Multiscale Entropy in the Spatial Context of Cities

Martin Barner

11987 words

1st September, 2015

first supervisor: Dr. Elsa Arcuate
second supervisor: Miss Clementine Cottineau

This dissertation is submitted in part requirement for the MSc Smart Cities and Urban Analytics in the Centre for Advanced Spatial Analysis, Bartlett Faculty of the Build Environment, UCL

I hereby declare that this dissertation is all my own original work and that all sources have been acknowledged. It is 11987 words in length.

Abstract

Considering a city as a complex system that emerges from individual decisions and activities, we expect for the city's entropy to increase with respect to the choices of locations available within it. We propose a spatial measure of entropy that takes into account the considered locations' environments as their characteristic on multiple scales. We test the hypothesis using land use data of west London from 1875 to 2005 and show that entropy increases, and does so stronger across a wider spectrum of scales than fully random spatial distributions and distributions of simple order.

Contents

List of Tables

List of Figures

Acknowledgements

Thank you Dr. Elsa Arcuate, for, while training future sushi chefs, teaching me about science and the universe in general with the side effect of magnificent advice for this work. As you know, your dedication went far beyond what anyone could reasonably expect.

Thank you Clémentine Cottineau, for your unique perspective, your precision, your unexpected, indispensable, selfless and genuine support. Without you, the entropy within my writing would be far too high.

Thank you Mr. Robin Morphet, for your encouragement, for understanding my thoughts far better than I did myself, and for seeing value in my naivete.

I would also like to express my gratitude towards Dr. Kiril Stanilov for providing the unique data set, without which this work obviously would not have been possible.

Chapter 1

Introduction

Urban science reached a new paradigm that recognises that cities emerge from individual behaviours that slowly carve the city over hundreds of years into the complex structures that they are [3][4][5][6][7]. These structures are not only efficient - they are sustainably adapting to revolutions and disruptions in society and technology [8]. The relationship between the activities of individuals in the city, and the slow, constant change in the city's shape are hard to relate, because the spatial and temporal scale of individual activities and their impact on the built environment are incommensurable.[9]
A solution to a conceptually similar problem can be found in an entirely different field - thermodynamics. The large scale properties and slow dynamics of a gas or liquid "emerges" from movements and interactions of individual particles.Boltzmann developed a quantity that is concerned with the probability of certain configurations on the scale of individual particles: Entropy [10]. The principles behind entropy led to new insights in a broad range of fields [11][12][13][14][15], and cities are no exception [16].
This work aims to develop an interpretation of entropy for cities that links insights from complexity sciences to the fundamental principles of entropy as consistently as possible. Based on these considerations, we reframe entropy for cities in a way that differs from existing approaches in two main aspects:

1. A common approach is to literally translate the geographical space as the equivalent to the thermodynamic "phase space". That means that increasing entropy relates to an increasingly even distribution of elements in space. In contrast, we no longer consider the spatial distribution of elements in the geographical space as the central element of entropy, but the distribution of locations across their characteristics. The view proposed here is rooted in considering urban form as an emergence of individual activities - its entropy should reflect on the entropy in people's activities, which can be represented by the entropy of the properties of the locations in which these activities can take place.

2. The proposed method differs from the reviewed approaches in the way space and spatial relationships are taken into consideration. From theories of cities and complexity, it is inferred that an inherent aspect of the characteristics of locations is given by their spatial relation to all other locations. Consequently, entropy is evaluated over the characteristics of locations determined by the characteristics of their spatial surroundings on multiple scales.

These core aspects of the offered interpretation of entropy lead to associate high entropy with a complex spatial order of elements - the “disorder” often associated with entropy is not a spatial “disorder”; it lies within the diversity of possible activities within a complex spatial order.

The proposed method is tested on data which looks at the dynamic change in the distribution of buildings with different functions in West London from 1875 to 2005. To tie in with the distinctive elements of the proposed method, the tested hypothesis is given by two highly interconnected points: First, that entropy increases over time, and does so stronger than for random spatial distributions or distributions of simple spatial order. Second, while random or ordered spatial distributions may display higher entropy of locations at specific scales, the real system is expected to develop relatively high entropy of locations over a wide range of scales. The results of the experiment show that in West London from 1875 to 1960, entropy increases across all scales, and does so stronger than corresponding random or ordered spatial distributions. Also, it is confirmed that while random and ordered distributions have high entropy only on specific scales, the spatial distribution in West London tended towards a higher entropy on a wider range of scales, resulting in a configuration that offers a larger variety of locations across all scales. After 1960, the total change in the built environment stagnates. From 1985 to 2005 entropy slightly decreases based on the total area covered by each category, independent from the spatial configuration.

This dissertation consists of 8 chapters. Chapter 1 is the present introduction.

Chapter 2 gives an introduction to the literature considering the two theoretical foundations this paper is built upon - cities as complex systems and entropy in statistical mechanics - and builds a connection between the two. It is explained that cities emerge from individual behaviours, and that entropy gives an indication of the probability of global properties of a system based on the constellations of its smaller individual parts. Then the two concepts are linked, proposing an interpretation of entropy for cities that see activities of individuals as the interacting parts, and the characteristics of the emerging built environment as the global properties.

Chapter 3 discusses the role of the spatial context for entropy, based on the interpretation of entropy in chapter 2. It is argued that locations in the city and their characteristics are the medium for activities and that therefore the entropy of locations determines the entropy of activities. Then the spatial surroundings of each location are identified as an intrinsic property of their characteristic. From this it is concluded that the spatial surroundings of locations should be included into a measure of entropy for cities.

Chapter 4 discusses how the reflections made in the chapters 2 and 3 link to complexity in spatial patterns. It first describes the ambiguous relationship between entropy and complexity and contradicting views in the literature, because high entropy is viewed as a sign of either very high complexity or its absence. This chapter then aims to clarify what entropy means for complexity, and suggests that for cities, the inclusion of the spatial context can bridge their paradoxical relationship.

Chapter 5 reviews existing approaches to measuring entropy in cities with focus on the spatial aspect. It is concluded that the inclusion of the spatial aspect is insufficient for our purposes, and a refined approach to measure entropy in cities can be valuable.

Chapter 6 lays out the details of the proposed measure of entropy in cities. From the origins of entropy in thermodynamics, a measure is developed that incorporates the surroundings of locations in a multidimensional and continuous way, and evaluates entropy on multiple scales.

Chapter 7 describes the details of the methodology in the application of the new approach. Entropy is measured in West London based on three categories of building functions for seven moments in time from 1875 to 2005. The analysis shows a steady increase in entropy over the full time period, and a significantly larger entropy than a fully random or fully ordered distribution across space. These results suggest that the structure developed towards a spatial configuration that maximises entropy on multiple scales, leading to a larger diversity of locations considering their spatial surroundings.

Chapter 8 gives concluding remarks. Conclusions are drawn from the empirical results and reflections on possible practical implications on planning are made.

Chapter 2

Entropy and Self-Organisation in Cities

2.1 Cities as Complex Systems

Viewing cities as complex systems brought a fundamental shift in our understanding of how cities function, grow, and change over time[17][3][4][18]. City systems and systems of cities display many characteristics of complex systems such as certain scaling laws [19], fractal [5] and multifractal [20][21] structures that have been studied extensively. Although "there is no concise definition of a complex system" [22], they usually are phenomena that emerge from the actions and interactions of smaller parts, while the complexity of the emergent phenomenon exceeds the sum of the individual parts it is composed of. One such system are "ants who are able to undertake complex tasks such as building bridges or farms" [22] without central control. In a chapter named "The Myth of the Ant Queen" Johnson states that a city "self-organizes out of millions of individual decisions, a global order built out of local interactions"[23, p.38]. These decisions and interactions shape the city quite literally: "They are patterns of human movement and decision-making that have been etched into the texture of city blocks, patterns that are then fed back to the [. . .] residents themselves, altering their subsequent decisions" [23, p.40]. Allen [24] sees the city and the individual acting and interacting agents as co-evolving: "the locations and behaviours of the actors are mutually inter-dependent" [24, p.1], and the dynamics driving changes in physical urban patterns are rooted in "the actions and responses of agents that are involved in the city" [24, p.1].

According to Portugali, the material components that make up the city and the people in it form a dually complex system in various ways: on the one hand, people using the city, "by means of their interaction - among themselves, with the citys material components and with the environment - transform the artifact city into the complex artificial system city" [6]. On the other hand, the shape and order of the city's material components then limits the behaviour of the agents [6].

It is worth to investigate the principles behind entropy, as it establishes a connection between the probability of the system to reach a certain global state and the possible configuration of the constellation of its smaller parts.

2.2 Entropy in Statistical Mechanics

Originating in thermodynamics, entropy is widely applied to various fields including information theory [11], finance [12], image processing [13], meteorology [14] and increasingly in various social sciences

[15], and plays a crucial role in urban modeling [16]. Entropy delivers a conceptual framework to describe the relation between individual microscopic elements and global properties of a system. As elaborated in the following, entropy can give an indication of the probability of some macroscopic properties based on stochastic principles concerning constellations of its microscopic elements.
Let us consider a system consisting of two coins that can each either be in the state “head” or in the state “tale”. For a macroscopic condition defined as “one coin shows head, one coin shows tail” there are two possible microscopic states (either coin one or coin two displaying head). A different macrostate, “all coins show head”, only allows a single microstate. The macrostate ‘one coin shows head, one coin shows tail” is therefore twice as likely. For real systems with a large number of elements, the difference in probability of different macrostates is far greater: it is for example almost infinitely unlikely that all molecules in a room have the exact same velocity.
In thermodynamics, the macrostate of a closed system containing identical particles is defined by macroscopic conditions such as the total energy, volume and number of particles in the system. A single unique constellation of all particles’ momentum and location is a microstate. In technical terms, a microstate can be described with a vector that contains the exact location and momentum of all particles in the system. This vector can be assigned a point in a high dimensional space, generally referred to as the phase-space [25].
Each macrostate covers a certain area of the phase-space, and a system in a specific macrostate can only take on the microstates contained in the area of the phase-space covered by the given macrostate. The probability of a macrostate depends on the number of equally probable microstates possible under the conditions of the macrostate [26], expressed as the volume in the phase-space covered by the macrostate [27, p. 322]. Over time, a system naturally tends towards the macrostate that allows the largest number of microstates, a macrostate of higher entropy.

2.3 Interpreting Entropy for the Urban Context

Here, an attempt is made at drawing a coherent interpretation of entropy in cities that includes the abstract relationship between microstates and macrostates and acknowledges the complex system approach to urban dynamics.
Batty (2010) recognises a “literature gap” and states that “the connections between entropy maximizing, substantive interpretations of entropy measures, and the longer-term dynamics of how equilibrium distributions are reached and maintained have not been well developed. [. . .] I sketch the need for new and different entropy measures that enable us to see how equilibrium spatial distributions can be generated as the outcomes of dynamic processes that converge to a steady state.”[28]. Furthermore, “an entirely new research agenda can be formulated with respect to the substantive meaning of entropy and related energy measures that tie these quantities back to their more fundamental thermodynamic origins.”
In order to reframe entropy for cities, the metaphoric meaning of microstates and macrostates must be carefully defined so that their abstract relationship is maintained in a meaningful way. Complexity theory suggests that the meaningful acting and interacting agents in the urban system are the people in it [6]. Consequently, a microstate is here interpreted as a description of all individuals’ exact features, behaviours, interactions and activities. By definition a macrostate is a set of conditions that contains a specific spectrum of microstates [26]. If microstates are combinations of individual behaviours, the macrostate of a city includes the conditions that define and limit which combinations of individual behaviour are possible in the city. The factors limiting the potential individual behaviour according to [6] are all material components of the city. These include buildings, streets, infrastructure,

institutions, and physical and non-physical networks. Certainly, there are also biological, physical, technological, societal and other constrains on how people can or can not use the city at a given moment in time.

The natural tendency of systems towards macrostates of larger entropy would then imply for this interpretation of entropy, that cities, over time, tend to increase the variety of possible combinations of activities and uses by its inhabitants. Linking back to Portugali's interpretation of dual complexity in cities [6], the material components of the city at the same time emerge from and limit the individuals' behaviour. This implies that if the material components allow more individual behaviour constellations, there are also more behaviour constellations they could alternatively have emerged from. Therefore, it is more likely to occur than a structure that allows very few of these constellations, and simultaneously has very few constellations of individual behaviour it could alternatively have emerged from.

As a trivial illustration, imagine a city in which all material components are highly specialised and exactly identical, in a way that strongly limits the variety of possible uses of the space. Clearly, the material components do not define the exact activities within them, but nonetheless set limits to what is possible. It is however highly unlikely that all individuals have the so similar behaviour, because there are only very few configuration that displays that global characteristic. In fact, the more diverse the constellation of individual behaviours within the given physical, technological, societal and human constrains is, the more likely it is as well. This is the case because there are considerably more possible internal configurations (permutations) for a distribution of behaviour with a larger diversity.

Following our hypothesis, we expect that cities with a low variety in the choice of locations available are unlikely to emerge, and if implemented by force through policy, planning and design, unlikely to last, leading to an increase of entropy in cities over time.

2.4 Far From Equilibrium

In a system with an inalterable macrostate, entropy measures the maximum entropy possible in the system with respect to that macrostate. If the system reached the maximum possible entropy, it is in equilibrium. Only if that equilibrium is reached, the entropy calculation yields valid results [29, p.40].

It has become apparent that "cities are far-from-equilibrium, as indeed are all human systems" [4, p.4]. Concerning entropy, a fundamental difference between a hypothetical closed particle system and a city is that unlike particles, the agents in a city are capable of actively altering the macrostate of the system. Before equilibrium is reached, the macroscopic conditions change, and with it the largest possible entropy that defines the equilibrium is no longer the same. Furthermore, the city is not a closed system. The city tends towards an ever changing equilibrium, leaving it in a constantly far from equilibrium state. "It has been shown that the equilibrium or steady state is the one with the maximum entropy, and [...] the system out of equilibrium is most likely to move in the direction of increasing entropy" [30]. At the same time, "a priori entropy is only defined for a system at equilibrium" [31]. To resolve this, we can assume this "transformation quasi-static if it takes place over a time much larger than that needed to re-establish equilibrium" [31]. Hence the methods applied here are only valid with the assumption that the fast dynamics of individual behaviour and activity changes are much faster than the slow dynamics of changes in the macroscopic constrains of the built environment. Then we can "define the entropy at each instant because the system is infinitesimally close to equilibrium" [31].

2.5 Limitations of a Metaphor

Inevitably, the translation of entropy for cities is a radical oversimplification. Jacobs brings up the unavoidable fact that "statistical people are a fiction [...]. Real people are unique, [...] and are not interchangeable in the least" [17, p.178]. Clearly, our notion of entropy in spatial structure is not concerned with how efficient or practical or suitable different spatial configurations are for the various aspects of human life. Additionally, as only the macrostates are observed, the actual nature of the interactions themselves is not taken into account. For example, there is reasonable evidence that people in groups tend to assimilate their behaviour, and do so stronger with larger spatial proximity [32]. However, from a purely theoretical point of view these factors are incorporated in the macrostate's conditions. They are, strictly speaking, part of the macroscopic conditions limiting potential behaviour constellations. Once taken into account as macrostatic conditions, the validity of the statistical distribution will be restored. An omnipresent incompleteness is critical to the observation of entropy. Observing different macrostatic conditions leads to different results, considering insufficient macrostatic conditions can result in decreasing or any highly distorted perception of entropy [33]. It leaves us with the conclusion that a careful definition of macrostates is crucial to the interpretability and validity of the measured entropy.

When applying entropy to cities, it remains a metaphor that must not be over-interpreted. Ultimately, entropy remains a descriptive quantity of spatial configurations.

Chapter 3

Entropy in a Spatial Context

3.1 Places and Activities

In chapter 2 it has been argued that in order to understand how the city emerges from individual decisions and interactions, it is necessary to observe the range of possible combinations of its uses. As the physical places in the city are the medium in which these activities take place [6], it is here assumed that places of different characteristics also enable different activities, while similar places tend to host similar activities. The characteristics analysed here are the surrounding building's functions at different scales. What kind of places exist is in return driven by activities taking place, shaping the characteristics of places over long periods of time.

3.2 The Spatial Context

Given these assumptions, in order to assess the entropy with respect to potential uses, the diversity of places plays a crucial role. Within the reviewed literature, it appears that what is so far almost entirely overlooked in measurements of entropy and diversity in cities (see chapter 5) are considerations on how the spatial relationship of places and locations affect their own characteristic. What is argued here can be outlined with a trivial example: The function and possible uses of the same distinct building is fundamentally different if it is located on a street with small family homes or on a high street, in an industrial zone or a quiet residential area, in a city's center or in the desert, or in any other imaginable environment of any scale. Two in themselves identical places can be different in the way that the spatial relationship to all other places fundamentally alters the nature of activities and interactions of individuals possible in a place of specific characteristics.

These spatial dependencies are a common thread in spatial sciences: First and foremost to mention is Tobler's first law of geography, stating that "everything is related to everything else, but near things are more related than distant things" [34]. Specifically for the case of cities, this interconnectedness and interdependencies on different scales was expressed by Christopher Alexander as part of his argument against tree like urban structures of modernist planning. Alexander understood cities as connected systems of overlapping sets and subsets. "It is not true that these systems exist in isolation, cut off from the other systems in the city. The different systems overlap one another, and they overlap many other systems besides" [35]. With reference to the structure of social relationship structures, he concludes that disjoint neighbourhoods can not be an adequate description for a city, because it ignores how the smaller elements of the system interact individually across neighbourhoods.

Alexander also relates to Jane Jacobs who follows a similar perspective. In her eyes, everything in the city is interconnected through random and purposely made interactions, forming a network of organized complexity from streets, over neighbourhoods and districts to cities as a whole [17].
A more practical and quantitative hint is given by the success of spatial interaction models [36][37][38]. In order to accurately predict for example business incomes, these models must take into account the precise spatial relationships of competing businesses, population densities and incomes of all locations in the considered system [39]. This means that the spatial relationship and characteristics of all other locations in the system determine whether a specific use (in this case a commercial center of a given size) is possible in a specific location. Further literature on the effect of the spatial context on local activities is given by agglomeration economics [40] and neighbourhood effects [41].

3.3 Entropy in a Spatial Context

The essence of these examples, relevant to the entropy of locations reflecting on the entropy of possible uses, is rather intuitive: Imagine a part of any size from a city that is well known to you, and imagine it ripped out of its surroundings. Place for example Soho in London in an entirely different environment, and the way it is and can be used will be entirely different. An office within a central business district enables different functionality than the exact same office surrounded by family homes. A measure of entropy needs to sufficiently incorporate the complex relationship between locations, and the way the characteristics and types of potential uses of each location are altered based on the location's surroundings.

Chapter 4

The Spatial Context Linking Entropy and Complexity

The following chapter reconsiders the relationship between entropy and complexity, and recognises entropy in a spatial context as a possible bridge between the two. It first mentions some contradictory approaches to entropy as a measure of complexity. Then, a more consistent relationship between randomness, predictability and order is given. State interdependencies are recognised as a crucial factor, and finally it is explained how the spatial context links entropy and complexity by incorporating state interdependencies of spatial proximity.

