

Sea-land connectivity in the global complex network: the case of United Kingdom

Bárbara POLO MARTÍN¹ and César DUCRUET²

Abstract

The main objective of this research is to analyse the connectivity of cities in a coupled network made of planar (railways) and non-planar (maritime) topologies. It takes the state of the network during the period 1880- 1925, namely the context of the First Globalization wave (1880-1914), when trade and urban development were closely tied to progress in communications systems and especially steam propulsion. Edges represent intercity physical infrastructure on land, and inter-port ship voyages at sea. We test several hypotheses in terms of inter-network specialization and urban hierarchies. Main results underline a crucial influence of railway proximity on vessel traffic volume. A case study of the United Kingdom also reveals that the networks are highly interdependent, while the combination of infrastructures is closely associated with city size and urban growth. We discuss main results in the light of network science, spatial science, maritime history, and transport research.

Keywords

Coupled networks; globalization; hinterlands; ports; maritime transport; multilayer networks; multigraph; railway networks; steam shipping; urban network

¹ French National Centre for Scientific Research (CNRS) & UMR 7235 EconomiX

² French National Centre for Scientific Research (CNRS) & UMR 7235 EconomiX

1. Introduction

Industrialization and related urbanization began in the midst of 18th century, but accelerated significantly during the 19th and early 20th centuries due to the set of circumstances during this period: technological innovations, social changes, and political institutions that increasingly favoured economic growth. Before 1880, industrialization was based on a prescribed division of labour: most jobs were dedicated to smaller tasks, and people repeated the same task indefinitely. After 1880, industrialization relied more heavily on mechanization to increase output and maximize profits. The development of the modern electrical grid in the early 1880s facilitated such technological advances. Mass production after the turn of the 20th century only exacerbated this effect. As a result, total output in the early 20th century was higher than in 1880. Besides, all these changes, together with sanitary improvements in cities, led to a population growth over the same period.

Industrialization, transport and urbanization have been recurrently studied as an interdependent growth process. Industrialization in Europe - specifically in the United Kingdom (Crafts and Mills, 2004)-, and later the United States (Kim, 2005) began in the early nineteenth century with the reorganization of manufacturing from artisanal shops to the first factories with steam power in a few industries and the emergence of new ways of transport (Oldlyzko, 2010; Marnot, 2015). During the second half of the nineteenth century, manufacturing activity increased in scale, became more mechanized, and spread to many industries. This consequence has been renamed by historians as “steam effect” or even “revolution”.

However, the development of this effect differed depending among European countries. Some of them witnessed an early development of the manufacturing sector– United Kingdom the first one of this list since the invent of James Watt-, and particularly in the second half of the nineteenth century, leading to the development of transport networks and a major increase in urban population. Others developed infrastructures before the industrial sector (Rietveld and van Nierop, 1995). That is the case of the Netherlands, where ports acted as transshipment points of international and intercontinental scale, and consequently, the country developed railway lines to connect merchandises inland. In the latter case, freight transport growth occurred before industrial development and population growth.

Growth or decline of cities in this context, especially port cities, has been well documented by historians, such as across the Atlantic (Konvitz, 1994) and in Asia (Murphey, 1969). Geographers soon proposed spatial models of transport network development in developing and developed countries under the concept of “port system” (for a recent review, see Ducruet and Notteboom, 2023). The main idea behind this concept is the long-term traffic concentration among ports situated in proximity along a given coast or “range”, especially at times of innovation diffusion in the transport sector. The development of cities next to railway stations has been studied as well by a variety of geographers in the cases of France, Spain, Italy (Brons, Givone and Rietveld, 2009; Mojica and Marti-Henneberg, 2011; Alvarez, Franch and Martí-Henneberg, 2013; Baron, 2015), Portugal (da Silveira et al. 2011), United Kingdom (Schwartz, Gregory and Thevenin, 2011; Stanilov, 2012), and the Netherlands (Koopmans, Rietveld and Huijg, 2012), to name but a few.

Beyond local differences, a relative consensus has been reached among scholars about the importance of both ports and railways in the rapid development of cities during the late 19th and early 20th centuries (Bretagnolle, 2015). Ports in particular had to “race for constant adaptation” to keep their position in an increasingly competitive environment (Marnot, 2005). Attracting maritime trade became more and more dependent on their ability to connect inland markets efficiently. Such dynamics had the effect of expanding the hinterland boundaries of successful gateways (e.g., Rotterdam and the Ruhr hinterland), at the expense of numerous, less-equipped nodes, like French ports for instance (Merger, 2004).

But, in order to better understand these effects, scholars somewhat neglected the existence of inter-network externalities, and the combined role of networks on urban development. The proposed research reconstituted the edges composing the global railway network in 1925 using geomatics, on the basis of printed historical maps. Port nodes and urban nodes are attributed to this network using additional urban databases and maritime atlases, together with population data. On the other hand, the global maritime network is constructed using Lloyd’s List printed data on inter-port vessel movements, which was extracted via an OCR software developed by the team. The preliminary results obtained in this specific case seem to confirm the hypothesis presented previously: proximity to transport is generally considered to have influenced population distribution, especially after the advent of the railway network. To have a hybrid network is more beneficial than having one or none. Negative effects were reported for regions that were not favourable to transport or for areas that had already had a dense network some years ago.

The paper is structured as follows: a brief literature review about cities and multiplex networks is presented in Section 2. Then, Section 3 describes the cartographic model along with an analytical analysis of its influence over degree nodes and clustering coefficient distributions. Section 4 presents the main results of the analysis of hybrid networks and population. Finally, conclusions about the effect of infrastructures over nodes are presented.

2. Cities in interconnected (transport) networks

2.1. Inter-network externalities and urban development

Specialists of urban networks originally dedicated their efforts to study the degree of interconnectedness among cities connected by numerous linkages of different nature and scale, in the 1950s and 1960s. The concept of urban network has changed depending on the approach and the scale adopted (see Peris et al., 2018; Derudder, 2019; Derudder and Neal, 2019 for useful and recent reviews). Cities have been studied in general terms as a way of organization, joined by economy, culture or politics. Numerous schools of thought emerged as well as different paradigms regarding urban networks, from the local-national to the global, and from graph theory in geography to complex network in physics for instance (Peris, 2016; Derudder, 2019). Due to limited data availability about flows, transport networks have often been the main material to study urban networks, considering their topology as well as their spatial or non-spatial structure (Barthelemy, 2003; Gastner and Newman, 2004; Barthélemy and Flammini, 2009). Since the 2000s, geographers and scientists of other domains have

increasingly used complex network approaches to analyze graphs in a geographical context (Ducruet and Lugo, 2013; Ducruet and Beauguette, 2014; Guimera et al., 2005; Krings et al., 2009; Kaluza et al., 2010; Neal and Ronzeblat, 2021).

What characterizes the current literature about cities and transport networks is the absence of node attributes such as socio-economic features (e.g., population, employment, value added) and the specialization of the approach on one network only. Recent research, however, has been done on the global maritime network in the age of steam, notably examining the relationship between maritime connectivity, technological innovation, and urban development (Ducruet and Itoh, 2022), but leaving aside the land-based network and ignoring inland cities. It did confirm, however, that the spread of steam shipping had been closely related with city size. Some parent works investigated such dynamics in other contexts, such as inter-network externalities among ports, canals, and roads in England between 1760-1890 (Bogart, 2014; Bogart, You, Álvarez, Satchell and Shaw-Taylor, 2020), or the combination of airline and other transport networks in Southeast Asia in recent years (Dai and Derruder, 2014 or 2016). Some scholars participated in the increased quality of urban network visualizations, but did not address the relationship between centrality and local population (Chapelon, 2006; Nelson, 2008; Lambert et al., 2013).

