Landscape Ecol (2019) 34:1849–1858
https://doi.org/10.1007/s10980-019-00876-x

RESEARCH ARTICLE

Calculating spatial configurational entropy of a landscape mosaic based on the Wasserstein metric

Yuan Zhao · Xinchang Zhang

Received: 7 September 2018 / Accepted: 28 June 2019 / Published online: 9 August 2019
© Springer Nature B.V. 2019

Abstract
Context Entropy is an important concept traditionally associated with thermodynamics and is widely used to describe the degree of disorder in a substance, system, or process. Configurational entropy has received more attention because it better reflects the thermodynamic properties of physical and biological processes. However, as the number of configuration combinations increases, configurational entropy becomes too complex to calculate, and its value is too large to be accurately represented in practical applications.
Objectives To calculate the spatial configurational entropy of a landscape mosaic based on a statistical metric.
Methods We proposed a relative entropy using histograms to compare two ecosystems with the Wasserstein metric, and used six digital elevation models and five simulated data to calculate the entropy of the complex ecosystems.
Results The calculation and simulation showed that the purposed metric captured disorder in the spatial landscape, and the result was consistent with the general configurational entropy. By calculating several spatial scale landscapes, we found that relative entropy can be a trade-off between the rationality of results and the cost of calculation.
Conclusions Our results show that the Wasserstein metric is suitable to capture the discrepancy using complex landscape mosaic data sets, which provides a numerically efficient approximation for the similarity in the histograms, reducing excessive expansion of the calculated result.

Keywords Boltzmann entropy · Configurational entropy · Landscape mosaic · Shannon entropy · Wasserstein metric

Y. Zhao
School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China
e-mail: giszy@qq.com;
zhaoyuan6@mail.sysu.edu.cn

X. Zhang
School of Geographical Sciences, Guangzhou University, Guangzhou 510006, China

X. Zhang (✉)
College of Environment and Planning, Henan University, Kaifeng 475004, China
e-mail: eeszxc@mail.sysu.edu.cn

Introduction and motivation

Entropy is an important concept traditionally associated with thermodynamics, and is widely used to describe the degree of disorder in a substance, system, or process. In ecology, entropy is often used to express the diversity or uncertainty of a landscape mosaic. There are two definitions of entropy, such as thermodynamic entropy and information-theoretical entropy.

Thermodynamic entropy is usually denoted by S of a physical system in the statistical thermodynamics established by Boltzmann and Willard Gibbs in the 1870s. Information theoretical entropy is usually expressed as H, and was developed by Claude Shannon and Ralph Hartley in the 1940s.

Shannon's entropy is more widely used in practice. Research in many fields has focused directly or indirectly on information entropy. According to a review by Vranken et al., the most widely used concept of entropy in ecology is concentrated in three indices (Vranken et al. 2015): the Shannon diversity index (Vajda et al. 1949), the Simpson diversity index (Simpson 1949), and the Brillouin index (Brillouin 1956). These indicators are an expansion of Shannon's entropy. Information entropy is the core of the calculations for these indicators.

Entropy of interest in the geographic arena has been used to evaluate the fragmentation and spatial heterogeneity of geographical phenomena (Batty 1976; Gatrell 1977; Foody 1995; Tobler 1997; Goodchild 2003).

More and more studies (Li and Huang 2002; Leibovici et al. 2014; Leibovici and Birkin 2015) have shown that Shannon's entropy only considers the number of symbols of each type, and the spatial arrangement of these symbols is completely neglected, and a set of metrics using information entropy should consider spatial neighbors. Related cases are more common in ecology, economics, and geography, particularly in urban geography and regional science (Batty 1976; Snickars and Weibull 1977; Feldman and Crutchfield 2003; Bogaert et al. 2005; Leibovici 2009; Li et al. 2016).

The first attempts to integrate some specific spatial properties in the measure of Shannon's entropy was suggested by Li and Reynolds (1993). The main idea was to quantify a measure of contagion and to which extent regions of a given class are adjacent to regions of another class, allowing the evaluation of the degrees of juxtaposition and aggregation of the categorical data.

Following a similar idea, in the field of cartography, Bjorke proposed a strategy to consider spatial properties of a map for entropy computations (Bjoke 1996). He provided a definition of positional entropy, which considers all occurrences of map entities as unique events. Bjorke used conditional entropy to estimate the similarity of two maps to calculate positional entropy. Neumann used an entropy measure based on the degree of the vertices in a graph, showing all connections within a map among point, line, and area symbols and obtained the "topological information content" of the map (Neumann 1994). Leibovici et al. developed severalspatial entropy indices to characterize the distribution of some instances in space and time (Leibovici et al. 2014).