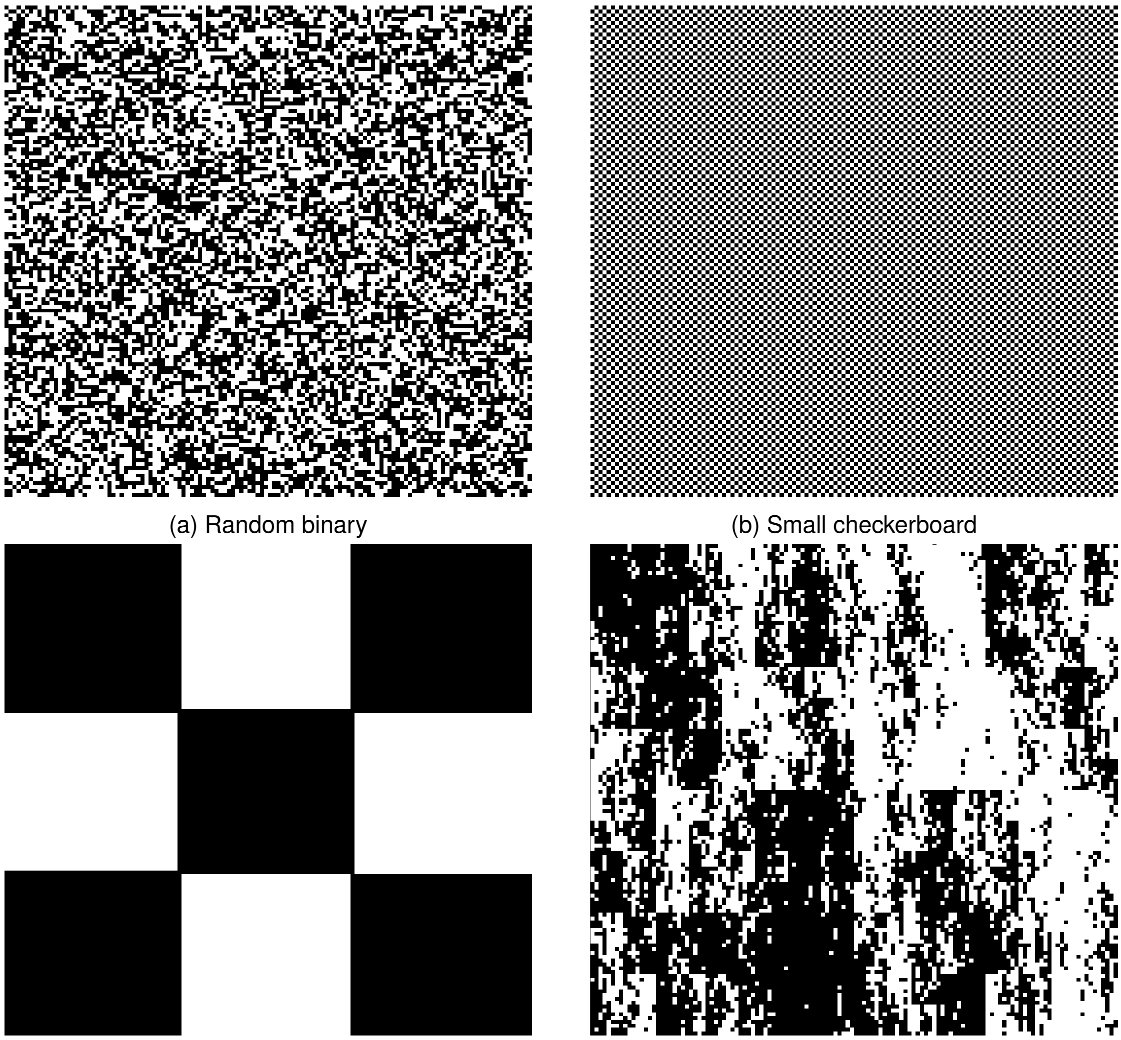
4.1 The Entropy Complexity Paradox

Shalizi [42] gives an insight into the incoherent understanding of the relationship between entropy and complexity in the literature. From the perspective of complex systems as the result of self-organisation, attempts have been made to associate complexity with low thermodynamic entropy [43][44], regarding the occurring order as higher complexity than the original randomness. This is paradoxical to a common understanding of unpredictability as a feature of complexity, that suggests that complexity displays high entropy. Garland et al. [45] regard fully unpredictable signals such as white noise as "fully complex" in contrast to fully ordered signals such as strictly periodical signals. This view is also adopted by Batty et al. [46] for the context of cities. Costa et al. conclude that "if one supposes that greater entropy is characteristic of greater complexity, such results are profoundly misleading."[2]. Arguably, processes are not complex simply because they are fully random.

4.2 Randomness, Predictability and Order: What is Complex?

In fact, these positions are less contradictory than it seems. By intuition, complexity requires "meaningful structural richness" [47]. This is neither given in entirely random signals, nor in signals of simple order. The confusion with respect to entropy is related to the conception of order, complexity, predictability and randomness as each on one side of two opposing extremes: Order and predictability on the side of low complexity, and randomness and unpredictability on the side of high complexity. This is not an accurate description. Costa et al. see the need to make clear that "Neither completely predictable (e.g., periodic) signals, which have minimum entropy, nor completely unpredictable (e.g., uncorrelated random) signals, which have maximum entropy, are truly complex, since they can be described very compactly.” [2]. High entropy is a necessary, but not sufficient condition for complexity. For cities, Batty et al. [46] mention a “dilemma” between entropy and complexity:

> “When something is entirely ordered, hence completely predictable, it is no longer complex, and we would expect the measure to be low relative to a situation of extreme unpredictability. However, we can also argue that something that is entirely ordered has great complexity in that to simply hold this order together in far-from-equilibrium systems, considerable intricacy in structure and relations is required. This is in contrast to a situation where events are completely unpredictable, hence in one sense quite disordered and possibly of low complexity in that random forces dominate.”



(a) Random binary

(b) Small checkerboard

(c) Large checkerboard

(d) Binary multiplicative cascade

Figure 4.1: Binary grid spatial patterns

A more practical description of a range of different levels of complexity can be proposed as following: In the extreme of low complexity we find entirely random processes without any kind of order because they lack any “meaningful structural richness” [47] as in figure 4.1a. Then we find systems with simple order (figure 4.1b and 4.1c), still not complex at all because they are simple and highly predictable. Next on the scale there are systems that can be described as complex to some extent: they are ordered in a multitude of overlaying ways, not merely the results of entirely random processes but hard to predict because of the multiplicity of their order - an example might be a multiplicative cascade as in figure 4.1d. Finally, at the extreme end of high complexity there are systems that display an intricate richness of overlaying, dynamically changing structural order so versatile and multifaceted it becomes impossible for us to understand or disentangle. These highly complex systems are not the result of fully random processes, but are entirely unpredictable and can appear random and entirely disordered to us, because their order is too complex to be visible. This leads to an ambiguous relationship between entropy and complexity, as high thermodynamical entropy can be found in both minimally and maximally complex structures: “A completely ordered system with a small entropy or a completely disordered system with maximum entropy is the least complex.” [48]

4.3 Entropy and Complexity reunited

All of the above relationships between entropy and complexity apply to measures of entropy that assume no interactions and interdependences between the observation’s states. A solution to the entropy ambiguity problem has been proposed by Zhang [49] with the phase space approach to entropy. In principle, entropy is calculated as the integral of the entropy on different scales of coarse-graining. The multiscale entropy is able to bridge the gap between entropy and complexity because it evaluates “long-range correlations in space and time” [48]. Figure 4.2 illustrates this on two dimensional binary patterns: The total occurrences of different configurations are listed for each scale. Here, Configurations are only distinguished by the total number of black cells, and entropy H is calculated according to $H = \sum_{i=1}^{n} p_i \log(p_i)$ for all n probabilities of the different configurations. On the scale of a single grid cell, there are only two options, black or gray. The scale 2x2 potentially has four options, with one to four black cells, and so forth. Depending on the spatial configuration of the eight black cells in each pattern, entropy differs on different scales.

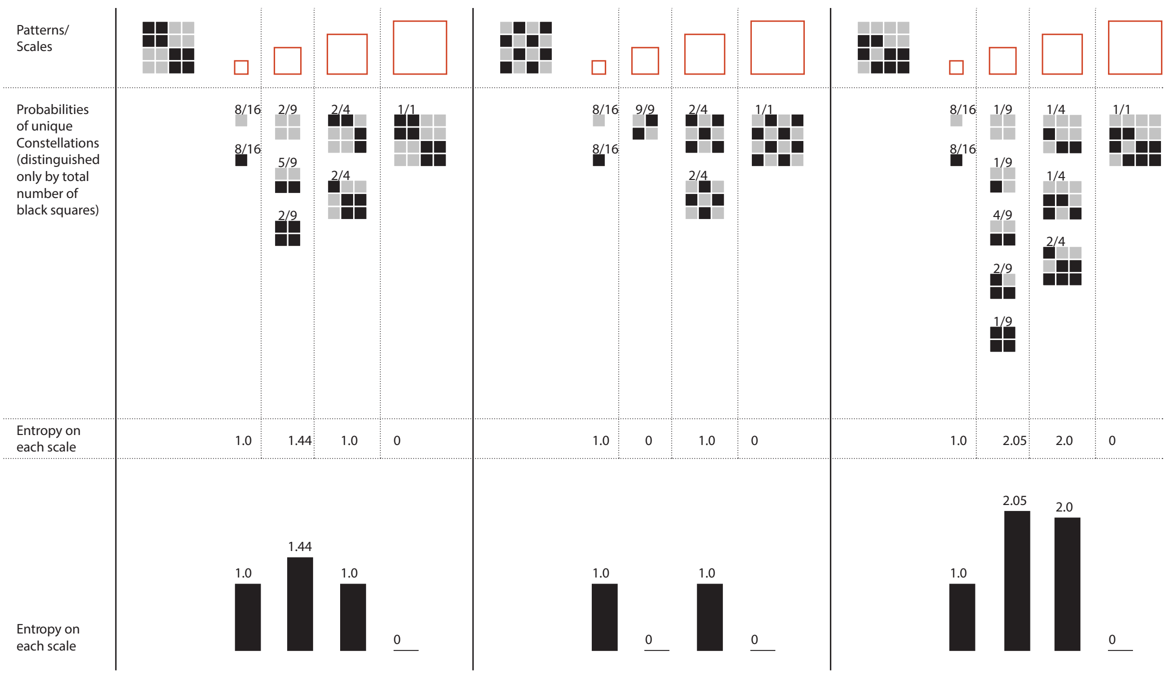
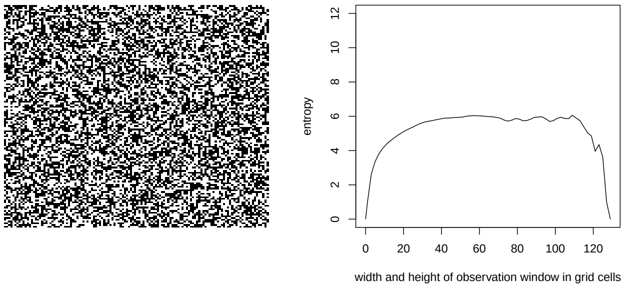


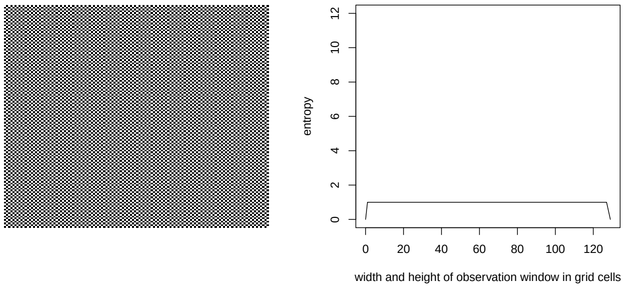
Figure 4.2: Manually counting microstate probabilities in a 2D binary grid

The coarse graining procedure literally calculates the state of a single moment in time (or a point in space) from the immediate values at different temporal (or spatial) proximities before calculating the entropy. Before the entropy of observations at specific spatiotemporal moments is calculated, the immediately observed values of observations at other spatio-temporal moments are included into the representation of each observations state. Long-range correlations can be phrased as interdependence of states: When considering observations at different proximity as inherently defining each observations state, entropy and complexity becomes consistent. Figure 4.3 examines the multiscale entropy of binary grid patterns according to the process in figure 4.2.[1]

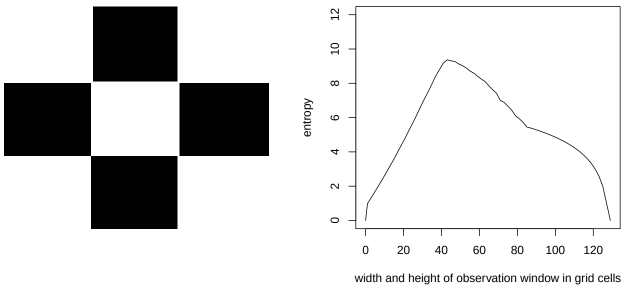
[1] This method treats almost identical configurations and entirely different configurations as equally different and therefore assigns overly high entropy to random patterns on large scales



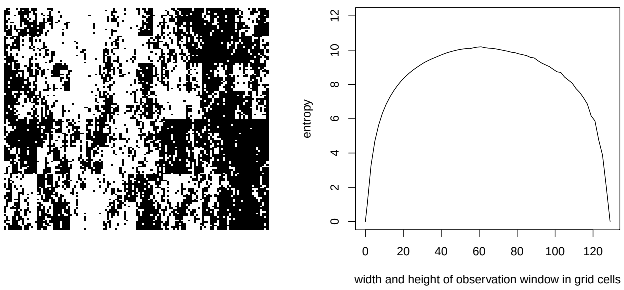
(a) Random binary



(b) Small checkerboard



(c) Large checkerboard



(d) Binary multiplicative cascade

Figure 4.3: Multiscale entropy in binary grid spatial patterns

4.4 The Spatial Context between Entropy and Complexity

This state dependence is exactly analogous to entropy in a spatial context described in chapter 3. Entropy in a spatial context uses the immediate state of surrounding observations as part of each location's state, which is exactly what methods that coherently link entropy to complexity do. A multiscale coarse-graining of building's functionality in cities based on [50] and [49] would mean not only to consider the entropy of the smallest observed elements, but of environments of different scales: For example not only considering the entropy of building as themselves, but also of groups of buildings in blocks, streets, neighbourhoods and districts.

In summarising chapter 2, it becomes apparent how here the concepts of complexity, emergence, entropy are linked coherently:

First, the interpretation of entropy in cities is focused on the notion of cities as complex systems emerging from individual behaviour; spatial structures with higher entropy of characteristics can host a higher entropy of activities and, on the assumption that these individual activities are the source of the emergence of the city, more likely to emerge.

Second, the focus on the entropy of locations in terms of the entropy of activities they can host leads to the notion of entropy in a spatial context: When measuring urban entropy based on an interpretation rooted in understanding cities as emerging complex phenomena, each location's state needs to include its surroundings. Concurrently, research in measures of complexity has revealed that this exact multiscale state dependency itself is in fact the link between complexity and entropy.

Chapter 5

Existing Approaches to Entropy in Cities

5.1 Entropy as a Metaphor

The term “entropy” in the context of cities is used to describe a broad range of different ideas. Ever since the introduction of entropy and maximum entropy methods by Wilson [16] to the field of urban sciences, they play a crucial role in urban modelling, especially in spatial interaction models. Although clearly the general concepts are linked, there are basic differences in their application.
A collection of research including [51] explores the urban ecosystem as an metabolism that exchanges energy, recourses and entropy with its environment. It is a perspective that tries to understand the real thermodynamic processes in cities by observing resources that enter and exit the urban metabolism.
Here, we are not concerned with real physical thermodynamic quantities. Much more, entropy and its general statistical principles serve as a metaphor. A range of recent thoughts and applications have been developed that are similar to what is of interest here. A selection of these approaches is discussed in the following with focus on the spatial aspect.

5.2 Aspatial Entropy and the Checkerboard Problem

Depending on the interpretation of microstates, macrostates, observations and the dimensions of the phase space, some approaches to entropy in cities are almost entirely independent from the spatial configuration.
Gudmundsson et al. [1] measure the entropy of street networks. Entropy is measured separately considering two characteristics of the street segments, their length and their orientation. The macrostates are defined by the distribution of street segments across the possible value of each characteristic. In a sense, it is a measure of how diverse the many locations (in this case streets) in a city are with respect to a specific characteristic globally, almost entirely independent from their spatial distribution.
Ceccato et al. [52] point out that “Shannon entropy is aspatial”. That is, when applied to characteristics of different locations, a priori the spatial relationship is not considered.
Assuming the streets of a city were uniformly distributed across all possible orientations, this would result in the largest possible entropy. It is not differentiated between a situation where the streets of different orientation are distributed randomly across space, or if they are sorted spatially. Figure 5.1 shows six hypothetical street networks with the same number of vertical, horizontal and diagonal

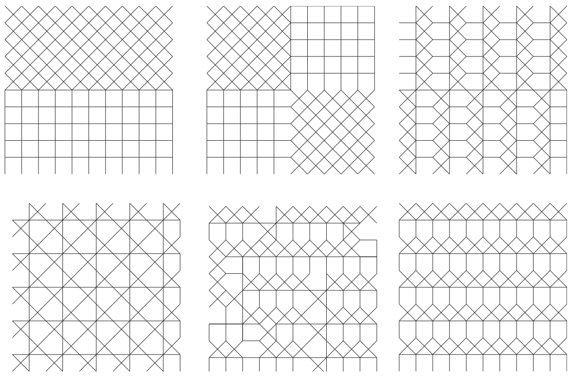


Figure 5.1: 6 Street networks with identical entropy according to [1]

street segments that because of the same distribution across the possible orientations of streets display the same entropy if evaluated according to the method applied in [1]. It raises the question of different possible interpretations which of the networks in figure 5.1 should be assigned the highest or lowest degree of order, and in how far the spatial configuration of elements can and needs to be considered. In the following this is identified as a central problem that appears consistently throughout measures of spatial segregation.

Even though this paper is not concerned with segregation directly, the mathematical concepts are similar. There has been a thorough discussion of the problem of sufficiently capturing spatial relations. A group of segregation indices reviewed in [53] mainly rely on the information included in the cumulative proportion of one of two categories in given zones (the segregation curve). These measures include the Gini Index [54], The dissimilarity index D [55], eta-squared [56] and others. The spatial aspect of these indices is limited to the distribution across arbitrary zones, while disregarding the constellations within each zone as well as the spatial relationship between the zones. Duncan [53] names the dissimilarity index D as sufficiently including the information given by any other of the proposed indices. It is defined as

$$D = \frac{1}{2} \sum \left| \frac{N_{1_i}}{N_1} - \frac{N_{2_i}}{N_2} \right| \tag{5.1}$$

where N_{1_i} and N_{2_i} represent the population of two groups in zone i. The equation illustrates that the information contained can not express any significant spatial relationships, and Duncan points out that it is incapable of capturing any actual "spatial pattern" crucial to segregation, as segregation can occur in different spatial configurations such as "a "Black Belt," [...] scattered "islands" or "pockets""[53]. Taeuber 1969 even suggests that "any single index of such a complex phenomenon is an arbitrary over-simplification of reality" [57]

White [58] proposes an index that "explicitly incorporates the spatial relationships among the geographic parcels" [58], and Morgan [59] proposes a distance based index as well. In these measures it becomes clear that the geographical distance between the individual elements as well as between the zones they are aggregated in is, at least for segregation, the central feature of the spatial relationship.

Logan [60] compares the dissimilarity Index D [55], the p* measure of interaction probabilities [61], and eta squared [56] and concludes that each of them relates to a different spatial aspect of segregation. In a similar way Massey and Danton [62] review and group twenty segregation indices with respect to the specific spatial aspect they measure into five categories, namely evenness, exposure, centralisation, concentration and clustering [62].

Wong [63] discusses similar patterns to those in figure 4.2 and 4.3, and identifies the problem of ignorance towards the large variety of possible spatial configuration as a problem of scale sensitivity: The segregation indices are sensitive to the scale of aggregation, where on different scales different aspects of the spatial configuration are measured. Concerning an accurate measure of segregation, Wong sees this as a problem - closely related to the modifiable areal unit problem [63] - to us, the ability to observe overlaying patterns on different scales presents itself as an opportunity and is adopted as a central part of the method proposed in chapter 6.

Theil [64] develops a measure of segregation based on information entropy [11], and develops a way to coherently aggregate entropy on different scales. It acknowledges different maximum entropies in zones of different compositions. Theil gives an example of segregation in schools: If there is high segregation between districts and therefore within a single district there is a total unevenness of different groups, the lowest segregation possible between the schools in that district is not an equal fraction of people from different groups in each school, but a fraction identical to the uneven composition of the district.

The more advanced measures of segregation do capture the spatial context to some degree that may be sufficient to evaluate segregation. They are without exception specialised to measure in different ways the deviation from an "ideal" state, which is the entirely even distribution of elements across the geographical space. When this is translated back to entropy, it creates fundamental implications on how entropy can be interpreted: Zones or locations become the discrete events the observations in space can represent. This seems necessary in order to properly incorporate space into the equation. In the following section, the implication of this interpretation of entropy are discussed in detail.

5.3 Geographic Space and Phase Space

A common practice in urban entropy analysis to overcome the inital non-spatiality of entropy is the literal interpretation of geographical space as the central element of the phase space. This involves splitting the geographic area of a city into zones and then evaluating how equally certain elements are distributed across these zones. In these cases, the largest observable entropy and disorder appears when all observed elements are spread equally across space.

Based on Theils's entropy-decomposition, Batty proposes an approach to spatial entropy and "how space can be introduced into the formulation" [28]. The spatial aspect is included as the phase-space; the different events that can occur in form of observations are different locations, while the elements inhabiting that space themselves are the observations. Notable considerations are made on "how entropy can be decomposed and nested to capture spatial variation at different scales" to account for "groups of objects that might pertain to some higher level of organization" [28].

Batty's [65] approach to spatial entropy on the example of population densities assumes that "systems in which the population is spread out evenly have maximum entropy and maximum disorder and constitute the situation that would emerge when the system has no constraints [. . .] and all persons can live where they want. Given enough time, people would spread out evenly in the absence of any reason for locating in any particular place." [28] This assumed even distribution of people across space holds true only for the way events, observations and probabilities are defined in this context.

