A Persistent Homology Model of Street Network Connectivity

Padraig Corcoran & Christopher B. Jones
School of Computer Science & Informatics
Cardiff University
Wales, UK.

August 20, 2021

Abstract

We propose a novel model of street network connectivity which uses a method from the field of applied topology entitled persistent homology. The output from this model is a pair of density functions which model the relative strength and frequency of connected components and cycles in the network. In this context, strength is a function of street type, such as motorway or residential, with more significant street types providing greater connectivity. The pair of density functions output from the model are easily interpreted and provide novel insights into the connectivity properties of different street networks. We demonstrate the usefulness of this model through an analysis of UK and USA city street networks. This analysis identifies tangible similarities and differences in the connectivity of different cities plus ways in which the connectivity of individual cities' might be improved.

1 Introduction

Cities are important phenomena in our society. Currently more than half of the world's population live in urban areas with this percentage projected to grow to two-thirds by 2050 (Ritchie, 2018). Furthermore, there is a strong positive correlation between a country's degree of urbanization and Gross Domestic Product (GDP) per capita (Henderson, 2003). A major activity in any city is the transportation of people and goods. The efficiency of this activity is strongly influenced by the quality of the underlying street network. As such, modelling the quality of a city's street network represents an important research problem where the outputs from such models are commonly used to inform city planning (Labi et al., 2019). In practice the quality of a given street network is difficult to formally define let alone measure. However, one important feature of any high quality street network is a high level of connectivity. Tal and Handy (2012) define street network connectivity as a measure of the quantity of the connections in the network and in turn the directness and multiplicity of routes within the network. It has been demonstrated both theoretically and empirically that street networks with a greater degree of connectivity exhibit greater transportation efficiency (Knight and Marshall, 2015).

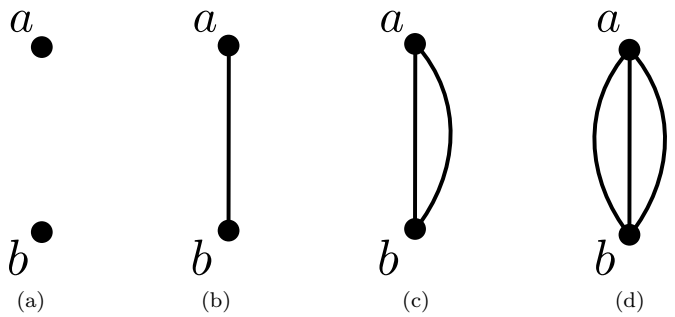


Figure 1: A network containing two locations a and b is displayed in (a). One, two and three connections are added between a and b in (b), (c) and (d) respectively.

Existing models of street network connectivity are mainly based on summary statistics of simple geometrical or topological features of the network. For example, the mean number of intersections per square mile/kilometer. In this work we propose a novel model of street network connectivity where connectivity is modelled in terms of connected components and cycles in the network. These features in turn model the existence and multiplicity of connections between locations respectively. To illustrate this consider the simple network displayed in Figure 1(a) which contains two locations a and b. The two locations form two distinct connected components and therefore the connectivity between the locations is poor. If a single edge is added between the locations to form a single connected component, as illustrated in Figure 1(b), the connectivity between the locations is greater than that in the previous figure. If a second edge is added between the locations to form a cycle, as illustrated in Figure 1(c), the connectivity is again greater. Finally, if a third edge is added between the locations to form another cycle, as illustrated in Figure 1(d), the connectivity is again greater. Therefore, if all edges provide equal connectivity, the connectivity of a network can be modelled in terms of the number of connected components and cycles in the network.

However, in the context of a street network all edges do not provide equal connectivity. Instead, the strength of connectivity provided by an edge is a function of its semantics or type, which varies. For example, an edge corresponding to a motorway provides greater connectivity than an edge corresponding to a secondary street. Therefore, when modelling connectivity in terms of connected components and cycles, it is necessary to consider the types of streets which compose these features. To achieve this goal the model proposed in this work uses a method from the field of applied topology entitled persistent homology (Edelsbrunner and Harer, 2010). The output from this model is a pair of density functions which model the relative frequency and strength of connected components and cycles in the street network. These two density functions are easily interpreted and provide novel insights into connectivity properties of the street network in question.

The layout of this paper is as follows. Section 2 reviews related works on modelling street network connectivity and applied topology. Section 3 describes

the proposed model of street network connectivity. Section 4 demonstrates the usefulness of this model with respect to interpreting and providing insight into the connectivity properties of a given street network, recommending ways to improve this connectivity, and identifying street networks with similar and dissimilar connectivity properties. This analysis is performed with respect to sets of UK and USA city street networks. Finally, section 5 draws conclusions from this work and discusses possible directions for future research.

2 Related Works

In this section we review related works on modelling the connectivity of city street networks. We focus exclusively on the problem of modelling intra city connectivity as opposed to inter city connectivity. A review of works in the latter category can be found in (Mansury and Shin, 2015). In this section we also briefly mention some related works on applied topology.

Connectivity is a well studied concept in graph theory (Gross et al., 2013). Since street networks are naturally modelled as graphs, graph theoretic measures of connectivity can easily be applied to street networks. Two commonly used such measures are vertex- and edge-connectivity which equal the minimum number of vertices and edges respectively whose removal disconnects the graph. Other graph theoretic measures include alpha (α) which equals the number of cycles in the graph divided by the maximum possible number of cycles, beta (β) which equals the number of edges divided by the number of vertices, and gamma (γ) which equals the number of edges divided by the maximum possible number of edges (Weber, 2016). Despite the fact that these measures are commonly applied to street networks (Sahitya and Prasad, 2019), they are limited by the fact that they do not consider the geometrical or spatial nature of street networks. Consequently, many measures of connectivity have been proposed specifically for street networks.

Dill (2004) defined the following set of street network connectivity measures drawn from multiple fields: maximum city block length, maximum city block size, city block density which equals the mean number of city blocks per square mile/kilometer, street intersection density which equals the mean number of intersections per square mile/kilometer, street density which equals the mean total length of street segments per square mile/kilometer, connected node ratio which equals the number of street intersections divided by the number of intersections plus cul-de-sacs, link node ratio which equals the number of street segments divided by the number of intersections and cul-de-sacs, grid pattern which is a binary indicator of whether the street network exhibits a grid pattern commonly associated with high connectivity, route directness which equals the mean network distance divided by straight-line distance and effective walking area which equals the mean percentage of land parcels within a specified network distance of a given location.

The US Green Building Council proposed a street intersection density measure which is related to that defined by Dill (2004) but uses some heuristics when counting the number of intersections (Stangl and Guinn, 2011). Specifically, intersections leading to isolated areas and intersections in these areas are not counted. Stangl and Guinn (2011) subsequently highlighted some limitations with this measure and instead proposed a measure similar to route directness

proposed by Dill (2004). Stangl (2012, 2019) later extended this measure by integrating a number of heuristics to make the measure more robust.

Peponis et al. (2008) proposed two measures of street network connectivity entitled metric reach and directional distance. Metric reach equals the mean network length reachable from a given location in the network. Directional distance equals the mean number of direction changes required to navigate between two locations in the network. In a related work, Ellis et al. (2016) evaluated a number of street connectivity measures. This included a novel measure entitled directional reach which equals metric reach proposed by Peponis et al. (2008) with the addition of a constraint on the number of possible direction changes. Directional reach is related to the concept of visual connectivity whereby locations which require fewer turns to navigate between are considered more connected (Hajrasouliha and Yin, 2015).

Knight and Marshall (2015) evaluated three measures of street network connectivity entitled intersection density, street density and connectivity index. The first two measures are equal to those of the same name proposed by Dill (2004). The connectivity index equals the number of street segments divided by the number of intersections. The authors found all three measures to be correlated with irrelevant features of area and geometry, and therefore not to be reliable measures of connectivity.

Stangl (2015) examined a number of the city block based connectivity measures defined by Dill (2004). The author identified some limitations of these measures and proposed an alternative block based connectivity measure. The measure in question is entitled block section and equals the mean maximum straight-line distance between any two points on the boundary of an area enclosed by streets.

In this article we propose a novel model of street network connectivity based on persistent homology. Feng and Porter (2020) previously used persistent homology to model street networks. However their model does not model connectivity but instead models the shape of the areas enclosed by streets. There exist a number of previous works which also considered the application of methods from the field of applied topology to geographical data. Ahmed et al. (2014) proposed a model for determining local differences between street networks using persistent homology. This model is used to recognize changes in street networks over time and to assess the quality of map construction algorithms. Dey et al. (2017) proposed a model for inferring a street network from GPS data using discrete Morse theory. Corcoran and Jones (2016, 2017, 2018) proposed models of spatial temporal phenomena based on persistent homology. Corcoran (2019) proposed a method for performing generalization of geographical data using persistent homology.

3 Model of Street Network Connectivity

As described in the introduction to this article, we propose a novel model of street network connectivity where connectivity is modelled in terms of the relative strength and frequency of connected components and cycles. This is achieved by applying a method from the field of applied topology entitled persistent homology to the street network in question.

To motivate the use of this method consider the street network in Figure 2

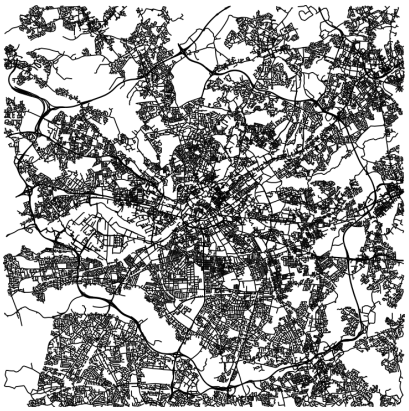


Figure 2: The street network for Manchester city is displayed where streets are represented by black lines.

corresponding to Manchester city which forms a running example in this section. Figures 3(a), 3(b), 3(c), 3(d), 3(e) and 3(f) display subsets of this network containing only those streets with type equal to motorway, trunk, primary, secondary, tertiary and unclassified respectively. These are the six most significant street types in our model ranked from most to least significant. These individual networks have distinct topological features and in turn different levels of connectivity. For example, the motorway network in Figure 3(a) consists of one larger and two smaller connected components where the larger connected component contains a single cycle. This combination of few connected components and a cycle indicates a relatively high level of connectivity. The trunk network in Figure 3(b) contains three connected components where the larger connected component contains many cycles indicating a relatively high level of connectivity. On the other hand, the secondary network in Figure 3(d) contains many connected components with few cycles indicating a relatively low level of connectivity.