However, the applicability of Shannon's entropy in these studies has been questioned, because there is no confirmed relevance between Shannon's entropy and the spatial pattern or spatial processing. For example, Vranken et al. considered that no thermodynamic interpretation of information theory is relevant and that information entropy is merely a formal parallel to thermodynamic entropy (Vranken et al. 2015). Li et al. reported that it is necessary to explore whether statistical information entropy can be used to effectively measure map information (Li et al. 2016).

Some researchers have returned to Boltzmann's entropy, also called configurational entropy, along the basic principles of thermodynamics/statistical physics. Cushman reported that landscape ecologists should consider the relationship between landscape ecology and thermodynamics (Cushman 2015). He believed that it is necessary to define structural entropy of a landscape mosaic to express a baseline to compare landscape patterns. Cushman's idea basically follows Boltzmann's entropy. The difference between the two definitions is that they are both an interpretation of the same mathematical framework with different terms.

It is necessary to consider the original definition of entropy when we talk about Boltzmann's entropy. Historically, Boltzmann proposed using $S \propto \ln W$ to indicate the degree of system disorder. Around 1900, Planck improved the formalization of entropy proposed by Boltzmann and described it as $S = k \ln W$; where k is the Boltzmann constant, and S is the macrosystem entropy value, which is a measure of the degree of molecular motion or disorder. W is the number of possible microstates. The higher the W, the higher the degree of disorder.

Notably, Boltzmann's entropy value corresponds directly to the number of states W, and it is an absolute value rather than a relative value. Boltzmann suggested using a distribution of microscopic states. He assumed that each state of system particles is equal, and that the number N is an ideal gas, where N_i is the state in case of i, and the overall expression of the state is diverse.

In the definition of Boltzmann's entropy, the potential trouble is W. In the microscopic case, we likely know the specific value of W, because it can accurately evaluate the state of the system from microscopic observations. However, the value of W for a given very large system, even in logarithmic form, causes numerical overflow, so the value cannot be expressed exactly.

Cushman used ecological terminology to give Boltzmann's entropy equation spatial significance by transforming the number and state of thermodynamic particles into the number and classification of landscape categories (Cushman 2016). In this way, Boltzmann's entropy is converted to landscape mosaic configurational entropy. The difference between the two is that the microscopic features of thermodynamic entropy are difficult to observe and measure, whereas the microscopic features of spatial configurational entropy are relatively easy to detect. Therefore, entropy is converted to calculate the spatial configuration.

Gao et al. extended the calculation to a spatial multi-scale representation of the landscape gradient model (Gao et al. 2017). They focused on splitting the landscape lattice into smaller units and recalculated the number of configurations for each unit by repeating the above steps. The advantage of their method is that the spatial scale is more suitable for calculating landscape gradient entropy. However, several unsolved problems are present in their work.

The first is that calculations of the microscopic state are limited to small units, which means that the relationship of particles may be ignored in the entire space. The second is that the expression of entropy is dependent on calculating power and multisets, resulting in intensive computing and numerical overflow.

The challenge to calculate configurational entropy is that the number of points becomes large, and the size of states grows exponentially, so the combination with a huge number of microstates may result in failure.

Cushman also realized that calculating configurational entropy is too complex. He proposed a method to estimate the edge in a landscape configuration using the shuffle permutation, the central limit theorem, the normal probability density function, and the microstate ratio (Cushman 2018).

Although we question the suitability of Shannon's entropy for spatial configuration, it is undeniable that Shannon's entropy is more useful because of its simpler form. We may imitate the definition of Shannon's entropy and rethink configurational entropy from a statistical point of view.

The main objective of this study was to propose an appropriate metric that can replace absolute configurational entropy and meet the requirements of mathematical axioms.

This article is organized as follows. First, we explore the structure of configurational entropy and outline a strategy to convert the permutation combination in configurational entropy into a probability distribution problem. We then propose a method using the Wasserstein metric to measure configurational entropy of a system. We carried out a simulation and experimental validation of our proposed strategy with finite samples. Finally, we present a discussion and provide some concluding remarks.