When it comes to understand the structure and evolution of multiplex networks - nodes connected by two or more links of a different nature, we have to deepen the review towards other domains (D'Agostino and Scala, 2014; Garas, 2016). Measuring and analyzing intersections between networks has been and continues to be a major challenge for researchers and practitioners (Ducruet and Berli, 2018). In that sense, we want to shed light on how the different transport networks are interacting, if there is a relationship between network diversity, urban hierarchy, and technological advance. We consider that technological innovation should impact more strongly the cities with multiple types of networks, in other words, inter-network externalities.

2.2 Research question and related hypotheses

From the recreation of a global maritime network and a railway network, we can focus on the main goal of this paper: to examine how land, sea, and land-sea connectivity correlate with urban population, for port cities but also for cities in general. It is therefore an examination of the nature of urban hierarchy and specialization, based on the hypothesis that the efficient combination of railways and shipping is more likely to profit urban growth than when considered apart. The independent influence of each network, ports and railways, on cities, also tackles the connection between the two networks, and whether and when railway expansion has improved port development or vice versa and how they interact together with population. If the networks are centralized and coupled in a hierarchical way, high-degree railway nodes will also be high-degree maritime nodes. Thus, the whole is more vulnerable than the independent parts (Parshani et al., 2010; Vespignani, 2010). This is less the case for a randomly-connected network, which will be more robust.

It echoes previous research on multilayer networks, such as the air-sea global network (Ducruet et al., 2011), with the originality to consider both planar and non-planar topologies. Existing research, such as on the sea-road combination (Ducruet and Berli, 2018), remained

static, with a vast majority of other works being more qualitative by their focus on actors and firms rather than network architecture (see Woodburn, 2013). After presenting the main procedures for modelling a global land-sea network, we propose an application to United Kingdom ports, railways, and cities as a first step in this direction. We have chosen the United Kingdom since it is the cradle of the Industrial Revolution, being the first country to develop steam power, and since it is an island, making it easier to analyze by avoiding an arbitrary “cut” of the transport links with the rest of the territory.

3. Data and methodology

3.1. Constructing the hybrid network and creation of attributes

The construction of the sea-rail global network starts with the cartography of the respective maritime (1880-1925) and railway (1925) global networks. The port is the focal node connecting the two networks along coastlines and rivers. Our objective is to shed new light on how the interconnected networks had an impact on the growth of nearby cities, compared with other cities and those connected by only one type of network.

For accomplishing this, first of all, we reconstituted the global railway network in 1925, digitalizing historical maps using QGIS³. In this software, a manual work has been done to recreate the railway edges over the Open Street Map layer, together with two types of nodes: stations related to cities near railways, and intermediate junctions like crossroads. Thus, the railway network is defined in this article by the infrastructure; while the maritime network is made of inter-port vessel voyages (see Figure 1).

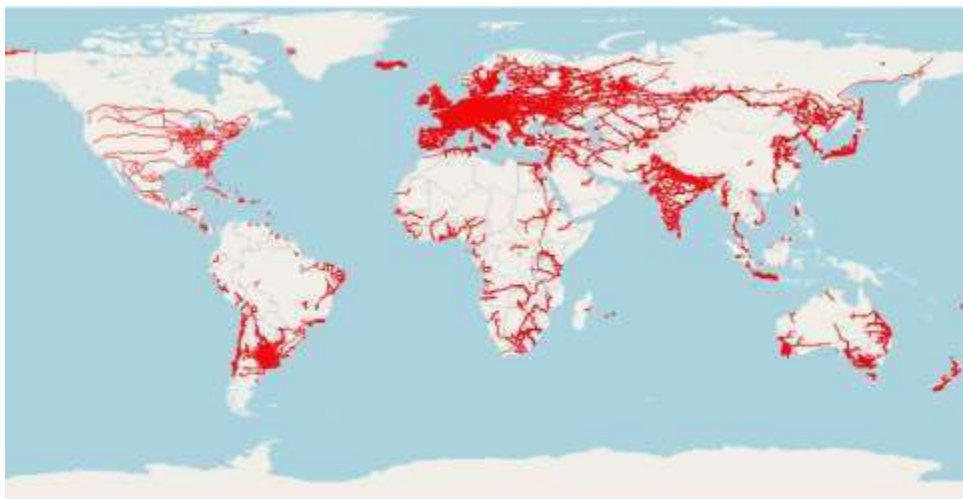


Figure 1: The global railway network in 1920

We then collected the shipping data of ports and vessel calls. For that purpose, we used the Lloyd's List corpus, specifically the Lloyd's Shipping Index, which provides inter-port vessel movements on a daily basis for the world fleet between 1880 and nowadays (see Figure 2).

³ Maps from different international institutions and libraries have been used, such as David Rumsey Library, Library of Congress or Bibliothèque National de France.

From that list, we took into account mainly two types of data: ports, host cities, the number of vessel calls, and the type of ship (steam or sail), between 1880 and 1925 (see Figure 3).

Reg.	STEAMER	Flag	Tons	From	For	Latest Reports
•	Haidrus	No	576	Newport Apr 8	Villa Real Pd	Oitavos Apr 14
•	Heighington (W.H.S.Nav.)	Br	1827	Oran Apr 2	Hampton Roads Pd	Gibraltar Apr 3
•	Helm	No	1916	Lisbon Mar 12	Puerto Plata Ar	Apr 5
•	Helm (Jacobson)	No	757	—	Bangkok Ar	Mar 31
•	Helmburg (Hunao)	Ge	2674	La Plata July 19	Hamburg Ar	St Vincent (CV) Aug 4—In
•	Helmdal	No	761	Clyde Apr 21	Gagnoa	(pt Nov 18)
•	Helmland I	Sw	568	Manchester Apr 15	Gothenburg Pd	Dunnet Hd Apr 18
•	Helmland II	No	209	Fredrikstad	—	Pd Hanstholm Apr 10
•	Helms	No	270	Rotterdam Apr 14	Fredrikstad	—
•	Heinrich	No	1797	Fort de Franco Apr 4	New York	Apr 15
•	Heinrich & Robert	Da	180	Tees Apr 14	Copenhagen Pd	Elsinore Apr 17
•	Heinrich & Robert	Ru	1124	Trading Black Sea & Asov		

Figure 2: Extract from the Lloyd’s List, 1915



Figure 3: Cities and transport networks

To avoid the visualization of vessel movements between ports as straight lines (Ducruet and Bunel, 2018), a recreation of the worldwide meshing thus had to be made. This has been possible thanks to the maritime grid previously created. Such a model allowed identifying possible improvements depending on the evolution of routes. The Grass software helped to solve the problem of representing three-dimensional features in a two-dimension surface, and to connect the grid within the Pacific Ocean, due to the focus on Europe in most maps using the Mercator projection (Figures 4 and 5).

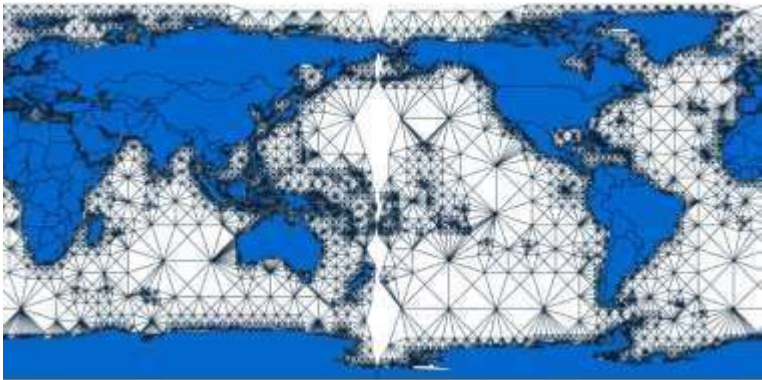


Figure 4: Connection mesh problem



Figure 5: Resolving the unconnected mesh problem

After that, we calculated the distance of each segment and created a topological network with the worldwide mesh. To that mesh, ports were added using 1.5 degree tolerance, to connect ports on islands, since it is optimized for the mainland coast. From this, the study of connectivity, accessibility and centrality measures become possible. The essential elements that compose such a spatial system are the urban centers/ports and the connections between them, that is, in topological terms, the nodes and arcs of the graph. The nodes are constituted by the ports, which allows including all the urban and semi-urban centers in the topological network. Arcs are defined by the layout of the ports that connect the different nodes of the network.