The above connectivity analysis of each street type in isolation gives a potentially misleading local perspective of the overall street network connectivity. For example, although the secondary network has a relatively low level of connectivity when considered in isolation, when considered in the context of other street types to which it is connected, it may have a relatively high level of connectivity. That is, the many connected components in this network may be strongly connected by streets of different types. Therefore, it is necessary to perform a more global connectivity analysis which considers all street types jointly. A naive solution would be to compute a simple union of all street types and perform a connectivity analysis of the resulting network. However, this approach fails to model street type which is an important feature of connectivity. For example, despite the fact that the motorway network contains only a single cycle, this single cycle can potentially provide greater connectivity than the large number of cycles in the trunk network. In this work we propose a novel model of street network connectivity which uses persistent homology to overcome this challenge and jointly consider all street types in an appropriate manner.

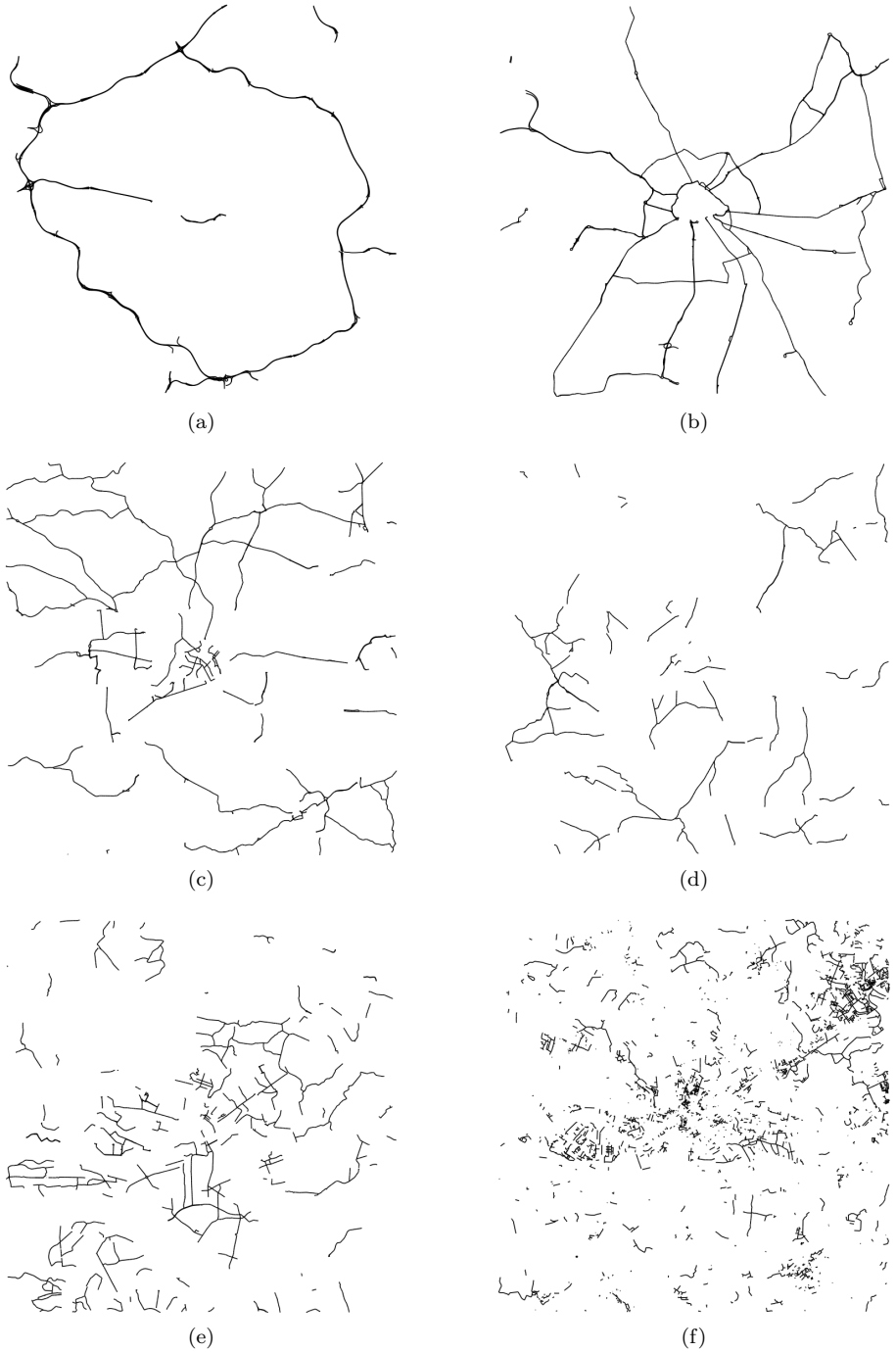


Figure 3: Figures (a), (b), (c), (d), (e) and (f) display subsets of the Manchester street network in Figure 2 containing only those streets with type equal to motorway, trunk, primary, secondary, tertiary and unclassified respectively.

The proposed model contains the following four computational steps. In the first step, a graph based representation of the input street network is constructed. In the second step, a representation of this graph entitled a filtration is constructed where this representation encodes street type information. In the third step, the persistent homology of this filtration is computed which returns a pair of mathematical objects called persistence diagrams. The first persistence diagram models the number and strength of connected components while the second persistence diagram models the number and strength of cycles. In the final step, these persistence diagrams are transformed into a corresponding pair of density functions which can be easily interpreted. This contrasts with existing models of street network connectivity, reviewed in the related works section, which output a single value indicating the level of connectivity. The density functions output from the proposed model provide greater information and in turn insight into the connectivity of the street network in question.

In the following four subsections we describe each of the above four computational steps in more detail. Note that, the descriptions in these subsections do not venture too deeply into the corresponding underlying mathematics. This presentation is intended to make the content accessible to non-mathematicians and more specifically those without a working knowledge of algebraic topology. We advise a reader seeking a more indepth description to consult the corresponding references in each subsection or the textbooks by Edelsbrunner and Harer (2010) and Ghrist (2014).

3.1 Graph Construction

In this step a graph based representation of the input street network is constructed. Toward this goal, the corresponding street network data is obtained from OpenStreetMap (OSM) which is a crowdsourcing project for geographical data (Boeing, 2017). The set of street types in OSM ordered from most to least significant are *motorway*, *motorway_link*, *trunk*, *trunk_link*, *primary*, *primary_link*, *secondary*, *secondary_link*, *tertiary*, *tertiary_link*, *unclassified* and *residential*. This ordering is specified by OSM on the OSM wiki which also contains a short description of each street type [1]. A street type containing link in its title refers to a type of street which has a short length and connects two streets. For the purposes of this work, all streets of such a type are assigned their corresponding parent type. For example, a street of type motorway_link is assigned the type motorway. Given this, we define the set T to be the set of OSM street types less those which contain link in their title.

Given OSM data corresponding to the input street network, we represent the street network as an edge labelled undirected graph $G = (V, E, L : E \to T)$, where V is the set of vertices corresponding to road intersections and dead-ends, E is the set of edges corresponding to street segments connecting these vertices, and L is a mapping from edges to the street type of the corresponding road segment (Corcoran and Mooney, 2013). The graph representation corresponding to the Manchester street network is displayed in Figure 2 where street segments and hence graph edges are represented by black lines. Figures 3(a), 3(b), 3(c), 3(d), 3(e) and 3(f) display subgraphs of this street network containing only those edges with type equal to motorway, trunk, primary, secondary, tertiary

[1] `https://wiki.openstreetmap.org/wiki/Key:highway`

and unclassified respectively.

3.2 Filtration Construction

Given a graph representation $G = (V, E, L : E \to T)$ of the input street network, in this step we construct a filtration of G which is a representation that encodes street type information.

A filtration of G is a sequence of $n+1$ graphs $G_0, G_1, \ldots, G_n$ which satisfy Equation 1 where $\subseteq$ is the subgraph relation. Toward constructing a filtration of graph G, we first define a filter function $f : E \to \mathbb{R}$ in Equation 2. This function maps each edge in G to a real value where edges corresponding to more significant street types are mapped to smaller values.

$$\emptyset = G_0 \subseteq G_1 \subseteq \cdots \subseteq G_n = G \tag{1}$$

$$f(e) = \begin{cases} 1.0 & L(e) = \text{motorway} \\ 2.0 & L(e) = \text{trunk} \\ 3.0 & L(e) = \text{primary} \\ 4.0 & L(e) = \text{secondary} \\ 5.0 & L(e) = \text{tertiary} \\ 6.0 & L(e) = \text{unclassified} \\ 7.0 & L(e) = \text{residential} \end{cases} \tag{2}$$

Let $\mathcal{E}$ denote the set of subsets of E, and $\mathcal{G}$ denote the set of subgraphs of G. Let $S : \mathcal{E} \to \mathcal{G}$ denote the map from a subset of E to the corresponding edge induced subgraph. That is, a graph containing only that subset of edges plus all adjacent vertices. For example, the subgraphs $S(\{e : e \in E, L(e) = \text{motorway}\})$ and $S(\{e : e \in E, L(e) = \text{trunk}\})$ corresponding to the Manchester street network are displayed in Figures 3(a) and 3(b) respectively.

Let $C : \mathbb{R} \to \mathcal{G}$ be the map defined in Equation 3 which returns sub-level sets of G which are subgraphs of G. For example, the subgraphs $C(1)$, $C(2)$, $C(3)$, $C(4)$, $C(5)$ and $C(6)$ corresponding to the Manchester street network are displayed in Figures 4(a), 4(b), 4(c), 4(d), 4(e) and 4(f) respectively. Given the map C, we define a filtration of G to be the sequence of $n+1$ graphs $G_0, G_1, \ldots, G_n$ where $G_i = C(i)$. The first five graphs in this filtration for the Manchester street network are the null graph followed by the subgraphs displayed in Figures 4(a), 4(b), 4(c), 4(d), 4(e) and 4(f) respectively.

$$C(a) = S(\{e : e \in E, f(e) \leq a\}) \tag{3}$$

The above filtration definition has the property that edges corresponding to more significant street types appear earlier and persist for longer in the sequence of graphs. For example, edges corresponding to motorways appear in G_1 and persist for n graphs of the sequence. On the other hand, edges corresponding to primary appear in G_3 and persist for $n-2$ graphs of the sequence.