Spatial configurational entropy

Following the definition of Boltzmann's entropy, the definition of configurational entropy is: (Cushman 2018; Gao et al. 2017; Cushman 2016)

$$S = \ln W \tag{1}$$

where W is the number of configuration permutations. If all configurations are distributed throughout a completely asymmetrical structure containing N units, then the number of individual configurations is easy to determine as the number of permutations of N objects divided by the number of permutations within the indistinguishable objects. Specifically, configurational entropy can be written as:

Specifically, configurational entropy can be written as:

$$S = \ln \frac{N!}{N_1!N_2!\cdots N_m!} \tag{2}$$

$$= \ln N! - \sum_{i=1}^{m} \ln N_i! \tag{3}$$

where N is the total number of configurations, m is the number of classes, and $N_1, N_2, \cdots, N_m$ is the number of each class. In addition, $N_1 + N_2 + N_3 + \cdots = N$.

Unfortunately, the problem with this definition for spatial or image analysis is obtaining the same value for every different spatial configuration (Fig. 1).

Nevertheless, some of the same elements may be concentrated in continuous space. This continuity can be considered a repetition of spatial configuration. Under the constraint of spatial continuity, the entropy of the system will decrease more than the random distribution.

The original definition of configurational entropy ignores the role of space. The value must be recalculated for spatial configuration to meet the needs.

Specifically, if class i appears in the consecutive area indexed as k, and the number of classes in each area is r_{ij} (Fig. 1), then entropy must be divided by

$$\prod_{j=1}^{k} r_{ij}! \tag{4}$$

Assume the number of classes is m, then all repeats due to spatial continuity are calculated as:

$$\prod_{i=1}^{m}\prod_{j=1}^{k} r_{ij}! \tag{5}$$

Further,

$$\ln\prod_{i=1}^{m}\prod_{j=1}^{k} r_{ij}! = \sum_{i=1}^{m}\sum_{j=1}^{k} \ln r_{ij}! \tag{6}$$

In summary, spatial configurational entropy can be rewritten as:

$$\ln W = \ln N! - \sum_{i=1}^{m} \ln N_i! - \sum_{i=1}^{m}\sum_{j=1}^{k} \ln r_{ij}! \tag{7}$$

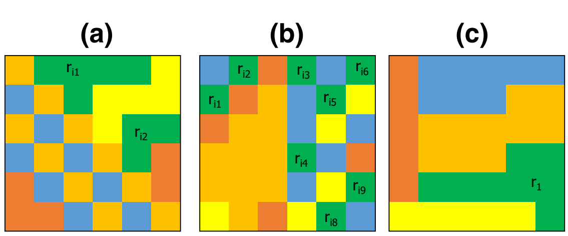

Fig. 1 The traditional configurational entropy values of a spatial configuration for different landscapes are the same

From configuration to statistics

It is necessary to examine the result before using Equation (7). For example, as spatial dimension increases from 10×10 to 1000×1000 , the number of differential permutations increases from 364 to 12815518. In fact, as the dimension and number of classes increases, the value of configurations rapidly becomes intractably large.

According to the analysis in the second section, the entropy value is determined by repetition of the elements. Focusing on the second part of configurational entropy, we can infer that:

$$\sum_{i=1}^{m} \ln N_i! = \ln N_1! + \ln N_2! + \cdots \ln N_m! \tag{8}$$

$$= (\ln 1 + \ln 2 + \cdots + \ln N_1) \tag{9}$$

$$+ (\ln 1 + \ln 2 + \cdots + \ln N_2) + \cdots \tag{10}$$

$$+ (\ln 1 + \ln 2 + \cdots + \ln N_m) \tag{11}$$

In Eq. (11), we temporarily do not know the relationship between $N_1, \cdots$, and N_m. When m is equal to 1, N_m is equal to N. To ensure that all cases are covered, we choose the longest sequence to express Eq. (11). It can be rewritten as:

$$\sum_{i=1}^{m} \ln N_i! = a_1 \ln 1 + a_2 \ln 2 + \cdots + a_N \ln N \tag{12}$$

$$= (a_1, a_2, \ldots, a_N)\ldots(\ln 1, \ln 2, \ldots, \ln N) \tag{13}$$

where $1 \le m \le N$, and $a_1, a_2, \ldots, a_N$ is the coefficient of $\ln 1, \ln 2, \ldots, \ln N$. The value of $\sum_{i=1}^{m} \ln N_i!$ is determined by the series $a_1, a_2, \ldots, a_N$. Actually, $a_1, a_2, \ldots, a_N$ is a histogram of $\ln 1, \ln 2, \ldots, \ln N$. It is undeniable that many of the coefficients in the series will be equal to zero, but this does not affect the true form of the histogram.

In the same way, the third part of Eq. (7) can also be written as:

$$\sum_{i=1}^{m}\sum_{j=1}^{k} \ln r_{ij}! = (b_1, b_2, \cdots, b_N) \cdot (\ln 1, \ln 2, \cdots, \ln N) \tag{14}$$

where $b_1, b_2, \ldots, b_N$ is the coefficient of $\ln 1, \ln 2, \ldots, \ln N$.