We measured for each edge the distance as cost, and disabled certain shortest paths that were not realistic for the maritime transport of goods, such as certain rivers (i.e., Volga, Rhine, Danube), the Panama Canal (opened in 1914 onwards), the Dead Sea, and the Arctic. As the exported vector of the mesh is not continuous – nobody can go around the world continuously, it has "extremes"-, we repeated the previous steps with the inverse vector. For example: America's continent first at one end and then at the opposite.

Once we settled up the mesh with nodes and edges, centrality measures became possible with R in order to obtain centrality. It indicates the greater or lesser structural complexity of the network, which is directly linked to the number of nodes, arcs and their spatial arrangement (Kansky, 1963). The centrality, accessibility or connectivity measures allow analyzing the spatial organization of the network, so the nodes establish a hierarchy based on the ease of access of each one to the rest of the nodes of the graph. The concept of accessibility is diverse and there are several measures to assess it. In general, it is defined as "the sum of the relative opportunities for contact and spatial interaction from the system as a whole" (Calvo Palacios, 1993). However, the location of the network must be interrelated with other agents such as distance, costs, and the time required.

In this work, only the distance has been considered as a function of the nodes and arcs of the network, without taking into account other variables, so that, in subsequent studies, the

method can be improved. The initial step for this analysis is the preparation of an accessibility matrix, where the topological distance is reflected by the shortest path between the nodes of the graph. Several accessibility measures are also deduced from this matrix.

In R, we used Lloyd's data to get a list of unique paths, always keeping in mind that if there are A - B and B - A records we aggregate them into a single undirected path. After obtaining the list, we use Grass to calculate the shortest path of each route using the two maritime networks and the distance travelled as cost. From each pair of shortest paths - for the same route- we discarded the longest path obtained, leaving us with a single path (for example we make sure that a LA-Shanghai route runs directly along the Pacific, not through Magallanes - South Africa - Indonesia)

We transferred the obtained shortest paths to R and used them as links, together with the ports as nodes, to build an undirected graph with the igraph package. The last step was to calculate the centralities, which included degree, betweenness, closeness, eigenvector, edge betweenness - absolute and normalized. Finally, we performed the representation in GRASS using:

- A measure of centrality of nodes (ports) obtained in R.
- The geometry of the maritime paths obtained on the grid in GRASS, smoothed with the drake's method to obtain less angular lines.
- The number of journeys on a certain route obtained from the Lloyds database (vessel calls)

The result is the possibility of create maps, in this article with the example of the United Kingdom (Figure 6), for which the hybrid network is represented using different values. Traffic and distance are used to calculate the centrality or importance of different nodes – such as railway stations or ports-, but that could be implemented with other variables such as freight or velocity.

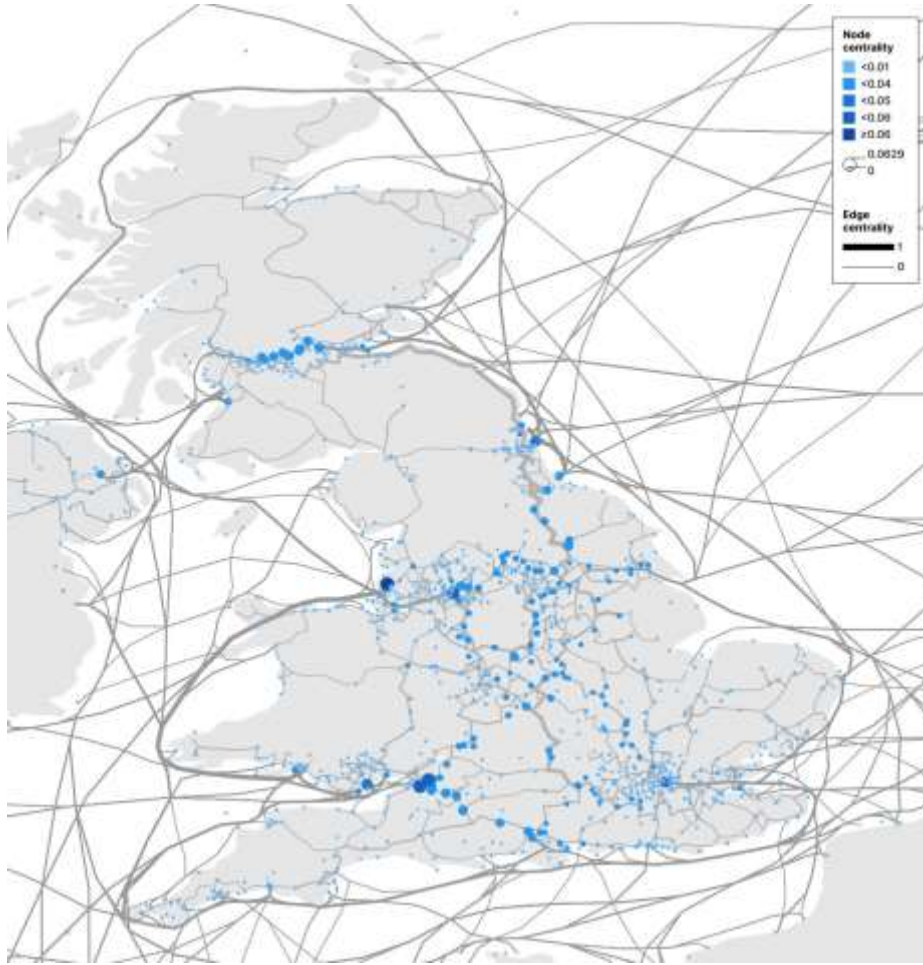


Figure 6: Map of the hybrid UK network in 1925

3.2 Statistical tests

Once we had data related to centralities of cities with ports, with railway stations and both of them, and data population over time (1880-1925) ⁴, it was possible to study the impact of hybrid networks on city size and city growth. We started with simple analyses and we increased the complexity in the following sections.

In order to explain the physical connectivity between infrastructures and between cities and infrastructures –in other words, inter-network connectivity–, different assessments were made: comparison of distance between networks, assortativity method, Pearson correlation, and ratio measures.

The first one allowed studying the relationship between maritime network and railway network, whereas assortativity is a simple and practical way to evaluate how several networks or layers are mutually interdependent. Pearson correlation allowed knowing how many cities grew up as consequence of having a network or a hybrid one. To have evidence that there is a relationship between centrality increase and growth in the population, we have conducted a hypothesis test with the following null and alternative hypotheses:

⁴ The main database used for urban population is Populstat. It was completed by national archives for the UK

Null hypothesis: $p=0$

Alternative hypothesis: $p \neq 0$

Following common statistical notation, we used the Greek letter ρ to denote the population Pearson correlation coefficient. The null hypothesis tells that the two variables are not correlated. The alternative hypothesis says that the two variables are correlated, that there is some linear relationship, either positive or negative. The not-equal sign in the alternative hypothesis implies that this is a two-tailed test, so either positive or negative Pearson correlation coefficients significantly far away from zero will result in the rejection of the null hypothesis.