3.3 Persistent Homology Computation

As stated in the motivation at the beginning of this section, we wish to model street network connectivity in terms of connected components and cycles while

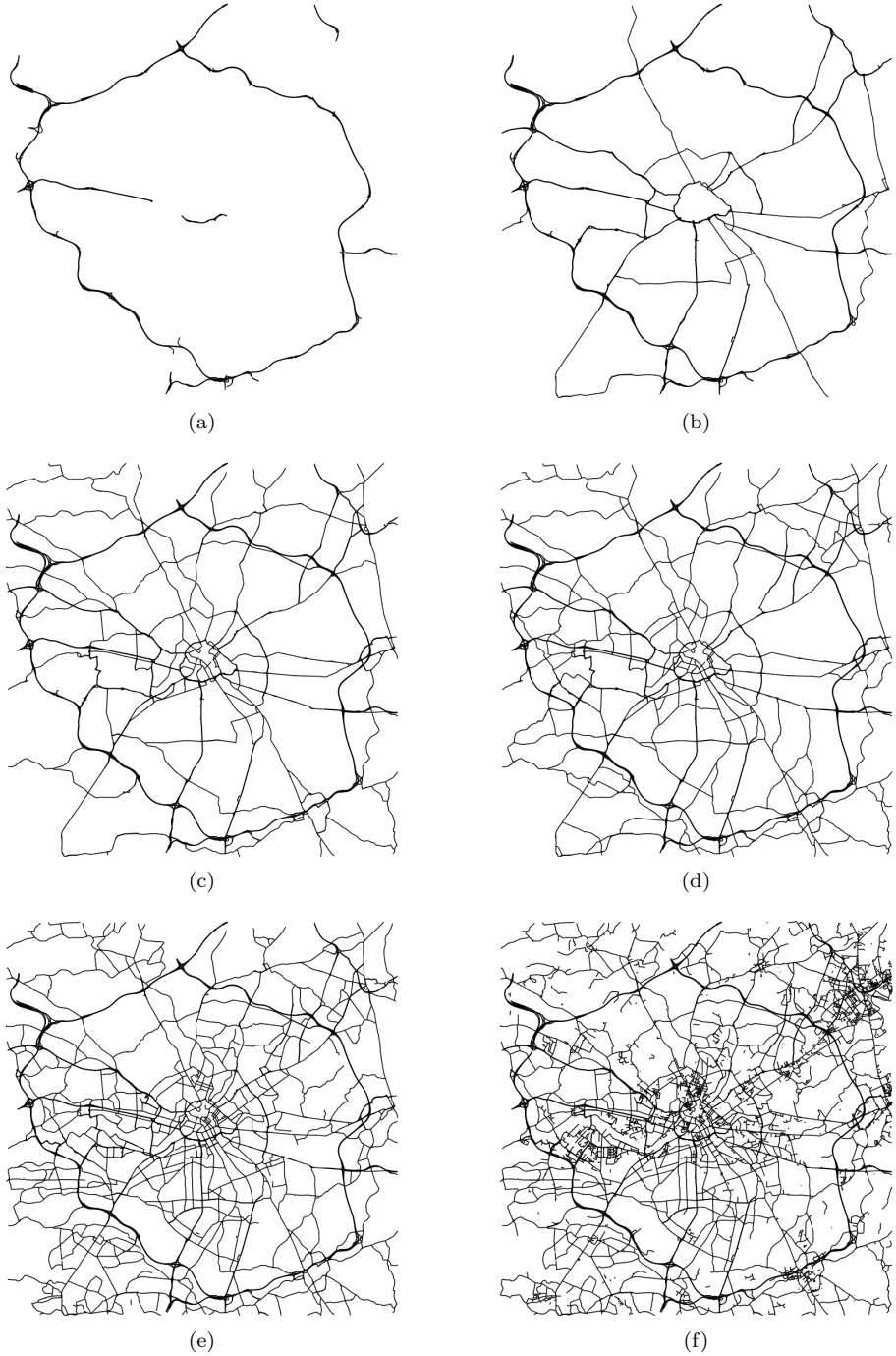


Figure 4: The subgraphs $C(1)$, $C(2)$, $C(3)$, $C(4)$, $C(5)$ and $C(6)$ corresponding to the Manchester street network of Figure 2 are displayed in (a), (b), (c), (d), (e) and (f) respectively.

considering that more significant streets provide a greater level of connectivity. We achieve this by computing the persistent homology of the filtration described in the previous subsection. Broadly speaking, this computation will model the existence and strength of connected components and cycles in the network. More specifically, persistent homology computes a pair of mathematical objects entitled persistence diagrams. A persistence diagram is a multiset of pairs of points in the extended real line ($\mathbb{R} \cup \infty$) which satisfy the condition that $p < q$ for each element (p, q) (Edelsbrunner and Harer, 2010). One of the persistence diagrams models the existence of connected components in the filtration. Specifically, an element (p, q) in this persistence diagram indicates that a connected component appeared and subsequently disappeared at the graphs G_p and G_q respectively in the filtration. The other persistence diagram models the existence of cycles in the filtration. Specifically, an element (p, q) in this persistence diagram indicates that a cycle appeared and subsequently disappeared at the graphs G_p and G_q respectively in the filtration. If a connected component disappears it is because it has merged with another connected component. As explained below, in this study cycles always persist to the end of the entire filtration; if a cycle merges with another cycle it will result in a further cycle in addition to itself. Note that, for a given element (p, q) in a persistence diagram, the value $q - p$ is known as the persistence of the element in question. The persistence can be regarded as the lifetime of the respective connected component or cycle which first appears at duration or 'time' point p and disappears at time point q. In the literature on persistence diagrams p and q are also referred to as points of birth and death of the respective connected component or cycle.

If two connected components merge causing exactly one of them to disappear, the one which appeared latest in the filtration is the one which disappears. This is called the *elder rule* (Otter et al., 2017). If a connected component or cycle appears at graph G_p in the filtration but does not subsequently disappear it is represented by an element (p, ∞). In this work we replace all elements of the form (p, ∞) with (p, u) where u is an upper bound. We use a value of 8 for this upper bound which is a valid upper bound because G_7 is the final element in the filtration. Note that, performing such a replacement is a commonly used approach for dealing with the challenges a value of ∞ presents (Adams et al., 2017). Since the filtration in this work is constructed by only adding edges, any cycles which appear will not disappear. Therefore, the second value of all elements in the cycle persistence diagram will have a value equal to u.

The utility of these two persistence diagrams can be understood by considering that connected components and cycles which appear earlier and persist for longer are connected by more significant streets and in turn are more strongly connected. We first consider the case of connected components. The fewer number of connected components which persist for longer, the more well connected the corresponding street network is. To illustrate this, consider again the filtration displayed in Figure 4 for the Manchester street network. The network in Figure 4(a) contains three connected components. However these become connected to other components in the next graph of the filtration in Figure 4(b) and in turn these connections persist for a large fraction of the filtration indicating a high level of connectivity between the connected components in question. In contrast, if components that appear early in filtration do not quickly become connected by other streets, this indicates a poorer level of connectivity, reflected by a larger number of components which persist for longer. We next consider the

case of cycles. The greater number of cycles which persist for longer, the more well connected the corresponding street network is. To illustrate this, consider again the filtration displayed in Figure 4 for the Manchester street network. A large number of cycles appear early in the filtration in Figure 4(b) and hence persist for a larger fraction of the filtration, as (unlike connected components) once they appear they cannot subsequently disappear. This indicates that these cycles provide a high level of connectivity. A large number of cycles also appear later in the filtration in Figures 4(e) and in turn persist for a small fraction of the filtration. This indicates that these particular cycles provide relatively less connectivity.

The persistence diagrams corresponding to connected components and cycles for the Manchester street network are displayed in Figures 5(a) and 5(b) respectively. Appearance and disappearance values are represented by the horizontal and vertical axes respectively. Many of the elements in these persistence diagrams are equal and therefore for the purpose of visualization a small amount of Gaussian noise has been added to each element. The above figures are a very standard way to visualize persistence diagrams apart from the addition of Gaussian noise. However the persistence diagrams in this work contain many elements making them difficult to interpret. For example, the persistence diagrams in Figures 5(a) and 5(b) contain 258 and 9,899 elements respectively, most of which cannot be distinguished visually.

To overcome this difficulty we propose an alternative representation of persistence diagrams entitled persistence densities which can be more easily interpreted. Let $[1, 7]$ denote the closed integer interval between 1 and 7 and let $[0, 1]$ denote the closed real interval between 0 and 1. Given a persistence diagram D, the corresponding persistence density is the map $\tau : [1, 7] \to [0, 1]$ defined in Equation 4 where $|D|$ denotes the number of elements in D (i.e. the number of connected components or cycles).

$$\tau(i) = \frac{|\{(p, q) : (p, q) \in D, q - p = i\}|}{|D|} \tag{4}$$

Figures 6(a) and 6(b) display the connected component and cycle persistence densities respectively for the Manchester street network represented using histograms. The x-axis in these histograms represents the interval between appearance and disappearance, i.e. the period of persistence. The y-axis in these histograms represents the corresponding density of these persistence values as defined by Equation 4. These histograms model the relative frequency of different persistence values and can be easily interpreted. For example, one can see in Figure 6(a) that for the street network in question, most connected components have a persistence value of one.

The proposed persistence density representation is similar to the persistence image representation proposed by Adams et al. (2017). Broadly speaking, a persistence image is an image representation of a persistence diagram where pixel values equal a weighted sum of nearby elements. After much experimentation, the authors decided to use the proposed persistence density representation instead of the persistence image representation for the following reasons. Firstly, the authors found it difficult to visually differentiate between different persistence images because most contain peaks at the same set of locations where the heights of these peaks are in many cases not significantly different. On the

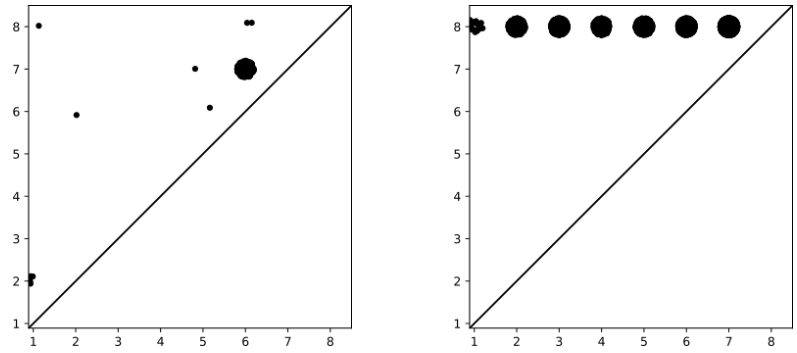


Figure 5: The persistence diagrams corresponding to connected components and cycles for the Manchester street network of Figure 2 are displayed in (a) and (b) respectively. Appearance and disappearance values are represented by the horizontal and vertical axes respectively.

other hand, since persistence densities are visualized using histograms it is much easier to visually differentiate between them. Secondly, the persistence image representation requires that a scale or smoothing parameter value be specified. It is not clear how best to select this parameter value. On the other hand, the persistent density representation does not require any parameter values to be specified.