In terms of entropy, in this particular interpretation each person is an observed element, and the zone they are in describes their state. "What is significant about this interpretation is its direct connections to thermodynamic entropy" [28], and in fact this is the most literal translation from thermodynamic entropy possible that still includes space: The main parameter despite the momentum that describes a single particle's state is its location, so here, the parameter that describes a single person's state is the index of the zone a person is in. For this approach, The geographical location of individuals defines literally the system's position in the phase-space. Batty et al. [46] extend these ideas and apply them to London's population densities and London's street network.

A careful interpretation of what this entropy means is required. Batty's approach includes a far more thoughtful consideration than depicted here, including interactions and global constraints [28]. Nonetheless, the primary definition of a particles state is given by its sole geographical position, and the integral over the phase space in the thermodynamic entropy is replaced by the integral over the geographical space. An entirely different interpretation might be equally valid:
instead of associating maximum entropy with an equal probability of people to be in any of the considered locations which results in locations with identical population density, maximum entropy can be associated with an equal probability of the different choices of individuals, resulting in an equal probability of the locations hosting these activities. Even if there is no information about the properties of different zones, they are at least characterised by the density of the elements in question.
In general terms of entropy, an additional event (accessible discrete state for individual observations) with associated observations increases entropy. The question here is what differentiates two events. In [28][46] and [65] an additional zone of any characteristic increases entropy, as long as there are observations (people) in it, even if the zone itself is identical to all other zones - it is considered an additional event, because the geographic location differentiates them. Grounded in our interpretation of entropy in cities we intuitively expect a different consequence: If individuals shape the city based on a high entropy of individual activities, entropy does not refer to if people could be in zones with different labels, but if there are zones with different characteristics.

5.4 Complex Spatial Relations

Even if, for now, one temporarily accepts the given specific definition of microstates, macrostates, observations and events in the proposed spatial entropy, it can still be argued that the inclusion of the spatial aspect insufficiently incorporates the complex spatial relationship of interacting elements in space. This becomes obvious when looking at patterns of the highest possible entropy which is, for the spatial entropy measures reviewed so far, a pattern of perfect uniform distribution of probability across space as in fig. 4.3a
If the spatial relation between the elements were considered relevant to the state of the element, spatially shuffling a pattern should have an effect on the entropy. Imagine cutting and switching two random patches of the same size repeatedly in the uniform pattern in fig. 4.3a - the spatial distribution would still be uniformly random, displaying maximum entropy in the geographical phase-space. From a certain perspective, this is a highly useful attribute; it makes entropy decomposable and consistent throughout scales. However, if the characteristics of these places and their surroundings matter, although for a uniform spatial distribution in each zone of a given size the diversity of elements is maximised, these elements at the same time have almost identical spatial environments - each element in a zone is surrounded by the same 'perfect mix" of elements. This actually makes the elements in each zone more similar than they could be. If these spatial relationships were taken into account, a complex pattern of differently sized clusters can display a higher entropy than a uniform spatial distribution, because although this leads to uneven probabilities of elements of different immediate characteristics in zones, the elements can at the same time be more different with respect to their spatial environment.

So far, two mutually exclusive principles have been laid out: Either, entropy can be calculated over the probabilities of locations to have certain characteristics, which mostly ignores the spatial distribution within and between the considered locations. Or alternatively, entropy can be calculated over the probabilities of elements to be in certain locations (zones), which excludes taking into account individual characteristics of these locations. Both approaches fundamentally differ in the definition of microstates, macrostates, events, observations and the spatial patterns associated with high and low entropy.

The majority of approaches undertaken so far in one way or another primarily consider spatial entropy as a question of how disordered the physical elements in space are, disregarding that high spatial disorder may also incorporate very low disorder in the choices of locations available in the cities. If, for example, buildings of different functions were distributed evenly across space, the only available larger scale environment for offices, homes, shops, parks and everything else in the city would be the same mixed-use environment; in that respect, the amount of choices of places given would be much lower than in a city with a complex patterns of locations in a multitude of different environments of different scales; interweaved patches of very different size of parks, residential areas with different types of homes, sometimes mixed, sometimes segregated.

Chapter 6

Method

From here on, this paper is concerned with developing a measure of entropy for cities that incorporates intrinsic spatial relations in a meaningful way. In parallel to the general interpretation of entropy for urban systems given in chapter 2, the mathematical foundations of entropy are reframed for an urban system. Starting from the original equations for entropy in thermodynamics and statistical mechanics over more recent entropy estimators, a practical implementation is developed as consistent as possible with the ideas from chapters 2 and 3.

6.1 Counting a City's Microstates

In thermodynamics, the Boltzman formula defines the entropy S as the logarithm of the "number of ways (W) [...] by which a given macroscopic state can be realized"[66], multiplied by the Boltzman constant k_B [66]:

$$S = k_B \ln W \tag{6.1}$$

If all microstates are equally probable, this can be written as [67, p.44]:

$$S = k_B \ln \Omega \tag{6.2}$$

"where Ω is the total number of microstates available to the system"[67, p.44]. For microstates of unequal probability, the Gibbs entropy can be seen as the generalisation of the Boltzman formula to the case of microstates with unequal probabilities p_i [68]:

$$S = -k_B \sum p_i \ln p_i \tag{6.3}$$

All microstates can be allocated a location in the phase space. If the phase space is discrete, it is mathematically possible to count the positive integer number of accessible microstates. For a continuous phase space, the Gibbs entropy H_G can be calculated through the Liouville function W_N, the "probability density in the full phase space of the system"[69, p.392 f.]

$$H_G = \int W_N \log W_N dr \tag{6.4}$$

where dr are the "phase-volume elements in the full phase space" [69, p.393]. The relationship between the Boltzmann entropy and the Gibbs entropy and their meaning is far more complex than depicted here [69], and their precise interpretations still controversially debated [70]. Details are further discussed in section 6.2.

Reframing this for the context of cities, the exact behaviour and activities of all individual agents in an urban system correspond to the state of individual particles in a thermodynamic system. For two locations whose observed characteristics are similar to a certain degree, we assume that the activities taking place in these locations are equally similar.

6.2 Microstates with Spatial Dependencies

In chapter 3 we identify the need to include the surroundings of a location into the description of its characteristics instead of considering all individual locations independent from each other. In fact the abstraction of this issue is an essential aspect that occurs between the Boltzmann entropy and the "more commonly used (and misused) entropy of Gibbs".[71]. The necessity to include the spatial context into a measure of entropy appears again from a different perspective when we examine the details of the Boltzmann and the Gibbs entropy.

"The most essential restriction of Boltzman's theory was the assumption of weak interactions between the particles. This assumption could be removed by [...] Josia Willard Gibbs [...]. He introduced a high-dimensional phase space, the so-called Γ-space." [72, p.61]. This new phase-space allowed the "generalization of the entropy to interacting systems" [72, p.62]. There are now two very different definitions of the phase space which can lead to confusion. There is a phase space that defines the microstate of individual particles, which is refered to as the μ-space. Additionally, there is the phase space that defines the microstate of a system as a whole, including all individual states of all individual particles together, which is refered to as the γ-space [73]. In the Gibbs entropy in formula (6.3), p_i refers to the probability of the system to be in a microstate defined in the γ-space, as a unique constellation of all individual states of all particles in the system together. It is possible to instead use the sum over probabilities of microstates in the μ-space, but with a critical implication: Boltzmann's entropy calculation over the μ-space "failed to take into account interactions between particles" [74]. *Dependencies among the individual states are ignored.* This means that the system is treated as if the state of individual particles in no way affect the state of other particles in the system.
Transferring this to agents in the urban system, each agent has a position in the μ-space defining its current activities. The city as a whole, the exact combinations of all individual agents' position in the μ-space is then given by a unique position in the γ-space. Therefore, for a system of N agents that are each in a D dimensional μ-space, there is a γ-space with ND dimensions.
Part of our hypothesis here is that the individual states of activities or locations are not independent, but in fact *spatially dependent*. The differentiation between the two phase-space definitions is at the core of our implementation of the spatial context into the original notion of entropy as discussed in section 3. Rephrasing this for the entropy of the characteristics of locations in the urban system, calculating their entropy based on the probability distribution of the isolated characteristics of each location implies the assumption that any interactions between these elements are neglectable.
All existing approaches to entropy in cities reviewed in chapter 5 calculate entropy over the probability distribution in the μ-space instead of ensembles of all individual states as microstates (γ-space), inheriting the same assumption that the interactions and flows between them are neglectable. The reflections made so far call for a refined practical method to assess entropy.

6.3 A Practical Implementation

A practical approach is developed to include these considerations into a measure of entropy in a spatial, urban context. To begin with, the observations considered are locations in space. For an equal consideration of the geographical space of the system, a grid with measurement points in equal distances is laid over the observed area. Each of these points represents an observation in entropy terms

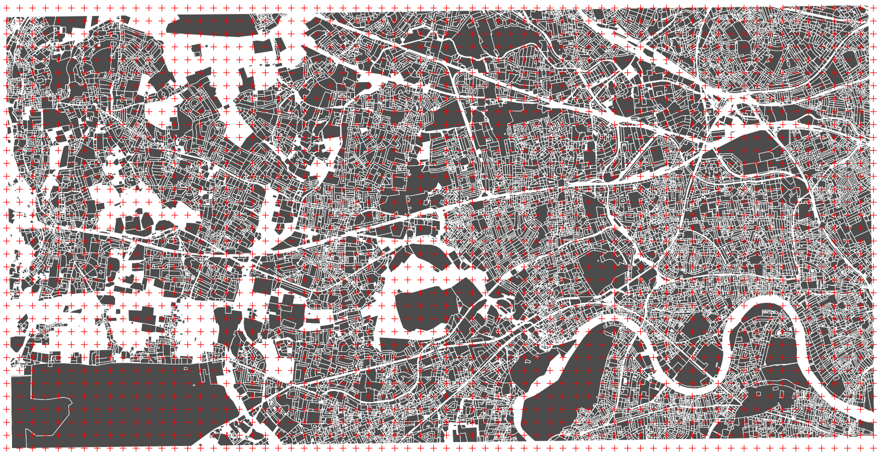


Figure 6.1: Grid of observation points with 300 m intervals

This leads to a larger number of observations for a higher resolution of the grid. Unlike in a discrete phase-space, in the continuous phase space, a low number of observations gives a worse picture of the density distribution, but not necessarily a lower value for the entropy. Nonetheless, a low grid resolution can give an inaccurate estimate of the probability density, leading to an underestimation of the entropy.

We required the representation of the characteristics of these locations to incorporate their spatial surroundings. The characteristics considered in this analysis are the functional categories of the buildings. Ensuing from the idea that material and immaterial flows and interactions in space "constitute activity at locations" [3, p.79], and that spatially, the flows between locations to some degree depend on geographical distance [16], we use the distance between observed locations and elements as the central variable defining their spatial relation. Certainly one could explore further the effects of flows and networks on spatial entropy. Here, for a given radius r from the geographical location of each observation, the total area covered by buildings of each functional category is taken as one dimension in the phase space for each category. The individual state of an observation at a given radius r is then defined by a value for each category,

$$v_c = \sum_{n=1}^{N} A(n) \tag{6.5}$$

where N is the total number of buildings of category c whose centroid (arithmetic mean) lies within the given radius r of the given observation point, and $A(n)$ the area of each building n, gives the value v_c for a dimension of the category c in the phase space. Clearly, as we assume distance as the main factor of the spatial relationships, elements within different accessibility based on the distance

from the observation form different aspects of a location' characteristic. Hence, we must consider the constellation of elements for a range of different radii. This means that the environment of a location at different scales - in a crude analogy one could say the characteristics of the nearby buildings, the street, the neighbourhood, the district and the city as a whole - are all considered as different aspects of a location's characteristics.
Entropy is then calculated for different distances r separately. This gives the entropy of the system depending on the scale at which the surroundings are taken into account as part of the observations' characteristics as $H(r)$. Clearly, this is far from a consistent or accurate transition into the γ-space, which would require an inclusion of all observations into a single microstate of the system. Hypothetically, it would give a slightly closer representation of the distribution in the γ space to define a unique dimension in the phase space for each category for each considered distance, resulting in $N_r N_c$ dimensions when N_r radii and N_c functional categories are considered. Such a large number of dimensions is not feasible for a practical implementation, because the entropy estimation applied requires at least $2^{N_c N_d}$ observations [75], which would require pushing the grid resolution beyond the resolution of the data and lead to computational difficulties.

A similar issue was recognised by Zhang [49] in his phase space approach to measuring complexity that was further developed by Costa et al. [2][76][77][50] who utilise these ideas for time series. This "has become a prevailing method to quantify the complexity of signals. It has been used successfully" [78] in numerous fields of research.
In Costa et al.'s approach, a time series is coarse grained at different scales and the values are averaged in partitions of different sizes as show in figure 6.2. Then, for each scale of aggregation, the sample entropy is calculated, with the goal to be able to observe order on various scales. [2].

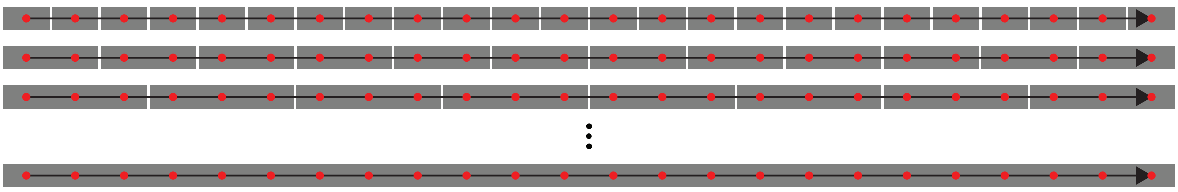


Figure 6.2: Multiscale entropy analysis coarse-graining procedure in [2]

Here, we follow a similar strategy with a few adjustments. First, the coarse graining steps are not taken by the number of included nearest neighbours but with respect to the distance from the observation point, as unlike in the described time series approach the nearest neighbouring elements can have very different distances and therefore very different meaning. Secondly, while in the time series analysis non-overlapping windows are considered, here the coarse graining is performed around all observations. An important aspect that is inherited is the handling of the edges of the system. While at the lowest scale the number of observations is equal to the original sample size, the number of considered observations is strictly decreasing until the coarse graining reaches the size of the system with a single observation and minimum entropy [49]. The increasing scale of the observations' characteristics must not be confused with an increase of the size of the whole system that is expected to have larger entropy. The proposed measure is consistent with higher entropy for larger systems because the largest possible observation scale at which entropy is minimum is larger.

For the analysis, an efficient computational implementation is needed to estimate the entropy for this continuous multidimensional representation of observations that we base on the Gibbs Entropy given in equation (6.4). For this case there is a variety of entropy estimators such as Voronoi diagram based estimators [79]. Here, we follow the computationally efficient kd-partitioning approach proposed by Stowell et al. [75].

While Stowell et al. [75] refer to differential information entropy, under the assumption that the phase space defined is a sufficient approximation of the γ space, we can equally apply the computational methods to the Gibbs equation (6.4), where W_N is the probability density function in the γ-space [69, p.393], so that the Gibbs entropy can be written as

$$H_G = -\int_x f(x) \log f(x) dx \tag{6.6}$$

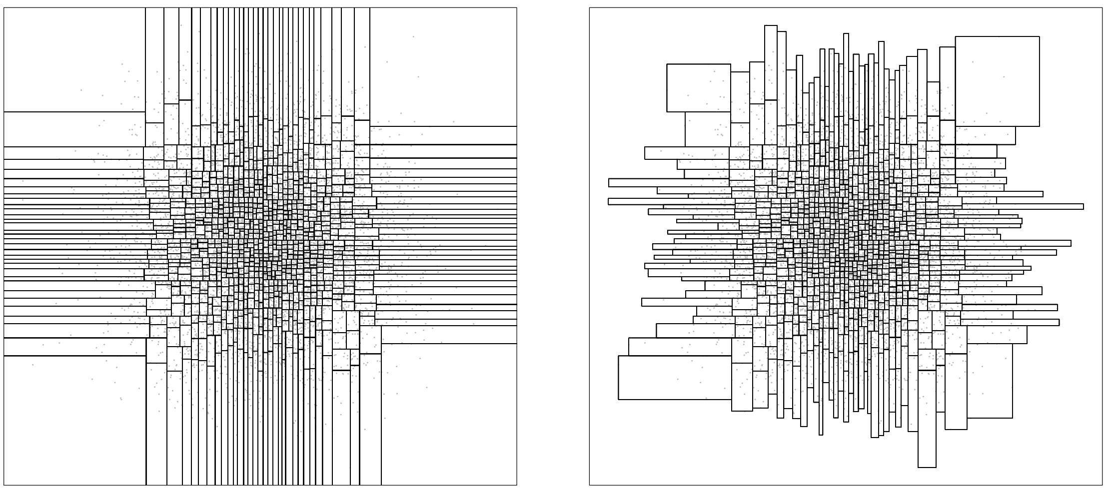
where $f(x)$ represents the probability density function [75]. In principle, based on the observation sample, the phase space is recursively split into quantiles along each dimension to create cells A_j so that "We may construct an approximation $f_A(x)$ having the same probability mass in each cell as $f(x)$, but with a uniform density in each cell [...]. An empirical density estimate can then be made:" [75]

$$\hat{f_A}(x) = \frac{n_j}{N\mu(A_j)} \qquad j s.t. x \in A_j \tag{6.7}$$

Where N is the total number of observations, n_j the number of observations within cell A_j, while "$\mu(A_j)$ is the D-dimensional volume of A_j. [...] The within-cell uniformity of $\hat{f_A}(x)$" [75] enables the entropy estimator

$$\hat{H} = \sum_{j=1}^{m} \left(\frac{n_j}{N} \log \left(\frac{N\mu(A_j)}{n_j} \right) \right) \tag{6.8}$$

A bias that occurs in this method is that the cells on the edges of the phase space have infinite volume, which makes it unsuitable for high dimensional data. Stowell et al. [75] suggest two solutions. Either, one simply ignores these regions, or the volume of the edge regions can be limited to the extreme values of the observations within that region. Here the second alternative is applied. Figure 6.3a shows the kd-partitioning of a two dimensional normal distribution, and figure 6.3b the same partitioning, but with limited edge cells.



(a) Infite edge cells (b) limited edge cells

Figure 6.3: Kd partitioning of a 2D normal distribution

The partitioning along each dimension is repeated for a minimum number of levels according to equation 6.10, and until a sufficient uniformity in each cell is reached, tested with equation 6.9, where $med_d(A_j)$,$min_d(A_j)$ and $max_d(A_j)$ correspond to the median, minimum and maximum value in dimension d of the observations in the phase-space partition A_j respectively. The maximum number of partitions possible is equal to the number of observations.

$$\left|\sqrt{n_j}\frac{2 \cdot med_d(A_j) - min_d(A_j) - max_d(Aj)}{max_d(A_j) - min_d(A_j)}\right| > 1.96 \qquad \text{uniformity test} \tag{6.9}$$

$$L_N > \frac{1}{2}\log_2 N \qquad \text{minimum number of levels } L_N \tag{6.10}$$

Continuous entropy differs from discrete entropy, as it is based on the values of the probability density distribution. Because of this, in opposition to discrete entropy, the result is not independent from the phase space coordinate system and its units [80]. Depending on the basic units, the result of the continuous entropy can even be shifted to take negative values [81][82]. In order to maintain it as a meaningful quantity for real systems and the possible microstate configurations, continuous entropy "must be referenced to units of the underlying random variable" [82]. We substitute $\mu(A_j)$ by $\mu(A_j)' = \frac{\mu(A_j)}{\mu(A_{min})}$, where A_{min} should be set to a value smaller than the resolution of the recorded data, and must satisfy the condition $A_{min} \leq A_j$ for all j. The same A_{min} must be used for all compared systems.

On different scales, the phase space has a very different maximum extent. The maximum volume of a D dimensional phase space at radius r is given by equation (6.11), while in reality, a uniform distribution across the entire phase space is impossible if the observed system is finite in size, because the observations are spatially dependent from each other.

$$A = (\pi r^2)^D \tag{6.11}$$

An attempt can be made to normalise entropy in relation to the largest possible entropy given the extent of the phase space at each scale with the Kullback-Leibler divergence or relative entropy D_{KL}. For two distributions P and Q, it "aims to quantify how 'close' two probability distributions are"[83, p.35]. However, relative entropy is not symmetric, and deeply rooted in information theory, which makes its exact interpretation outside information theory unintuitive.

In this chapter, we constructed an entropy estimator that can handle multidimensional continuous spatial data, and includes the observations' surroundings on multiple scales, which can now be applied to real world data.