The metric for configurational entropy

If we directly calculate spatial configurational entropy using Eqs. (13) and (14), the problem of entropy overflow remains unresolved. We are more interested in the series in Eqs. (13) and (14) because it is simply a histogram.

A histogram is a representation of a data distribution, and an estimate of the distribution of variables. A histogram can be used to estimate the probability distribution by depicting the empirical frequency.

For example, consider a landscape consisting of a 3×3 lattice. There are 9! ways to arrange the cells in this landscape. Different combinations constitute various possibilities. When there are a large number of classes and repeatability is low, the configurational entropy value tends to be high (Fig. 2a), and when there are a small number of classes and repeatability is high, the configurational entropy value tends to be low (Fig. 2c).

Figure 2 shows that the series $\ln 1, \ln 2, \ldots, \ln 9$ represents the repetition of a configuration. When the histogram tends to be a Dirac-delta distribution, the value of configurational entropy reaches the maximum. In contrast, when the histogram tends to be uniform, the value of configurational entropy reaches the minimum.

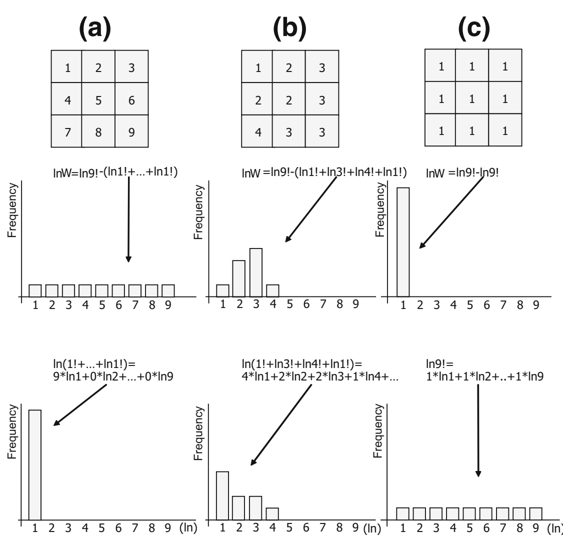

Fig. 2 Histogram of configurational entropy coefficients. The upper panels express the relationship between frequency and classification, while the lower panels express the relationship between frequency and the logarithmic sequence

In a sense, we can consider a histogram as a map of configurational entropy. However, using a histogram as a measure is neither convenient nor needed. To compare two histograms, we can use the Wasserstein metric to represent their similarities (Villani 2008).

In mathematics, the Wasserstein metric between the distributions u and v is:

$$l_1(u, v) = \inf_{\pi \in \Gamma(u,v)} \int_{\mathbb{R} \times \mathbb{R}} |x - y| \mathrm{d}\pi(x, y) \tag{15}$$

where $\Gamma(u, v)$ is the set of (probability) distributions on $\mathbb{R} \times \mathbb{R}$ whose marginals are u and v on the first and second factors respectively.

If U and V are the respective cumulative distribution functions of u and v , this distance also equals to:

$$l_1(u, v) = \int_{-\infty}^{+\infty} |U - V| \tag{16}$$

Intuitively, if two distributions are treated as two mounds, the Wasserstein metric is the minimum cost of having one mound to another. Computationally, two distributions are required to have the same amount to calculate the Wasserstein metric (just like two mounds have the same amount).

It is worth noting that the Wasserstein metric is a distance function defined between probability distributions. Unlike the transitional Euclidean distance, the Wasserstein metric is a natural way to compare the probability distributions of u and v, where one variable is derived from the other by small, non-uniform perturbations (random or deterministic). The Wasserstein metric defines the geometry over the space of probability measures using principles from optimal transport theory. It satisfies the mathematical requirements for the distance definition, meaning that it can be used to compare differences (Deza and Deza 2016; Díaz-Varela et al. 2016).

Calculating Spatial configurational entropy using the Wasserstein metric

With the Wasserstein metric, we measured a landscape mosaic and considered how to use this metric to calculate configurational entropy. Our strategy can be divided into three parts.

First, the histogram is map of a landscape mosaic configuration. According to the entropy equation, it describes the frequency of states. When the configuration changes, the shape of histogram also changes accordingly. When we want to know the change in configurational entropy, we can use the histogram to detect the change in entropy of a system.

