Index for the Consistent Measurement of Spatial Heterogeneity for Large-Scale Land Cover Datasets

Jing Yu [1],*, Shu Peng [2], Weiwei Zhang [3] and Shun Kang [4]

[1]   College of Geoscience and Surveying Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China

[2]   National Geomatics Center of China, Beijing 100830, China; pengshu@ngcc.cn

[3]   School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou 215011, China; zhangweiwei@usts.edu.cn

[4]   School of Electrical and Electronics Engineering, Hubei Polytechnic University, Huangshi 435003, China; 226005@hbpu.edu.cn

*   Correspondence: tbp150202035@student.cumtb.edu.cn; Tel.: +86-156-5292-8340

Abstract: Recognizing land cover heterogeneity is essential for the assessment of spatial patterns to guide conservation planning. One of the top research priorities is the quantification of land cover heterogeneity using effective landscape metrics. However, due to the diversity of land cover types and their varied distribution, a consistent, larger-scale, and standardized framework for heterogeneity information extraction from this complex perspective is still lacking. Consequently, we developed a new Land Cover Complexity Index (LCCI), which is based on information-theory. The LCCI contains two foundational aspects of heterogeneity, composition and configuration, thereby capturing more comprehensive information on land cover patterns than any single metric approach. In this study, we compare the performance of the LCCI with that of other landscape metrics at two different scales, and the results show that our newly developed indicator more accurately characterizes and distinguishes different land cover patterns. LCCI provides an alternative way to measure the spatial variation of land cover distribution. Classification maps of land cover heterogeneity generated using the LCCI provide valuable insights and implications for regional conservation planning. Thus, the LCCI is shown to be a consistent indicator for the quantification of land cover heterogeneity that functions in an adaptive way by simultaneously considering both composition and configuration.

Keywords: land cover heterogeneity; landscape metrics; complexity; information theory

1. Introduction

Recognizing land cover spatial heterogeneity is crucial for ecological process modeling, spatial pattern understanding, and environmental change analysis [1–4]. Land cover heterogeneity is a key concept of land system science, a discipline that has long focused on regional structure and patterns. The heterogeneity of land cover can be quantitatively described in different forms, such as fragmentation [5], diversity [6], connectivity [7], and complexity [8]. Recently, some studies have focused on land cover heterogeneity in terms of land surface parameterization and land cover classification quality [9,10]. The need for consistent and accurate information on land cover heterogeneity to support large-scale geospatial applications has been increasingly acknowledged and emphasized [11]. Therefore, it is necessary to extract standardized land cover heterogeneity information at fine resolution and large scales to meet the requirements of scientists and policy-makers.

One of the priorities in the quantitative assessment of land cover heterogeneity is the development and the utilization of various landscape metrics, which are generally particularly well suited for use at the regional scale. In practice, different landscape metrics broadly fall into one of two categories, non-spatial and spatial, which are widely used to quantify heterogeneity composition and configuration, respectively [12]. For instance, Riitters uses a landscape mosaic metric to identify the United States land cover heterogeneity composition to support the Forest Service Renewable Resources Planning Act Assessment [13]. The EEA (European Environment Agency) is using an effective mesh size metric to quantify the degree of European land cover fragmentation for various human activity planning applications and the sustainable conservation of nature [14]. Although landscape metrics are helpful for extracting regional land cover heterogeneity information, measuring land cover heterogeneity at large scales is still a difficult task. This is due to several factors, including the correlation and the robustness of different metrics and the comprehensive evaluation of multiple metrics [15]. Further, choosing the appropriate index and constructing a complete quantization scheme for measuring large-scale land cover heterogeneity remains a challenge.

Entropy is another concept that has been introduced to measure land cover heterogeneity. It has long been considered as an excellent tool to estimate the complexity of a system [16–19]. Shannon's entropy-based indices have been widely used to quantify landscape heterogeneity in space and time [20–22]. In particular, Shannon's diversity index (SHDI) was one of the most common indicators [23,24]. However, the Shannon's-based entropy lacks the configurational information due to the fact that it only considers one-dimensional information, thus it cannot effectively describe land cover heterogeneity [25,26]. Although several former attempts have been made to extend the Shannon entropy for measuring the spatial complexity, these methods based on the distance were relatively complex and varied [26–30]. Recently, two classes methods have been developed to apply Boltzmann entropy in landscape ecology. Cushman proposed the first-ever idea that a landscape mosaic can be computed using the Boltzmann entropy, which adopted the total edge (TE) of a landscape mosaic to measure macrostate [31,32]. The efficiency problem limited its application in real landscapes due to the large number of possibilities [33]. Another method to compute Boltzmann entropy of a landscape gradients was proposed by Gao, adopting a hierarchical perspective to define macrostate and microstate [34]. The configurational entropy calculated by Boltzmann entropy can distinguish different land cover patterns and provide additional perspectives to understand the relationship of complexity and entropy. The characteristic of Boltzmann entropy is the capability to capture the composition and the configurational information of a system, which appears to be more suitable for landscape quantification [35]. A general and effective method for both landscape mosaic and landscape gradients based on Boltzmann entropy would be desirable for development.

Land cover heterogeneity can be characterized by relatively few components, and each one can be quantified by suitable independent metrics [36]. Two main components were recognized: a more heterogeneous land cover structure, which is an area with various proportions of different cover types, called compositional heterogeneity, and a more complex spatial structure called configurational heterogeneity [37]. In ecogeographic studies, the separation evaluation of land cover heterogeneity components, especially diversity and fragmentation, are fundamental in biodiversity and environmental research at different time scales. However, as the purpose of this study is to characterize heterogeneity and identify the heterogeneity distribution, we do not consider multiple landscape metrics to construct indicator sets; instead, we adopt an adaptive method to combine land cover composition and land cover configuration for comprehensive heterogeneity information. As the utilization of various metrics may possibly yield redundant information, limiting our approach to two metrics ensures the most effective interpretation of both structure and patterns. In other words, the more complex the calculations are, the more difficult the interpretation of the land cover heterogeneity is [38].

In this study, a readily applicable measure is proposed to address the lack of a consistent and standardized framework for heterogeneity information extraction at large spatial scales. Specifically, information-theoretical metrics were employed for fusing a consistent indicator, the Land Cover Complexity Index (LCCI), for large-scale land cover heterogeneity quantification at 1 km resolution. The main objectives of this study are as follows: (1) to describe a methodology suitable for quantifying the characteristics of large-scale land cover heterogeneity; (2) to build a database of continent land cover heterogeneity for large-scale geospatial sampling and ecological assessment; and (3) to discover the heterogeneous distribution characteristics of the different continents.

This paper is organized as follows. In Section 2, we illustrate the inconsistency problems in large-scale quantifications of land cover heterogeneity and give a solution. Section 3 introduces the key concepts for the construction of methodology in this paper. In Section 4, we present the quantification results and compare LCCI with different single metric approaches using path analysis. Section 5 provides a summary of our results and discussion and includes our conclusions.

2. Spatial Heterogeneity of Land Cover: Problems and Solution

2.1. Inconsistency Problems at Large Scales

Spatially continuous coverage of earth observatory data and the fast development of geoinformation technologies encourage land cover heterogeneity research with the aim of gaining more robust and continuous information. This facilitates further understanding of ecological processes and monitors the distribution of our natural resources and their dynamics. Moreover, land cover heterogeneity is becoming increasingly helpful in large-scale geospatial research, especially for surface parameterization and space sampling. Therefore, the scientific measurement of large-scale heterogeneity requires new quantitative methods. Due to the complexity and the variation of land cover types in a large area, using a single index for measuring the heterogeneity results in low heterogeneity values for areas where the heterogeneity is actually high.