Chapter 7

Analysis

7.1 Data: West London 1875 - 2005

The dataset used in the analysis was originally built and provided by Stanilov et al. [84]. It covers 200 square kilometers, spanning 20km from east to west, from London's green belt in the west to the west end hyde park, and roughly 10km from north to south [84] (see figure 7.1).

Figure 7.1: Map of the study area in the context of greater London

It was chosen based on the combination of its large temporal extend and detailed information about land-use. The data provides detailed maps in seven time steps about every 20 years: 1875,1895,1915,1935,1960,1985 and 2005. It includes information on each building's function in one of the 32 categories shown in table 7.1, also including open sites such as parks. Figure 7.2 to figure 7.8 show the full area with the original categories for all time steps.

Stanilov et al. [84] provide detailed information on the collection process of the data that involved digitising historical maps from Ordnance Survey (OS) on the scale 1:2,500. The land use classes were categorised manually: "The process of land use classification involved the interpretation of building footprints from the OS maps; verification of building type (for buildings still in existence) in Microsoft Virtual Earth (now Bing Maps 2D and 3D) and Google Street View; and cross-referencing the results with several land use databases for Greater London" [84], while for some categories label information was given in the OS maps.

Table 7.1: Functional Categories in the data

APT	apartments	GAR	garrages
APTH	high rise apt	IND	industrial
COT	cottages	UTL	utilities
DET	detached housing	WAT	water
DETH	high density detached	GEN	mixed/commercial
MEW	mews	RET	big box retail
SDT	semi-detached housing	LDG	lodges/hotles
TER	terraced housing	INS	institutional
INSL	large institutional	FRM	farm structures
O	office	NRS	tree nurseries
RRS	rail stations	PRK	parks
CHR	religious	REC	recreational
OLD	old fabric/mix	SCH	schools
AGR	alotment gardens	AIR	airports
CEM	cemeteries	CLR	cleared
EST	land estates	STA	stadia

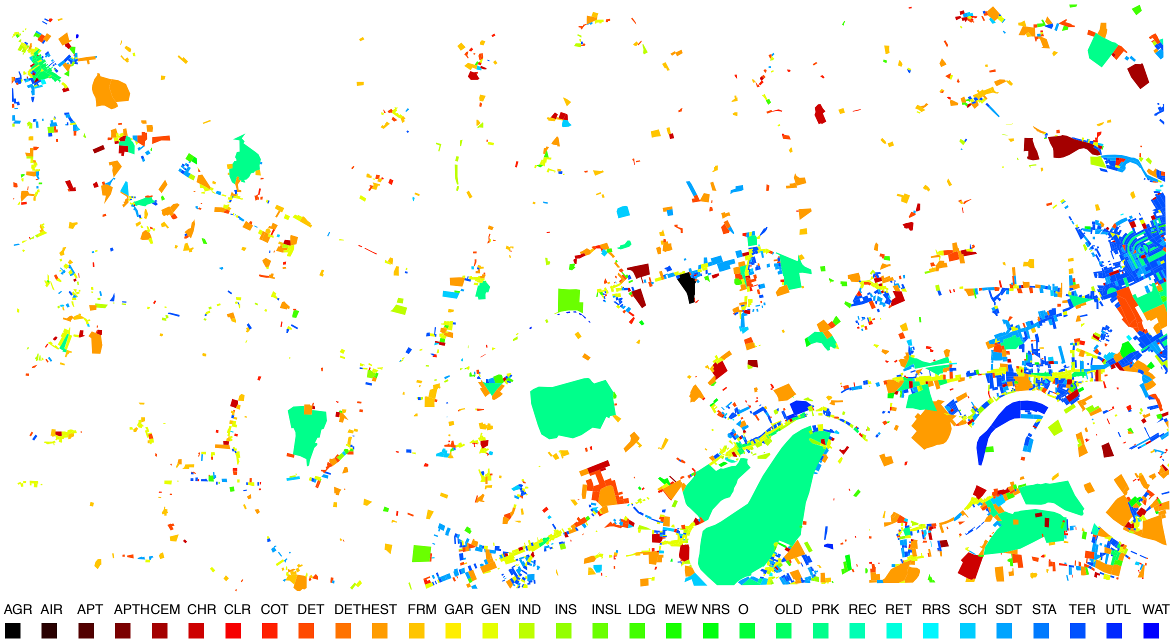


Figure 7.2: Map of west London by category 1875

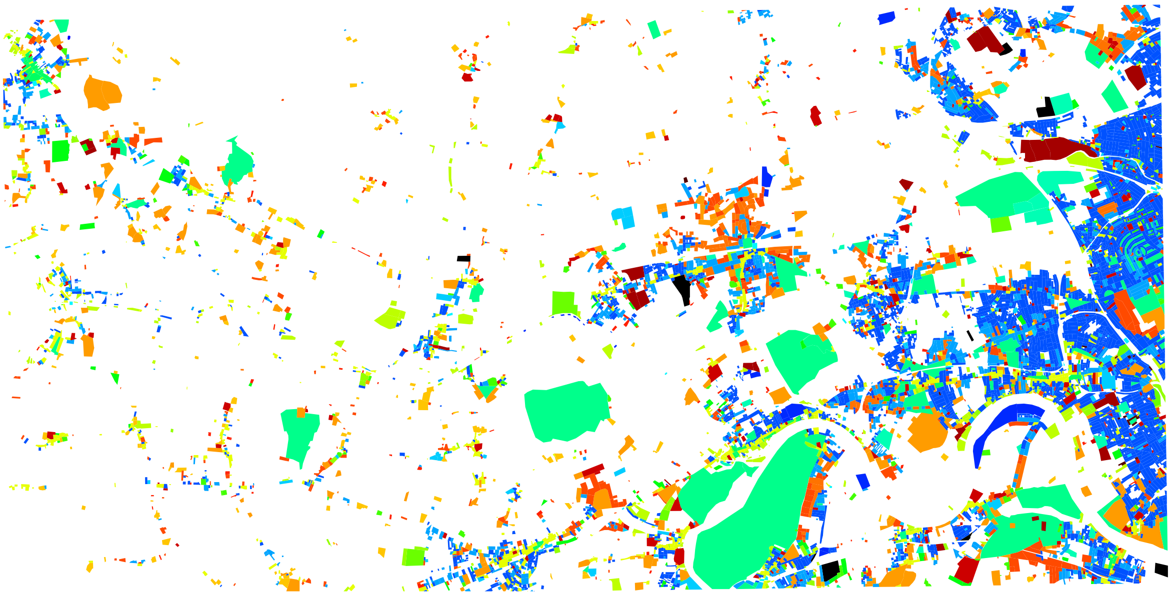


Figure 7.3: Map of west London by category 1895

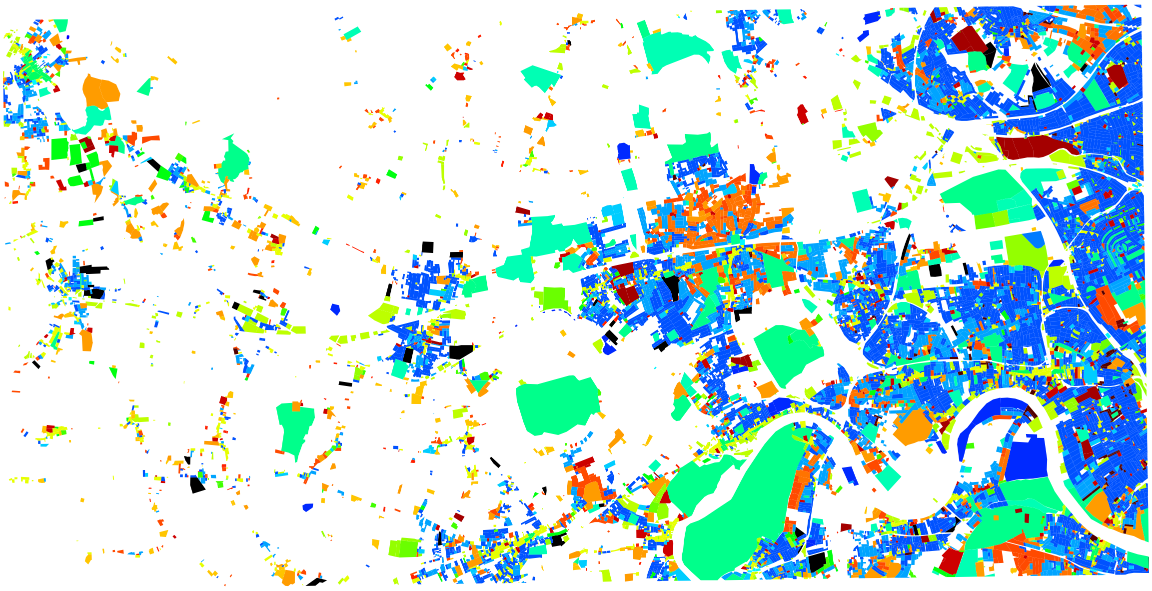


Figure 7.4: Map of west London by category 1915

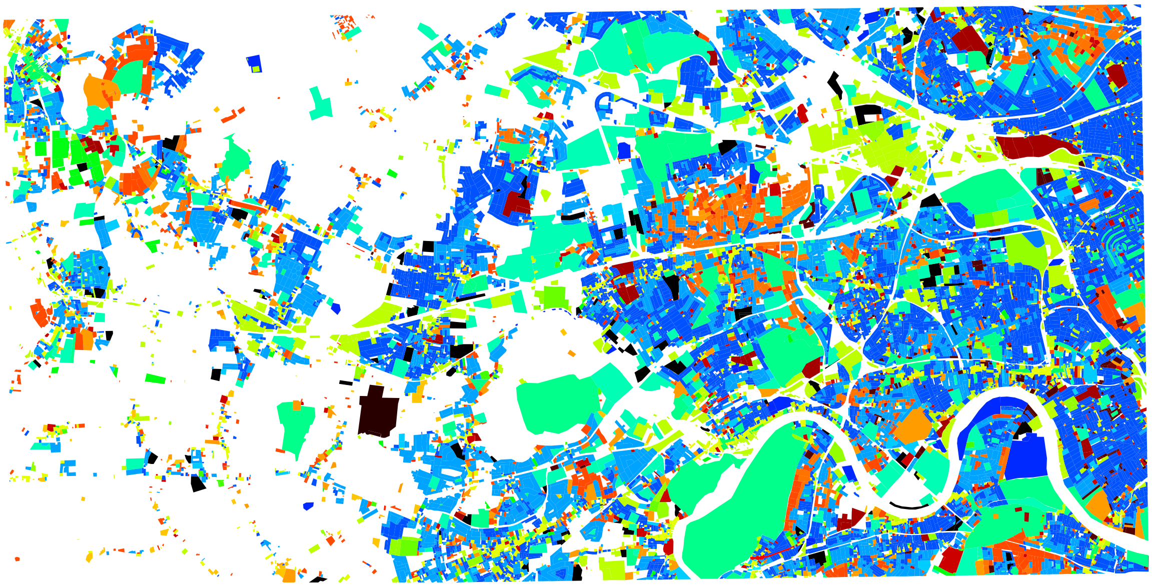


Figure 7.5: Map of west London by category 1935

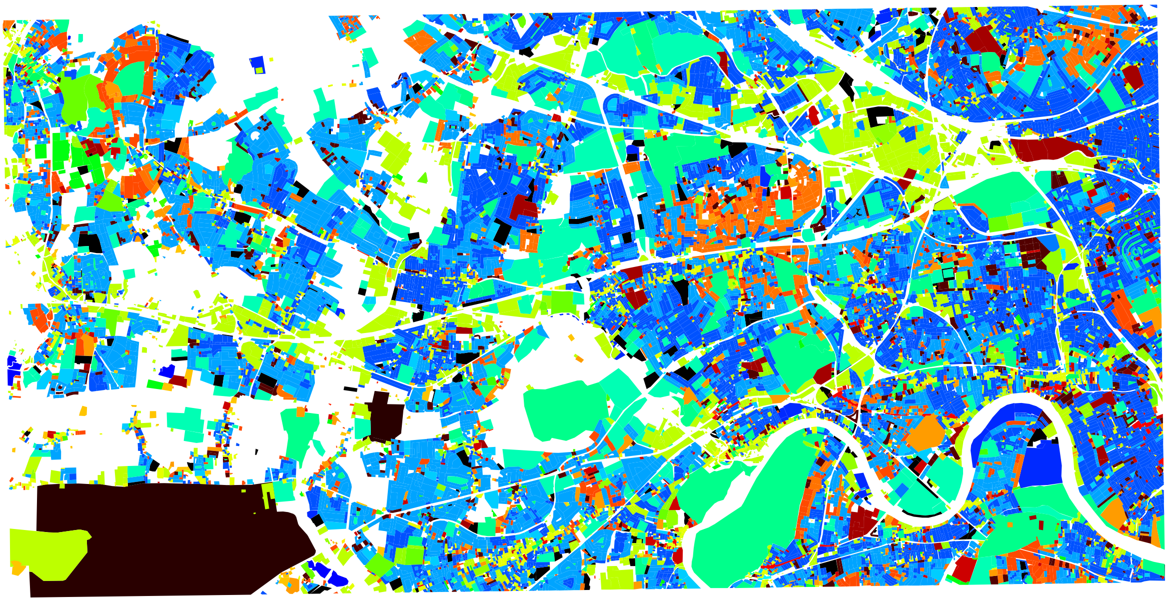


Figure 7.6: Map of west London by category 1960

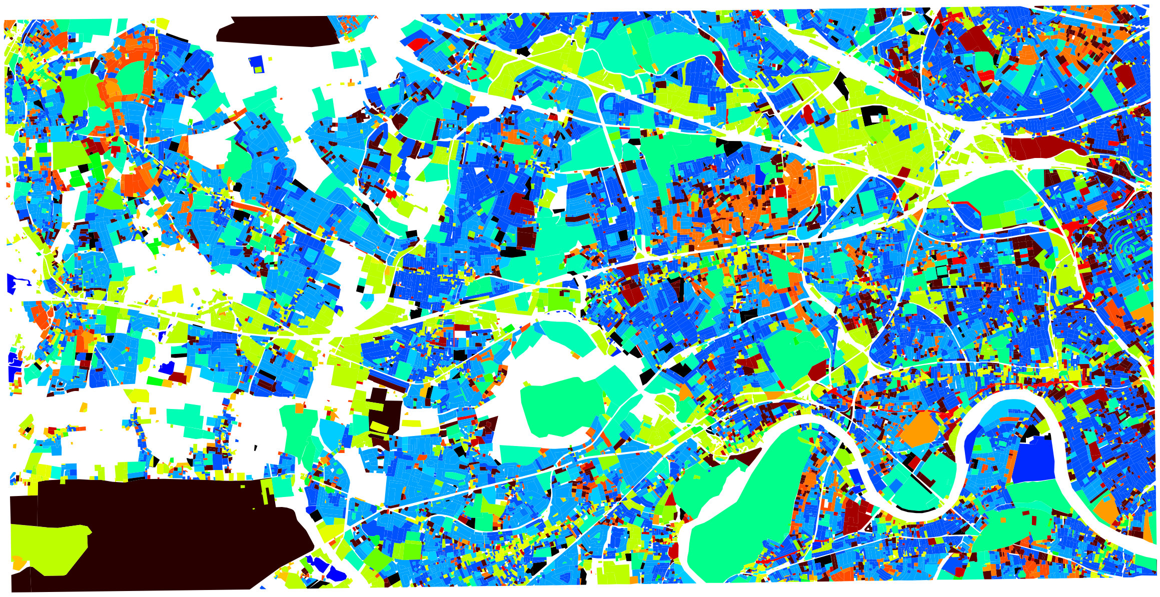


Figure 7.7: Map of west London by category 1985

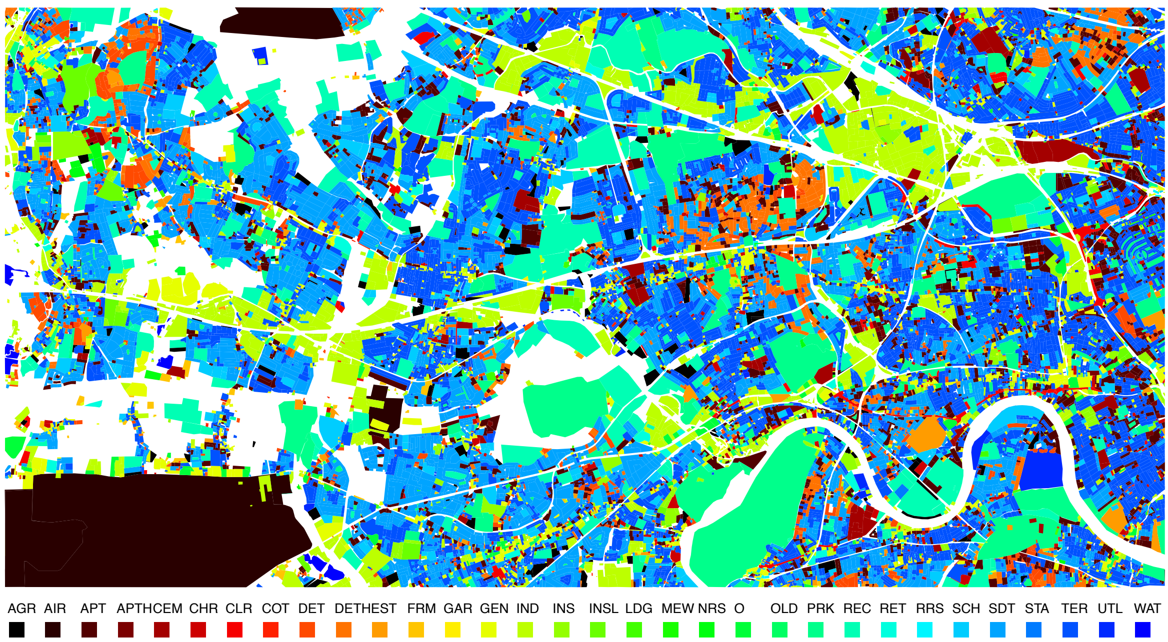


Figure 7.8: Map of west London by category 2005

Rooted in the original purpose of the collection of the data, the study area was selected with regard to specific qualities. “in many respects, this slice has the most variety of any of the possible slices through the metropolis in that its diversity of land uses is far greater than slices which run out from the centre to the east, north or south or any other compass points that might be chosen”[84]. Stanilov et al. [84] evalutate it to be a “representative sample of Londons metropolitan fabric” [84], but point out differences between different parts of London on the scale of the study area. As in this study the diversity of land uses is itself at the focus of attention, the biased choice of the study area towards the section with the largest variety may propagate to make it less representative of London for this works purpose.

7.2 Data Preparation and Processing

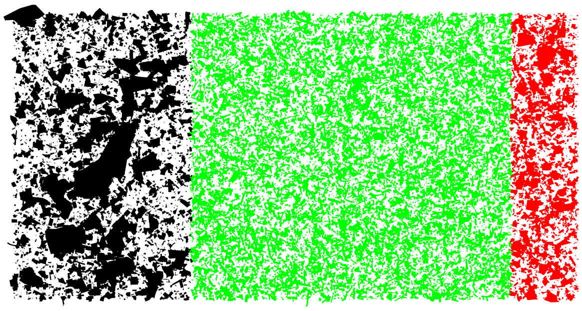
The functions are grouped into three main groups: “living”, “working” and “leisure” (see table 7.2). The chosen categories broadly reflect classifications used by geoinformation sciences [85], urban planning theory [17] and practice [86], altered to cover a broader range of uses. “Leisure” is taken in the broadest sense of activities not related to workplace or home, including most categories not covered by the former two. A virtual fourth category appears in the open space not covered by any of these three categories, equally entering the calculation. This simplification is unavoidable within the given time frame, as the number of considered dimensions heavily affects the computational effort.