Second, calculating the Wasserstein metric requires a reference. The Dirac delta distribution is a good choice. When a landscape mosaic is in the most chaotic state, its combined entropy is the largest. At this time, according to Eq. (13), the histogram formed by its state tends to be a Dirac delta distribution. It is like finding the center of a reference system, and the Dirac delta distribution can be used as a reference point for calculating spatial configurational entropy. Thus, the Dirac delta distribution is proposed to be v, which indicates that the system is in the most chaotic state. The u content can be filled by the histogram in a specified configuration.

Third, the calculated Wasserstein metric is used as a similarity, indicating how close the configurational entropy of a landscape mosaic is to the maximum value.

Now we look back at examples in the third section. Assume v is the baseline histogram and u is the same as v; then, the Wasserstein metric is 0 (Fig. 3a).

However, in Fig. 3b, the two histograms are significantly different due to the increased repeatability of the configuration. In Fig. 3c, the state represented by u is the highest order and the similarity between u and v reaches the maximum.

Notably, the metric calculated here is essentially the statistical similarity of the distribution of parameters in the second and third parts with respect to the specified distribution.

When the Wasserstein metric is 0, the selected parameters are exactly the same as in the specified distribution. For example, if the reference we selected corresponds to the largest value and the calculated Wasserstein metric is 0, it means that the parameters are mapping to the distribution with the maximum value. For consistency with the definition of entropy, the similarity w_c must be adjusted to $1 - W_c$. For the same reason, the Wasserstein metric in the third part also needs to be calculated.

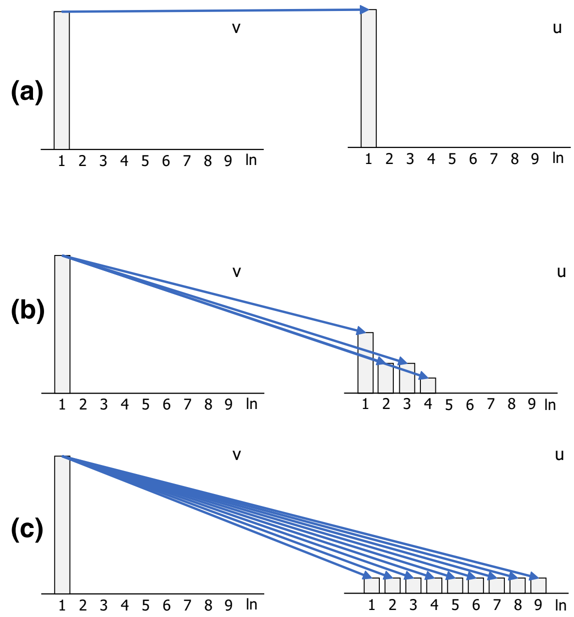

Fig. 3 The Wasserstein metric expresses the cost for the case in which the block on the left is transported to the right

Finally, we define relative configurational entropy as:

$$W_{\text{dist}} = (1 - W_c)(1 - W_s) \tag{17}$$

where W_c represents the similarity of the class parameters, and W_s represents the similarity of the spatial parameters.

For convenience, we have organized these ideas into an algorithm.

Algorithm 1 Calculating spatial configurational entropy.

Require: The raster set of a landscape mosaic, R_{data}; The set of a classification system for current raster data, C_{data};
Ensure: The spatial configurational entropy using Wasserstein metric of current landscape mosaic, W_{dist};
1: Extracting the unit classification C_{data} of the landscape mosaic R_{data}, and processing it as a histogram H_{data};
2: Constructing the Dirac delta distribution H_{Dirac} and initialize its parameters;
3: Calculating the Wasserstein distance (W_c) of H_{data} and H_{Dirac};
4: Calculating the spatial repeatability and representing it as a histogram H_{space};
5: Calculating the Wasserstein distance (W_s) of H_{space} and H_{Dirac};
6: **return** $W_{dist} = (1 - W_c)(1 - W_s)$;

According to the proposed algorithm, we calculate the spatial configurational entropy in Fig. 1. The results show that the relative configurational entropy is consistent with the absolute value, but it has the advantage of normalization and satisfying the distance axiom.

In fact, we are more concerned that the value of absolute entropy will explode in complex cases, while the relative configurational entropy remains stable within the range of 0–1. This finding will be discussed in the next discussion section.

Experimental validation

This section validates the proposed landscape metric by examining whether it is consistent with absolute configurational entropy and captures the spatial pattern. The original definition and the Wasserstein metric was used for a comparative analysis.

We choose six digital elevation models (DEMs) from the NASA Space Shuttle Radar Terrain Mission data (http://srtm.csi.cgiar.org/), and the size of each DEM was 800 × 800 pixels.