As shown in Figure 1a, there are three areas in the world with different degrees of heterogeneity. Region a and region b have the same composition, but due to differences in their configuration, the degree of heterogeneity of region b is greater than that of region a. Similarly, region b and region c have comparably complex configurations, but region c is more diverse than region b. Thus, the heterogeneity of region c is higher than that of region b. However, using the traditional configuration landscape metrics such as the SHDI and edge density (ED) index shown in Figure 1b, the difference is indistinguishable. These circumstances result in inconsistent heterogeneity values for large-scale quantifications of land cover heterogeneity that persist even if other landscape metrics are utilized. The fundamental reason is that the composition and the configuration are not considered simultaneously. To quantify land cover heterogeneity more accurately, thereby meeting the demand for large-scale applications, efficiently combining the composition and the configuration indices is critical.

In other words, due to the failure of factoring in both composition and configuration, the land cover heterogeneity quantified using a single indicator is not consistent with the true degree of heterogeneity. Over time, the land cover type of an area may be converted to a different land cover type, and in such cases, the composition of the land cover data increases, whereas the configuration remains the same. The change may thus actually enhance the heterogeneity of this area, and the quantized value needs to reflect this change. For instance, in Figure 2, the type of conversion increases the heterogeneity of the land cover, but using only one configuration index fails to capture the difference. Adopting two different indices simultaneously may resolve the problem but is not suitable for large-scale monitoring.
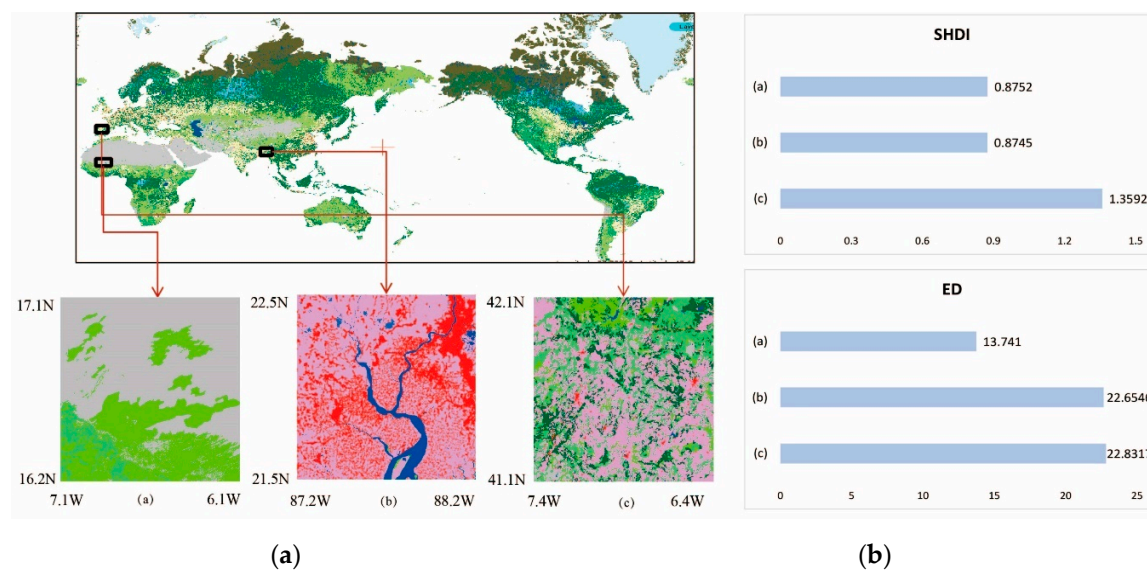
Figure 1. Land cover heterogeneity inconsistencies. (**a**) Examples of regions with different degrees of heterogeneity, (**b**) inconsistent results of land cover heterogeneity values, calculated using the traditional Shannon's diversity (SHDI) and edge density (ED) indices.
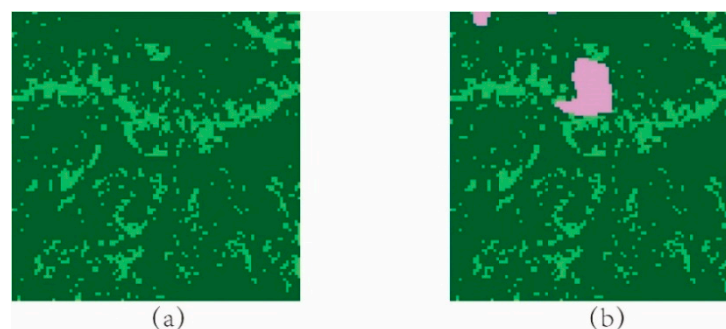

Figure 2. Example of a change in land cover heterogeneity resulting from land cover type conversion: (**a**) Local GlobeLand30 for the year 2000; (**b**) heterogeneity change seen in the year 2010.

2.2. Solution: Land Cover Complexity Index (LCCI) Design

With LCCI, we aim to construct a consistent indicator of land cover heterogeneity. Land cover composition and configuration express the complexity of land cover categories and adjacencies/spatial distributions, respectively. A comprehensive measurement of land cover heterogeneity should, therefore, account for both. To achieve this, we (i) constructed a co-occurrence histogram to express both the composition and the configuration in a single histogram, (ii) quantified two entropy-based indices called marginal entropy and conditional entropy, (iii) calculated the relative mutual information to measure the difference between the two entropy-based indices, and (iv) combined the entropy-based indices to create the LCCI. Simultaneously considering different aspects of heterogeneity can consolidate the results and enhance robustness. Moreover, entropy-based indices not only record the correlation between configuration and composition, but, based on our empirical tests, also capture details of the complexity of small patch distribution. Accounting for the asymmetry of data confidence, we combined the information-theoretical metrics into a complex fusion technique, which is based on the dependence between random variables.

The flow chart in Figure 3 details the heterogeneity quantification process.
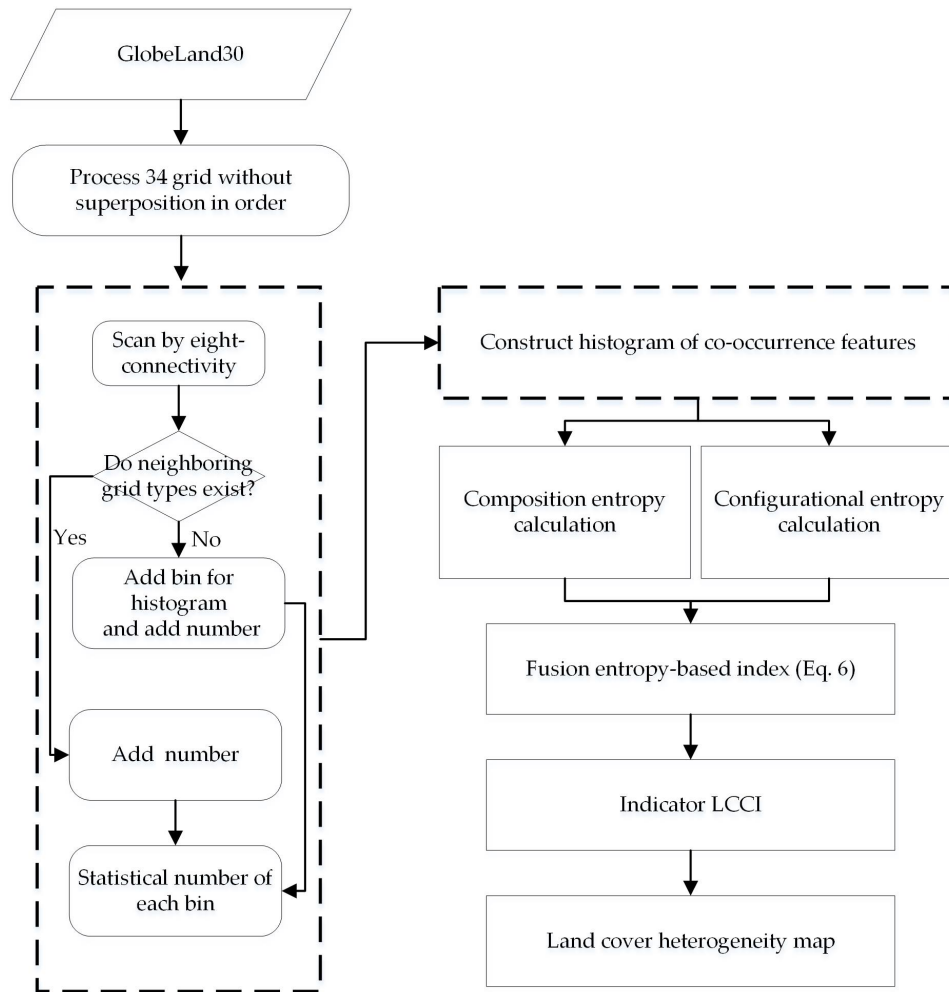
Figure 3. Flowchart of the heterogeneity information extraction procedure.