The best way to understand the correlation between city size and centrality is to normalize the results. From the results obtained, we have been able to calculate the ratio in R and put it in a graphic way in GRASS. Maps allow seeing if all the previous steps made correspond with reality.

After all these assessments, and in order to confirm the hypotheses proposed in this research, different complex methods were applied. First of all, and due to the large number of variables, we decided to run two different multivariate analyses as exploratory analyses, a hierarchical clustering and a Principal Components Analysis (PCA), to know better the correlation between the variables and which ones explain better city size. In order to get better results in the clustering and the PCA, a logarithmic transformation of the variables was applied. It was used the equation $\text{sign}(x) * \log(\text{abs}(x))$ to be able to apply it to negative numbers (for negative growth).

Variable	Unit	Type
Population Growth	% of thousand people	Quantitative
Population Slope	Regression over population growth	Quantitative
Type of infrastructure		Qualitative
Maritime betweenness mean	betweenness	Quantitative
Maritime betweenness slope	slope	Quantitative
Maritime degree mean	mean	Quantitative
Maritime degree slope	slope	Quantitative
Railway betweenness mean	mean	Quantitative
Railway betweenness slope	slope	Quantitative
Railway betweenness mean	mean	Quantitative
Railway degree slope	slope	Quantitative
Combined betweenness mean	mean	Quantitative
Combined betweenness slope	slope	Quantitative
Combined degree mean	mean	Quantitative
Combined degree slope	slope	Quantitative

Table 1 Variables calculated for the study

Given a large set of observations characterized by the information of variables (see Table 1), we set ourselves the challenge of being able to classify them in such a way that the individuals belonging to a group (cluster) as similar to each other as possible. For that, we chose k-means. In the same way, we ran a PCA that diminishes the complexity of datasets while preserving data covariance. This analysis allowed discovering interrelationships among the data and according to the results, propose the most appropriate statistical analyses. As well, a cluster technique was made over the PCA results to check which variables explain population data. Also, the PCA graph was rescaled in logarithmic scale.

Besides, we calculated a Correlation matrix and the Kaiser-Meyer-Olkin (KMO) Sample Adequacy Measure. The KMO measure, which oscillates between 0 and 1, allowed us to know that a factorial model was possible. Kaiser (1974) recommends not accepting a factorial model if the KMO is less than 0.5.

After that, we ran a Linear Regression of urban population every five years. For that purpose, we used Ordinary Least Squares (OLS), a model specified in urban economics. The OLS is a type of global regression model that observe the (non)spatial relationships between the set of control and response variables with the fundamental assumption of homogeneity and spatial non variability. In that case we used it to analyze infrastructure levels and their effects on future population changes (some examples of research made by with this model were Rietveld and van Nierop, 1995; Tayman, Smith and Rayer, 2011; Duranton and Puga, 2014; Ward and Gleditsch, 2018; Oshan et al., 2019; Bogart, You, Alvarez, Satchell and Shawn Taylor 2020; Sadigov, 2022).

As well, an OLS was applied to cities with centralities, which were divided by categories: cities, cities with port, cities with railway station, and cities with a hybrid network. To have clearer results, cities without population data were excluded from the study, resulting at the end in 6 cities in the first category, 155 with railway stations, –only 89 have centrality, the rest obtained zero centrality, and 36 with multimodal infrastructure.

Our baseline specification is a cross-section growth equation as follows:

$\Delta \ln pop_{ij, 1925-1880} = \beta * Station_{1851i} + \gamma \cdot xi + cj + \varepsilon_{ij}$ where the dependent variable $\Delta \ln pop_{ij, 1925-1880}$ is the difference in log 1925 and log 1880 population for unit i in country j.

The main variable, *station* or port, is an indicator that equals 1 if unit i has one station or port and zero otherwise. In other words, the control group is all cities with station or port between 1880 and 1925. The idea is that cities grew up more if they had an infrastructure next to them. Due to positive net migration, having a railway station or a port in a unit is predicted to cause its population to grow more than in units without railway access all else equal.

The control vector *xi* always includes first nature characteristics and the natural log of unit population density every five years to capture effects of base year levels and prior trends. In preferred specifications, second nature characteristics and fixed effects *cj* are added as controls. The standard errors are always clustered on cities. All results were compared in a graphic way to better see how much each class of city grew.

The final step was to calculate the CAGR (Compound Annual Growth Rate), which expresses the growth of a variable regarding different years. Usually, this algorithm is

applied to economics, but in our case, it is very useful to calculate the growth of population between 1880 and 1925. The results allow us to compare them with the results obtained from the OLS based on population and the OLS based on centrality.

4. Results

4.1. Network-level analysis

This section reports the main results of the studies in terms of the impact of railway and port networks on cities in the case of United Kingdom. As previously mentioned, the analysis of connectivity and centrality allows us to know the structure of the networks, in an isolated and combined way. The map (figure 5) shows the results of the different indices as well as the disaggregation of the network. In all cases, it is accepted that during the analysis of connectivity, the sections pass through a maximum of 5 nuclei with respect to the ports in a radius of 10 km on the map, attending the average shown in different studies about the distance between stations, ports and residencies during the first half of the 19th century (Bogart et al., 2020).

In many cities of the United Kingdom, the degree of connection is almost null in the maritime network (990 cities - our database has 1072 cities- and with negligible values. Regarding the alpha index, the value equal to zero corresponds to incoherent graphs - and shows similar results to the previous one, that is, the values of the port network prevail. In summary, the structure of the network shows the high connectivity in the case of the integration of the port's set infrastructures in coastal areas; however, when each network is considered independent of the rest, the degree of cohesion decreases notably. That occurs when analyzing a country that is detached from the whole. On the other hand, inland network created from the railway, results of connexions were better: 586 cities connected.

The average spatial centrality, obtained through calculation algorithms, makes it possible to know the position of the network nodes (Ducruet and Marnot, 2017). The minimum value shows the most central point of the network and the mean accessibility indicates the distances and spatial relationships of the different nodes. In the maps, the centrality with respect to the central formed is observed in size. On the contrary, the marginality of the central nodes of United Kingdom is clearly distinguished, although in this aspect it is necessary to remember the closed nature of the graph, since it does not consider -because in some cases it does not exist- the relationships between all the parts of the land, and the conditions of the physical environment. At the same time, this cartography allows us to appreciate the proximity located in the example. Large cities are peripheral because of their link to ports (figure 6).

4.2 Inter-network connectivity

One first attempt to evaluate the relationships between the two networks is to measure their physical connectivity. A query calculated the distance between each port and the closest railway station in the worldwide railway network. In Figure 7, should it be for sail

vessels or steamers, the number of calls is inversely proportional to the distance. The largest number of calls occurs within ports, in the class of zero kilometers. The farther away from railways, the lower is port activity. This proximity has not much changed over the study period, as the different classes of distance remain separated from each other in similar ways between 1880 and 1925. The spatial distribution of steamer traffic among classes is slightly clearer than for sailing ships, and the gap between 5-9.9km and 10-24.9km is wider for steam than for sail.