The selected DEM images are from the eastern and western regions of China. We roughly arranged these DEMs according to the degree of chaos in the local ecological landscape. The results are shown in Fig. 4.

The purposed metric successfully captured disorder in the landscape, and the Wasserstein metric was consistent with the absolute entropy calculation.

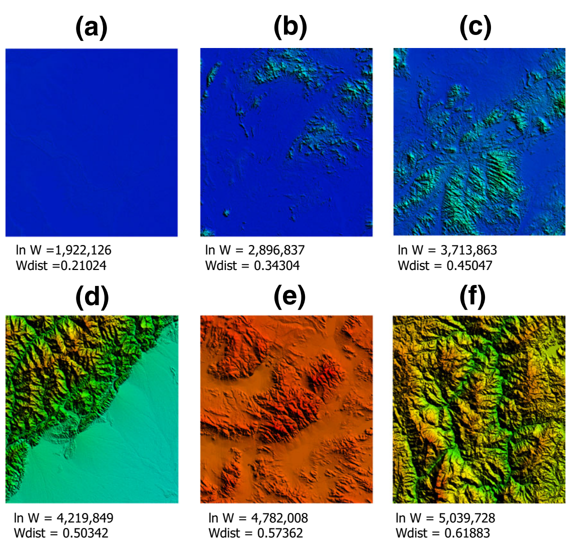

Fig. 4 Traditional configurational entropy and the Wasserstein metric were calculated according to the digital elevation models for the six regions of China, where $\ln W$ represents traditional configurational entropy and W_{dist} means the relative value

However, is the proposed metric suitable for the calculated entropy for different spatial configurations? Five simulated data with the same configurational entropy were used to demonstrate the case, and the results are shown in Fig. 5.

When the landscape class was unique, the entropy value was the lowest (Fig. 5). As the number of landscape classes increased, the entropy value increased quickly. According to the original calculation, the different spatial configurations can obtain the same results, but the proposed method can produce different results.

Discussion

Spatial arrangement

Configurational entropy provides detailed information on spatial arrangements. As described in the original definition of configurational entropy, the method for reduction is to find a repetitive configuration. Without it, there are $n!$ unique permutations to arrange the unit of a landscape represented as a lattice of n units.

As a general rule, the entropy of a spatial pattern of randomness is very high. In the extreme case, the entropy of a spatial pattern with complete aggregation is the lowest, while a spatial pattern that is completely dispersed has the highest entropy.

Many metrics can be used to describe entropy but the spatial arrangement is a challenging problem throughout a wide range of disciplines.

The original definition of configurational entropy does not consider the spatial arrangement, so we can embed it into the definition. Although our proposed solution effectively distinguishes the entropy of different spatial configurations, the arrangement method may be confusing. It prioritizes the different

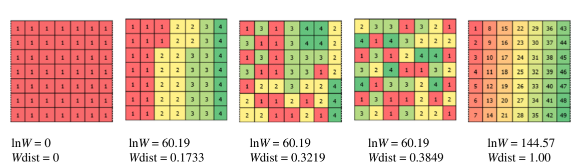

Fig. 5 The overall landscape sequence is organized by complexity to represent the relationship between spatial configuration and entropy. The traditional metric of the three middle landscapes is equal, but a significant difference was calculated by the Wasserstein metric

contiguous spaces, rather than considering factors, such as the area and perimeter of the spatial features. However, there is no consensus on how to introduce spatial arrangement in configurational entropy or how to define repetition. We followed the principle of expressing repetitiveness as a permutation and combination problem. If repeatability is defined by other means, it may be far from the definition of configurational entropy and it may become more difficult to maintain the characteristics of configurational entropy and the appropriate extension.

It is necessary to develop a common form to unify the spatial arrangement metrics into the configurational entropy formula. Previous studies used neighbor counts of different quantities to construct entropy related to metric information (proportion of areas using Voronoi regions of labels), topological information (no entropy but average number of neighbors in a Delaunay graph), and thematic information (using Voronoi regions adjacency counts) (Li and Huang 2002).

Cushman considered that the total edge length landscape metric is an appropriate measurement (Cushman 2016, 2018). Gao et al. (2017) proposed adding up all of the permutations for each multiset to obtain the microstate. In fact, the objective of all metrics is to find spatial repetitiveness.

Therefore, in future research, it will be necessary to develop a suitable indicator that describes spatial configuration in a configurational entropy formula.

Multi-scale effect of spatial configurational entropy

A landscape can be described using different scales of observation. For example, we know that configurational entropy at a variable spatial scale is completely different. At this time, the normalized configurational entropy allows us to know the level of landscape chaos at each scale.