To facilitate the application of data mining methods, it is useful to define a distance or metric between different street networks represented using the proposed model. There exists a number of distance measures on the space of persistence diagrams such as the Bottleneck and Wasserstein distances (Edelsbrunner and Harer, 2010). However these distances are biased with respect to the number of elements with significant persistence in the corresponding persistence diagrams. That is, any two persistence diagrams having significantly different numbers of elements with significant persistence will generally have a large corresponding distance. Since larger cities will have graph representations with a greater number of edges, their corresponding persistence diagrams will have a corresponding greater number of elements with significant persistence. This in turn results in a bias where cities with similar sized street networks are determined to be more similar. This is undesirable because we would prefer a distance measure which is a function of street network connectivity alone. To overcome this challenge we instead use the Wasserstein distance on the space of persistence densities (Peyré et al., 2019). Intuitively, this distance measures the cost of transforming one density function into another. Since the density function is normalized such that its codomain sums to 1.0, this provides a distance measure which is unbiased with respect to street network size.

4 City Connectivity Analysis

This work proposes a novel model of street network connectivity. The usefulness of any model can only be defined with respect to performing a given task. In

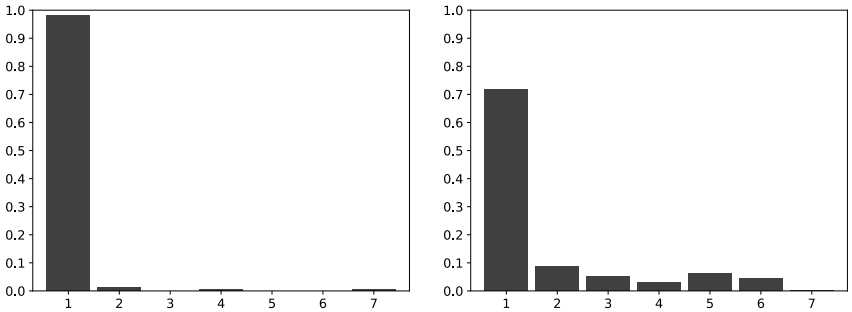


Figure 6: The connected component and cycle persistence densities for the Manchester street network are displayed in (a) and (b) respectively. The vertical axis represents density values while the horizontal axis represents the interval between appearance and disappearance (the persistence value).

this section we demonstrate the usefulness of the proposed model with respect to performing the tasks of interpreting and providing insight into the connectivity properties of a given street network, recommending ways to improve this connectivity, and identifying street networks with similar and dissimilar connectivity properties. We perform these tasks with respect to the set of UK city street networks and a subset of USA city street networks.

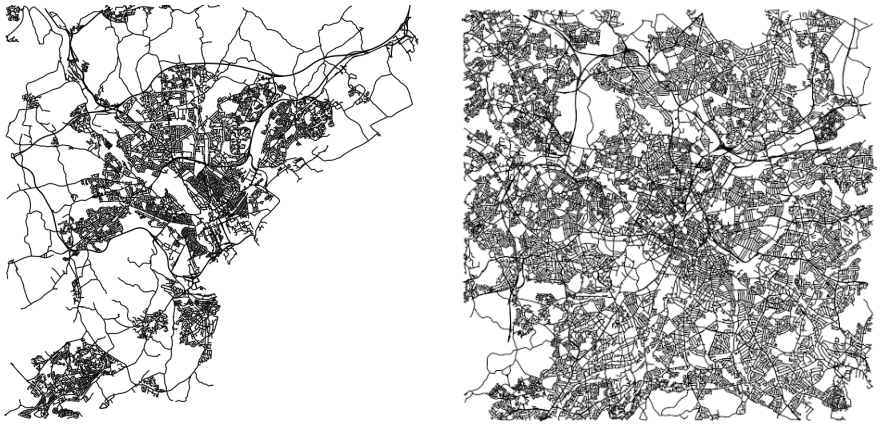
The remainder of this section is structured as follows. In section 4.1 we present details of the UK and USA street networks considered and the process used to construct these networks. In section 4.2 we demonstrate the usefulness of the proposed model with respect to performing the downstream tasks mentioned above.

4.1 Street Networks

The first set of street networks we considered is a set of street networks corresponding to all UK cities for which there are 66. All UK cities have official boundaries [2]. However, we found these boundaries to be inconsistent with respect to what they contained. For example, some city boundaries contain the corresponding suburban street network while others do not. This is a consequence of the fact that many city boundaries were defined before the street network in question fully developed and are a function of political influences. To overcome this challenge, for each UK city we selected a location in the city center and extracted the street network within a ten kilometers bounding box centred at this location. We found this approach to consistently return the desired street network.

The second set of street networks we considered is a set of street networks corresponding to a subset of USA cities. In order to have a representative sample of cities we used the subset of cities proposed by Angel and Blei (2016). In this work the authors used a random stratified sampling procedure to select a subset

[2] `https://data.gov.uk/dataset/7879ab82-2863-401e-8a29-a56e264d2182/major-towns-and-cities-december-2015-boundaries`



(a) (b)

Figure 7: The street networks for the UK cities of Cardiff and Birmingham are displayed in (a) and (b) respectively where street segments are represented by black lines.

of 40 cities from the set of all 242 USA cities that had populations of 100,000 or more in the year 2000. We defined the boundary of each city to be the boundary of the corresponding Metropolitan Statistical Area (MSA) in which it is contained (Arribas-Bel and Sanz-Gracia, 2014). MSAs are defined by the United States Office of Management and Budget (OMB) as core areas containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core [3].

The street networks for both the UK and USA cities were obtained from OpenStreetMap, which is a crowdsourcing project for geographical data, using the OSMNX software library (Boeing, 2017). Tables 1 and 2 display the names of the UK cities and USA MSAs respectively plus the number of vertices and edges in the corresponding graph representations. Figures 2, 7(a) and 7(b) display the street networks for the UK cities of Manchester, Cardiff and Birmingham respectively. Figures 8(a) and 8(b) display the street networks for the USA MSAs of Boston-Cambridge-Newton, MA-NH and Columbia, SC respectively.

4.2 Analysis

The analysis of the connected component and cycle persistence densities for a given street network provides insight into its connectivity properties. A connected component persistence density which is more skewed to the left indicates a greater level of connectivity. Specifically, such skewness indicates that most elements in the corresponding persistence diagram have smaller persistence. That is, most parts of the street network get connected earlier in the filtration where these connections correspond to more significant streets. On the other hand, a cycle persistence density which is more skewed to the right also indicates a greater level of connectivity. Specifically, such skewness indicates that most elements in the corresponding persistence diagram have larger persistence. That

[3] `https://www.census.gov/programs-surveys/metro-micro/about.html`

City	$\|V\|$	$\|E\|$	City	$\|V\|$	$\|E\|$
Aberdeen	9,348	11,870	Liverpool	31,743	39,405
Armagh	2,112	2,624	London	46,878	61,858
Bangor	3,301	3,937	Manchester	44,780	55,325
Bath	8,782	10,117	Newcastle	29,719	35,326
Belfast	18,286	21,884	Newport	10,610	12,523
Birmingham	37,797	46,776	Newry	3,623	4,270
Bradford	30,804	36,468	Norwich	12,902	14,875
Brighton	6,564	8,616	Nottingham	23,364	27,382
Bristol	23,236	27,922	Oxford	7,573	8,851
Cambridge	7,884	9,096	Perth	2,999	3,556
Canterbury	6,185	7,245	Peterborough	10,174	11,594
Cardiff	14,970	17,951	Plymouth	10,038	11,926
Carlisle	4,421	5,218	Portsmouth	13,631	16,309
Chelmsford	6,322	7,508	Preston	13,636	15,978
Chester	9,224	10,844	Ripon	2,081	2,509
Chichester	5,206	6,142	St Albans	12,919	15,513
Coventry	13,502	16,132	St Asaph	6,084	7,094
Derby	12,482	14,471	St Davids	335	410
Derry	5,200	5,983	Salford	40,974	50,474
Dundee	7,276	9,039	Salisbury	2,864	3,386
Durham	11,581	13,505	Sheffield	22,081	26,950
Edinburgh	15,055	18,505	Southampton	18,306	21,020
Ely	3,284	3,631	Stirling	5,563	6,593
Exeter	8,299	9,658	Stoke-on-Trent	16,465	19,700
Glasgow	30,266	38,512	Sunderland	18,727	22,716
Gloucester	10,529	12,083	Swansea	10,352	12,342
Hereford	3,637	4,246	Truro	3,193	3,808
Inverness	4,792	5,455	Wakefield	20,568	23,725
Hull	14,151	16,322	Wells	3,655	4,286
Lancaster	5,081	6,241	Westminster	46,938	61,863
Leeds	29,393	35,470	Winchester	5,792	6,640
Leicester	18,604	22,358	Wolverhampton	25,229	30,387
Lichfield	11,079	12,832	Worcester	7,810	9,049
Lincoln	7,995	9,251	York	7,724	8,780
Lisburn	11,002	13,111			

Table 1: This table displays the names of the 66 UK cities considered plus the number of vertices ($|V|$) and edges ($|E|$) in the corresponding graph representations.

Metropolitan Statistical Area	$\|V\|$	$\|E\|$
New York-Newark-Jersey City, NY-NJ-PA	414,171	58,777
Los Angeles-Long Beach-Anaheim, CA	256,499	358,592
Chicago-Naperville-Elgin, IL-IN-WI	263,540	383,581
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	190,291	272,286
Miami-Fort Lauderdale-Pompano Beach, FL	163,900	242,164
Dallas-Fort Worth-Arlington, TX	303,428	430,668
Boston-Cambridge-Newton, MA-NH	167,664	231,761
Washington-Arlington-Alexandria, DC-VA-MD-WV	232,169	303,521
Detroit-Warren-Dearborn, MI	156,080	225,515
Houston-The Woodlands-Sugar Land, TX	263,633	368,016
Salt Lake City, UT	41,084	55,026
San Francisco-Oakland-Berkeley, CA	81,789	111,203
Cleveland-Elyria, OH	56,504	79,345
Pittsburgh, PA	118,614	158,033
Portland-Vancouver-Hillsboro, OR-WA	97,364	127,851
Virginia Beach-Norfolk-Newport News, VA-NC	58,522	77,547
Sacramento-Roseville-Folsom, CA	89,105	116,801
Kansas City, MO-KS	118,997	165,857
Columbus, OH	79,508	109,027
Austin-Round Rock-Georgetown, TX	85,495	11,5397
Hartford-East Hartford-Middletown, CT	41,484	56,222
El Paso, TX	35,911	52,111
Omaha-Council Bluffs, NE-IA	50,696	74,427
Albuquerque, NM	56,353	76,762
Grand Rapids-Kentwood, MI	42,099	56,527
Columbia, SC	51,325	66,669
Des Moines-West Des Moines, IA	31,969	46,063
Spokane-Spokane Valley, WA	26,906	37,749
Pensacola-Ferry Pass-Brent, FL	27,426	36,640
Jackson, MS	44,356	56,669
Shreveport-Bossier City, LA	26,585	35,514
Asheville, NC	39,766	46,531
Tallahassee, FL	21,230	27,989
Manchester-Nashua, NH	17,062	24,275
Portland-South Portland, ME	37,028	46,736
Norwich-New London, CT	11,621	15,340
Kennewick-Richland, WA	14,413	19,493
Greensboro-High Point, NC	38,629	49,717
Pueblo, CO	10,141	14,281
Tyler, TX	11,887	16,275

Table 2: This table displays the names of 40 USA Metropolitan Statistical Areas considered plus the number of vertices ($|V|$) and edges ($|E|$) in the corresponding graph representations.