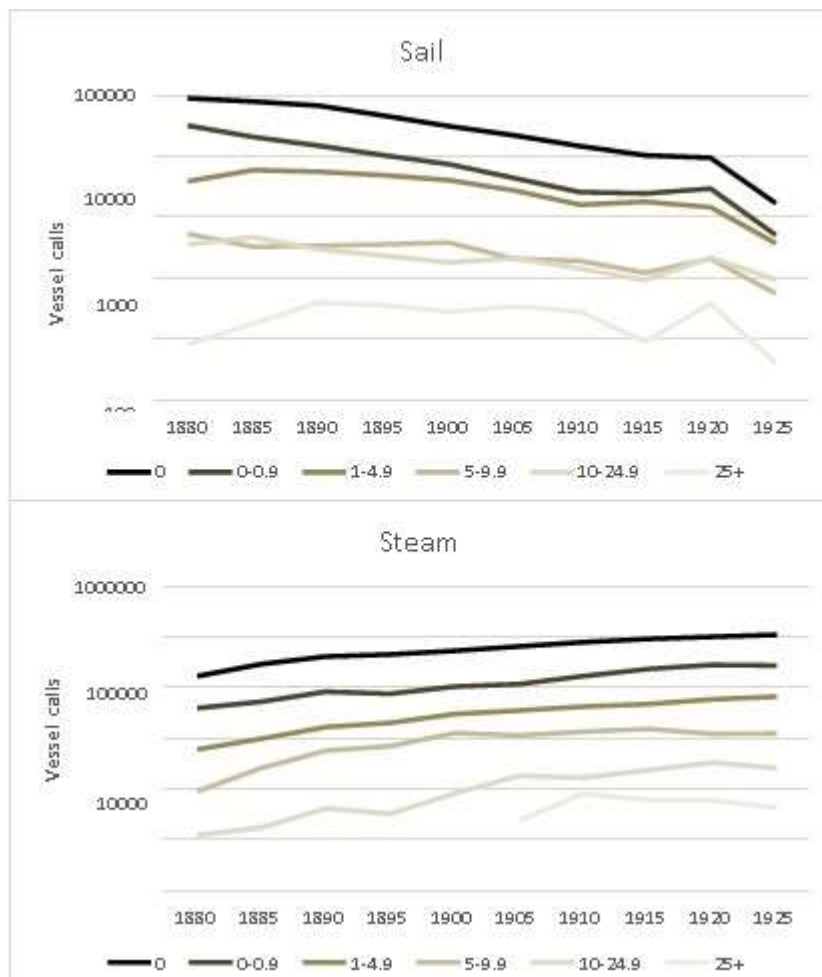


Figure 7: Vessel traffic and distance to railways by type of ship, 1880-1925

A second step is to apply the method proposed by Parshani et al. (2010), close to the idea of assortativity, in the case of the global airline and maritime networks. For the United Kingdom only, we considered the degree centrality (number of links of each node) in each network and calculated the Pearson correlation between them. As seen in Figure 8, the so-called inter degree-degree coefficient oscillated between 0.50 and 0.65 along the period. This is a very high score overall, and the evolution does not show a specific trend. At the level of UK, the principal maritime and railway nodes are thus relatively the same locations, and they connect the respective networks in similar ways. This also implies that there is a high level of geographic overlap, since the two networks are spatial.

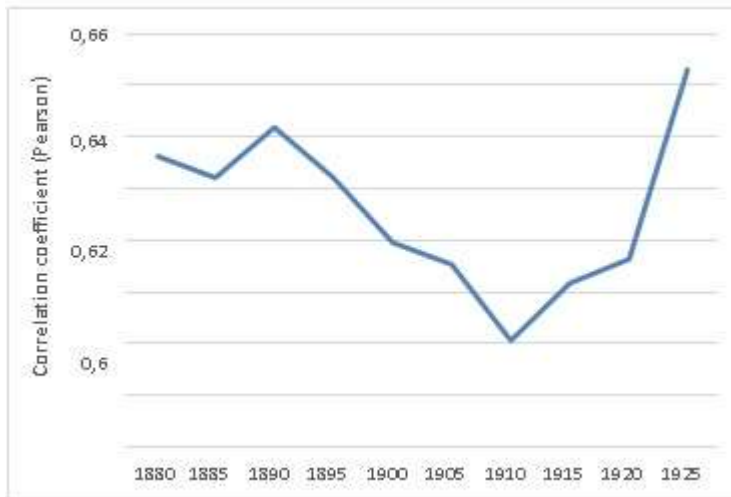


Figure 8: Inter degree-degree coefficient in the UK sea-rail network

In order to know how this inter-network connectivity interacts with population growth, the next step was to calculate correlation measures between population and degree of centrality. The results obtained seemed to agree to the previous one. The overlap between hybrid networks and population increase in many cases (89 over 195 cities showed a strong correlation), but in others, there is a negative correlation between an indicator for having railway lines pass/ports through a unit between 1880 and 1925 and an increase in population (figure 9).



Figure 9: heatmap of correlation between population and centrality

The reason could be that from 1880 to 1925, the distance to the railway line had an effect on land-cover change in the United Kingdom. Once the boom of railway had reached its maximum (1850s), city size seems to have a stable growth and maintain a nucleation

pattern until 1915 (Stanilov, 2012). Surprisingly, proximity to ports seems to impede urban uses as residency in some cases. This could be due, according to the authors, to the fact that rail does not enable within-city displacements but rather serves "long-distance interurban" commutes (Luo and Wei, 2009).

Due to this last item, and in order to have a clearer vision, we have calculated the ratio between combined centrality and single centrality. This step has been especially useful for comparing areas that are not uniform in size or population. In Table 2 we show the first 20 biggest cities in each period with population and ratio scores for maritime and railway centrality. The retained centrality is the betweenness centrality, a global accessibility measure.

City	1880	score maritime	score railway	name	1900	score maritime	score railway	name	1920	score maritime	score railway
London_GBR	3815	810.1	3.58	London_GBR	4537	1033.4	4.04	London_GBR	4485	712.1	3.15
Liverpool_GBR	552	1132.0	15.88	Glasgow_GBR	762	17394.0	4.54	Glasgow_GBR	1034	30258.7	5.92
Glasgow_GBR	512	2621.9	4.45	Liverpool_GBR	702	988.2	11.59	Birmingham	922	-	1.19
Birmingham	401	-	1.67	Birmingham	653	-	1.30	Liverpool_GBR	803	1079.8	13.28
Manchester_GBR	342	18457.0	1.58	Manchester_GBR	544	2437.4	4.15	Manchester_GBR	730	2000.0	4.65
Leeds	309	-	1.76	Leeds	429	-	2.26	Sheffield_GBR	491	-	1.39
Sheffield_GBR	284	-	1.63	Sheffield_GBR	409	-	1.24	Leeds	458	-	2.49
Edinburgh_GBR	228	244.1	6.35	Belfast_GBR	349	3539.5	205.42	Edinburgh_GBR	420	544.8	22.01
Belfast_GBR	208	3067.2	164.31	Bristol_GBR	329	3734.4	6.59	Belfast_GBR	415	2583.5	196.07
Bristol_GBR	207	3999.4	6.05	Edinburgh_GBR	305	1504.4	34.00	Bristol_GBR	377	5882.2	6.59
Bradford	183	-	1.99	Bradford	280	-	4.68	Bradford	291	-	4.80
Hull_GBR	154	434.6	26.62	Hull_GBR	240	390.5	36.71	Hull_GBR	287	284.5	4.80
Stoke-on-Trent	153	-	26.62	Nottingham	240	390.5	36.71	Newcastle upon Tyne	275	284.5	42.15
Newcastle upon Tyne	145	6957.3	6.90	Newcastle upon Tyne	215	889.4	8.74	Nottingham	263	635.1	7.85
Dundee_GBR	143	2716.4	6.90	Stoke-on-Trent	215	-	0.85	Portsmouth_GBR	247	3675.4	4.71
Portsmouth_GBR	128	0.0	4.00	Leicester	212	-	1.71	Stoke-on-Trent	240	-	1.03
Leicester	122	-	1.79	Portsmouth_GBR	188	22900.0	3.92	Leicester	234	-	1.64
Sunderland	116	-	4.00	Bolton	168	-	0.15	Plymouth_GBR	210	955.0	2.53
Nottingham	112	-	1.81	Cardiff	164	782.2	5.83	Cardiff	200	781.7	2.53
Oldham	11	-	2.16	Dundee_GBR	161	3698.3	3.32	Croydon	191	-	0.16

Table 2: Ratio scores

As an example, London in 1880 is 810 times more central in the combined network than in the maritime network alone. By comparison for the same year, modal specialization and geographic location explain for a large part the much higher score of certain other cities, especially in Scotland (Glasgow, Dundee) and the north of England (Manchester, Newcastle). Such cities are thus very central in the railway network, as their combined centrality overwhelms maritime centrality. In other words, the moderate score goes to cities that are well positioned in both networks, such as Liverpool and Hull. For the railway score and all years, Belfast surpasses all other cities for the prominence of combined versus railway centrality. This expresses the special case of being in Northern Ireland, i.e. with a limited railway accessibility compared with "mainland" UK. Maritime centrality thus appears as a vital complement to railway centrality, to palliate the relative peripherality of the place. It is followed by Liverpool in 1880, and a number of eccentric cities such as Hull, Dundee, Newcastle, Sunderland, and Edinburgh.