When the spatial scale changes, the classification system on the ecological landscape will also change. Here we used the reclassification method to adapt the spatial change. According to the scale, we divided the landscape into 12 levels to explore the relationship between configurational entropy and scales. Lastly, we calculated the metric separately at each level, and the results are shown in Fig. 6.

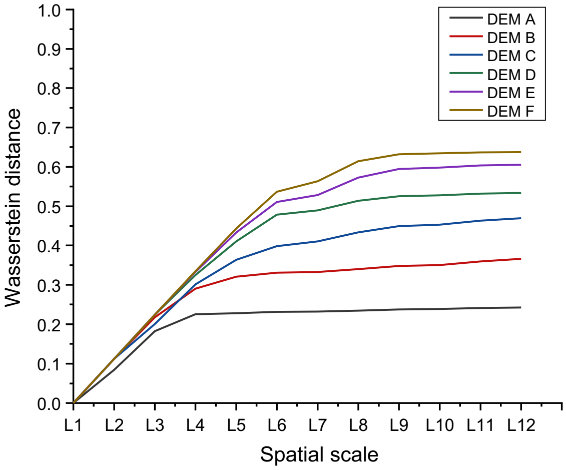

Fig. 6 The Wasserstein metric was significantly different for the configurations. It varied from 0 to separate maximum value at different scales in the L1–L12 range. Numerical comparisons can be made in different configurations and scales because the Wasserstein metric is normalized. Close values indicate that two landscape mosaics have similar configurations

In response to this problem, we constructed four sets by shuffling actual data and calculated the Wasserstein metric. The results are shown in Fig. 7.

As shown in Fig. 7, the entropy of the shuffled data decreased faster as the scale became smaller. The results tell us that we should determine the appropriate range to guarantee the availability of the suggested metric.

The results show that, when the spatial scale relative to ecological landscape is too large or small, the effect of spatial configurational entropy on scale is different. When the scale is too small, the diversity of the ecological landscape is neglected, resulting in too small spatial configurational entropy. This will make it difficult to compare different ecosystems.

Conversely, when the spatial scale is too large, the diversity of the ecological landscape is stabilized, but it will lead to unnecessary calculation costs. Therefore, to prevent the affecting comparison of landscape systems at the same scale, we recommend testing some samples to determine the most appropriate spatial scale. This approach can be a trade-off between the rationality of results and the cost of calculation.

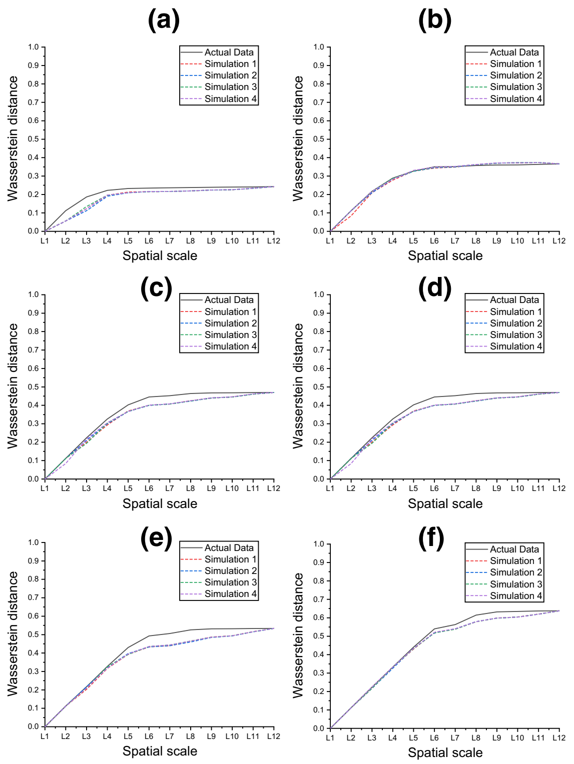

Fig. 7 Comparison of the Wasserstein metric between the real data and four sets of shuffled data. **a–f** Represent the order of actual data from simple to complex

Conclusions

In this study, we have introduced a novel application of the Wasserstein metric to quantify spatial configurational entropy of a landscape mosaic. It was demonstrated that the Wasserstein metric is suitable to capture the discrepancy between different spatial configurations. The normalized Wasserstein metric was used in the analysis of actual landscape data and performed well.

Although the main objective was to develop a metric that better represented configurational entropy using complex landscape mosaic data sets, we also note that the Wasserstein metric provided a numerically efficient approximation for the similarity in histograms, reducing excessive expansion of the calculated results.