3. Adaptive Method for Index Fusion

3.1. Entropy-Based Fundamental Index of Spatial Heterogeneity Measurements

To obtain the diversity (composition) and the adjacency (configuration) characteristics of land cover heterogeneity simultaneously, a bivariate approach for the analysis of two fundamental variables is needed [39]. For the purpose of describing each land cover unit mathematically, we extracted a bivariate co-occurrence histogram featured in color from the gray-level co-occurrence matrix for the subsequent calculations (Figure 4). Each bin of the co-occurrence histogram is a land cover type adjacency feature extracted from two neighboring grids, called *Pij*. We adopted an eight-connectivity as the adjacency rule and to distinguish bins that have the same land cover type but a different neighboring order as *Pij* and *Pji*. Basic information-theoretical metrics applied to quantify land cover heterogeneity are as follows [16]:

$$H(y) = -\sum_{j=1}^{n} P_j log P_j \tag{1}$$

$$H(y|x) = -\sum_{i=1}^{n} \sum_{j=1}^{n} P_{ij} log P_{i \rightarrow j} \tag{2}$$

$$H(x,\ y) = \sum_{i=1}^{n} P_{ij} log P_{ij} \tag{3}$$

where $H(y)$ is the standard Shannon entropy calculated based on pairs of cells, not single cells; $H(y|x)$ is the conditional entropy based on the joint probabilities Pij and the second-order probabilities $P(i{\rightarrow}j) = Pij/Pi$; and $H(x,y)$ is the joint entropy computable directly from Pij, which equals $H(y)$ plus $H(y|x)$, measuring the overall complexity of the land cover pattern. These three metrics describe composition entropy, configuration entropy, and joint entropy of the land cover patterns, respectively. Many previous studies have shown that land cover composition and configuration are highly correlated, but the nature of the correlation, i.e., its linearity, has not been determined [40]. The ability of the mutual entropy to capture the dependence or the relevance between land cover composition and configuration has recently led to attempts to employ it in landscape complex ordering and pattern classification [41]:

$$I(y,\ x) = H(y) - H(y|x) \tag{4}$$

$$U = I(y,\ x)/H(y) \tag{5}$$

where $I(y,x)$ is the uncertainty of variable y reduced by knowing variable x and U the relative mutual information, which measures the difference between composition and configuration. Certain land cover patterns have high diversity (composition) and low fragmentation (configuration), which can result in high values of U. In this study, we adopted it as a measure of compositional confidence.
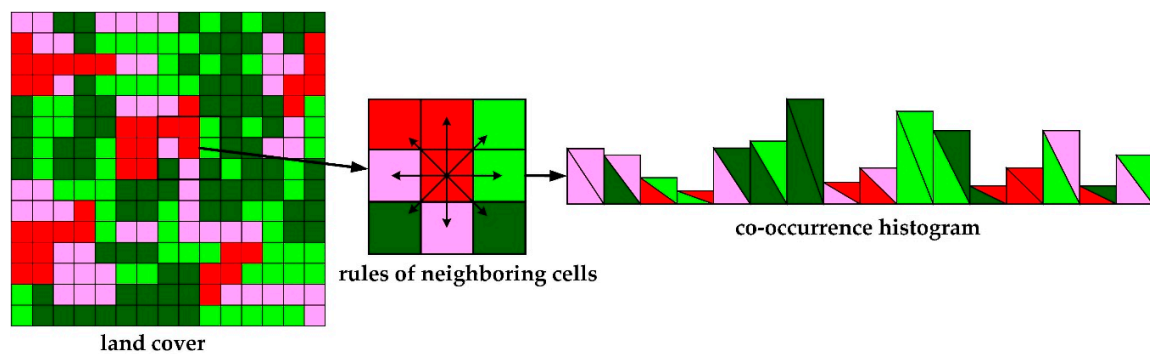

Figure 4. Examples of co-occurrence histogram extracted from land cover.

3.2. Fusion of Entropy-Based Indices

A fusion method was adopted to combine the two entropy-based indices that describe the different aspects of heterogeneity, $H(y)$ and $H(y|x)$, into one final indicator. As information theory was used as a consistent framework, normalization of the index values to ensure a comparable range for the indices was not necessary.

The fusion process was designed because the reliability of the two indices needs to be adjustable to accommodate intricate land cover patterns. Indices may be imperfect; therefore, the expert should consider their imperfections, thereby specifying a partial trust [42]. In the fusion process, indices with partial trust are weighted proportionally to their trust in the result. Diverse composition may lead to a complex configuration, and both composition entropy and configurational complexity generally function as trust indicators for building the final indicator. However, in some real world situations, the land cover has a high diversity value but a geometrically simple pattern, which causes the composition index to fail as trust indicator (the heterogeneity is low whereas the diversity value is high). Furthermore, both indices fail to capture small patch patterns. To improve the accuracy of the heterogeneity quantification in these situations, we introduced a relative mutual information parameter U to judge the trust of the indicators. Empirical studies revealed that U is an appropriate reference for measuring the difference between composition and configuration [41]. We applied the following rule: both $H(y)$ and $H(y|x)$ are considered to express the different characters of land cover heterogeneity. However, if the U value is far greater than 0, we assume that $H(y)$ does not accurately describe the land cover heterogeneity and the actual heterogeneity is lower than described by the

composition layer. We define $(1 - U)$ as an adaptive weight to adjust this difference at a 1 km scale based on expert experience and a large number of experiments. If U is close to 0, the land cover pattern is more complex than the single layer result indicates.

$$LCCI = (1 - U) * H(x, y) \qquad (6)$$

The fusion method results in higher values for intricate land cover patterns. In this case, $H(y)$ and $H(y|x)$ can be equally trusted, and U may be used as relative information gain to increase the value of land cover complexity because it measures the interaction of composition and configuration.

4. Experiment and Analysis

4.1. Datasets and Quantization Scheme

A GlobeLand30-2010 dataset with a fine resolution of 30 m was used in this study to extract the characteristics of land cover heterogeneity. The dataset comprises ten first-level classes, namely cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bare land, and permanent snow/ice, for the years 2000 and 2010. In this study, we used the 2010 map (see http://www.globallandcover.com for the detailed legend). The maps were extracted from Landsat and HJ-1 satellite images through a pixel-object-knowledge-based (POK-based) approach, with preliminary validation results for the overall classification accuracy of greater than 80% in 2010.