Table 7.2: Grouping categories

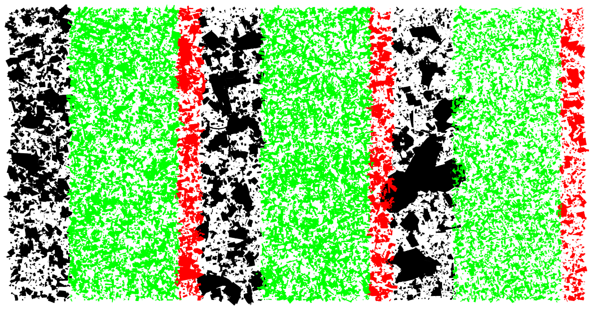
Working		**Living**	
INS	institutional	APT	apartments
INSL	large institutional	APTH	high rise apt
O	office	COT	cottages
GAR	garages	DET	detached housing
IND	industrial	DETH	high density detached
		MEW	mews
		SDT	semi-detached housing
		TER	terraced housing
		LDG	lodges / hotels
Leisure		**Omitted**	
AGR	alotment gardens	UTL	utilities
CEM	cemeteries	RRS	rail stations
EST	land estates	AIR	airports
FRM	farm structures	CLR	cleared
NRS	tree nurseries		
PRK	parks		
REC	recreational		
CHR	religious		
WAT	water		
STA	stadia		
SCH	schools		
GEN	mixed/commercial		
RET	big box retail		
OLD	old fabric/mix		

All parts of the analysis are performed in R [87] with the integrated development environment RStudio [88]. The spatial data, provided in Shapefile format (.shp), is imported as a spatial polygons data frame using the rgdal package [89], and transformed from the global positioning system (GPS) into the Universal Transverse Mercator coordinate system (UTM), allowing simpler euclidean distance calculations. The centroids and total area of the polygons are calculated. A regular grid of observation points is laid over the entire area in 300m intervals (2278 in total), and a distance matrix between all polygons and all observation points created. With the distance matrix, the total area covered by each function group in different radii from each observation point is calculated and stored in a matrix with a single value for each category at each radius. The radii used here are 300m, 400m, 500m, 600m, 1000m, 1500m, 2000m and 2500m. For each year, the kd-partitioning entropy and the relative entropy are calculated as explained in chapter 6, with $(A_{min} = 1m)^D$ with $D = 3$ for three dimensions, representing the different functions.

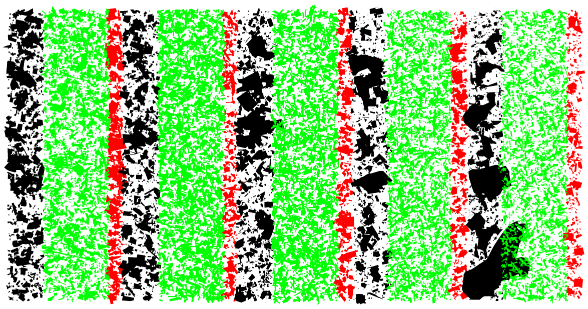
For comparison, a fully random case and a case of simple order are evaluated as well. The polygons are redistributed randomly, uniformly across the study area as in figure 7.9e for all time steps individually, Each building's area is set to the mean area of buildings of its category, which preserves the total amount of building area of each category. The result of the randomised case is averaged over 50 repetitions of the analyis. For the ordered case, the east axis is divided into partitions proportional to the total area of buildings of each category, and the polygons distributed uniformly in the respective partition (see figure 7.9a to 7.9d). We intend to confirm with these baseline configurations our expectation that while random spatial distributions and spatial distributions of simple order display high entropy on specific scales, the real distribution in its complex spatial order should display a larger entropy across a wider spectrum of scales.



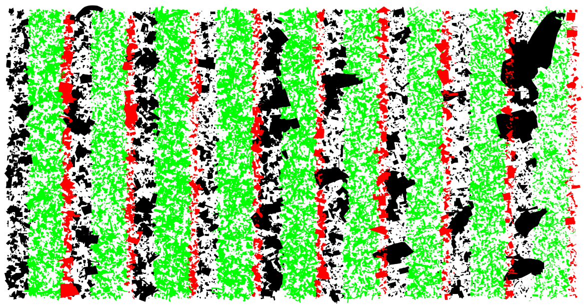
(a) ordered in 1 partition



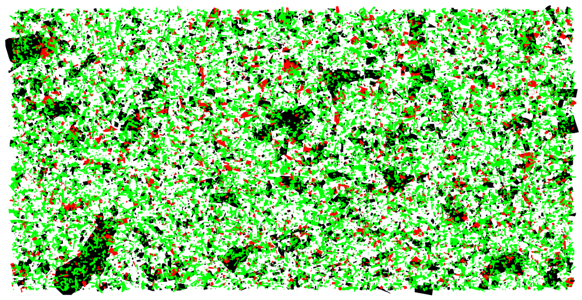
(b) ordered in 3 partitions



(c) ordered in 5 partitions



(d) ordered in 9 partitions



(e) randomised

Figure 7.9: Maps of reorganised spatial distribution: 2005 polygons by category

The code for the computational implementation into R can be found in the appendix 9.1.

7.3 Results

The entropy of the west London data set is shown to be mostly increasing with time and scale. The entropy of the original data is evaluated in the context of the randomised and ordered complements. It is observed that under simultaneous consideration of multiple scales, the increase in entropy in the real spatial distribution of buildings exceeds the increase in entropy of the randomised and ordered versions.

Figure 7.10 displays the entropy measured for each year in the original data set, plotted over the radii r taken into account as the surroundings defining each observation's characteristics.

As in the example in 4.2 and the multiscale entropy of the patterns in 4.3, from the way the method is constructed, entropy first increases with scale, as for larger radii the absolute difference between the values describing the observations can be larger, and then, because the observed system is finite in size, decreases from a certain scale upwards, because the observations overlap more and more with increasing scale. Therefore, the general increase, tipping point and decline depend on the size of the observed system and have no particular relevance to the evaluation of the entropy.

From 1875 to 1960, entropy clearly increases over time for all scales (except between 1935 and 1960 on scale 2500m). Between 1960 to 1985 entropy stagnates, and between 1985 and 2005 a decrease of 0.2% or less appears on all scales. The stagnation and decrease after 1960 is equally observed after randomising and reordering the building's locations, indicating that this effect is due to the total area of buildings of different categories in the whole study area rather than their spatial distribution within it. Figure 7.11 displays the same data as 7.10, but with time as the x-axis.

	300m	400m	500m	600m	1000m	1500m	2000m	2500m
1895	28.59	23.06	19.69	16.14	9.34	6.42	5.60	4.74
1915	16.53	12.95	9.23	9.09	5.59	4.59	4.42	3.58
1935	22.09	15.97	13.25	9.92	5.08	3.54	3.07	2.91
1960	10.14	6.03	4.10	2.87	1.19	0.41	0.37	-0.14
1985	1.09	1.02	0.42	0.27	0.04	0.24	0.31	0.53
2005	-0.07	-0.18	-0.49	-0.49	-0.14	-0.15	-0.00	-0.38

Table 7.3: Change in entropy compared to previous year in %

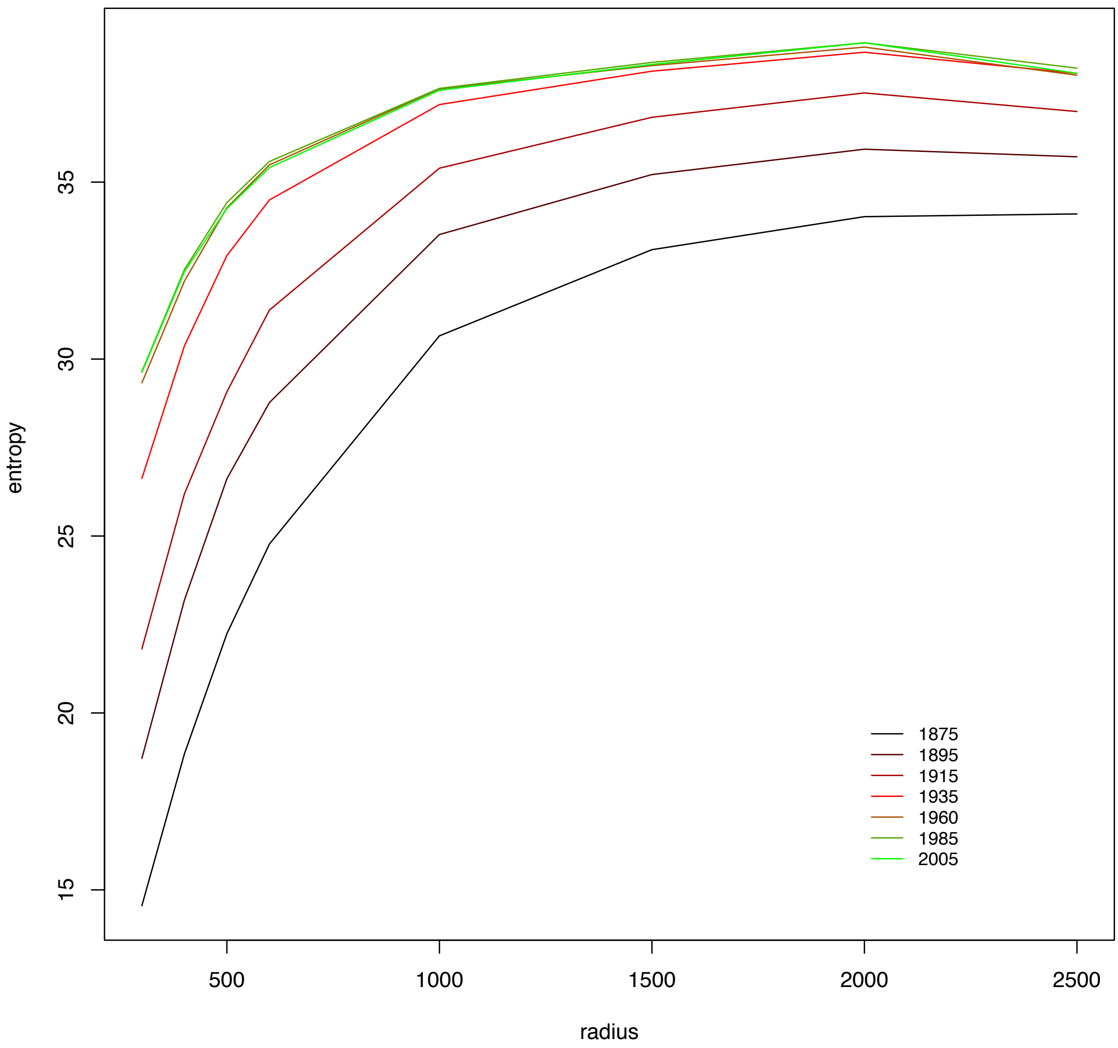


Figure 7.10: Entropy in west London along the observation radius, for 1875-2005

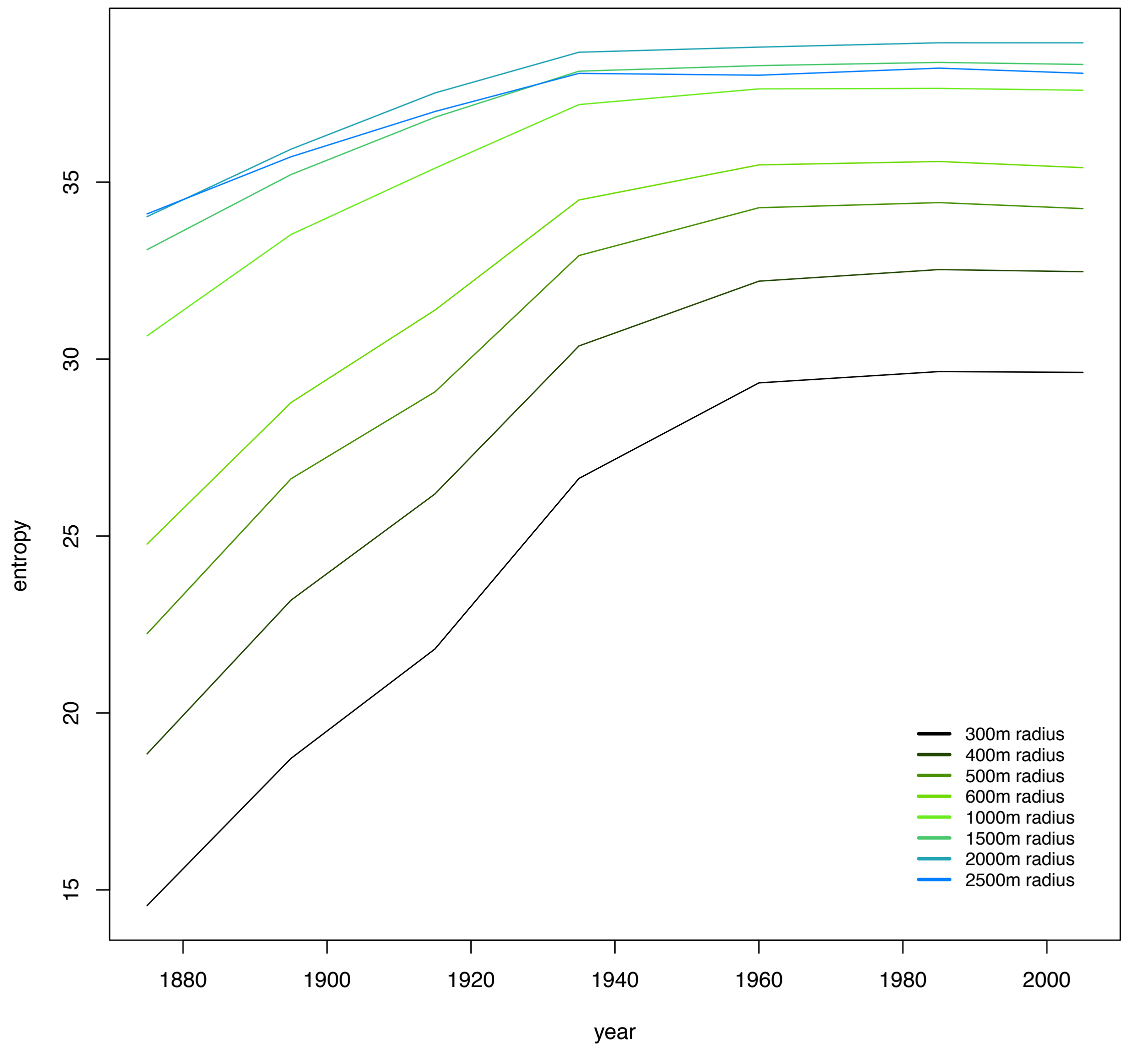


Figure 7.11: Entropy in west London over time, over observation radii

So far, it is unclear in how far the observed change in entropy is due to the spatial configuration, or simply governed by a change in the total amount of building area of different categories. In order to gain a clearer picture, a comparison to a randomised and ordered spatial distributions as described in section 7.2 is performed.

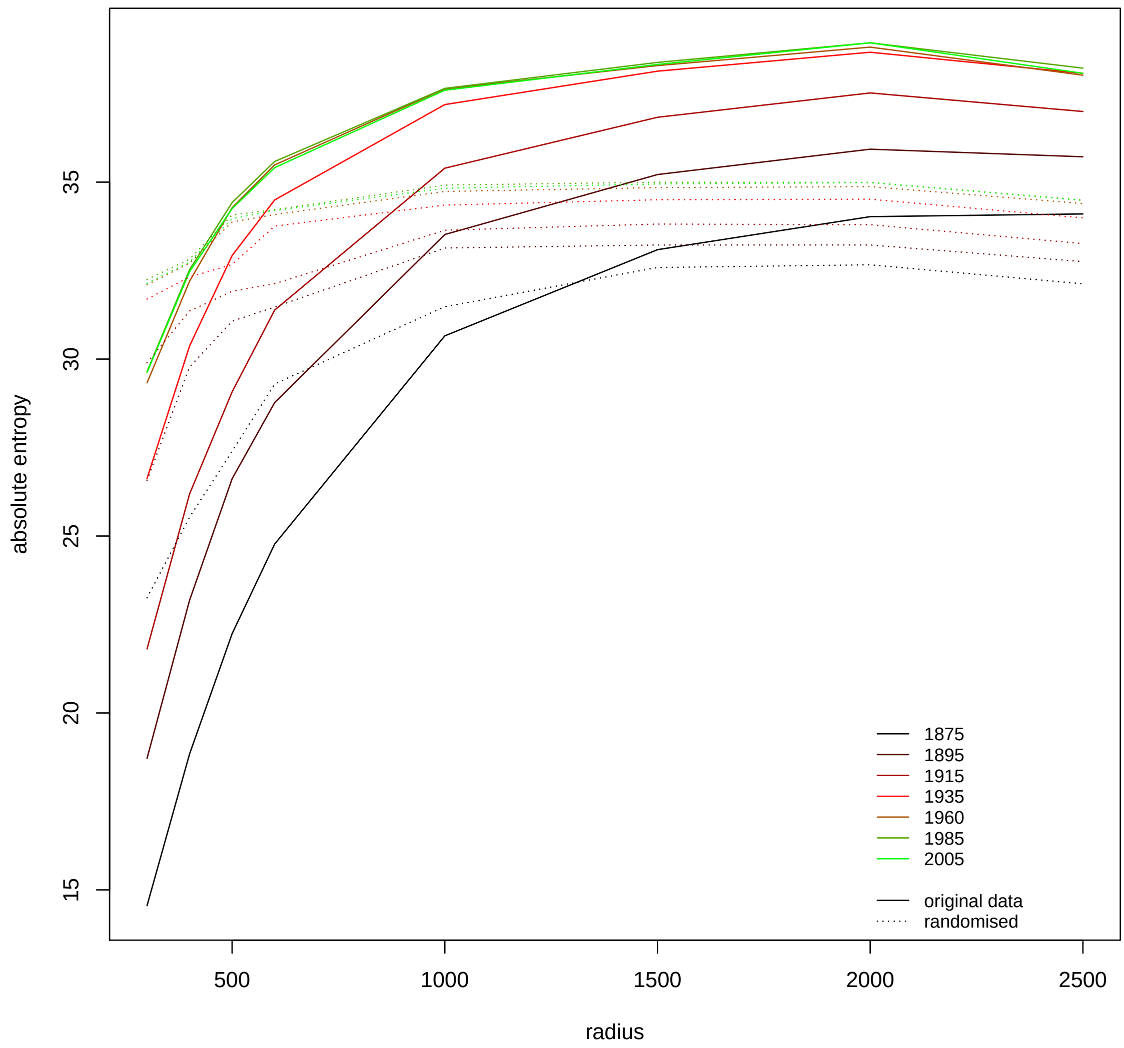
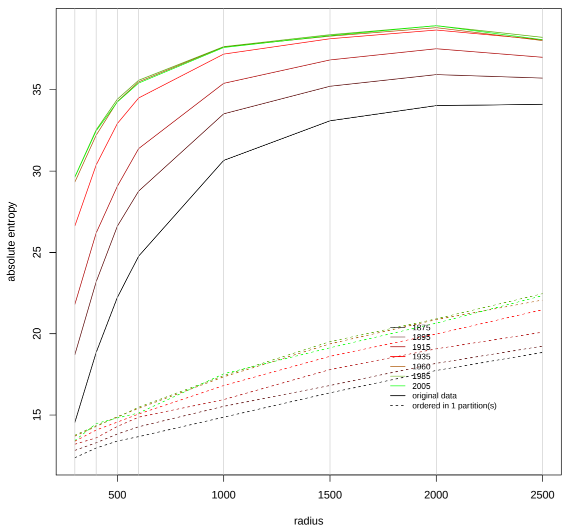
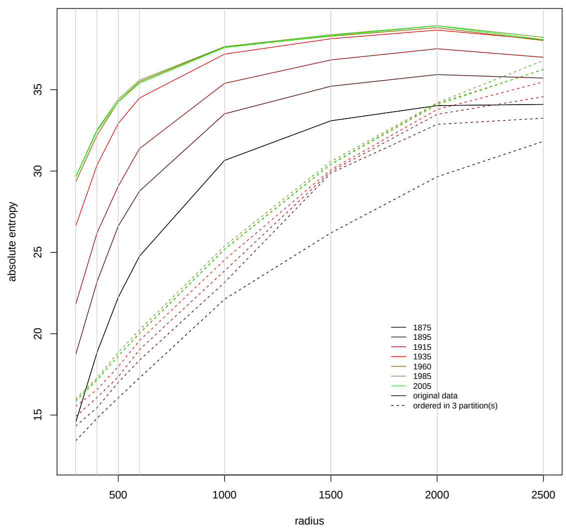


Figure 7.12: Entropy in London and the random case over the observation radius, for 1875-2005

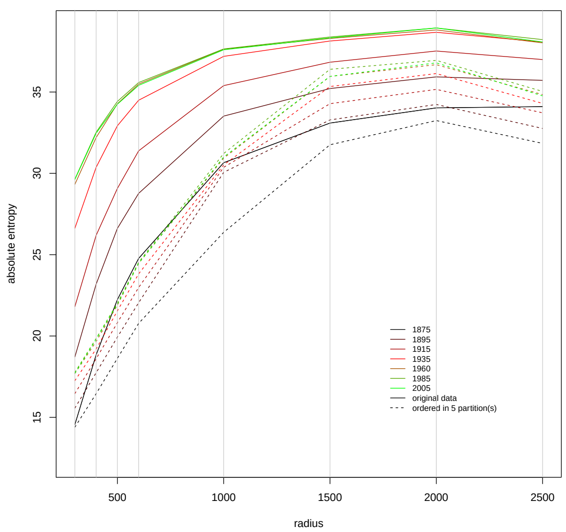
Figure 7.12 shows the entropy of the original an randomised data in context. For all time steps, the randomised data has a higher entropy on the lowest considered scale of observation, but a lower entropy on the highest scale of observation. What changes over time is that the crossing point, the scale at which the entropy of the original data exceeds the entropy of the randomised case, moves towards lower scales over time. This clearly demonstrates the part of our hypothesis concerning randomised spatial distributions. While the randomised spatial distribution shows higher entropy on low scales, the real spatial distribution tends towards higher entropy across a wider range of scales.