A next step would be to consider a more detailed study of a specific application. For example, we could have considered more statistical distribution distances, such as the energy metric to make configurational entropy more reasonable. We could have embedded the spatial effects in the configurational entropy expression to explore the relationship between spatial scale, reclassification, and configurational entropy.

Acknowledgements We thank anonymous reviewers for their constructive comments. This research was supported by the National Natural Science Foundation of China (Grant No. 41431178), the Natural Science Foundation of Guangdong Province in China (Grant No. 2016A030311016), the National Administration of Surveying, Mapping and Geoinformation of China (Grant No. GZIT2016-A5-147) and the Research Institute of Henan Spatio-Temporal Big Data Industrial Technology (Grant No. 2017DJA001).

References

Batty M (1976) Entropy in spatial aggregation. Geogr Anal 8(1):1–21

Bjoke JT (1996) Framework for entropy-based map evaluation. Am Cartogr 23(2):78–95

Bogaert J, Farina A, Ceulemans R (2005) Entropy increase of fragmented habitats: A sign of human impact? Ecol Indic 5(3):207–212

Brillouin L (1956) Science and information theory. Academic Press, New York

Cushman SA (2015) Thermodynamics in landscape ecology: the importance of integrating measurement and modeling of landscape entropy. Landscape Ecol 30(1):7–10

Cushman SA (2016) Calculating the configurational entropy of a landscape mosaic. Landscape Ecol 31(3):481–489

Cushman SA (2018) Calculation of configurational entropy in complex landscapes. Entropy 20(2984):298. https://doi.org/10.3390/e20040298

Díaz-Varela E, Roces-Díaz JV, Álvarez álvarez P (2016) Detection of landscape heterogeneity at multiple scales: use of the quadratic entropy index. Landscape Urban Plan 153:149–159

Deza E, Deza MM (2016) Encyclopedia of distances. Springer, New York

Feldman DP, Crutchfield JP (2003) Structural information in two-dimensional patterns: entropy convergence and excess entropy. Phys Rev E 67(5):051104

Foody GM (1995) Cross-entropy for the evaluation of the accuracy of a fuzzy land cover classification with fuzzy ground data. ISPRS J Photogramm Remote Sens 50(5):2–12

Gao P, Zhang H, Li Z (2017) A hierarchy-based solution to calculate the configurational entropy of landscape gradients. Landscape Ecol 32(6):1–14

Gatrell AC (1977) Complexity and redundancy in binary maps. Geogr Anal 9(1):29–41

Goodchild MF (2003) The nature and value of geographic information. Foundations of geographic information science. Taylor and Francis Group, London

Leibovici DG (2009) Defining spatial entropy from multivariate distributions of co-occurrences. Springer, Berlin, pp 392–404

Leibovici DG, Birkin MH (2015) On geocomputational determinants of entropic variations for urban dynamics studies. Geogr Anal 47(3):193–218

Leibovici DG, Claramunt C, Le Guyader D, Brosset D (2014) Local and global spatio-temporal entropy indices based on distance-ratios and co-occurrences distributions. Int J Geogr Inf Sci 28(5):1061–1084

Li H, Reynolds JF (1993) A new contagion index to quantify spatial patterns of landscapes. Landscape Ecol 8(3):155–162

Li Z, Liu Q, Gao P (2016) Entropy-based cartographic communication models:evolution from special to general cartographic information theory. Acta Geod Et Cartogr Sinica 45(7):757–767

Li ZL, Huang PZ (2002) Quantitative measures for spatial information of maps. Int J Geogr Inf Sci 16(7):699–709. https://doi.org/10.1080/13658810210149416

Neumann J (1994) The topological information content of a map an attempt at a rehabilitation of information theory in cartography. Cartogr Int J Geogr Inf Geovisualization 31(1):26–34

Simpson EH (1949) Measurement of diversity. Nature 163:688

Snickars F, Weibull JW (1977) A minimum information principle: theory and practice. Reg Sci Urban Econ 7(1–2):137–168

Tobler W (1997) Introductory comments on information theory and cartography. Cartogr Perspect 26(27):5341–5357

Vajda S, Shannon CE, Weaver W (1949) The mathematical theory of communication. Bell Syst Techn J 27(4):379–423

Villani C (2008) Optimal transport: old and new, vol 338. Springer Science & Business Media, Berlin

Vranken I, Baudry J, Aubinet M, Visser M, Bogaert J (2015) A review on the use of entropy in landscape ecology: heterogeneity, unpredictability, scale dependence and their links with thermodynamics. Landscape Ecol 30(1):51–65

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.