The quantification of land cover heterogeneity was carried out within ArcGIS version 10.1 [43] using Python scripts (http://www.python.org). The GlobeLand30 data were first partitioned into regular 1 km × 1 km units at smaller computational cost than the basic units for the following quantification and analysis. These 1 km × 1 km squares were determined empirically for large-scale landscape analysis [44], and each unit overlapped by approximately 34 × 34 pixels of GlobeLand30. These simple units made up a cell array, and each unit contained complex content. In the next step, we employed these units of blocks as elementary units for each nation's heterogeneity aggregation. Subsequently, we extracted composition, configuration, and complexity information separately from the units with valid data. Finally, the land cover complexity results were mapped for intuitive perception.

Real land cover data at two scales, local and continental, were selected for the validation of our entropy-based fusion model, which was performed by examining the ability of the LCCI to quantify the complexity characteristics of land cover. Because the fusion method aims to quantify the complexity of land cover, we compared the performance of the LCCI with two indices proposed to measure two fundamental aspects of land cover heterogeneity. Furthermore, we compared the performance of the LCCI with one of the commonly used indices that is strongly correlated with it to validate the superior performance of the LCCI regarding the extraction of comprehensive land cover heterogeneity information.

4.2. Validation of LCCI

In total, 36 types of real land cover patterns from different parts of the globe were used as the evaluation dataset to test the consistency of our model for the quantification of land cover heterogeneity. Evaluation data used for heterogeneity quantification usually had the following two characteristics: first, they represented different degrees of heterogeneity and thus complex land cover patterns; second, their difference in heterogeneity could be captured by the naked eye. For each of the 36 pattern types, we calculated the LCCI and the two most commonly used heterogeneity metrics: ED as the configurational metric and SHDI as the compositional metric. For each indicator, we sorted the values into quintiles (20-percentile), referred to as classes 1–5. Class 1 represents the initial 20-percentile, expressing the lowest heterogeneity of land cover. Class 5 represents the last 20-percentile, expressing the highest complex distribution of land cover. Figure 5 shows the tiles that fell within that quintile for each indicator, and they are shown in order of increasing value. The evaluation dataset is marked using the ordering label of $H(y)$.
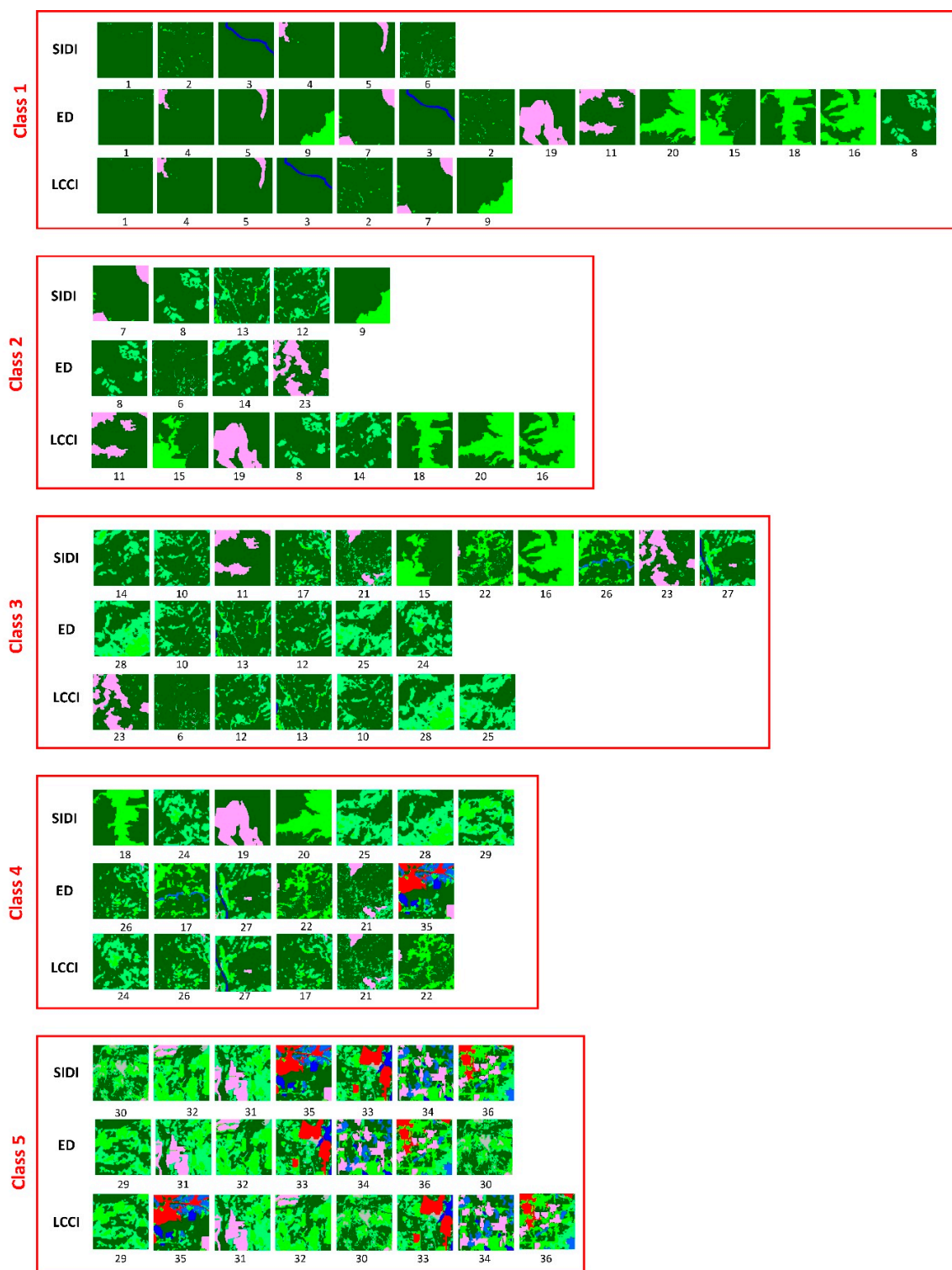
Figure 5. Evaluation dataset assigned to different land cover heterogeneity levels based on three metrics.

The evaluation data illustrated in Figure 5 displays distinctly different heterogeneities of land cover for the three different metrics. Based on visual inspection, SHDI seemed to increase with increasing diversity of the land cover distribution. In addition, different land cover heterogeneities had similar SHDI values (see class2, #12 and #9). Edge density was selected to contrast SHDI not only because it measures the spatial configuration of heterogeneity but also because it extracts complex boundary information, similar to the co-occurrence histogram, which connects different ecological interactions

with a variety of mosaic types. Based on visual inspection, the increase of ED seemed to correspond with an increase in the complexity of the land cover configuration heterogeneity. The land cover distribution changed from simple to complex, but units with similar values and classified according to similar degrees of heterogeneity did, in some cases, show different degrees of heterogeneity (see class1, #2 and #19). Several such discrepancies are apparent in Figure 5.

Thus, heterogeneity information derived from SHDI and ED did not fully capture the inherent complexity characteristics (see class1, where ED was predominantly classified by percentile class1, and class4, where SHDI showed obviously different patterns). LCCI, on the other hand, was able to distinguish the different land cover patterns more distinctly (when comparing LCCI, SHDI, and ED in each classification, LCCI classified each percentile homogeneously). These results show that the LCCI performed better than the SHDI and the ED indices, proclaiming its suitability for quantifying land cover heterogeneity. It captures richer heterogeneity information than single indices by considering both land cover composition and configuration information. Meanwhile, the purpose of this validation test is not to discredit the single indices but rather to show that they only partially quantify land cover heterogeneity information.