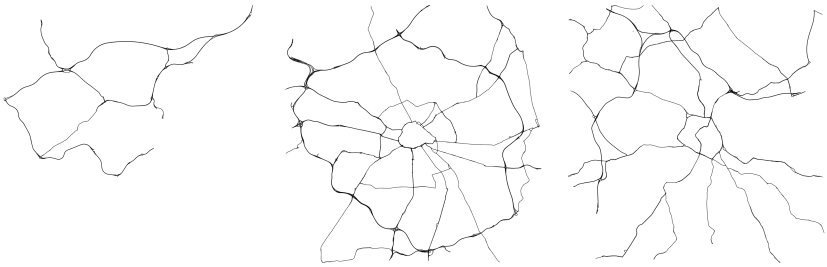
(a) (b)

Figure 8: The street networks for the USA MSAs of Boston-Cambridge-Newton, MA-NH and Columbia, SC are displayed in (a) and (b) respectively where street segments are represented by black lines.

is, different parts of the street network get connected by multiple paths that create cycles (i.e. circuits), earlier in the filtration where these paths correspond to more significant streets.

To illustrate the insights which this analysis can provide with respect to street network connectivity properties consider the connected component and cycle persistence densities for the UK cities of Cardiff, Manchester and Birmingham displayed in Figure 10. The connected component persistence densities for Manchester and Birmingham are more skewed to the left than that of Cardiff. Notably, Manchester and Birmingham both have significantly more connected components with a persistence value of 1. This indicates that Cardiff has a number of areas which are not connected by significant streets. Therefore, the connectivity of Cardiff could be improved by connecting these areas by such streets. The cycle persistence densities for Manchester and Birmingham are more skewed to the right than that of Cardiff indicating higher levels of connectivity. Notably, both Manchester and Birmingham have significantly more cycles with persistence values of 6 and 7. This is a consequence of the fact that, unlike Cardiff, both Manchester and Birmingham are both inland cities with significant roads surrounding them. This fact is illustrated in Figure 9 which displays the set of streets with type equal to motorway or trunk for the cities of Cardiff, Manchester and Birmingham. Motorway or trunk are the most significant street types and we can see that Manchester and Birmingham each contain more cycles formed from these types than Cardiff.

To contrast the above analysis with that which can be achieved using traditional models of street network connectivity consider Table 3 which displays the



(a) (b) (c)

Figure 9: The set of streets with type equal to motorway or trunk for the cities of Cardiff, Manchester and Birmingham are displayed in (a), (b) and (c) respectively.

intersection density for each UK city. The intersection density for a given street network equals the mean number of street network intersections per square kilometer. This is a commonly used model of street network connectivity where greater density is considered to indicate greater connectivity (Dill, 2004). The intersection densities for the cities of Cardiff, Manchester and Birmingham are 101, 311 and 263 respectively. The model therefore indicates that Manchester has the greatest connectivity while Cardiff has the worst. However, the model does not provide any insight for how to improve connectivity beyond increasing the number of intersections. For example, it does not indicate what types of streets should be added to increase intersection density or between which areas these streets should be constructed.

For the UK and USA street networks we computed the pairwise Wasserstein distances between the corresponding sets of connected component and cycle persistence densities. Using these distances, we subsequently performed hierarchical single-linkage clustering of the UK and USA street networks to obtain dendrogram representations (Everitt et al., 2011). Figures 11(a) and 11(b) display the dendrograms for the UK street networks with respect to connected component and cycle persistence densities respectively. Figures 12(a) and 12(b) display the dendrograms for the USA street networks with respect to connected component and cycle persistence densities respectively. Examining these dendrograms allows us to systematically determine cities that have more- or less-similar connectivity properties. To further assist in this analysis, for each set of pairwise distances we computed a corresponding representation of each city as a point in $\mathbb{R}^2$ using the t-SNE manifold learning technique (Maaten and Hinton, 2008). Figures 13(a) and 13(b) display the t-SNE representations for the UK street networks with respect to connected component and cycle persistence densities respectively. Figures 14(a) and 14(b) display the t-SNE representations for the USA street networks with respect to connected component and cycle persistence densities respectively.

To evaluate the stability of the above dendrograms and t-SNE representations with respect to the choice of Wasserstein distance, we recomputed these representations using the energy distance (Rizzo and Székely, 2016). Figures 15(a) and 15(b) display the dendrogram representations in question for the UK street networks with respect to connected component and cycle persistence den-

City	Intersection Density	City	Intersection Density
Aberdeen	68	Liverpool	223
Armagh	15	London	354
Bangor	22	Manchester	311
Bath	56	Newcastle	198
Belfast	124	Newport	71
Birmingham	263	Newry	24
Bradford	203	Norwich	82
Brighton	50	Nottingham	152
Bristol	157	Oxford	48
Cambridge	50	Perth	20
Canterbury	40	Peterborough	64
Cardiff	101	Plymouth	67
Carlisle	29	Portsmouth	91
Chelmsford	42	Preston	89
Chester	60	Ripon	14
Chichester	35	St Albans	89
Coventry	91	St Asaph	39
Derby	81	St Davids	3
Derry	34	Salford	286
Dundee	51	Salisbury	20
Durham	76	Sheffield	152
Edinburgh	104	Southampton	117
Ely	21	Stirling	36
Exeter	53	Stoke-on-Trent	112
Glasgow	219	Sunderland	129
Gloucester	65	Swansea	68
Hereford	23	Truro	21
Inverness	29	Wakefield	134
Hull	90	Wells	23
Lancaster	35	Westminster	354
Leeds	198	Winchester	36
Leicester	126	Wolverhampton	172
Lichfield	72	Worcester	51
Lincoln	50	York	48
Lisburn	73		

Table 3: The intersection density for each UK city street network where intersection density equals the mean number of street network intersections per square kilometer.

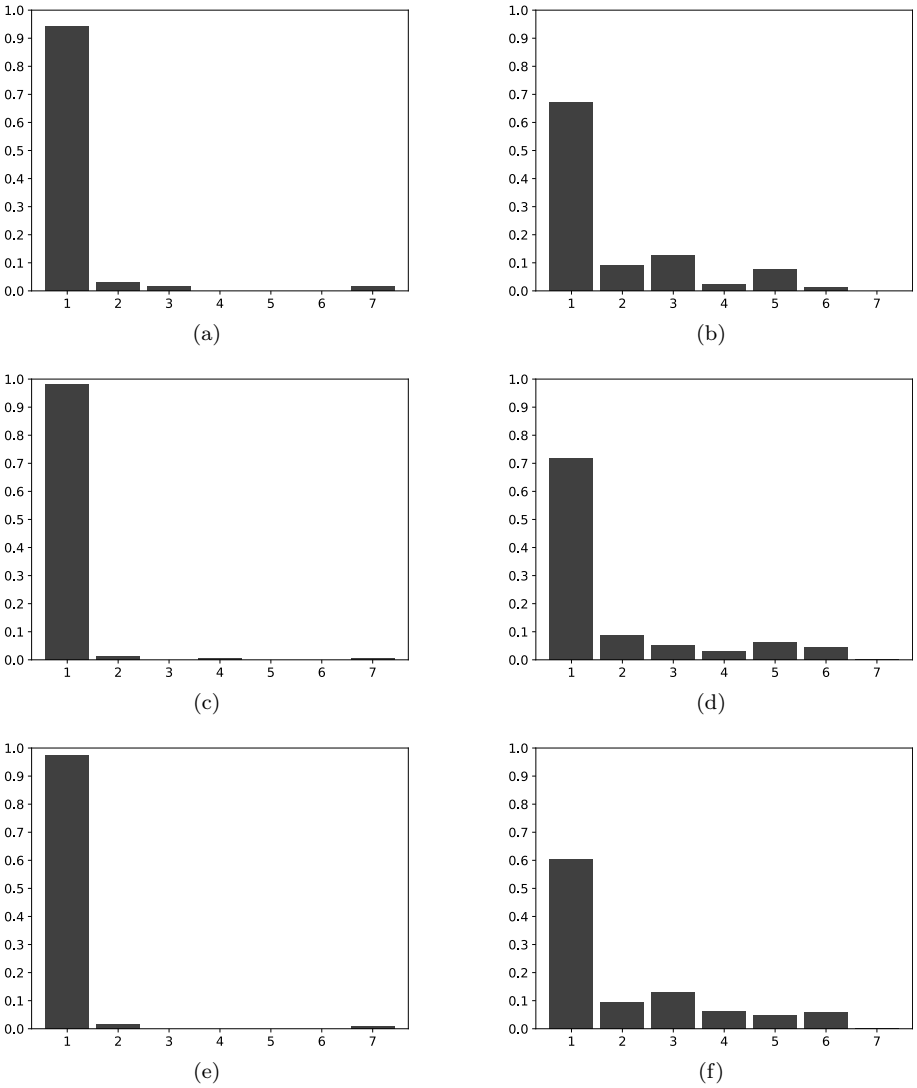
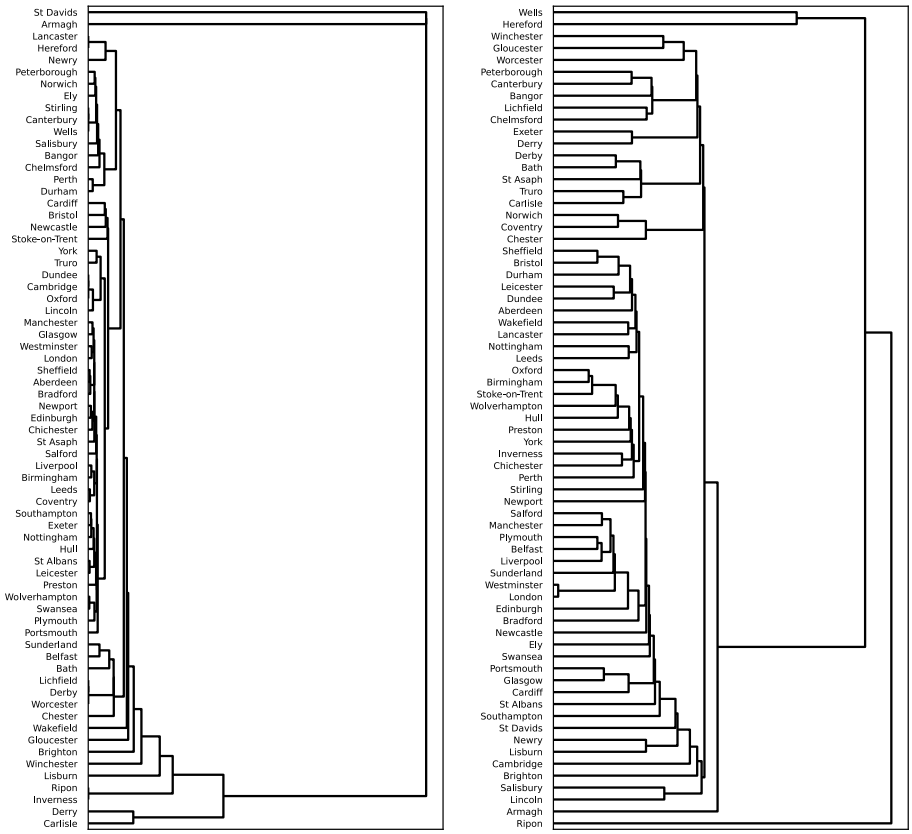