In 1900, Portsmouth stands out for the maritime ratio, which expresses the relative importance of railway. It is followed by Dundee, Belfast, Bristol, and Manchester. The same results are observable in 1920. For the railway ratio, the combined centrality shows that Liverpool, Edinburgh, Hull, Nottingham, and Newcastle greatly benefit from the

maritime network in 1900, with a similar order in 1920.

4.3 Intermodal connectivity and city size

4.3.1 Statistical testing

Our hypothesis was reinforced, in first place, by a Principal Components Analysis (PCA) and an Ordinary Least Squares (OLS) model. But, first of all, in order to know if it was possible to apply an OLS model, we did different suitability tests of variables, whereby we obtained a relevant Bartlett's Test of Sphericity (Table 3) and a KMO (Table 4) higher than 0.5. Besides, the result of the Reliability Analysis (Table 5) was 0,97, that revealed a higher agreement between items. These results indicated that it was possible to apply an OLS model.

Assumption Checks Bartlett's Test of Sphericity		
χ^2	df	p
2749	6	< .001

Table 3 Bartlett's Test of Sphericity

KMO Measure of Sampling Adequacy	MSA
Overall	0.719
Pop_Growth	0.633
Pop_Slope	0.852
combined_bet_mean	0.864
Port	0.833
Train	0.666
railway_deg_mean	0.658
maritime_bet_mean	0.637
combined_deg_mean	0.650
railway_bet_mean	0.685

Table 4 KMO Measure of Sampling

Adequacy Scale Reliability Statistics	
Cronbach's α	
scale	0.974

Table 5 Scale Reliability Statistics

4.3.2 Modal diversity and urban growth

The first exploratory analysis of variables allowed knowing how the created variables worked together (Figure 10). It is evident that for various combinations of variables, differentiated groups can be formed (Figure 11). Apparently, the growth slope – the regression made from the population growth- has better results with the railway degree mean – because of the lack of data related to railway traffic-, maritime betweenness mean and combine betweenness mean. Knowing this, it was justified to try to run a K-means cluster classification method. The results were three differentiated groups.

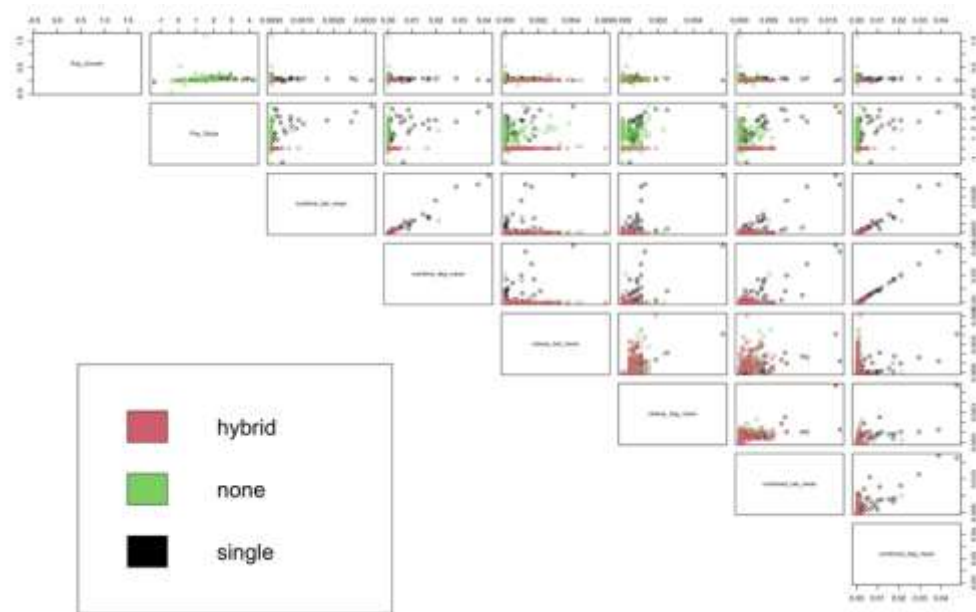


Figure 10 Exploratory analyses with variables

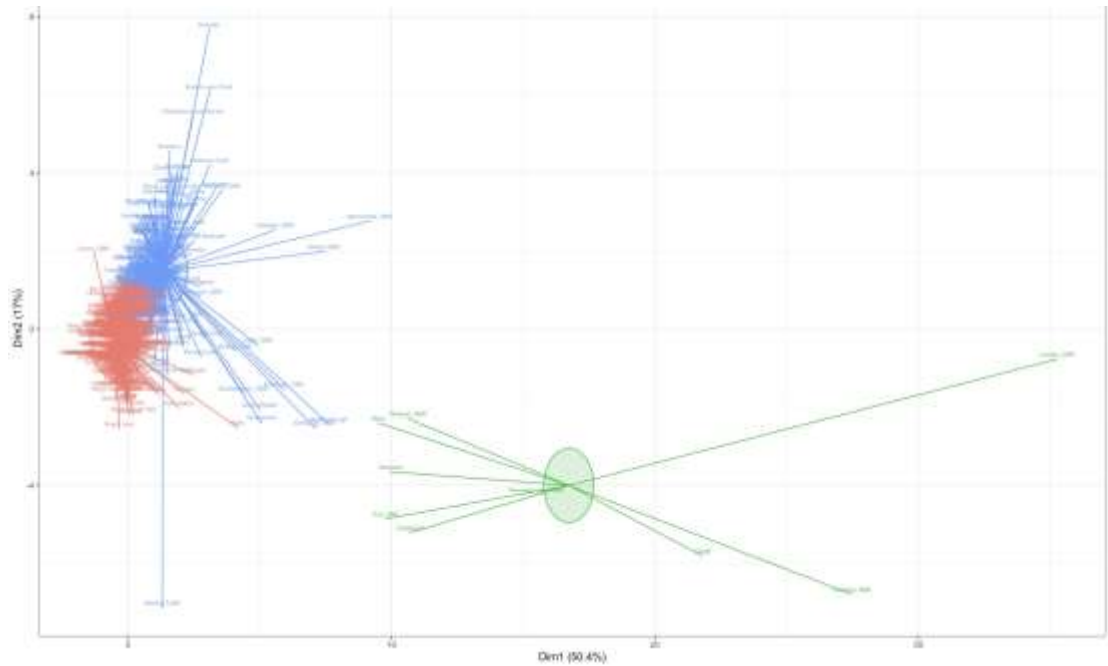


Figure 11 Clusters from variables

Although grouping variables is useful for analyzing aggregate data, the larger the number of variables, the more demanding the process. Therefore, PCA becomes an important tool to identify and use only the variables that contribute most to the variability in the data and confirm the hypotheses presented (figure 12)