4.3. Relationships between LCCI and Landscape Metrics

In addition to SHDI and ED, seven metrics related to patch shape, size, and connectivity were selected for path analysis in order to further evaluate the quality of LCCI. Landscape metrics selected include patch density (PD), largest patch index (LPI), patch cohesion index (COHE), aggregation index (AI), fractal dimension (FRAC), landscape division index (DIVISI), and splitting index (SPLIT). These metrics are commonly used as heterogeneity measures in regional and local scale studies [45–47]. The selected metrics were calculated by FRAGSTAT 4.2 at the same level as the landscapes. Path analysis was performed using the structural equation modeling module of Amos [48]. Path analysis can decompose the interaction between dependent and independent variables (correlations) into direct (path coefficient) and indirect effects (indirect path coefficient).

The correlation and path coefficients between LCCI and the nine landscape metrics are listed in Table 1. The correlation analysis results show that LCCI was strongly correlated with both ED and PD and moderately correlated with SHDI. In addition, LCCI was strongly negatively correlated with AI and COHE, two metrics classified as aggregation indices. The path analysis results indicate that LCCI had the highest direct path coefficient with ED, which means that a close relationship between LCCI and ED existed. The direct path coefficient between AI and LCCI was greater than 0.5, indicating that AI had a large direct effect on LCCI. ED is a configuration metric of land cover heterogeneity and measures the boundary abundance, and AI is an aggregation metric that measures the land cover complexity. Higher complexity of land cover was thus reflected by higher ED and AI values. The direct path coefficient of PD was low at 0.115, but the indirect path coefficient mediated by ED was high at 0.683. Similar coefficients were observed for SHDI, DIVISI, and SPILT mediated by ED. This performance of PD, SHDI, DIVISI, and SPILT implies that their influence upon LCCI through ED was relatively important. The direct path coefficient and the indirect path coefficient values of COHE, mediated by ED, to LCCI were −0.222 and −0.657, respectively, which suggests a negative relationship between COHE and LCCI. FRAC was weakly correlated with LCCI, as shown by the low values of less than 0.5 of the direct and the indirect path coefficients. LPI was negatively correlated with LCCI and all indirect path coefficients less than 0.5, which implies a weak relationship with LCCI.

4.4. Large-Scale Application: Example Africa

Based on the entropy-based fusion method, we developed a comprehensive index to quantify the complexity of land cover in local environments. The path analysis results shown in Table 1 suggest that ED had the closest relationship with LCCI. This strong collinearity was expected because both metrics are boundary-based. The moderate correlation of SHDI and LCCI makes SHDI a complementary index

able to capture the different aspects of land cover heterogeneity, showing strikingly disparate land cover patterns.

Table 1. Correlation and path coefficients between Land Cover Complexity Index (LCCI) and landscape metrics.

Path	Correlation	Direct Path Coefficient	Indirect Path Coefficients								
			PD	LPI	ED	COHE	SHDI	AI	FRAC	DIVISI	SPLIT
PD→LCCI	0.914	0.115	-	0.065	0.683	0.183	0.056	−0.541	−0.002	−0.108	0.002
LPI→LCCI	−0.755	−0.122	−0.061	-	−0.526	−0.185	−0.071	0.408	−0.003	0.189	−0.002
ED→LCCI	0.948	0.734	0.107	0.087	-	0.199	0.066	−0.579	0.000	−0.146	0.002
COHE→LCCI	−0.965	−0.222	−0.095	−0.102	−0.657	-	−0.071	0.514	−0.001	0.158	−0.003
SHDI→LCCI	0.87	0.08	0.081	0.108	0.607	0.197	-	−0.469	0.002	−0.177	0.002
AI→LCCI	−0.942	0.58	−0.107	−0.086	−0.733	−0.197	−0.065	-	0.000	0.143	−0.002
FRAC→LCCI	−0.009	0.011	−0.018	0.036	0.014	0.014	0.011	−0.009	-	−0.064	0.000
DIVISI→LCCI	0.758	−0.193	0.064	0.119	0.555	0.182	0.073	−0.430	0.004	-	0.002
SPILT→LCCI	0.844	0.003	0.073	0.100	0.530	0.207	0.063	−0.412	0.001	−0.147	-

PD: patch density; LPI: largest patch index; COHE: patch cohesion index; AI: aggregation index; FRAC: fractal dimension; DIVISI: landscape division index; SPLIT: splitting index.

The fusion LCCI not only captures the spatial difference of the land cover configuration but also takes the diversity of land cover type information into account, even for locations that are classified as having similar heterogeneities by the traditional landscape metric ED (Figure 6). The aggregation results at the country level suggest that the LCCI captures richer heterogeneity information than the single metric ED (see Table 2; South Africa and Somalia have similar fragmentation levels but different diversity levels, which result in different complexity levels). The aggregation measure indicates that consistent measures of land cover heterogeneity may be more suitable for large-scale conservation planning or spatial sampling parameterization.
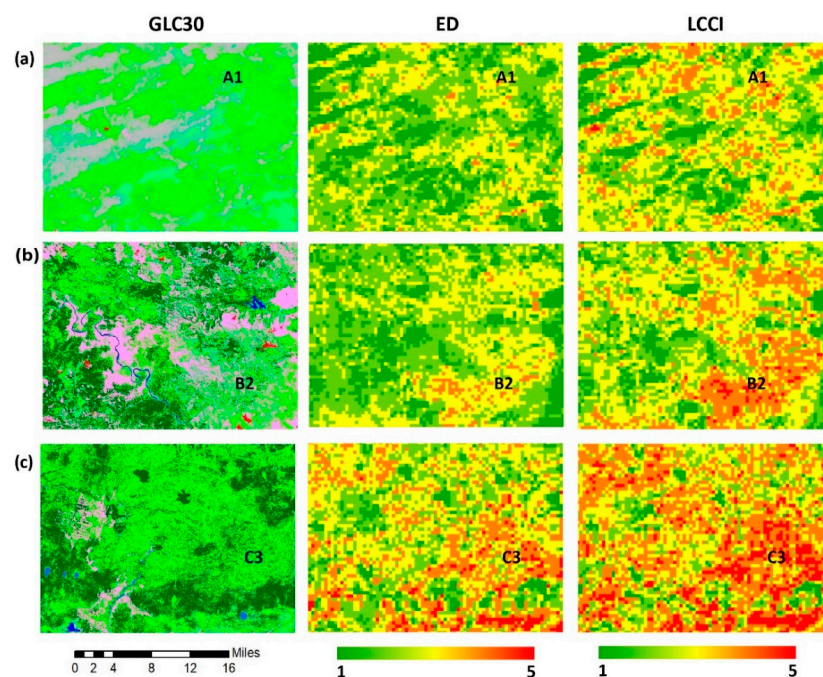

Figure 6. Examples of land cover heterogeneity captured by ED and LCCI. The panels from left to right are GlobeLand30, ED, and land cover complexity. (**a**) Desert and grassland mosaic regions in the west of the Sahel region (15°56′ N, 14°56′ W). The LCCI captures more patch detail of the land cover heterogeneity (A1); (**b**) diverse cropland region in East Africa (29°1′ S, 31°17′ E). The LCCI captures a large amount of heterogeneity because of the high land cover diversity (B2); (**c**) region with a landscape of forest fragmentation and grassland in Zambia (14°17′ S, 25°8′ E). The LCCI captures the extreme fragmentation feature of land cover better than ED (C3).

Table 2. Land cover heterogeneity of different countries measured by LCCI and ED.