Figure 10: The connected component persistence densities for the Cardiff, Manchester and Birmingham street networks are displayed in (a), (c) and (e) respectively. The cycle persistence densities for the Cardiff, Manchester and Birmingham street networks are displayed in (b), (d) and (f) respectively.



(a) (b)

Figure 11: Dendrograms for the connected-component and cycle persistence diagrams for the UK street networks are displayed in (a) and (b) respectively. Pairwise distances computed using the Wasserstein distance.

sities respectively. Figures 16(a) and 16(b) display the t-SNE representations in question for the UK street networks with respect to connected component and cycle persistence densities respectively. Comparing the dendrograms and t-SNE representations computed using the Wasserstein and energy distances we see that they are quite similar. For example, comparing dendrograms in Figures 11(a) and 15(a) we see that in both cases the distance between the St. Davids and Armagh street networks is determined to be relatively large.

Examining the dendrograms and t-SNE manifolds, we can identify cities which have similar and dissimilar persistence densities. For example, Birmingham and Liverpool have similar connected component persistence densities which are displayed in Figure 17. Both these persistence densities are heavily skewed to the left and, as discussed above, this indicates a high level of connectivity. This can be attributed to the fact that both are major UK cities with a large number of significant connecting streets. The connected component persistence densities for Armagh and Brighton are displayed in Figure 18. Both these cities have dissimilar connected component persistence densities to that

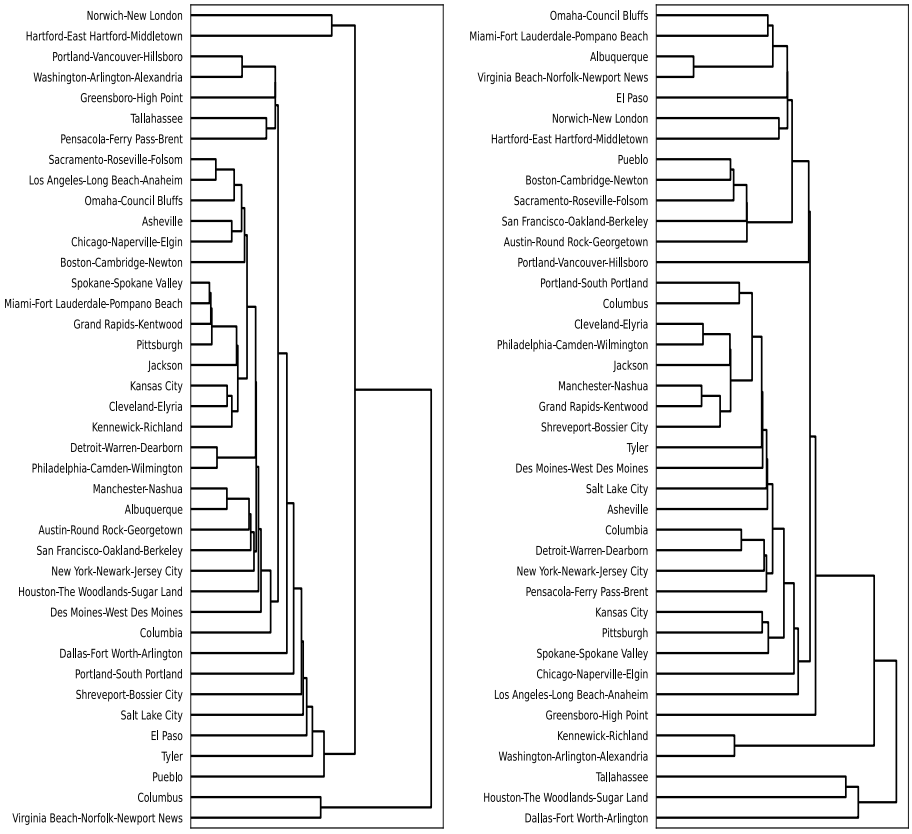
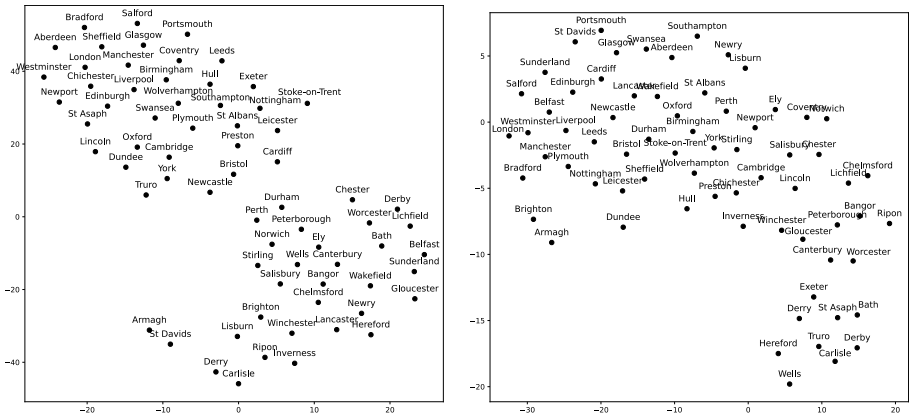
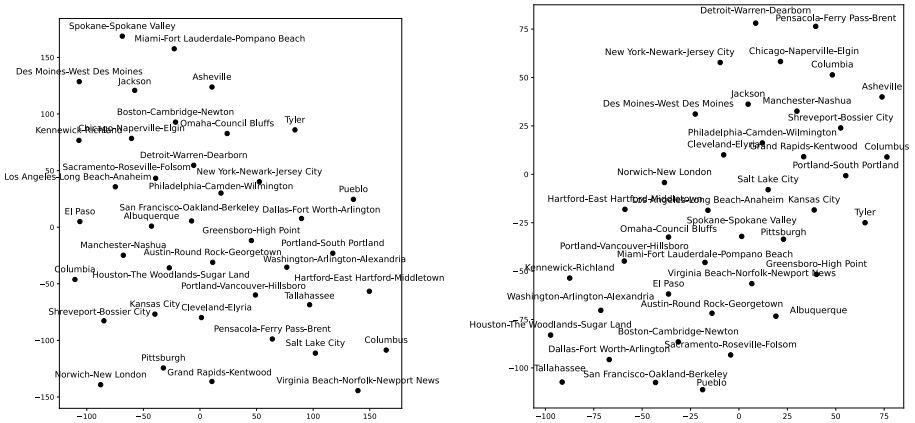


Figure 12: Dendrograms for the connected-component and cycle persistence diagrams for the USA street networks are displayed in (a) and (b) respectively. Pairwise distances computed using the Wasserstein distance.



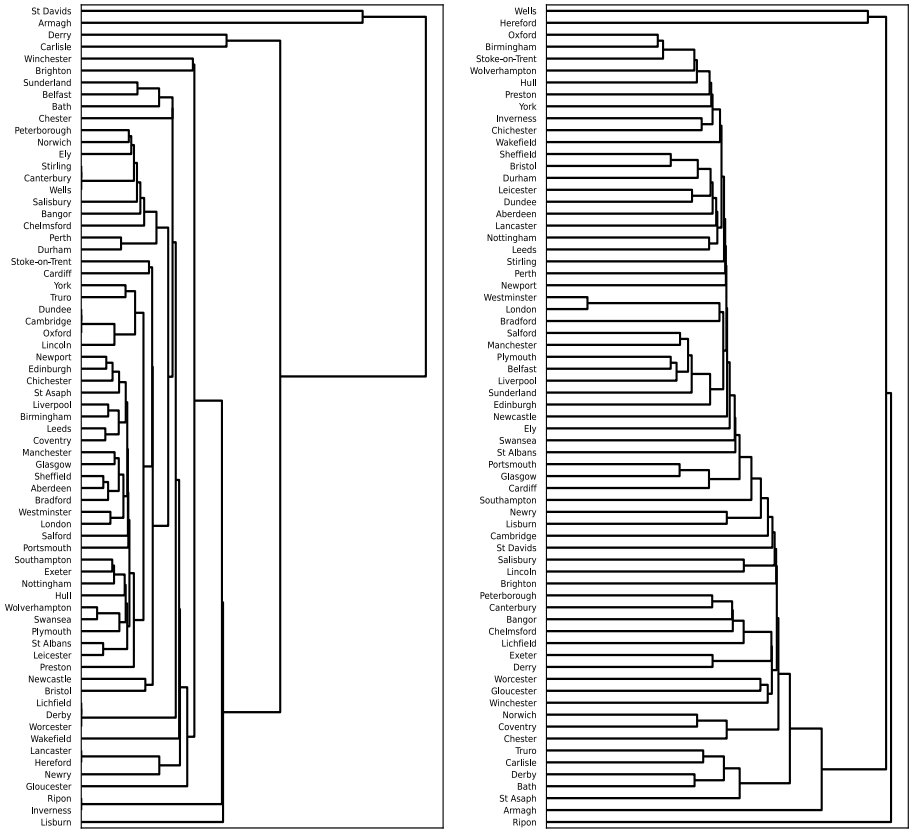
(a) (b)

Figure 13: t-SNE representation of each UK street network as a point in $\mathbb{R}^2$ with respect to connected component and cycle persistence densities are displayed in (a) and (b) respectively. Pairwise distances were computed using the Wasserstein distance.