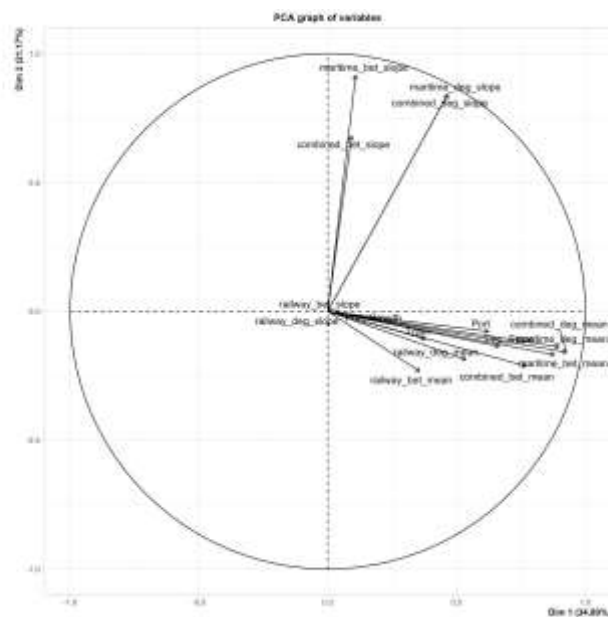


Figure 12: PCA

The PCA (Figure 12) let us know which variables were more suitable for running different analysis and confirming the proposed hypothesis. As it is shown, population growth, the slope of growth population, the presence of different infrastructures and different analysis of centralities are correlated. After that, the K-Means method was combined with PCA (figure 13). The two components showed account for 95% of the variability in the data, so using these two elements was sufficient.

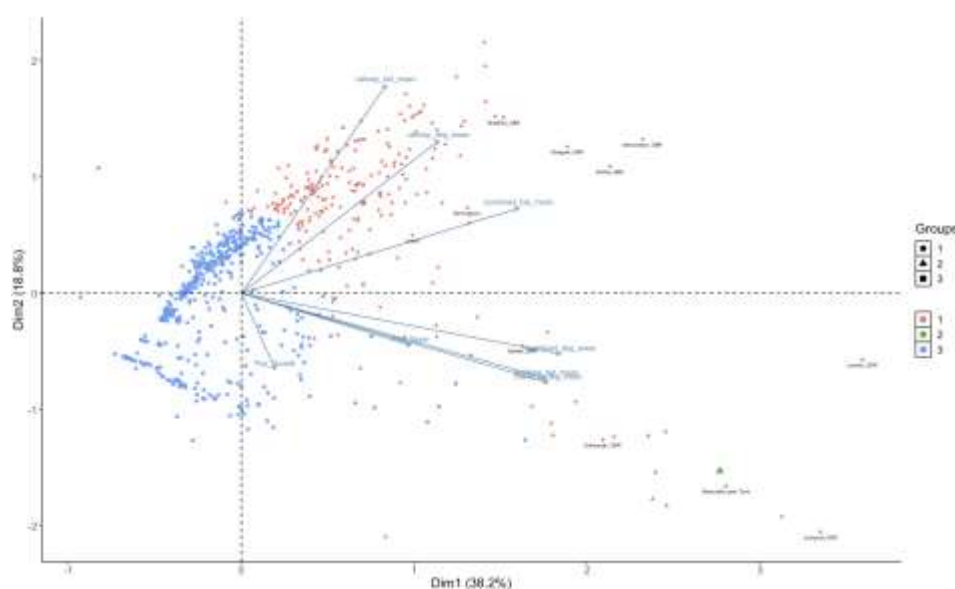


Figure 13 Clusters from PCA

It is clear from figure 13 that there is some degree of conflict between clusters 2 and 3, as there is some overlap between them. These representations are a mixture of clustering and PCA, where indicates how each variable is represented in each component. As it was obtained previously, population slope is influenced by three variables: railway degree mean, maritime betweenness mean and combined betweenness mean. These results were strengthened by the OLS model. The Linear Regression of cities' population every five years showed that structures of any kind – port or train- had a great impact on population slope.

Model Coefficients - Pop_Slope				
Predictor	Estimate	SE	t	p
Intercept ^a	-0.801	0.157	-5.09	< .001
combined_bet_mean	527.898	124.506	4.24	< .001
maritime_bet_mean	6788.352	843.881	8.04	< .001
railway_deg_mean	1609.603	325.395	4.95	< .001
railway_bet_mean	-390.780	207.529	-1.88	0.060
Type:				
multimodal – city	6.611	0.677	9.76	< .001

port – city	1.942	1.743	1.11	0.026
train – city	5.165	0.311	16.61	< .001
^a Represents reference level				

Table 5 Linear Regression from variables. Silhouette Method

As it is showed in the diagram (Figure 14), the mean of each OLS result shows a clear impact of infrastructures, especially the hybrid network, over city size and growth. It is possible to say that the combination of networks has a bigger influence than each mode isolated.

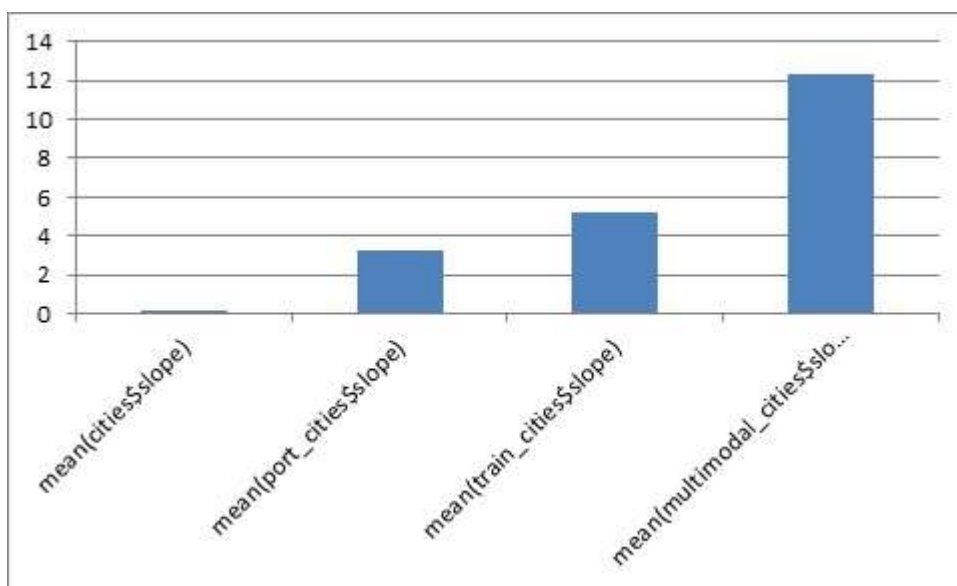


Figure 14 Diagram from OLS

The next step was to do a comparison between centrality of cities with these results. The outcomes were also illustrative: having a port or a port and a railway station had a great impact over population. On the other hand, having only a railway station seemed not to have an influence of any kind. An important result is that cities with infrastructure grew up more than deprived ones. Regarding the results of centralities Linear Regression, the first results were confirmed: the growth was stronger in cities with ports or railway and ports, compared with those which had only a railway infrastructure, which growth was slower during this period. The results were logical because data from vessel calls is available, whereas from railway movements of each city station don't. We are certain that with all data collected from that infrastructure, results would be even more positive (figure 15).