Name	ED	LCCI	Name	ED	LCCI
Egypt	4.911	0.017	Rwanda	132.235	0.436
Libya	6.174	0.021	Madeira	132.541	0.442
Sao Tome And Principe	7.506	0.026	South Africa	141.229	0.462
Western Sahara	11.070	0.036	Somalia	142.393	0.438
Mayotte	13.440	0.056	Madagascar	156.256	0.509
Equatorial Guinea	15.247	0.056	Angola	164.069	0.529
Algeria	23.089	0.075	Tanzania	171.451	0.545
Gabon	22.555	0.077	Canarias	167.345	0.554
Mauritania	26.133	0.083	Burundi	175.094	0.555
Seychelles	33.432	0.097	Ethiopia	174.560	0.555
Mauritius	33.937	0.126	Cote D' Ivoire	170.183	0.555
Chad	43.717	0.140	Burkina Faso	174.283	0.573
Liberia	41.498	0.148	Reunion	167.598	0.577
Comoros	45.797	0.160	Malawi	185.540	0.587
Congo	49.338	0.169	Gambia	190.583	0.595
Mali	70.920	0.227	Nigeria	187.229	0.610
Niger	78.068	0.242	Sierra Leone	185.612	0.613
Tunisia	81.901	0.262	Ghana	194.802	0.630
Cameroon	87.971	0.293	Senegal	211.942	0.633
Sudan	94.708	0.298	Benin	212.662	0.654
Djibouti	98.053	0.316	Guinea-Bissau	208.493	0.656
Uganda	105.541	0.345	Zambia	225.873	0.701
Congo	109.305	0.358	Central African	226.201	0.710
Kenya	115.615	0.374	Zimbabwe	228.842	0.716
Morocco	117.283	0.380	Swaziland	222.681	0.723
Lesotho	114.813	0.401	Guinea	226.978	0.736
Eritrea	127.472	0.411	Togo	238.801	0.770
Botswana	136.813	0.422	Mozambique	259.578	0.814
Namibia	133.939	0.431			

We identified five heterogeneity levels in Africa using the LCCI (Table 3). Variable heterogeneity levels reveal the uneven distribution of spatial patterns in Africa (Figure 7). Most high-heterogeneity areas are concentrated in natural or mixed areas. Artificial and agricultural area heterogeneity is low due to human-made planning. In general, Africa's land cover heterogeneity is not evident. On the coast of East Africa, the heterogeneity shows both high diversity and fragmentation, indicating that the coastal region has a complex pattern and more attention should be paid in both planning and monitoring.

Table 3. Heterogeneity levels of land cover in Africa.

| Complexity Index	Heterogeneity Level	Distribution Characteristics
<0.2	Very low	Area with dispersed forest and bare land areas.
(0.2, 0.6)	low	Mainly in grassland and cropland mixed areas.
(0.6, 1)	Median	Natural and semi-natural mixed areas show moderate fragmentation.
(1, 1.5)	High	Distributed in grassland and shrub mixed regions, grasslands, and forest mixed regions.
>1.5	Very high	A small number of areas distributed in a natural area with shrubs, grassland, and forest mosaic.
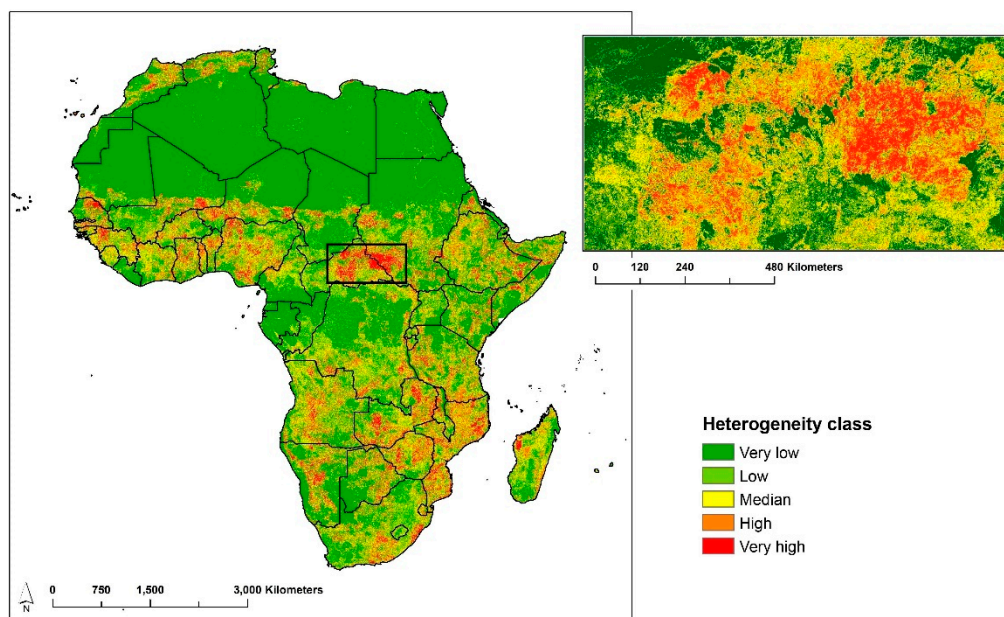
Figure 7. LCCI distribution for Africa. The map shows the high heterogeneity regions for grass-shrub and forest mosaic land covers in Central Africa.

5. Summary and Conclusions

Usually, land cover heterogeneity appears to be captured easily by landscape metrics. Numerous indices have been used to quantify land cover heterogeneity by describing features such as density, texture, size, and area. However, choosing an appropriate method for robust quantification on a global scale is still challenging because no single index can adequately take into account the whole spectrum of spatial characteristics [36]. In this study, a consistent indicator for large-scale land cover heterogeneity quantification was developed based on information theory. This measure effectively extracts more comprehensive information to distinguish the spatial variation of the land cover distribution at the continental level. Our experimental results suggest that the LCCI, a standardized and harmonized indicator, may be a good candidate parameter for large-scale geospatial sampling considerations of heterogeneity features [49–51]. One advantage of the LCCI is its consistent information theory framework because it eliminates the need to standardize, whereas landscape metrics are characterized by multiple value ranges and strong correlations between each other that necessitate the elimination of redundancy [52]. Furthermore, the moderate resolution of the land cover dataset can capture features at any scale, provided they are greater than 30 m, for analyses, and the temporal updateability and easy accessibility of the data should promote land cover heterogeneity data applications in environmental conservation.

Another important advantage of the entropy-based LCCI is that it extracts more abundant heterogeneity information compared with single landscape metrics by utilizing a fusion approach, thus successfully capturing information closer to the true heterogeneity of the surface. This is especially important for landscape ecology research. Within a sampling unit, the same land cover configuration may have higher LCCI values, identifying richer land cover type distributions and more complex arrangements. In such regions, edge effects may lead to unstable habitats. A recent study indicated that heterogeneous land cover mosaics may be represented as separate classes [4]. By measuring the heterogeneity using the LCCI, similar land cover patterns can be identified that offer valuable information to relevant developers.

A thorough evaluation and comparison for all quantification indices is beyond the purpose of this study. Instead, the SHDI index was chosen as a representative basic composition metric for validating the LCCI because of its computational simplicity and ease of interpretation. The results show that the SHDI does not always express the compositional complexity of the 1 km× 1 km units. This is

because it extracts not only the diversity of land cover types but also the evenness of the distributions. Therefore, small numbers of classes with even distributions may have high SHDI values, even though those regions are in actuality not high diversity regions. As a relative indicator for the assessment of heterogeneity change in the same region in different periods, it is still excellent. The diversity of land cover, however, should be fully explored through explicit classification of mosaic types in the future. For quantifying the configuration heterogeneity, the ED index is the most suitable when compared with the patch-based metrics, because it is easily computable, which is a big advantage. However, the ED index overlooks patch information, which results in underestimation of the heterogeneity. It is noteworthy that, although the LCCI is positively correlated with ED, PD, SHDI, and SPILT metrics, as shown by path analysis, the meanings of LCCI are different from them due to the fact that it quantifies land cover heterogeneity by incorporation of composition and configuration simultaneously. Because the LCCI provides comprehensive land cover heterogeneity information and thus more closely captures the actual degree of heterogeneity, we predict that the LCCI heterogeneity information can resolve fine-grained land cover variations.