(a) (b)

Figure 14: t-SNE representation of each USA street network as a point in $\mathbb{R}^2$ with respect to connected component and cycle persistence densities are displayed in (a) and (b) respectively. Pairwise distances were computed using the Wasserstein distance.



(a) (b)

Figure 15: Dendrograms for connected component and cycle persistence diagrams for the UK street networks are displayed in (a) and (b) respectively. Pairwise distances were computed using the energy distance.

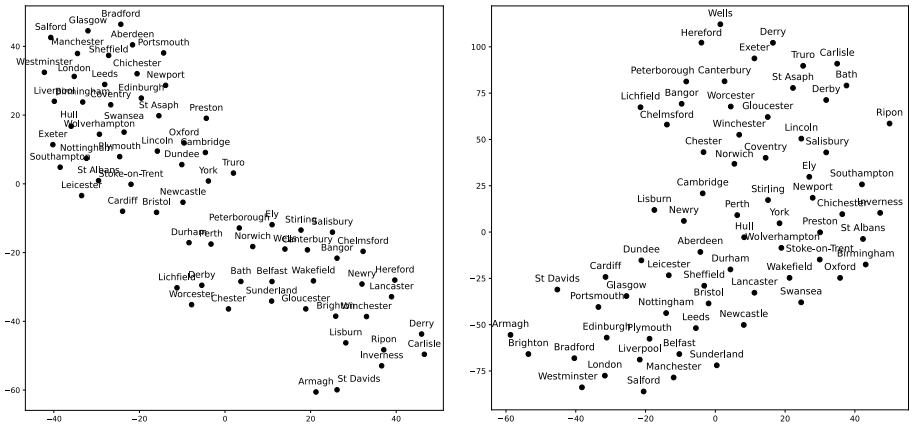


Figure 16: t-SNE representation of each UK street network as a point in $\mathbb{R}^2$ with respect to connected component and cycle persistence densities are displayed in (a) and (b) respectively. Pairwise distances were computed using the energy distance.

of Birmingham and Liverpool. Specifically, they are more skewed to the right. This can be attributed to the fact that both are smaller less significant cities with a lower number of significant connecting streets. In fact, Armagh is an extremely small city and this is reflected in the relative size of its graph representation displayed in Table 1. This analysis indicates that the connectivity of Armagh and Brighton could be improved by adding more significant connecting streets.

In order to compare the connectivity of the UK and USA street networks we computed the mean connected component and cycle persistence densities for the corresponding sets of street networks. These mean persistence densities are displayed in Figure 19. The mean UK connected component persistence density

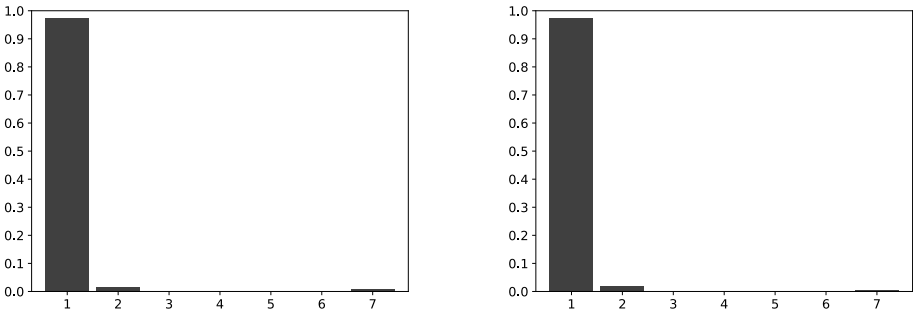


Figure 17: The connected component persistence densities for the Birmingham and Liverpool street networks are displayed in (a) and (b) respectively.

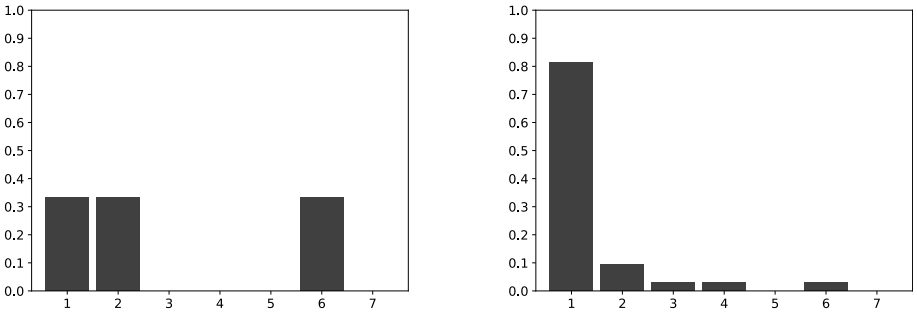


Figure 18: The connected component persistence densities for the Armagh and Brighton street networks are displayed in (a) and (b) respectively.

is more skewed to the left than that corresponding to the USA. Furthermore, the mean UK cycle persistence density is more skewed to the right than that corresponding to the USA. These two facts indicate that, on average the UK street networks have a greater level of connectivity than the USA street networks.

As discussed in section 4.1, in many cases the boundary of a city is ambiguous and difficult to define. To determine if the proposed model of connectivity is sensitive to the choice of city boundary we computed the connected component and cycle persistence densities for the city of Manchester for three different boundaries. Specifically, we considered boundaries corresponding to the 3.33, 6.67 and 10 kilometer bounding boxes centred at the city center. Figure 20 displays the connected component and cycle persistence densities corresponding to each of these bounding boxes. We can see that the persistence densities corresponding to the 6.67 and 10 kilometer bounding boxes are very similar. However, the persistence densities corresponding to the 3.33 kilometer bounding box are quite distinct. This result demonstrates that when attempting to model the connectivity of a given city, the proposed model is sensitive to the choice of corresponding boundary and therefore care must be taken when defining this. Note that, most models of street network connectivity also exhibit this sensitivity. For example, the intersection density values corresponding to the above bounding boxes are 494, 360, and 311 respectively. That is, as the bounding box increases the intersection density decreases. This is evident from viewing the Manchester city street network in Figure 2 where we see that the density of streets is greatest in the city center.

We extended the above experiment to the set of all UK cities by computing the mean persistence densities for these cities for 3.33, 6.67 and 10 kilometer bounding boxes. These persistence densities are displayed in Figure 21. We see that as the bounding box size increases the mean connected component persistence density becomes slightly more skewed to the left indicating a greater level of connectivity. Furthermore, we see that as the bounding box size increases the mean cycle persistence density becomes more skewed to the right, also indicating a greater level of connectivity. Both these results are expected because more significant street types and in turn connections generally exist outside the center of a city. This analysis provides a form of validation for the proposed

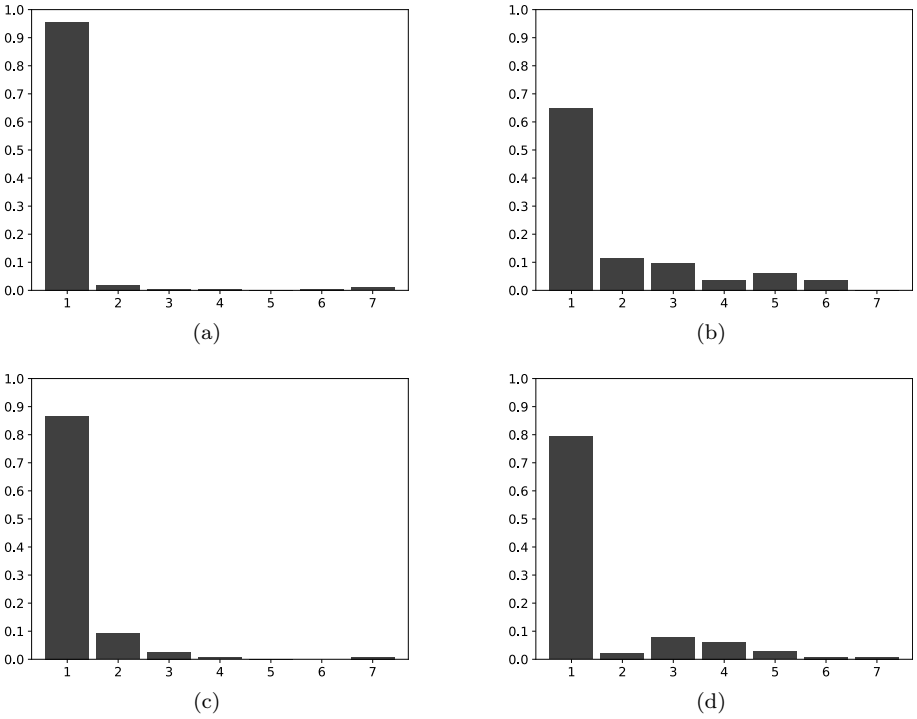


Figure 19: The mean connected component and cycle persistence densities for the set of UK cities are displayed in (a) and (b) respectively. The mean connected component and cycle persistence densities for the set of USA cities are displayed in (c) and (d) respectively.