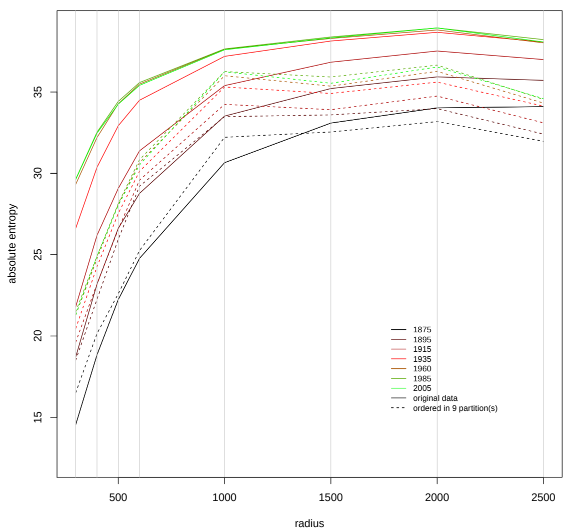
(a) ordered in 1 partition



(b) ordered in 3 partitions



(c) ordered in 5 partitions



(d) ordered in 9 partitions

Figure 7.13: Entropy in London compared to ordered spatial distributions over the observation radii

In contrast, spatial distributions of simple order have their largest entropy compared to the observed data on very different scales, depending on the scale at which the ordering pattern repeats itself (see number of partitions in figure 7.9). Figure 7.13 clearly displays the shift of the relatively highest entropy towards larger scales for simple orders with fewer repetitions. For the way the simple order is constructed here, if the frequency of the repetition of the pattern is increased infinitely, the distribution approximates the fully random distribution, shifting the relatively largest entropy to ever smaller scales.

In summary, the analyis of the original data, the randomised data and the ordered data with four different scales of pattern repetition are viewed together year by year in figures 7.14 to 7.20. Viewing them in the given order reveals the steady increase in total entropy and balance across the various scales of the original data, generally exceeding the increase in both the randomised and ordered configurations. Here we see how in contrast to the beginning in 1875, after the transformation of the built environment over the entire time period, the entropy of the spatial distribution is unmatched by any other tested configuration when considering all scales together.

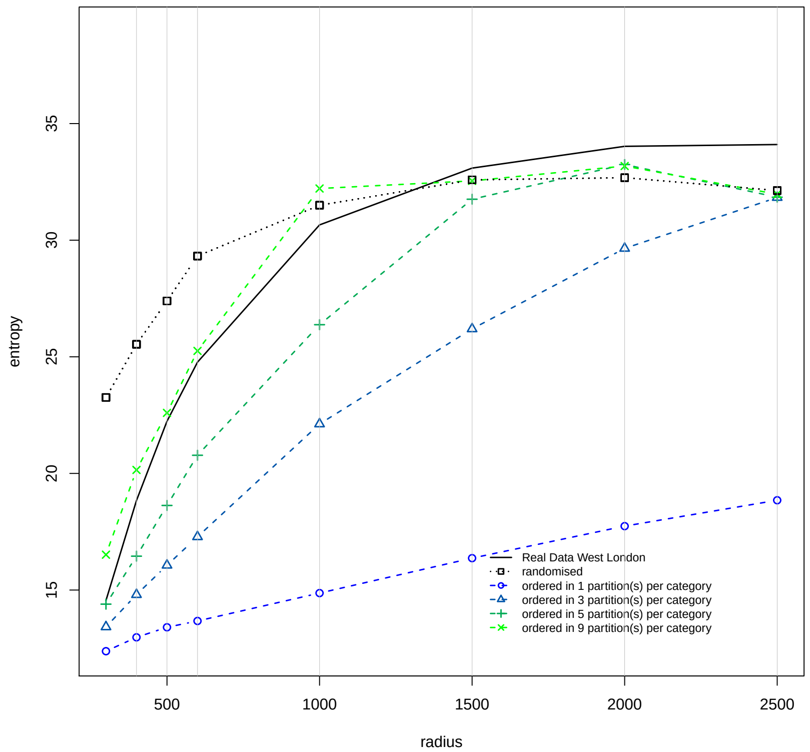


Figure 7.14: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1875

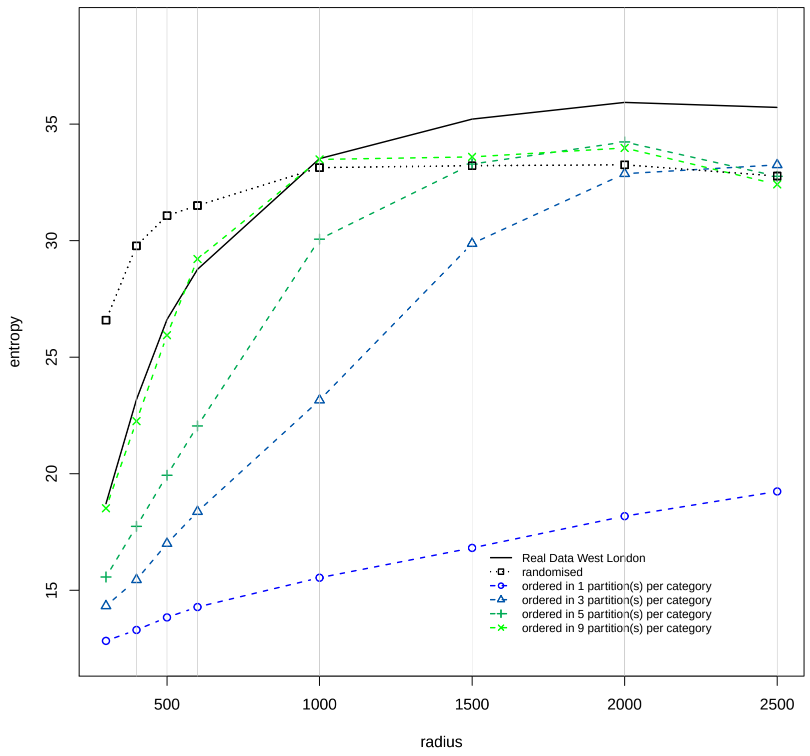


Figure 7.15: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1895

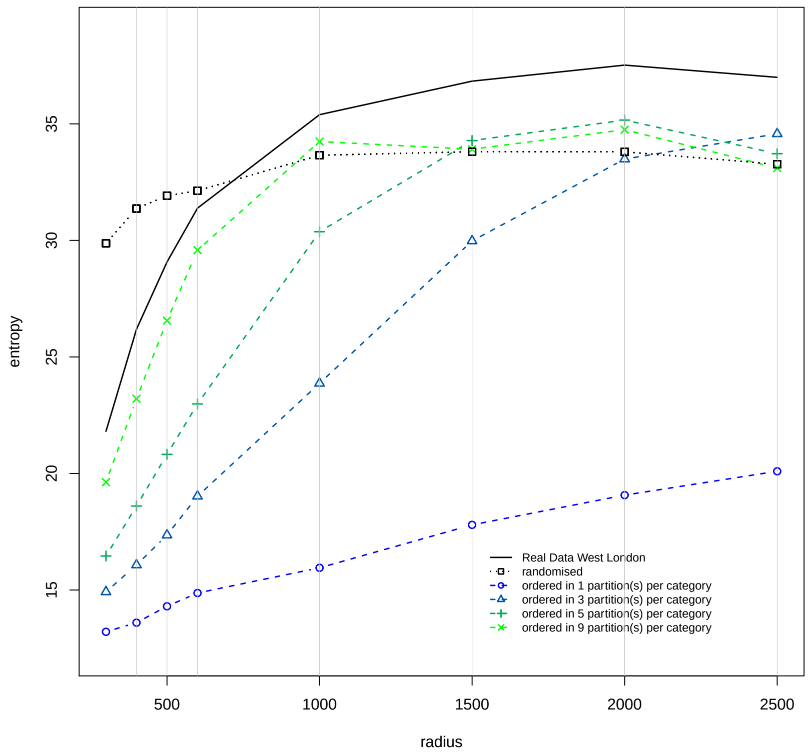


Figure 7.16: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1915

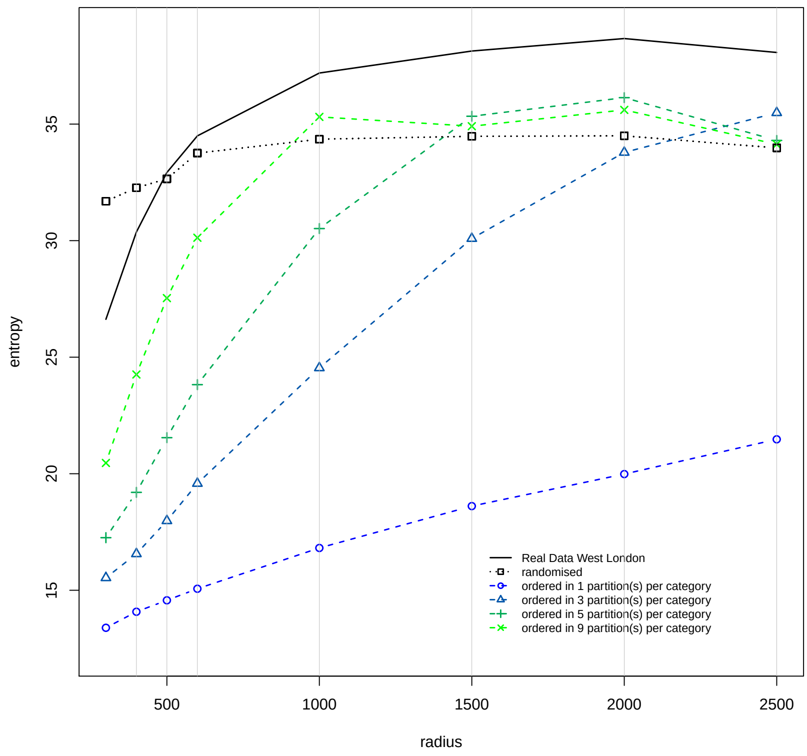


Figure 7.17: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1935

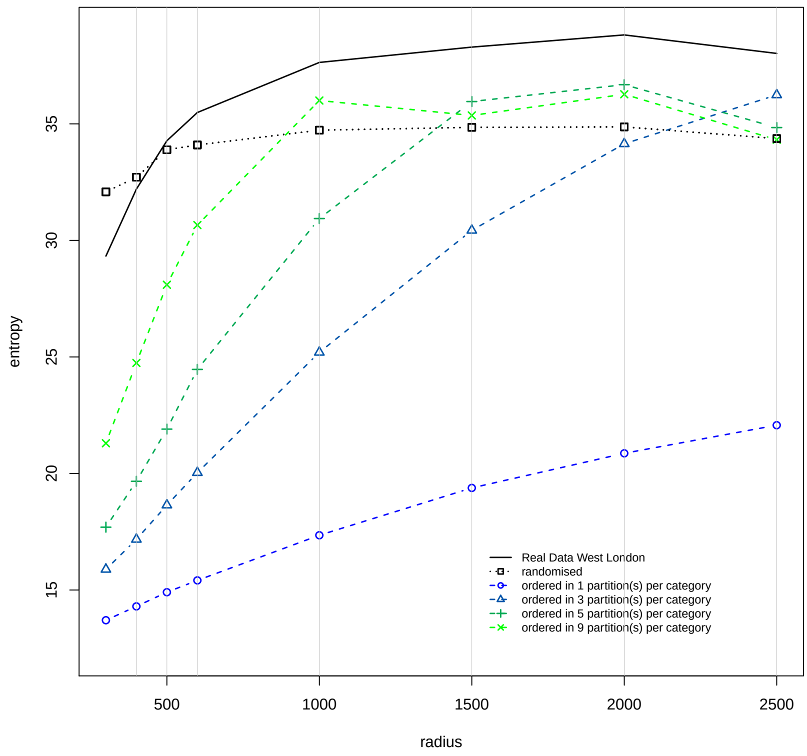


Figure 7.18: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1960

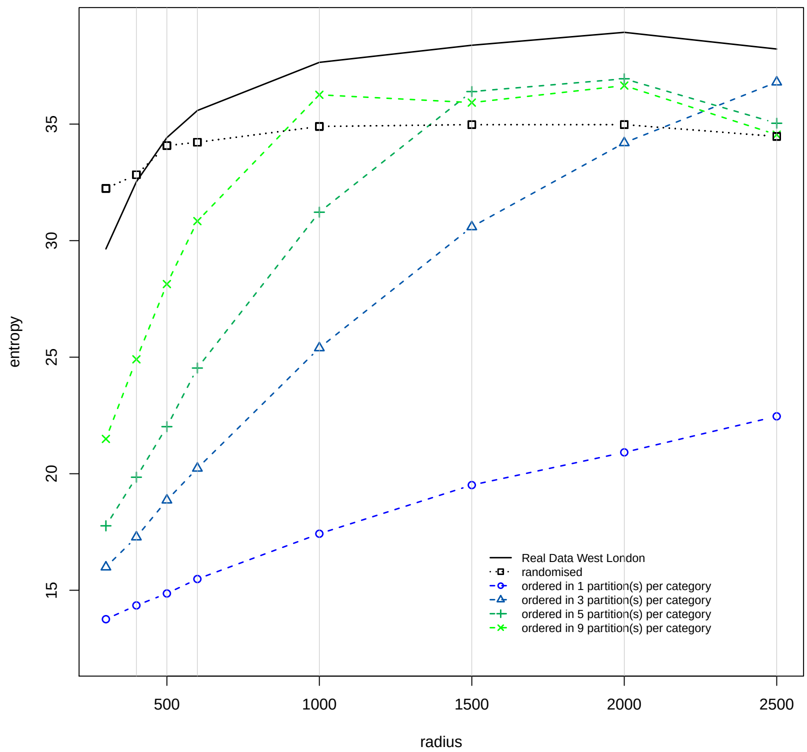


Figure 7.19: By year comparison: Absolute Entropy; Real, randomised and ordered data, 1985

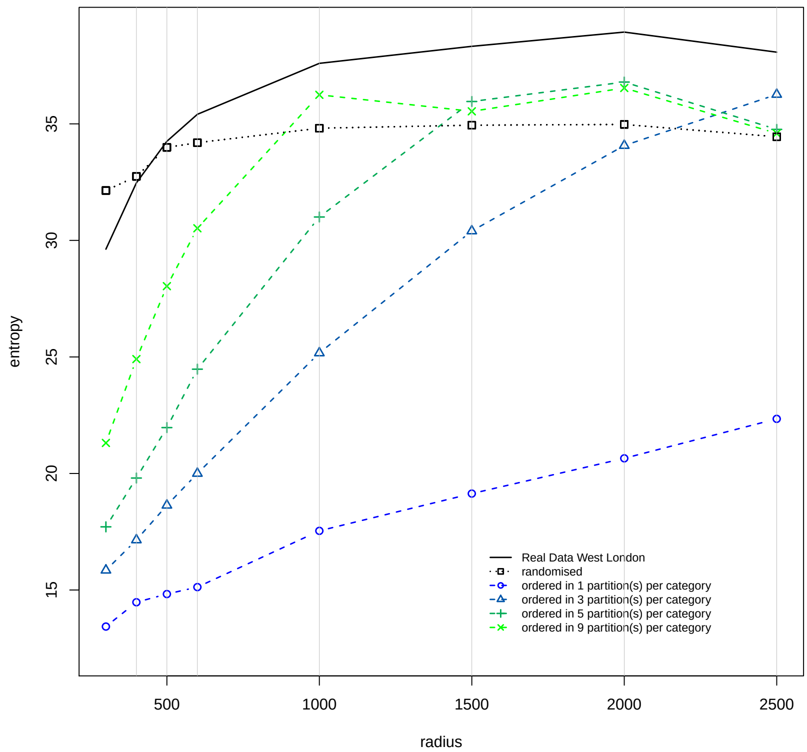


Figure 7.20: By year comparison: Absolute Entropy; Real, randomised and ordered data, 2005

Finally, it remains to evaluate what this means for the urban structure. High entropy on a low scale means that if we look at small patches of land, they are very different concerning the amount of area that falls into the different categories of use; there is a large variety of choice when it comes to very local environments. This could roughly correspond to small blocks being very different from each other. High entropy on a large scale means a large variety of of choice looking at large scale environments. This could for example vaguely correspond to districts, where you would find a large variety including mixed-use, residential, and business districts as very different large scale environments. It is shown that distributions that are random or follow a simple orderer do not achieve the amount of variety of locations across multiple scales the spatial distribution in west London provides. In fact, the distribution tends towards a higher and more balanced entropy over time. Recalling the patterns from figure 4.3, we see that this is only achieved by a complex spatial order.

Chapter 8

Concluding Remarks

This work was set out to develop a coherent interpretation of entropy for cities, transfer it into a practical method, and to evaluate the development of entropy in west London from 1875 to 1985. Any conclusions drawn from the quantitative analysis must be evaluated with the limitations concerning the scope and depth of this work in mind. The following section first outlines these limitations and further steps that could be taken. Then, the empirical results are linked to the theoretical concepts discussed. Furthermore, a perspective is given how each point may relate to practice in planning if they are supported by further evidence.

The scope of this work is limited not only in the depth and precision of the methodology, but unavoidably omits large parts of the literature although it clearly relates to the discussed topic. Among those the author is aware of, one finds analytical methods concerning fractals [90] and multi-fractals such as the box counting method to determine a structure's fractal dimension [91] which shows a strong resemblance to the proposed method. Furthermore, crucial aspects of the concept of entropy in thermodynamics such as the relation to energy and pressure are omitted, with focus only on a narrow aspect of the statistical principle. Another quite obvious aspect is that no considerations are made in how far certain distributions may or may not be efficient or suitable to any specific aspect of urban life, while urban form is governed by a multiplicity of dynamically changing mechanisms overlaying the principles discussed here, for example attracting and repulsing forces between specific functions [84]. In addition, unconsciously made assumptions have become apparent during the process of creating this work, indicating that many remained unnoticed. As an example can be given the assumption that a discrete categorisation of buildings into clearly separable functions is at all reasonable now, or relevant in the future.

Any generalisation to "universal" processes from the observations made in the small observed area in west London are speculations. The edges of the system are not ideally chosen, but given by the extend of the data. Infrastructure networks are ignored, and distance as the defining element of the relationship between places is an oversimplification.

Further steps could be undertaken by introducing a refined methodology incorporating a continuous approach to distance using a decay function. Analysis of data for multiple cities and aspects of urban morphology would be needed to reach a degree of generalisability. In the larger picture, the broader question of structural variety on multiple scales, and how it is morphed by networks, flows and interactions seems worth exploring further.

Despite these limitations, a careful attempt can be made to drawn conclusions from the observed. The analysis generally supports the hypothesis of increasing entropy of location's characteristics. It suggests that the system shifted towards an increase of entropy in the proposed spatially dependent

characteristic phase space rather than in the “geographical phase space”, increasing the variety of environments of multiple sizes simultaneously.
If it should emerge that the general perspective of multiscale entropy in characteristics of locations as a relevant factor in urban morphology can be supported by further and more generalisable evidence, one could draw from it noteworthy thoughts considering planning theory: If in fact cities have a “natural tendency” to provide a high entropy of location’s characteristics, one may consider it a useful quality of the urban system. Let us, for now, adopt Jane Jacobs perspective that “The point of cities is multiplicity of choice” [17]. Then, supported by the observations made in this paper, one might argue that the “best” distribution of functions is neither a total segregation of functions as proposed by modernist planners in the “Athens Charter” [92], nor an as even mix of functions as possible, advocated more recently by “Walkable Cities” theorists [93]. They both lead to sameness, the former to sameness among low scale environments, the later to sameness on large scale environments. The solution that provides the largest multiplicity of choice is given by a much more complex pattern.
Independent from what is practically considered “good” urban design - the total absence of “badly” designed places might be disadvantageous simply because they had offered opportunities for unforeseen uses. Obviously, these considerations are speculative. Nonetheless, they raise interesting questions. Any optimisation in design or planning can only be rooted in uses, actions, interactions, technologies and societal structures the planner observes in the environment, or can personally imagine as a planned alternative, narrowing the possible uses down to those the planner sees as the most comfortable or efficient or sustainable or in any other way preferable. This by its very nature excludes the unexpected, the innovative and the surprising - meanings and functions of urban spaces that may exceed a planners imagination.