Understanding the key role of the spatial scale is essential in geography analysis [53]. No optimal measurement scale exists because land cover patterns are naturally scale-dependent [13]. The scale should be selected based on at least one principle—the scale should be large enough to stand for one unit of landscape and reflect the heterogeneity features [54]. Previous studies have shown that a scale of 1 km^2 is useful for studying land cover heterogeneity at country and continental scales [55,56]. At the continental scale, we chose 1 km × 1 km square cells for this study, because this unit size is typical for the representation of local landscapes, supporting the subsequent analysis at the national level by aggregating the available metrics. In addition, the resolution of the selected size (1 km) allows for easy resampling of socioeconomic data (1 km) for future heterogeneity change and associated driving force research.

The GlobeLand30 dataset for the year 2010 was used for extracting the African heterogeneity characteristics. However, the accuracy of the GlobeLand30 leads to uncertainties of the heterogeneity characteristics, and, hence, misclassification of land cover data is unavoidable. Heterogeneity data extracted directly from remote sensing data may be a satisfying solution for research that requires high precision heterogeneity data. A recent study improved the land cover mapping accuracy by clustering the heterogeneity types of land cover, which helped to improve the classification accuracy of remote sensing-based land cover mapping [57]. In this study, we extracted heterogeneity information for spatial variation analysis. For environmental monitoring, obtaining the heterogeneity of each class is essential, and its success depends on the classification accuracy.

Overall, the LCCI is a novel indicator that can provide detailed information on land cover heterogeneity to support regional planning and ecological assessment. By integrating the occurrence of land cover differences between neighboring grids and information theory, we (i) propose the LCCI, a consistent scheme for the quantification of land cover heterogeneity and (ii) build a database of continent-scale land cover heterogeneity-elemental data for sustainable development monitoring and geographical analysis. Further, the performance of selected metrics at both regional and continental scales was evaluated, and the LCCI was found to enhance the robustness of land cover pattern characterization and distinction by combining both composition and configuration information. Meanwhile, our results show an improved accuracy compared with single metric approaches. We expect that our work will contribute to large-scale environmental sustainability monitoring and conservation planning by providing more direct data. Future work will attempt to apply our entropy-based index to the extraction of homogeneous land cover regions at multiple scales, which will simplify spatial statistics, increase their efficacy, and improve the meaningful to analyze.

Author Contributions: Writing—original draft, Jing Yu; Methodology, Jing Yu; Data curation, Shu Peng and Weiwei Zhang; Funding acquisition, Shu Peng; Investigation, Jing Yu and Shun Kang; Supervision, Shu Peng. All authors have read and agreed to the published version of the manuscript.

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

References

1. Skidmore, A.K.; Pettorelli, N.; Coops, N.C.; Geller, G.N.; Hansen, M.; Lucas, R. Environmental science: Agree on biodiversity metrics to track from space. *Nature* **2015**, *523*, 403–405. [CrossRef] [PubMed]
2. Tuanmu, M.-N.; Jetz, W.A. Global remote sensing-based characterization of terrestrial habitat heterogeneity for biodiversity and ecosystem modelling. *Glob. Ecol. Biogeogr.* **2015**, *24*, 1329–1339. [CrossRef]
3. Yu, Q.; Hu, Q.; van Vliet, J.; Verburg, P.H.; Wu, W. GlobeLand30 shows little cropland area loss but greater fragmentation in China. *Int. J. Appl. Earth Obs.* **2018**, *66*, 37–45. [CrossRef]
4. Verburg, P.H.; Neumann, K.; Nol, L. Challenges in using land use and land cover data for global change studies. *Glob. Chang. Biol.* **2011**, *17*, 974–989. [CrossRef]
5. Li, S.; Yang, B. Introducing a new method for assessing spatially explicit processes of landscape fragmentation. *Ecol. Indic.* **2015**, *56*, 116–124. [CrossRef]
6. Schindler, S.; von Wehrden, H.; Poirazidis, K.; Hochachka, W.M.; Wrbka, T.; Kati, V. Performance of methods to select landscape metrics for modelling species richness. *Ecol. Model.* **2015**, *295*, 107–112. [CrossRef]
7. Hernández, A.; Miranda, M.; Arellano, E.C.; Saura, S.; Ovalle, C. Landscape dynamics and their effect on the functional connectivity of a Mediterranean landscape in Chile. *Ecol. Indic.* **2015**, *48*, 198–206. [CrossRef]
8. Parrott, L. Measuring ecological complexity. *Ecol. Indic.* **2010**, *10*, 1069–1076. [CrossRef]
9. Chen, F.; Chen, J.; Wu, H.; Hou, D.; Zhang, W.; Zhang, J.; Zhou, X.; Chen, L. A landscape shape index-based sampling approach for land cover accuracy assessment. *Sci. China Earth Sci.* **2016**, *59*, 2263–2274. [CrossRef]
10. Tran, T.; Julian, J.; de Beurs, K. Land cover heterogeneity effects on sub-pixel and per-pixel classifications. *ISPRS Int. J. Geo.-Inf.* **2014**, *3*, 540–553. [CrossRef]
11. Cochran, F.; Daniel, J.; Jackson, L.; Neale, A. Earth observation-based ecosystem services indicators for national and subnational reporting of the sustainable development goals. *Remote Sens. Environ.* **2020**, *244*, 111796. [CrossRef] [PubMed]
12. Gustafson, E.J. Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems* **1998**, *1*, 143–156. [CrossRef]
13. Riitters, K.H. *Spatial Patterns of Land Cover in the United States: A Technical Document Supporting the Forest Service 2010 RPA Assessment*; Gen. Tech. Rep. SRS–136; Department of Agriculture Forest Service, Southern Research Station: Asheville, NC, USA, 2011; 64p.
14. Jaeger, J.A.G.; Madrinan, L.F. *Landscape Fragmentation in Europe*; Joint EEA-FOEN report; European Environmental Agency, Publications Office of the European Union: Luxembourg, 2011; pp. 20–46, ISSN 1725-9177. [CrossRef]
15. Dale, V.H.; Kline, K.L. Issues in using landscape indicators to assess land changes. *Ecol. Indic.* **2013**, *28*, 91–99. [CrossRef]
16. Nalewajski, R.F. Elements of Information Theory. In *Perspectives in Electronic Structure Theory*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 371–395. [CrossRef]
17. Proulx, R.; Parrott, L. Measures of structural complexity in digital images for monitoring the ecological signature of an old-growth forest ecosystem. *Ecol. Indic.* **2008**, *8*, 270–284. [CrossRef]
18. Vranken, I.; Baudry, J.; Aubinet, M.; Visser, M.; Bogaert, J. A review on the use of entropy in landscape ecology: Heterogeneity, unpredictability, scale dependence and their links with thermodynamics. *Landsc. Ecol.* **2015**, *30*, 51–65. [CrossRef]
19. Andrienko, Y.A.; Brilliantov, N.V.; Kurths, J. Complexity of two-dimensional patterns. *Eur. Phys. J. B* **2000**, *15*, 539–546. [CrossRef]
20. Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [CrossRef]
21. Li, H.; Reynolds, J.F. A new contagion index to quantify spatial patterns of landscapes. *Landsc. Ecol.* **1993**, *8*, 155–162. [CrossRef]
22. Bogaert, J.; Farina, A.; Ceulemans, R. Entropy increase of fragmented habitats: A sign of human impact? *Ecol. Indic.* **2005**, *5*, 207–212. [CrossRef]
23. Cousins, S.A.O.; Auffret, A.G.; Lindgren, J.; Trank, L. Regional-scale land-cover change during the 20th century and its consequences for biodiversity. *Ambio* **2015**, *44*, S17–S27. [CrossRef]