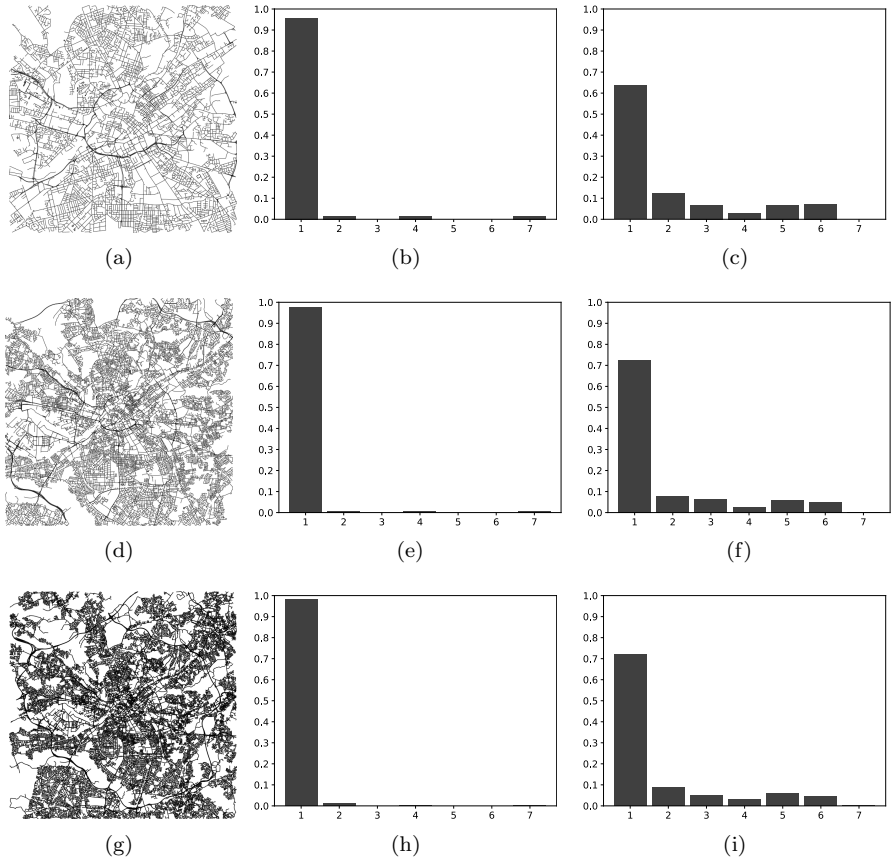


Figure 20: The Manchester street network corresponding to a 3,333 meter bounding box is displayed in (a) while the corresponding connected component and cycle persistence densities are displayed in (b) and (c) respectively. The Manchester street network corresponding to a 6,667 meter bounding box is displayed in (d) while the corresponding connected component and cycle persistence densities are displayed in (e) and (f) respectively. The Manchester street network corresponding to a 10,000 meter bounding box is displayed in (g) while the corresponding connected component and cycle persistence densities are displayed in (h) and (i) respectively.

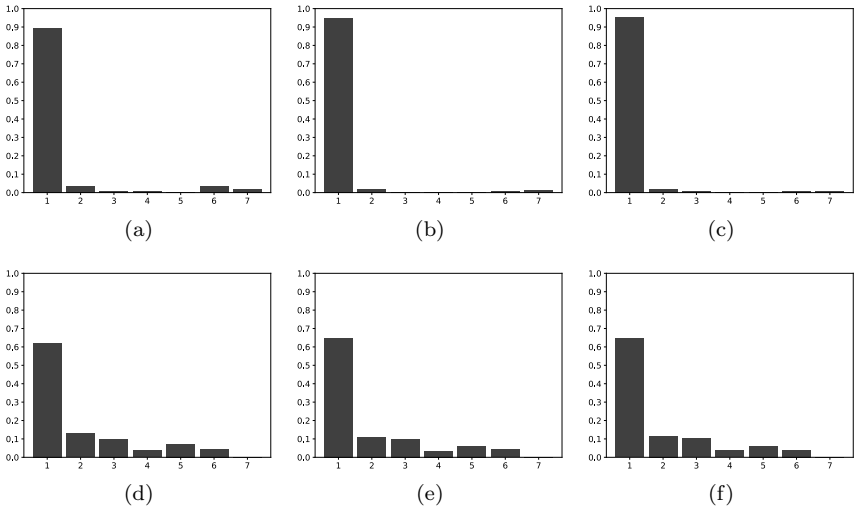


Figure 21: The mean connected component persistence density for the set of UK cities for 3.33, 6.67 and 10 kilometer bounding boxes are displayed in (a) and (b) and (c) respectively. The mean cycle persistence density for the set of UK cities for 3.33, 6.67 and 10 kilometer bounding boxes are displayed in (d) and (e) and (f) respectively.

model of street network connectivity.

5 Conclusions

In this article we propose a novel model of street network connectivity which is distinct from existing models in a number of ways. Firstly, in the proposed model connectivity is modelled in terms of the relative strength and frequency of connected components and cycles in the network. In existing models connectivity is generally modelled in terms of summary statistics of simple geometrical or topological features of the network. Secondly, the proposed model considers the type of different streets and different levels of connectivity they provide. The authors are unaware of any existing models of connectivity which consider such information. Finally, the proposed model does not represent connectivity using a single number. Instead connectivity is represented using a richer representation, specifically a set of persistence densities, which can be used to gain novel insights into the connectivity properties of different cities. Despite being a somewhat more complex representation, persistence densities can still be easily interpreted. This is particularly important given that the planning practice has traditionally relied on connectivity measures that can be intuitively understood and/or easily applied (Stangl and Guinn, 2011).

Despite the above achievements, the proposed model of connectivity has some limitations which provide opportunities for future research. Firstly, when attempting to model the connectivity of a given city, the proposed model is sensitive to the choice of corresponding boundary. One potential solution to this issue would be to not pick a single boundary but consider all possible

boundaries using some form of multi-dimensional persistent homology (Carlsson et al., 2009). Another approach is to provide more consistent methods of defining region types with regard for example to their inclusion or otherwise of outlying suburban residential areas and measures of population density and of employment (Arribas-Bel and Sanz-Gracia, 2014). Secondly, the proposed model is solely a function of the corresponding street network. However, in reality connectivity is a function of many other factors including population size and spatial distribution (Mansury and Shin, 2015). For example, if a large percentage of the city population live in the suburbs but work in the city center then the importance of connectivity between these two types of region must be weighted more highly.

References

Adams, H., Emerson, T., Kirby, M., Neville, R., Peterson, C., Shipman, P., Chepushtanova, S., Hanson, E., Motta, F., and Ziegelmeier, L. (2017). Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18.

Ahmed, M., Fasy, B. T., and Wenk, C. (2014). Local persistent homology based distance between maps. In *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 43–52.

Angel, S. and Blei, A. M. (2016). The spatial structure of american cities: The great majority of workplaces are no longer in CBDs, employment sub-centers, or live-work communities. *Cities*, 51:21–35.

Arribas-Bel, D. and Sanz-Gracia, F. (2014). The validity of the monocentric city model in a polycentric age: US metropolitan areas in 1990, 2000 and 2010. *Urban Geography*, 35(7):980–997.

Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. *Computers, Environment and Urban Systems*, 65:126–139.

Carlsson, G., Singh, G., and Zomorodian, A. (2009). Computing multidimensional persistence. In *International Symposium on Algorithms and Computation*, pages 730–739. Springer.

Corcoran, P. (2019). Topological generalization of continuous valued raster data. In *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 428–431.

Corcoran, P. and Jones, C. B. (2016). Spatio-temporal modeling of the topology of swarm behavior with persistence landscapes. In *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 1–4.

Corcoran, P. and Jones, C. B. (2017). Modelling topological features of swarm behaviour in space and time with persistence landscapes. *IEEE Access*, 5:18534–18544.

Corcoran, P. and Jones, C. B. (2018). Robust tracking of objects with dynamic topology. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 428–431.

Corcoran, P. and Mooney, P. (2013). Characterising the metric and topological evolution of openstreetmap network representations. *The European Physical Journal Special Topics*, 215(1):109–122.

Dey, T. K., Wang, J., and Wang, Y. (2017). Improved road network reconstruction using discrete morse theory. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pages 1–4.

Dill, J. (2004). Measuring network connectivity for bicycling and walking. In *83rd annual meeting of the Transportation Research Board, Washington, DC*, pages 11–15.

Edelsbrunner, H. and Harer, J. (2010). *Computational Topology: an Introduction*. American Mathematical Soc.

Ellis, G., Hunter, R., Tully, M. A., Donnelly, M., Kelleher, L., and Kee, F. (2016). Connectivity and physical activity: using footpath networks to measure the walkability of built environments. *Environment and Planning B: Planning and Design*, 43(1):130–151.

Everitt, B. S., Landau, S., Leese, M., and Stahl, D. (2011). *Cluster analysis 5th ed.* John Wiley, West Sussex, UK.

Feng, M. and Porter, M. A. (2020). Spatial applications of topological data analysis: Cities, snowflakes, random structures, and spiders spinning under the influence. *Physical Review Research*, 2(3):033426.

Ghrist, R. W. (2014). *Elementary Applied Topology*, volume 1. Createspace Seattle.

Gross, J. L., Yellen, J., and Zhang, P. (2013). *Handbook of graph theory*. CRC press.

Hajrasouliha, A. and Yin, L. (2015). The impact of street network connectivity on pedestrian volume. *Urban Studies*, 52(13):2483–2497.

Henderson, V. (2003). The urbanization process and economic growth: The so-what question. *Journal of Economic growth*, 8(1):47–71.

Knight, P. L. and Marshall, W. E. (2015). The metrics of street network connectivity: their inconsistencies. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 8(3):241–259.

Labi, S., Faiz, A., Saeed, T. U., Alabi, B. N. T., and Woldemariam, W. (2019). Connectivity, accessibility, and mobility relationships in the context of low-volume road networks. *Transportation research record*, 2673(12):717–727.

Maaten, L. v. d. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov):2579–2605.

Mansury, Y. and Shin, J. (2015). Size, connectivity, and tipping in spatial networks: Theory and empirics. *Computers, Environment and Urban Systems*, 54:428–437.

Otter, N., Porter, M. A., Tillmann, U., Grindrod, P., and Harrington, H. A. (2017). A roadmap for the computation of persistent homology. *EPJ Data Science*, 6:1–38.

Peponis, J., Bafna, S., and Zhang, Z. (2008). The connectivity of streets: reach and directional distance. *Environment and Planning B: Planning and Design*, 35(5):881–901.

Peyré, G., Cuturi, M., et al. (2019). Computational optimal transport: With applications to data science. *Foundations and Trends® in Machine Learning*, 11(5-6):355–607.

Ritchie, H. (2018). Urbanization. *Our World in Data*. https://ourworldindata.org/urbanization.

Rizzo, M. L. and Székely, G. J. (2016). Energy distance. *wiley interdisciplinary reviews: Computational Statistics*, 8(1):27–38.

Sahitya, K. S. and Prasad, C. (2019). Modelling structural interdependent parameters of an urban road network using gis. *Spatial Information Research*, pages 1–8.

Stangl, P. (2012). The pedestrian route directness test: A new level-of-service model. *Urban Design International*, 17(3):228–238.

Stangl, P. (2015). Block size-based measures of street connectivity: A critical assessment and new approach. *Urban Design International*, 20(1):44–55.

Stangl, P. (2019). Overcoming flaws in permeability measures: modified route directness. *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, 12(1):1–14.

Stangl, P. and Guinn, J. M. (2011). Neighborhood design, connectivity assessment and obstruction. *Urban Design International*, 16(4):285–296.

Tal, G. and Handy, S. (2012). Measuring nonmotorized accessibility and connectivity in a robust pedestrian network. *Transportation research record*, 2299(1):48–56.

Weber, J. (2016). The properties of topological network connectivity measures and their application to us urban freeway networks. *The Professional Geographer*, 68(3):485–495.