Bibliography

[1] A. Gudmundsson and N. Mohajeri, “Entropy and order in urban street networks,” *Nature Scientific Reports*, vol. 3, p. Article number: 3324, 2013.

[2] M. Costa, A. L. Goldberger, C. Peng, *et al.*, “Multiscale entropy analysis (mse),” 2000.

[3] M. Batty, *The new science of cities*. Mit Press, 2013.

[4] M. Batty, *Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT press, 2007.

[5] M. Batty and P. A. Longley, *Fractal cities: a geometry of form and function*. Academic Press, 1994.

[6] J. Portugali, “What makes cities complex?,” *Complexity, Cognition, Urban Planning And Design. Springer*, 2014.

[7] A. Wilson, *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis*. Taylor & Francis, 2014.

[8] P. Clark, *The Oxford Handbook of Cities in World History*. Oxford Handbooks in History, OUP Oxford, 2013.

[9] M. Batty, “Cities as complex systems: scaling, interaction, networks, dynamics and urban morphologies,” in *Encyclopedia of complexity and systems science*, pp. 1041–1071, Springer, 2009.

[10] L. Boltzmann, *Vorlesungen über Gastheorie*, vol. 1. JA Barth, 1896.

[11] C. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal, The*, vol. 27, pp. 379–423, July 1948.

[12] R. Zhou, R. Cai, and G. Tong, “Applications of entropy in finance: A review,” *Entropy*, vol. 15, no. 11, p. 4909, 2013.

[13] S. Gull and J. Skilling, “Maximum entropy method in image processing,” *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 131, pp. 646–659, October 1984.

[14] J. R. Holton and G. J. Hakim, *An introduction to dynamic meteorology*, vol. 88. Academic press, 2012.

[15] F. Ma, P. Lv, and M. Ye, “Study on global science and social science entropy research trend,” in *Advanced Computational Intelligence (ICACI), 2012 IEEE Fifth International Conference on*, pp. 238–242, IEEE, 2012.

[16] A. G. Wilson, *Entropy in urban and regional modelling*. Pion Ltd, 1970.

[17] J. Jacobs, *The death and life of great American cities*. Vintage, 1961.

[18] P. M. Allen, “Cities and regions as evolutionary, complex systems,” *Geographical systems*, vol. 4, pp. 103–130, 1997.

[19] L. M. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, “Growth, innovation, scaling, and the pace of life in cities,” *Proceedings of the national academy of sciences*, vol. 104, no. 17, pp. 7301–7306, 2007.

[20] Y. Chen, J. Wang, *et al.*, "Multifractal characterization of urban form and growth: the case of beijing," *Environment and Planning B: Planning and Design*, vol. 40, no. 5, pp. 884–904, 2013.

[21] Y. Chen and Y. Zhou, "Multi-fractal measures of city-size distributions based on the three-parameter zipf model," *Chaos, Solitons & Fractals*, vol. 22, no. 4, pp. 793–805, 2004.

[22] J. Ladyman, J. Lambert, and K. Wiesner, "What is a complex system?," *European Journal for Philosophy of Science*, vol. 3, no. 1, pp. 33–67, 2013.

[23] S. Johnson, *Emergence: The connected lives of ants, brains, cities, and software*. Simon and Schuster, 2002.

[24] P. M. Allen, "Cities: the visible expression of co-evolving complexity," in *Complexity Theories of Cities Have Come of Age*, pp. 67–89, Springer, 2012.

[25] A. Wehrl, "General properties of entropy," *Reviews of Modern Physics*, vol. 50, no. 2, p. 221, 1978.

[26] F. Reif, *Fundamentals of statistical and thermal physics*. McGraw-Hill, 1965.

[27] C. Callender, *The Oxford handbook of philosophy of time*. OUP Oxford, 2011.

[28] M. Batty, "Space, scale, and scaling in entropy-maximising," *Geographical Analysis*, vol. 42, 4, p. 395421, 2010.

[29] J. V. Michalowicz, J. M. Nichols, and F. Bucholtz, *Handbook of differential entropy*. CRC Press, 2013.

[30] P. Attard, *Thermodynamics and Statistical Mechanics: Equilibrium by Entropy Maximisation*. Elsevier Science, 2002.

[31] M. Le Bellac, F. Mortessagne, and G. G. Batrouni, *Equilibrium and non-equilibrium statistical thermodynamics*. Cambridge University Press, 2004.

[32] B. Latané, "The psychology of social impact.," *American psychologist*, vol. 36, no. 4, p. 343, 1981.

[33] E. T. Jaynes, "The gibbs paradox," in *Maximum entropy and bayesian methods*, pp. 1–21, Springer, 1992.

[34] W. R. Tobler, "A computer movie simulating urban growth in the detroit region," *Economic geography*, pp. 234–240, 1970.

[35] C. Alexander, "A city is not a tree," *Ekistics*, vol. 23, no. 139, pp. 344–348, 1967.

[36] A. G. Wilson, "A family of spatial interaction models, and associated developments," *Environment and Planning*, vol. 3, no. 1, pp. 1–32, 1971.

[37] K. E. Haynes and A. S. Fotheringham, *Gravity and spatial interaction models*, vol. 2. Sage publications Beverly Hills, 1984.

[38] D. J. Maguire, M. Batty, and M. F. Goodchild, *GIS, spatial analysis, and modeling*. Esri Press, 2005.

[39] B. Harris and A. G. Wilson, "Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models," *Environment and planning A*, vol. 10, no. 4, pp. 371–388, 1978.

[40] M. Fujita and J. Thisse, *Economics of Agglomeration: Cities, Industrial Location, and Globalization*. Cambridge University Press, 2013.

[41] G. C. Galster, "The mechanism(s) of neighbourhood effects: Theory, evidence, and policy implications," in *Neighbourhood effects research: New perspectives*, pp. 23–56, Springer, 2012.

[42] C. R. Shalizi, K. L. Shalizi, and R. Haslinger, "Quantifying self-organization with optimal predictors," *Physical Review Letters*, vol. 93, no. 11, p. 118701, 2004.

[43] S. Wolfram, "Statistical mechanics of cellular automata," *Reviews of modern physics*, vol. 55, no. 3, p. 601, 1983.

[44] I. Klimontovich, *Turbulent motion and the structure of chaos: a new approach to the statistical theory of open systems*.

[45] J. Garland, R. James, and E. Bradley, "Model-free quantification of time-series predictability," *Physical Review E*, vol. 90, no. 5, p. 052910, 2014.

[46] M. Batty, R. Morphet, P. Masucci, and K. Stanilov, "Entropy, complexity, and spatial information," *Journal of geographical systems*, vol. 16, no. 4, pp. 363–385, 2014.

[47] P. Grassberger, "Information and complexity measures in dynamical systems," in *Information dynamics*, pp. 15–33, Springer, 1991.

[48] H. C. Fogedby, "On the phase space approach to complexity," *Journal of statistical physics*, vol. 69, no. 1-2, pp. 411–425, 1992.

[49] Y.-C. Zhang, "Complexity and 1/f noise. a phase space approach," *Journal de Physique I*, vol. 1, no. 7, pp. 971–977, 1991.

[50] M. D. Costa and A. L. Goldberger, "Generalized multiscale entropy analysis: Application to quantifying the complex volatility of human heartbeat time series," *Entropy*, vol. 17, no. 3, pp. 1197–1203, 2015.

[51] Y. Zhang, Z. Yang, and W. Li, "Analyses of urban ecosystem based on information entropy," *Ecological Modelling*, vol. 197, no. 1, pp. 1–12, 2006.

[52] V. Ceccato and A. Karlström, "A new information theoretical measure of global and local spatial association," *European Regional Science Association, Barcelona*, vol. 29, 2000.

[53] O. D. Duncan and B. Duncan, "A methodological analysis of segregation indexes," *American sociological review*, pp. 210–217, 1955.

[54] J. Jahn, C. F. Schmid, and C. Schrag, "The measurement of ecological segregation," *American Sociological Review*, pp. 293–303, 1947.

[55] J. J. Williams, "Another commentary on so-called segregation indices," *American Sociological Review*, pp. 298–303, 1948.

[56] W. Bell, "A probability model for the measurement of ecological segregation," *Social Forces*, pp. 357–364, 1954.

[57] K. E. Taeuber and A. F. Taeuber, *Negroes in cities: Residential segregation and neighborhood change*. Atheneum, 1969.

[58] M. J. White, "The measurement of spatial segregation," *American journal of sociology*, pp. 1008–1018, 1983.

[59] B. S. Morgan, "An alternate approach to the development of a distance-based measure of racial segregation.," *American Journal of Sociology*, pp. 1237–1249, 1983.

[60] L. B. Stearns and J. R. Logan, "Measuring trends in segregation three dimensions, three measures," *Urban Affairs Review*, vol. 22, no. 1, pp. 124–150, 1986.

[61] S. Lieberson and D. K. Carter, "Temporal changes and urban differences in residential segregation: a reconsideration," *American Journal of Sociology*, pp. 296–310, 1982.

[62] D. S. Massey and N. A. Denton, "The dimensions of residential segregation," *Social forces*, vol. 67, no. 2, pp. 281–315, 1988.

[63] D. W. Wong, "Spatial dependency of segregation indices," *The Canadian Geographer/Le Géographe canadien*, vol. 41, no. 2, pp. 128–136, 1997.

[64] A. J. F. Henri Theil, "A note on the measurement of racial integration of schools by means of informational concepts," *The Journal of Mathematical Sociology*, vol. 1:2, pp. 187–193, 1971. http://dx.doi.org/10.1080/0022250X.1971.9989795.

[65] M. Batty, "Spatial entropy," *Geographical analysis*, vol. 6, no. 1, pp. 1–31, 1974.

[66] A. Kleidon and R. D. Lorenz, *Non-equilibrium thermodynamics and the production of entropy: life, earth, and beyond*. Springer Science & Business Media, 2005.

[67] V. Cápek and D. P. Sheehan, *Challenges to the second law of thermodynamics*. Springer, 2005.

[68] M. H. Ambaum, *Thermal physics of the atmosphere*, vol. 1. John Wiley & Sons, 2010.

[69] E. T. Jaynes, "Gibbs vs boltzmann entropies," *American Journal of Physics*, vol. 33, no. 5, pp. 391–398, 1965.

[70] R. H. Swendsen and J.-S. Wang, "The gibbs" volume" entropy is incorrect," *arXiv preprint arXiv:1506.06911*, 2015.

[71] J. L. Lebowitz, "Boltzmann's entropy and time's arrow," *Physics today*, vol. 46, pp. 32–32, 1993.

[72] W. Ebeling and I. M. Sokolov, *Statistical thermodynamics and stochastic theory of nonequilibrium systems*, vol. 8. World Scientific, 2005.

[73] R. Frigg and C. Werndl, "Entropy-a guide for the perplexed," *Probabilities in physics*, 2011.

[74] A. Giffin, *Maximum entropy: the universal method for inference*. ProQuest, 2008.

[75] D. Stowell and M. D. Plumbley, "Fast multidimensional entropy estimation by k-d partitioning," *IEEE Signal Processing Letters*, vol. 16, pp. 537–540, Jun 2009.

[76] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of complex physiologic time series," *Physical review letters*, vol. 89, no. 6, p. 068102, 2002.

[77] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale entropy analysis of biological signals," *Physical review E*, vol. 71, p. 021906, 2005.

[78] A. Humeau-Heurtier, "The multiscale entropy algorithm and its variants: A review," *Entropy*, vol. 17, no. 5, p. 3110, 2015.

[79] E. G. Miller, "A new class of entropy estimators for multi-dimensional densities," in *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, vol. 3, pp. III–297, IEEE, 2003.

[80] M. Hazewinkel, *Encyclopaedia of Mathematics: Coproduct Hausdorff Young Inequalities*. Encyclopaedia of Mathematics, Springer US, 2013.

[81] T. Bartz-Beielstein, M. Chiarandini, L. Paquete, and M. Preuss, *Experimental methods for the analysis of optimization algorithms*. Springer, 2010.

[82] J. Michalowicz, J. Nichols, and F. Bucholtz, *Handbook of Differential Entropy*. CRC Press, 2013.

[83] T. Blumentritt, *On Copula Density Estimation and Measures of Multivariate Association*. Reihe Quantitative Ökonomie, Eul, 2012.

[84] K. Stanilov and M. Batty, "Exploring the historical determinants of urban growth patterns through cellular automata," *Transactions in GIS*, vol. 15, no. 3, pp. 253–271, 2011.

[85] J. R. Anderson, *A land use and land cover classification system for use with remote sensor data*, vol. 964. US Government Printing Office, 1976.

[86] A. Harrison, "National land use database: land use and land cover classification," *O. Office of the Deputy Prime Minister, Ed., ed. London, UK: Queen's Printer and Controller of Her Majesty's Stationery Office*, p. 81, 2006.

[87] R Development Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0.

[88] RStudio Team, *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2015.

[89] R. Bivand, T. Keitt, and B. Rowlingson, *rgdal: Bindings for the Geospatial Data Abstraction Library*, 2014. R package version 0.9-1.

[90] B. B. Mandelbrot, *The fractal geometry of nature*, vol. 173. Macmillan, 1983.

[91] P. Addison, *Fractals and Chaos: An illustrated course*. Taylor & Francis, 1997.

[92] J. Holston, *The Modernist City: An Anthropological Critique of Brasilia*. University of Chicago Press, 1989.

[93] J. Speck, *Walkable City: How Downtown Can Save America, One Step at a Time*. Farrar, Straus and Giroux, 2012.

Chapter 9

Appendix

9.1 Code

Execution:

```
L1875<-readOGR("all_LU.shp", layer="lu__1875_all")
L1895<-readOGR("all_LU.shp", layer="lu_1895_all")
L1915<-readOGR("all_LU.shp", layer="lu_1915_all")
L1935<-readOGR("all_LU.shp", layer="lu_1935_all")
L1960<-readOGR("all_LU.shp", layer="lu_1960_all")
L1985<-readOGR("all_LU.shp", layer="lu_1985_all")
L2005<-readOGR("all_LU.shp", layer="lu_2005_all")

L1875<-spTransform(L1875, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L1895<-spTransform(L1895, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L1915<-spTransform(L1915, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L1935<-spTransform(L1935, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L1960<-spTransform(L1960, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L1985<-spTransform(L1985, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
L2005<-spTransform(L2005, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified

catcols<-colorRampPalette(c(rgb(0,0,0),rgb(0,1,0), rgb(0,0,1)), alpha = TRUE)(length(levels(L2005@data$LU)))
names(catcols)<-levels(L2005@data$LU)
years<-c("1875","1895","1915","1935","1960","1985","2005")
yearlty<-c(1,1,1,1,1,1,1)
yearcols<-colorRampPalette(c(rgb(0,0,0),rgb(1,0,0), rgb(0,1,0)), alpha = TRUE)(7)
spdfNames<-paste("L",years,sep="")
sdist<-300
radsNow<-c(0,300,400,600,800,1000,1500,2000,2500)

# original data
allTheYears<-array(dim=c(length(radsNow)-1,4,7))
for(thisLayer in c(1:length(spdfNames))){
  allTheYears[,,thisLayer]<-
    as.matrix(doItAll(spdf =get(spdfNames[thisLayer]),rads = radsNow,stepDistance = sdist),length(radsNow)-1,3)
}
#ordered distributions
slices<-c(1,3,5,9)
orderedTests<-array(dim=c(length(radsNow)-1,3,7,length(strips)))
for(orderSlice in c(1:length(slices))){
print(paste("slice",orderSlice,"/",length(strips)))
for(thisLayer in c(1:length(spdfNames))){
print(paste("year",thisLayer))
orderedTests[,,thisLayer,orderSlice]<-as.matrix(doItAll(spdf =get(spdfNames[thisLayer]),rads = radsNow,stepDistance = sdist,orderByCategories =
        slices[orderSlice]),length(radsNow)-1,3)
}
}
#random distributions (performs 20 in ca. 8 hours)
randomTests<-array(dim=c(length(radsNow)-1,3,7,20))
for(i in c(1:length(randomTests[1,1,1,]))){
for(thisLayer in c(1:length(spdfNames))){
  print(paste("test Number:",i,"- year",thisLayer))
  randomTests[,,thisLayer,i]<-as.matrix(doItAll(spdf =get(spdfNames[thisLayer]),rads = radsNow,stepDistance = sdist,randomise = T),length(radsNow)
        -1,3)
}
}
```