24. Chmielewski, S.; Chmielewski, T.J.; Tompalski, P. Land cover and landscape diversity analysis in the West Polesie Biosphere Reserve. *Int. Agrophys.* **2014**, *28*, 153–162. [CrossRef]

25. Gao, P.C.; Li, Z.L. Computation of the Boltzmann entropy of a landscape: A review and a generalization. *Landsc. Ecol.* **2019**, *34*, 2183–2196. [CrossRef]

26. Wang, C.; Zhao, H. Spatial heterogeneity analysis: Introducing a new form of spatial entropy. *Entropy* **2018**, *20*, 398. [CrossRef]

27. Claramunt, C. *A Spatial form of Diversity*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 218–231.

28. Li, Z.L.; Huang, P.Z. Quantitative measures for spatial information of maps. *Int. J. Geogr. Inf. Sci.* **2002**, *16*, 699–709. [CrossRef]

29. Altieri, L.; Cocchi, D.; Roli, G. A new approach to spatial entropy measures. *Environ. Ecol. Stat.* **2018**, *25*, 95–110. [CrossRef]

30. Gao, P.C.; Li, Z.L.; Zhang, H. Thermodynamics-Based Evaluation of Various Improved Shannon Entropies for Configurational Information of Gray-Level Images. *Entropy* **2018**, *20*, 19. [CrossRef]

31. Cushman, S.A. Calculation of Configurational Entropy in Complex Landscapes. *Entropy* **2018**, *20*, 298. [CrossRef]

32. Cushman, S.A. Calculating the configurational entropy of a landscape mosaic. *Landsc. Ecol.* **2015**, *31*, 481–489. [CrossRef]

33. Cushman, S.A. Editorial: Entropy in Landscape Ecology. *Entropy* **2018**, *20*, 314. [CrossRef]

34. Gao, P.C.; Zhang, H.; Li, Z.L. An efficient analytical method for computing the Boltzmann entropy of a landscape gradient. *Trans. GIS* **2018**, *22*, 1046–1063. [CrossRef]

35. Gao, P.; Zhang, H.; Li, Z. A hierarchy-based solution to calculate the configurational entropy of landscape gradients. *Landsc. Ecol.* **2017**, *32*, 1133–1146. [CrossRef]

36. Cushman, S.A.; McGarigal, K.; Neel, M.C. Parsimony in landscape metrics: Strength, universality, and consistency. *Ecol. Indic.* **2008**, *8*, 691–703. [CrossRef]

37. Fahrig, L.; Nuttle, W.K. Population Ecology in Spatially Heterogeneous Environments. In *Ecosystem Function in Heterogeneous Landscapes*; Lovett, G.M., Turner, M.G., Jones, C.G., Weathers, K.C., Eds.; Springer: New York, NY, USA, 2005; pp. 95–118. [CrossRef]

38. Llausàs, A.; Nogué, J. Indicators of landscape fragmentation: The case for combining ecological indices and the perceptive approach. *Ecol. Indic.* **2012**, *15*, 85–91. [CrossRef]

39. Riitters, K. Pattern metrics for a transdisciplinary landscape ecology. *Landsc. Ecol.* **2018**, *34*, 2057–2063. [CrossRef]

40. Riitters, K.H.; O'Neill, R.V.; Hunsaker, C.T.; Wickham, J.D.; Yankee, D.H.; Timmins, S.P.; Jones, K.B.; Jackson, B.L. A factor analysis of landscape pattern and structure metrics. *Landsc. Ecol.* **1995**, *10*, 23–39. [CrossRef]

41. Nowosad, J.; Stepinski, T.F. Information theory as a consistent framework for quantification and classification of landscape patterns. *Landsc. Ecol.* **2019**, *34*, 2091–2101. [CrossRef]

42. Bordogna, G.; Boschetti, M.; Brivio, P.A.; Carrara, P.; Stroppiana, D.; Weissteiner, C.J. Handling heterogeneous bipolar information for modelling environmental syndromes of global change. *Environ. Modell Softw.* **2012**, *36*, 131–147. [CrossRef]

43. ESRI, R. *ArcGIS Desktop: Release 10*; Environmental Systems Research Institute: Redlands, CA, USA, 2011.

44. Niesterowicz, J.; Stepinski, T.F. Regionalization of multi-categorical landscapes using machine vision methods. *Appl. Geogr.* **2013**, *45*, 250–258. [CrossRef]

45. Peng, J.; Wang, Y.; Zhang, Y.; Wu, J.; Li, W.; Li, Y. Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. *Ecol. Indic.* **2010**, *10*, 217–223. [CrossRef]

46. Tian, Y.; Jim, C.Y.; Tao, Y.; Shi, T. Landscape ecological assessment of green space fragmentation in Hong Kong. *Urban. For. Urban. Green* **2011**, *10*, 79–86. [CrossRef]

47. Reddy, C.S.; Sreelekshmi, S.; Jha, C.S.; Dadhwal, V.K. National assessment of forest fragmentation in India: Landscape indices as measures of the effects of fragmentation and forest cover change. *Ecol. Eng.* **2013**, *60*, 453–464. [CrossRef]

48. Thakkar, J.J. Applications of Structural Equation Modelling with AMOS 21, IBM SPSS. In *Structural Equation Modelling: Application for Research and Practice (with AMOS and R)*; Springer: Singapore, 2020; pp. 35–89. [CrossRef]

49. Mayaux, P.; Eva, H.; Gallego, J.; Strahler, A.H.; Herold, M.; Agrawal, S.; Naumov, S.; De Miranda, E.E.; Di Bella, C.M.; Ordoyne, C.; et al. Validation of the global land cover 2000 map. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1728–1737. [CrossRef]

50. Zeng, Y.L.; Li, J.; Liu, Q.H.; Li, L.H.; Xu, B.D.; Yin, G.F.; Peng, J.J. A Sampling Strategy for Remotely Sensed LAI Product Validation Over Heterogeneous Land Surfaces. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **2014**, *7*, 3128–3142. [CrossRef]

51. Wang, S.G.; Li, X.; Ge, Y.; Jin, R.; Ma, M.G.; Liu, Q.H.; Wen, J.G.; Liu, S.M. Validation of Regional-Scale Remote Sensing Products in China: From Site to Network. *Remote Sens.* **2016**, *8*, 980. [CrossRef]

52. Niesterowicz, J.; Stepinski, T.F. On using landscape metrics for landscape similarity search. *Ecol. Indic.* **2016**, *64*, 20–30. [CrossRef]

53. Li, H.; Peng, J.; Yanxu, L.; Yi'na, H. Urbanization impact on landscape patterns in Beijing City, China: A spatial heterogeneity perspective. *Ecol. Indic.* **2017**, *82*, 50–60. [CrossRef]

54. Freedman, D.; Diaconis, P. On the histogram as a density estimator: L 2 theory. *Probab. Theory Rel.* **1981**, *57*, 453–476.

55. Bunce, R.G.H.; Metzger, M.J.; Jongman, R.H.G.; Brandt, J.; de Blust, G.; Elena-Rossello, R.; Groom, G.B.; Halada, L.; Hofer, G.; Howard, D.C.; et al. A standardized procedure for surveillance and monitoring European habitats and provision of spatial data. *Landsc. Ecol.* **2007**, *23*, 11–25. [CrossRef]

56. Wrbka, T.; Schindler, S.; Pollheimer, M.; Schmitzberger, I.; Peterseil, J. Impact of the Austrian agri-environmental scheme on diversity of landscapes, plants and birds. *Community Ecol.* **2008**, *9*, 217–227. [CrossRef]

57. Zheng, Z.; Du, S.; Wang, Y.-C.; Wang, Q. Mining the regularity of landscape-structure heterogeneity to improve urban land-cover mapping. *Remote Sens. Environ.* **2018**, *214*, 14–32. [CrossRef]