G.E.; Eltahir, E.A.B.; Bras, R.L. On the sensitivity of drainage density to climate change. *Water Resour. Res.* **1998**, *34*, 855–862. [CrossRef]
63. Yang, H.-L.; Xiao, H.; Guo, C.; Sun, Y. Spatial-temporal analysis of precipitation variability in Qinghai Province, China. *Atmos. Res.* **2019**, *228*, 242–260. [CrossRef]
64. Yan, D.; Wang, K.; Qin, T.; Weng, B.; Wang, H.; Bi, W.; Li, X.; Li, M.; Lv, Z.; Liu, F.; et al. A data set of global river networks and corresponding water resources zones divisions. *Sci. Data* **2019**, *6*, 1–11. [CrossRef] [PubMed]
65. Han, X.; Fang, W.; Li, H.; Wang, Y.; Shi, J. Heterogeneity of influential factors across the entire air quality spectrum in Chinese cities: A spatial quantile regression analysis. *Environ. Pollut.* **2020**, *262*, 114259. [CrossRef]
66. Hartmann, J.; Moosdorf, N. The new global lithological map database GLiM: A representation of rock proper-ties at the Earth surface. *Geochem. Geophys. Geosyst.* **2012**, *13*, Q12004. [CrossRef]
67. Jenness, J.S. Calculating landscape surface area from digital elevation models. *Wildl. Soc. Bull.* **2004**, *32*, 829–839. [CrossRef]
68. Hodgson, M.E.; Gaile, G.L. A cartographic modeling approach for surface orientation-related applications. *Photogramm. Eng. Remote Sens.* **1999**, *65*, 85–95.
69. Grohmann, C.H.; Smith, M.J.; Riccomini, C. Multi-scale analysis of topographic surface roughness in the mid-land valley, Scotland. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1200–1213. [CrossRef]
70. Lindsay, J.B.; Newman, D.R.; Francioni, A. Scale-Optimized Surface Roughness for Topographic Analysis. *Geoscience* **2019**, *9*, 322. [CrossRef]
71. Feng, H.; Zou, B.; Tang, Y. Scale- and Region-Dependence in Landscape-PM2.5 Correlation: Implications for Urban Planning. *Remote Sens.* **2017**, *9*, 918. [CrossRef]
72. Wang, C.; Chen, Z.; Yang, C.; Li, Q.; Wu, Q.; Wu, J.; Zhang, G.; Yu, B. Analyzing parcel-level relationships between Luojia 1-01 nighttime light intensity and artificial surface features across Shanghai, China: A comparison with NPP-VIIRS data. *Int. J. Appl. Earth Obs. Geoinform.* **2020**, *85*, 101989. [CrossRef]
73. Zou, L. On a conjecture concerning the Frobenius norm of matrices. *Linear Multilinear Algebra* **2012**, *60*, 27–31. [CrossRef]
74. Yang, Y.Z.; Cai, W.H.; Yang, J. Evaluation of MODIS Land Surface Temperature Data to Estimate Near-Surface Air Temperature in Northeast China. *Remote Sens.* **2017**, *9*, 410. [CrossRef]
75. Cui, X.; Zhang, J.; Wu, X.; Hao, N.; Wang, Q. Dynamic Change of Land Cover of Qinling Mountains Based on MODIS NDVI. In Proceedings of the 2018 Fifth International Workshop on Earth Observation and Remote Sensing Applications (EORSA), Xi'an, China, 18–20 June 2018; pp. 1–5.
76. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence, Montreal, QC, Canada, 20–25 August 1995; Volume 2, pp. 1137–1143.
77. He, L.H.; Zhao, H. The fractal dimension of river networks and its interpretation. *Sci. Geogr. Sin.* **1996**, *2*, 124–128. (In Chinese)
78. Martz, L.W.; Garbrecht, J. Numerical definition of drainage network and subcatchment areas from Digital Elevation Models. *Computer Geosci.* **1992**, *18*, 747–761. [CrossRef]

79. Stein, J.L.; Hutchinson, M.F. A new stream and nested catchment framework for Australia. *Hydrol. Earth Syst. Sci.* **2014**, *18*, 1917–1933. [CrossRef]
80. Yao, Q.; Liu, K.-B.; Aragón-Moreno, A.A.; Rodrigues, E.; Xu, Y.J.; Lam, N.S. A 5200-year paleoecological and geochemical record of coastal environmental changes and shoreline fluctuations in southwestern Louisiana: Implications for coastal sustainability. *Geomorphology* **2020**, *365*, 107284. [CrossRef]