Functions:

```
 library("sp")
 library("rgdal")
 library("maptools")
 library("ggplot2")
 library("UScensus2010")
 library("geospacom")
 library("pdist")
 library("mgcv")
 library("abind")
areaInRadius<-function(spdf,mpoint,radius,category=NULL,cumulative=T,divideByDiskArea){

  divideByDiskArea<- F
  cumulative <- T

  if(cumulative){
    radAreaTotal<-(pi*(radius[2]^2))
  }else{radAreaTotal<-(pi*(radius[2]^2))-(pi*(radius[1]^2))}

  inRad<-inRadius(spdf,mpoint,radius,category,cumulative = cumulative)

  return(sum(
```

```
    round(spdf@data$area[inRad])
    # round sets A_min to (1m)^D
    ))
}


inRadius<-function(spdf=london, mpoint=46, radius=c(0000,2000), category=NULL, cumulative=T){
  if (length(category)==0){
    category<-levels(spdf@data$LU)
  }
if(!cumulative){
    inradius<-which(spdf@data$distTomPoints[,mpoint]>=radius[1] #distance larger than min radius
                    & spdf@data$distTomPoints[,mpoint]<radius[2]#distance smaller than max radius
                    & spdf@data$LU %in% category
    )

  }else{
    inradius<-which(spdf@data$distTomPoints[,mpoint]<radius[2]#distance smaller than max radius
                    & spdf@data$LU %in% category
    )
  }


  return(inradius)
}


 plotByCat<-function(spdf, whichToPlot=c(1:length(spdf))){

 plot(spdf[whichToPlot,],  col=catcols[as.character(spdf[whichToPlot,]@data$LU)], border=NA)
  par(xpd=TRUE)

}

 pointRep<-function(spdf,  mpoint, rads, cats, cumulative=T){
  rep<-array(dim = c(length(cats), length(rads)-1))
  colnames(rep)<-rads[1:length(rads)-1]
  rownames(rep)<-cats
  for(thisCat in c(1:length(cats))){
    for(thisRad in  c(1:length(rads)-1)){
      rep[thisCat, thisRad]<-areaInRadius(spdf = spdf, mpoint = mpoint, radius = rads[thisRad:(thisRad+1)], category = cats[thisCat], cumulative =
            cumulative, divideByDiskArea = F)
    }
  }
  return(rep)

}

groupCategories<-function(spdf){
  spdf@data$LUoriginal<-spdf@data$LU
  spdf@data$LU<-vector(mode = "character", length = length(spdf@data$LUoriginal) )
  spdf@data$LU[]<-"other"
  #working
  spdf@data$LU[which(spdf@data$LUoriginal== "INS"|
                          spdf@data$LUoriginal== "INSL"|
                          spdf@data$LUoriginal== "O"|
                          spdf@data$LUoriginal== "GAR"|
                          spdf@data$LUoriginal== "IND"
  )]<-"working"
  #living
  spdf@data$LU[which(spdf@data$LUoriginal== "APT"|
                          spdf@data$LUoriginal== "APTH"|
                          spdf@data$LUoriginal== "COT"|
                          spdf@data$LUoriginal== "DET"|
                          spdf@data$LUoriginal== "DETH"|
                          spdf@data$LUoriginal== "MEW"|
                          spdf@data$LUoriginal== "SDT"|
                          spdf@data$LUoriginal== "TER"|
                          spdf@data$LUoriginal== "LDG"
  )]<-"living"

  #buying
  spdf@data$LU[which(spdf@data$LUoriginal== "GEN"|
                          spdf@data$LUoriginal== "RET"|
                          spdf@data$LUoriginal== "OLD"
  )]<-"buying"


  #nature
  spdf@data$LU[which(spdf@data$LUoriginal== "AGR"|
                          spdf@data$LUoriginal== "CEM"|
                          spdf@data$LUoriginal== "EST"|
                          spdf@data$LUoriginal== "FRM"|
                          spdf@data$LUoriginal== "NRS"|
                          spdf@data$LUoriginal== "PRK"|
                          spdf@data$LUoriginal== "REC"|
                          spdf@data$LUoriginal== "WAT"
  )]<-"nature"

  #culture
  spdf@data$LU[which(spdf@data$LUoriginal== "CHR"|
                          spdf@data$LUoriginal== "SCH"|
                          spdf@data$LUoriginal== "STA"
  )]<-"culture"

  #infrastructure
  spdf@data$LU[which(spdf@data$LUoriginal== "UTL"|
                          spdf@data$LUoriginal== "AIR"|
                          spdf@data$LUoriginal== "RRS"
  )]<-"infrastructure"




  spdf@data$LU[which(spdf@data$LU== "nature"|
                          spdf@data$LU== "buying"|
                          spdf@data$LU== "culture"

  )]<-"leisure"


  #empty
  spdf@data$LU[which(spdf@data$LUoriginal== "CLR"
  )]<-"empty"
  spdf<-spdf[which(spdf@data$LU== "leisure" | spdf@data$LU== "living" | spdf@data$LU== "working"),]
```

```
  spdf@data$LU<-as.factor(spdf@data$LU)
  spdf@data$LU<-droplevels(spdf@data$LU)
  return(spdf)
}

reptovecs<-function(rep){
  transposed <- aperm(rep, c(2,3,1))
  vecs<-matrix(0,0,length(transposed[1,,1]))
  for(i in c(1:length(transposed[1,1,]))){
    vecs<-rbind(vecs,transposed[,,i])
  }
  vecs
  return(vecs)
}

prepShapeFile<-function(fileName="all_LU.shp",layerName="",randomise,orderByCategories=0,shapeFile="No Shapefile!"){
  if(is.character(shapeFile)){

    if(layerName==""){
    print("which layer? - run again with layerName attribute")
    print(ogrListLayers(fileName))
    return(0)
  }else{
    #load osm shapefile
    #convert to UTM
    london=readOGR("all_LU.shp", layer=layerName) #will load the shapefile to your dataset.
    london <- spTransform(london, CRS("+proj=utm +zone=30 ellps=WGS84")) # the utm  zone must be specified
  }}else{london<-shapeFile}

  london<-groupCategories(london)
  london@data$area<-areaPoly(london)
  london@data$centroid<-coordinates(london)
  if(randomise){
      london<-randomiseLocs(london)
    }
  if(orderByCategories>0){
    london<-orderLocs(london,orderByCategories = orderByCategories)
  }
  london@data$centroid<-coordinates(london)

    return(london)
  }

makePointGrid<-function(spdf,stepDistance=200){
  minmaxX<-c(min(spdf@data$centroid[,1]),max(spdf@data$centroid[,1]))
  minmaxY<-c(min(spdf@data$centroid[,2]),max(spdf@data$centroid[,2]))
  xvals<-seq(from = minmaxX[1],to = minmaxX[2],by = stepDistance)
  yvals<-seq(from = minmaxY[1],to = minmaxY[2],by = stepDistance)
  mPoints<-as.matrix(expand.grid(xvals,yvals))

  print(paste(nrow(mPoints), "measurement points in", stepDistance,"m intervals"))
  return(mPoints)
}

subsetInField<-function(allreps,mPoints, radius,minmaxX,minmaxY){
  infield<-which(
    (mPoints[,1] - radius)>minmaxX[1] &
      (mPoints[,1]+radius)<minmaxX[2] &
      (mPoints[,2] - radius)>minmaxY[1] &
      (mPoints[,2]+radius)<minmaxY[2]

  )

  sub<-allreps[,,infield]
  return(allreps)
}

KDEmultirad<-function(allreps,mPoints,rads,minmaxX,minmaxY){
  kdes<-vector()
  for(thisRad in c(1:(length(rads)-1))){
    print(paste("radius",thisRad, "/", length(rads-1)))
    sub<-subsetInField(allreps = allreps,mPoints = mPoints,radius = rads[thisRad+1],minmaxX = minmaxX,minmaxY = minmaxY)
    sub<-sub[,thisRad,]
    sub<-aperm(sub, c(2,1))
    thisKDE<-KDE(sub,plot = F,removeNA = T,makeRep = F,radiusForNorm = rads[thisRad+1])$H_kd
    kdes<-rbind(kdes,c(rads[thisRad+1],thisKDE))
  }
  return(kdes)
}

shuffleCategories<-function(spdf){
  spdf@data$LU<-sample(x = spdf@data$LU,size = length(spdf@data$LU),replace = F)
  return(spdf)
}
doItAll<-function(layerName="",spdf=NA, randomiseCats=FALSE,rads,stepDistance,randomise=F,orderByCategories=F){
  london<-prepShapeFile(layerName = layerName,randomise = randomise,orderByCategories = orderByCategories, shapeFile = spdf)
  mPoints<-makePointGrid(spdf = london,stepDistance = stepDistance)
  minmaxX<-c(min(london@data$centroid[,1]),max(london@data$centroid[,1]))
  minmaxY<-c(min(london@data$centroid[,2]),max(london@data$centroid[,2]))
    print("make distance matrix...")
  london@data$distTomPoints<- as.matrix(pdist(london@data$centroid,mPoints))
  allreps<-array(dim = c(length(levels(london@data$LU)),length(rads)-1,nrow(mPoints)))
  print("making representations - this might take a few minutes")
  for(i in c(1:nrow(mPoints))){
    allreps[,,i]<-pointRep(london,mpoint = i,rads = rads,cumulative = T,cats = levels(london@data$LU))
  }
  return(KDEmultirad(allreps = allreps,mPoints=mPoints,rads = rads,minmaxX = minmaxX,minmaxY=minmaxY))
}

whichRow<-function(mat,vec){
  return(which(apply(mat,1,function(x) identical(as.vector(x),vec))))
}

addKdTreeLevelND<-function(subspaces,level){
for (i in c(1:ncol(subspaces$data))){
subspaces<-addKdTreeLevel_single(subspaces,level = subspaces$level,dim = i)
subspaces$level<-subspaces$level+1
}
return(subspaces)
}

H_kd<-function(n,A,phaseSpaceVolume=NA,radiusForNorm,dims){
  #n: list with number of observations in each subhyperspace
  #A: hypervolume of each subhyperspace
  N<-sum(n)#number of observations
  m<-length(n)#number of subhyperspaces
```

```
  AfullPhasespace<-(pi*radiusForNorm^2)^dims
  A_min= 1
  H<-data.frame(relative=0,notRelative=0,newRelative=0)
  notCounted<-0
  for(j in c(1:m)){
    if(A[j]!=0 & n[j]!=0){
   # H$relative and H$newRelative were normalisations depending on the full phase space extend at each radius, but eventually were omitted due to
        theoretical inaccuracies.
    H$notRelative<- H$notRelative + (((n[j]/N))*log((N*A[j]/A_min)/n[j]))
    H$relative<- H$relative + ((n[j]/N)*log(n[j]*AfullPhasespace/(A[j]/A_min)/N))
    H$newRelative<-H$newRelative + (1/AfullPhasespace)*log(N/n[j]*(A[j]/A_min)/AfullPhasespace)
    }
    else{notCounted<-notCounted+1} # was used for debugging

  }

  return(c(H$notRelative,H$relative,H$newRelative))
}

addKdTreeLevel_single<-function(subspaces,dim,level){


  #get unique tree vectors describing each top level group
  uniques<-unique(subspaces$tree)


  for(i in c(1:length(uniques))){
    #for each group (subspace)..

    #..which rows belong to this group?
    #print(paste(i,"/",length(uniques)))
    thisgroupRows<-which(subspaces$tree==uniques[i])
    #get splitting value: median of all values in this group
    q<-median(subspaces$data[thisgroupRows,dim])

    if(is.na(subspaces$splits[[dim]][1])){
    subspaces$splits[[dim]]<-q   #get row indices that are in this group AND smaller than median
    }else{
      subspaces$splits[[dim]]<-c(subspaces$splits[[dim]],q)   #get row indices that are in this group AND smaller than median
    }

    lowersplit<-intersect(thisgroupRows,which(subspaces$data[,dim]<=q))
    uppersplit<-intersect(thisgroupRows,which(subspaces$data[,dim]>q))

    #change value of these
    subspaces$tree[lowersplit]<-paste(subspaces$tree[lowersplit],"0",sep = "")
    subspaces$tree[uppersplit]<-paste(subspaces$tree[uppersplit],"1")

    subspaces$highlimit[lowersplit,dim]<-q
    subspaces$lowlimit[uppersplit,dim]<-q

  }

subspaces<-getVandas(subspaces = subspaces)
#print("in kd tree single")
#print(paste("inkdtreesingle",head(subspaces$groups$vol)))
return(subspaces)
}

fixNaLimits<-function(subspaces){

for(dim in c(1:ncol(subspaces$data))){
#HIGH LIMITS:
#get the group ids that have an open upper limit
NAgroups<-unique(subspaces$tree[which(is.na(subspaces$highlimit[,dim]))])
#replace highlimit NA with max data in group of dim
if(length(NAgroups)!=0){
for(g in c(1:length(NAgroups))){
    subspaces$highlimit[which(subspaces$tree==NAgroups[g]),dim]<-max(subspaces$data[which(subspaces$tree==NAgroups[g]),dim])
  #print(paste("high; dim",dim,"group",g,"newhigh",min(subspaces$data[which(subspaces$tree==NAgroups[g]),dim])))
}}

#LOW LIMITS:
#get the group ids that have an open lower limit
NAgroups<-unique(subspaces$tree[which(is.na(subspaces$lowlimit[,dim]))])
#replace lowlimit NA with min data in group of dim
if(length(NAgroups)!=0){
for(g in c(1:length(NAgroups))){
  subspaces$lowlimit[which(subspaces$tree==NAgroups[g]),dim]<-min(subspaces$data[which(subspaces$tree==NAgroups[g]),dim])
  #print(paste("low; dim",dim,"group",g,"newlow",min(subspaces$data[which(subspaces$tree==NAgroups[g]),dim])))
}
}
}

      return(subspaces)

}

getVandas<-function(subspaces,fixNA=F){

  A_min<-1

  if(fixNA){

    subspaces<-fixNaLimits(subspaces)
  }

  diffs<-abs(subspaces$highlimit-subspaces$lowlimit)
  if(fixNA){diffs[which(diffs==0)]<-A_min}
  volumes<-apply(diffs, 1, prod)
  subspaces$volumes<-volumes
  uniquevols<-data.frame(subspaces$tree,subspaces$volumes)[which(!duplicated(subspaces$tree)),]
  counts<-data.frame(attributes(table(subspaces$tree))$dimnames[[1]],as.numeric(table(subspaces$tree)))
  vanda<-merge(counts,uniquevols,by.x=1,by.y=1)
  subspaces$groups$id<-vanda[,1]
  subspaces$groups$count<-vanda[,2]
  subspaces$groups$vol<-vanda[,3]
  return(subspaces)
}

subspaces.2Dplot<-function(subspaces,grid=T,allBlack=F){
  rgb.palette <- colorRampPalette(c("green","blue" ,"red"), space = "rgb")
  cols <- rgb.palette(length(unique(subspaces$tree)))
  cols<-sample(cols)

  if(allBlack){
```

```
  cols[]<-"grey"
  frame.plot=F
  }

  plot(subspaces$data[,1],subspaces$data[,2],col="grey",cex=.1,xaxt='n',yaxt='n', ann=FALSE)
  for(i in c(1:length(unique(subspaces$tree)))){
    this<-which(subspaces$tree==unique(subspaces$tree)[i])
    points(subspaces$data[this,1],subspaces$data[this,2],cex=0.3,pch=20,col=cols[i])
    #print(i)
  }
  if(grid){

    subspaces$lowlimit[which(is.na(subspaces$lowlimit))]<- -10000000
    subspaces$highlimit[which(is.na(subspaces$highlimit))]<- 10000000

    #abline(h=subspaces$splits[[2]],col="red")
    #abline(v=subspaces$splits[[1]],col="blue")
    #xlow to xhigh at ylow
    segments(x0 = subspaces$lowlimit[,1],y0 = subspaces$lowlimit[,2],x1 = subspaces$highlimit[,1],y1 =subspaces$lowlimit[,2])
    #xlow to xhigh at yhigh
    segments(x0 = subspaces$lowlimit[,1],y0 = subspaces$highlimit[,2],x1 = subspaces$highlimit[,1],y1 =subspaces$highlimit[,2])
    #ylow to yhigh at xlow
    segments(x0 = subspaces$lowlimit[,1],y0 = subspaces$lowlimit[,2],x1 = subspaces$lowlimit[,1],y1 =subspaces$highlimit[,2])
    #ylow to yhigh at xhigh
    segments(x0 = subspaces$highlimit[,1],y0 = subspaces$lowlimit[,2],x1 = subspaces$highlimit[,1],y1 =subspaces$highlimit[,2])

  }
}

makeSubspacesDataFrame<-function(data){
  subspaces<-list()
  subspaces$splits<-list()
  subspaces$data<-data
  subspaces$tree<-array("",dim=nrow(data))
  subspaces$splits[1:50]<-array()
  subspaces$highlimit<-array(NA,c(nrow(subspaces$data),ncol(subspaces$data)))
  subspaces$lowlimit<-array(NA,c(nrow(subspaces$data),ncol(subspaces$data)))
  subspaces$howManyNa<-vector()
  return(subspaces)
}

needAnotherLevel<-function(subspaces){


minimumLevelsPerDataPointReached<-subspaces$level >0.5*log2(nrow(subspaces$data))

if(subspaces$level >1){
 meds<-array()
 maxs<-array()
 mins<-array()
 for(dimension in c(1:ncol(subspaces$data))){
 for(i in c(1:length(subspaces$groups$id))){
 meds<-c(meds,median(subspaces$data[which(subspaces$tree==subspaces$groups$id[1]),dimension]))
 maxs<-c(maxs,max(subspaces$data[which(subspaces$tree==subspaces$groups$id[i]),dimension]))
 mins<-c(mins,min(subspaces$data[which(subspaces$tree==subspaces$groups$id[i]),dimension]))
 }
 uniformities<-abs(sqrt(length(which(subspaces$tree==subspaces$groups$id[i])))*((2*meds)-mins-maxs)/(maxs-mins))
 uniformities<-uniformities[!is.na(uniformities)]
 isUniformEnough<-uniformities >1.96
 isUniformEnough<-isUniformEnough[!is.na(isUniformEnough)]
 isUniformEnough<-all(isUniformEnough)
 }


enoughNoNAs <- (length(which(is.na(subspaces$groups$vol)))/length(subspaces$groups$vol))<0.1
singleObsSubspaces<-length(which(subspaces$groups$count==1))/length(subspaces$groups$count)
howmanyna<-length(which(is.na(subspaces$groups$vol)))/length(subspaces$groups$vol)
if(subspaces$level >40){
  #print(paste("More than fourty levels. this should not happen."))
  subspaces$howManyNa<-c(subspaces$howManyNa,howmanyna)
  return(FALSE)
}

if(singleObsSubspaces >0.5){
  #prevent wasting computation time by trying to split too many subspaces that already have only a single observation inside.
  return(FALSE)
}
}else{isUniformEnough<-FALSE
      enoughNoNAs<-FALSE}
if(enoughNoNAs & minimumLevelsPerDataPointReached & isUniformEnough){
return(FALSE)
}else{
return(TRUE)
}
}
KDE<- function(data,plot=T,removeNA=T,makeRep=F,radiusForNorm){
  A_min<-1

  subspaces<-makeSubspacesDataFrame(data)
  subspaces$level<-0
  subspaces<-getVandas(subspaces = subspaces,fixNA = F)
  while(needAnotherLevel(subspaces)){#split until average ov sqrt(N) data points in each partition
    subspaces<-addKdTreeLevelND(subspaces,subspaces$level)

  }

  #plot partitioning if plot==true (default false):
  if(plot){
    if(ncol(subspaces$data)==2){
    subspaces.2Dplot(subspaces,T)
    }else{
      #print(paste("Plotting ",ncol(subspaces$data)," Dimensional data is not my thing sorry"))
    }
  }

  subspaces<-getVandas(subspaces,fixNA = T)
  subspaces$groups$vol[which(subspaces$groups$vol<1 & subspaces$groups$n>0)]<-A_min

if(removeNA){
  n<-subspaces$groups$count[which(!is.na(subspaces$groups$vol))]
  A<-subspaces$groups$vol[which(!is.na(subspaces$groups$vol))]
  }
  subspaces$H_kd<-H_kd(n = n,A = A,phaseSpaceVolume = 1,radiusForNorm,dims= ncol(subspaces$data)) # phasespace volume is (max(d1)-min(d1)) * (max(
        d2)-min(d2)))... here all dimensions from 0 to 1 -> 1^d = 1

  return(subspaces)
```

```r
}

integrateScales<-function(hkd){
area<-0.0
for(i in c(1:nrow(hkd)-1)){
  area <- sum(area,((hkd[i+1,1]-hkd[i,1])*(hkd[i,2]+hkd[i+1,2])/2))
}
return(area)
}


randomiseLocs<-function(spdf){
   minmaxX<-c(min(coordinates(spdf)[,1]),max(coordinates(spdf)[,1]))
   minmaxY<-c(min(coordinates(spdf)[,2]),max(coordinates(spdf)[,2]))
   newLocations<-cbind(runif(n = length(spdf),min = minmaxX[1],max=minmaxX[2]),
                       runif(n = length(spdf),min = minmaxY[1],max=minmaxY[2]))
   locShift<-newLocations-coordinates(spdf)
   for(i in c(1:length(spdf))){
     cords<-slot(slot(spdf@polygons[[i]],"Polygons")[[1]],"coords")
     cords<-sweep(cords,2,locShift[i,],"+")
     slot(spdf@polygons[[i]],"labpt")<-newLocations[i,]
     slot(slot(spdf@polygons[[i]],"Polygons")[[1]],"coords")<-cords
   }
   for(i in c(1:length(levels(spdf@data$LU)))){
    spdf@data$area[which(spdf@data$LU==levels(spdf@data$LU)[i])]<-mean(spdf[which(spdf@data$LU==levels(spdf@data$LU)[i]),]@data$area)
   }
   return(spdf)
 }

orderLocs<-function(spdf,orderByCategories){
  for(i in c(1:length(levels(spdf@data$LU)))){
    spdf@data$area[which(spdf@data$LU==levels(spdf@data$LU)[i])]<-mean(spdf[which(spdf@data$LU==levels(spdf@data$LU)[i]),]@data$area)
  }
  howManySlices<-orderByCategories
  minmaxX<-c(min(coordinates(spdf)[,1]),max(coordinates(spdf)[,1]))
  minmaxY<-c(min(coordinates(spdf)[,2]),max(coordinates(spdf)[,2]))
   cats<-levels(spdf$LU)
   cats<-data.frame(cats, catarea= NA)
   for(thisCat in c(1:nrow(cats))){
     cats[thisCat,"catarea"]=sum(spdf[which(spdf$LU==cats[thisCat,1]),]$area)
   }
   cats$catproportion<-cats[,"catarea"]/sum(cats[,"catarea"])
   sliceProps<-rep(1/howManySlices,howManySlices)
   sliceMinXs<-minmaxX[1]+((minmaxX[2]-minmaxX[1])*(cumsum(c(0,sliceProps[-1]))))
   sliceMaxXs<-minmaxX[1]+((minmaxX[2]-minmaxX[1])*(cumsum(c(sliceProps))))
   for(slice in c(1:howManySlices)){
   cats$minX<-sliceMinXs[slice]+((sliceMaxXs[slice]-sliceMinXs[slice])*(cumsum(c(0,cats$catproportion[-length(cats$catproportion)]))))
   cats$maxX<-sliceMinXs[slice]+((sliceMaxXs[slice]-sliceMinXs[slice])*cumsum(cats$catproportion))
   for(thisCat in c(1:nrow(cats))){
    whichBuildingsInThisCat<-which(spdf$LU==cats[thisCat,1])
    howMany<-length(whichBuildingsInThisCat)
    whichInThisSlicesMin<-floor(howMany*(sum(c(0,sliceProps[1:slice])[-(length(sliceProps[1:slice])+1)])))#from zero to slice minus the last one
    whichInThisSlicesMax<-floor(howMany*(sum(sliceProps[1:slice])))#from first to slice
    whichBuildingsInThisCat<-whichBuildingsInThisCat[whichInThisSlicesMin:whichInThisSlicesMax]
    newLocations<-cbind(
        runif(
         n = length(whichBuildingsInThisCat),
         min = cats$minX[thisCat],
         max=cats$maxX[thisCat]),
        runif(n = length(whichBuildingsInThisCat),min = minmaxY[1],max=minmaxY[2]))
     locShift<-newLocations-coordinates(spdf[whichBuildingsInThisCat,])
     for(i in c(1:length(whichBuildingsInThisCat))){
       cords<-slot(slot(spdf@polygons[[whichBuildingsInThisCat[i]]],"Polygons")[[1]],"coords")
       cords<-sweep(cords,2,locShift[i,],"+")
       slot(spdf@polygons[[whichBuildingsInThisCat[i]]],"labpt")<-newLocations[i,]
       slot(slot(spdf@polygons[[whichBuildingsInThisCat[i]]],"Polygons")[[1]],"coords")<-cords
     }
  }
   }
  return(spdf)
 }

rarea<-function(r){
   return(pi*r^2)
 }
```