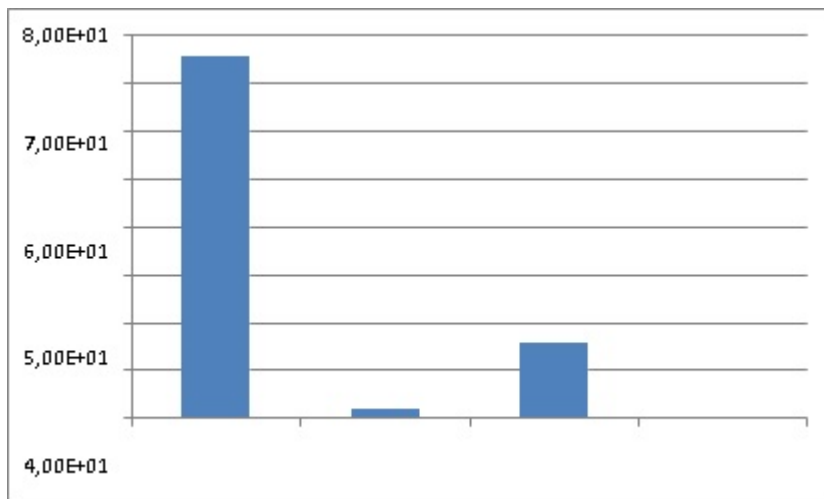


Figure 15 Centralities taking into account betweenness

All these results are supported by the Growth Annual Mean. We have calculated it with data population available and a Compound Annual Growth Rate (CAGR) algorithm for each available city (Figure 16). We have classified cities with infrastructures and without them (MAG_cities). Finally, only 195 cities with at least two consecutive years were available over the list of cities we have.

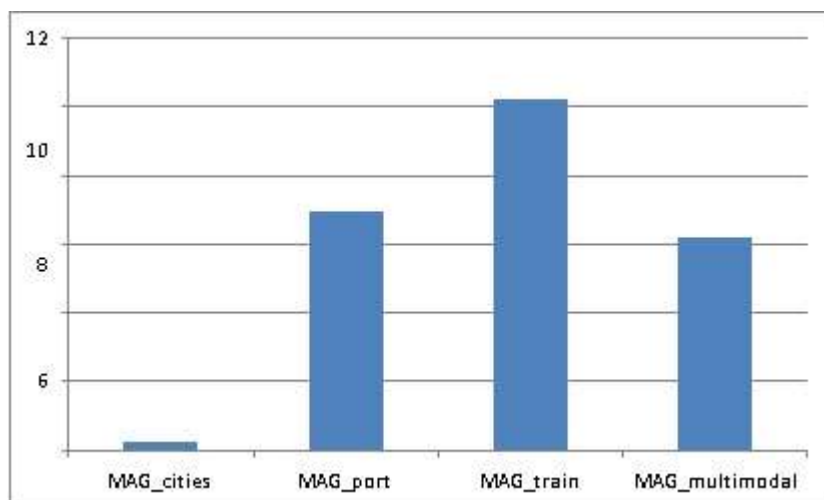


Figure 16 Mean Compound Annual Growth Rate

Conclusions

This paper systematically examined the influence of railroads and ports on cities in the age of steam. It examined how population changed over time as a result of the development of the various modes of transportation and the role of external factors such as traffic of ports. The study's contribution to the field is threefold. First, it provides a synthesis of recent empirical findings on cities, transport networks and diversity. Second, it zooms on the case of the United Kingdom, which by its situation as an island is more likely to witness a correlation between the observed layers. Third, it focuses specifically on the effects of a hybrid network on city size and growth.

Proximity to rail transport is generally considered to have influenced population distribution. Negative effects were reported for regions that were not favorable to transport or for areas that had already had a dense rail network for some time. Proximity to the railway has also encouraged conversion to residential areas and the development of higher density housing. However, the evidence on the role of rail in increasing employment density is inconclusive, suggesting that its success depends more on exogenous factors such as ports and the attractiveness of the area, which are particularly favorable in coastal areas.

As for the maritime network, studies generally suggest that the presence of or proximity to major ports is associated with conversion to urban land, increased employment density and commercial and industrial development. However, this is not always the case for residential purposes, suggesting that living in close proximity to these areas may be unattractive (Ducruet et al., 2022).

Of the studies that have examined both rail and port access, almost all of them found high coefficients for rail access compared to port access, regardless of the period studied. However, it should be noted that many studies focus mainly on the changes in the second half of the twentieth century and after, when the rail network is considered to have lost its initial influence.

As we have shown in this paper, the presence of any kind of infrastructure has a great impact over growth population. Econometric analyses have allowed us to analyse deeply the relationships between infrastructures and the connexion of a hybrid network with population size. The use of different quantitative variables confirms the presented hypothesis.

Further research may complement this work by an analysis of the combined global rail-sea network. The time period also should be extended to better observe long-term evolutions, with a snapshot of the global railway network in 1950, 1980, and 2010, together with port, maritime, and urban data.

References

ABDEL-RAHMAN, H., ANAS A. (2004) Theories of systems of cities. In: Henderson J.V., Thisse J.F. (Eds.), *Handbook of Regional and Urban Economics*, pp. 2293-2339.

ALVAREZ, E., FRANCH, X. and MARTÍ-HENNEBERG, J. (2013). Evolution of the Territorial Coverage of the Railway Network and its Influence on Population Growth: The Case of England and Wales, 1871–1931, in *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, pp. 175-191.

BARKE, M. (1986), *Transports and trade*, Oliver and Boyd, Edinburgh.

BARON, N. (2015). Railways and demographic change, in M. Finger and P. Messulam (eds). *Rail Economics, Policy and Regulation in Europe*, Elgar, chapter 2., pp. 22-38.

BARTHÉLEMY, M. (2010). "Spatial Networks." *Physics Reports*, 499: 1– 101.

BERRY B.J.L. (1964) Cities as systems within systems of cities. *Papers in Regional Science*, 13(1): 147 163.

BOGART, D. "The Transport Revolution in Industrializing Britain," in Floud, Roderick, Jane Humphries, and Paul Johnson (eds.) *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation, 1700–1870*. Cambridge University Press, 2014.

BOGART, D; YOU, X.; ÁLVAREZ, E.; SATCHELL, A.E.M. ; SHAW-TAYLOR, L. (2020) Railways, divergence, and structural change in 19th century England and Wales.

BRETAGNOLLE A. (2015) City systems and maritime transport in the long term. In: Ducruet C. (Ed.), *Maritime Networks: Spatial Structures and Time Dynamics*, Routledge Studies in Transport Analysis, Routledge: London and New York, pp. 27-36.

BRONS, M; GIVONE, M; and RIETVELD,P (2009). Access to railway stations and its potential in increasing rail use, *Transportation Research Part A: Policy and Practice* Volume 43, Issue 2, February 2009, pp. 136-149

BURTON, I. (1963). The quantitative revolution and theoretical geography, *Canadian Geographies*, vol. 7, issue 4, p.. 151-162.

CAMAGNI R.P., SALONE C. (1993) Network urban structures in northern Italy: elements for a theoretical framework. *Urban Studies*, 30(6): 1053-1064. Capello R. (2000) The city network paradigm: measuring urban network externalities. *Urban Studies*, 37(11): 1925-1945.

CHAPELON, L. (2006). "Accessibility as a Marker of Disparities of Influence among Port Cities in Europe." *Cybergeo: European Journal of Geography*, 345.

CRAFTS, N., and MILLS, T.C. (2004) Was 19th century British growth steam-powered?: the climacteric revisited, in *Explorations in Economic History*, Volume 41, Issue 2, pp. 156-171.

DA SILVEIRA, L. E., ALVES, D., LIMA, N. M., ALCANTARA, A., and PUIG, J. (2011). Population and railways in Portugal, 1801–1930. *Journal of Interdisciplinary History*, 42(1), 29–52.

DERUDDER B., NEAL Z. (2019) Uncovering links between urban studies and network science. *Networks and Spatial Economics*, 18: 441-446.

DERUDDER, B. (2019). "Network Analysis of 'Urban Systems': Potential, Challenges, and Pitfalls." *Tijdschrift voor Economische en Sociale Geografie*.

DUCRUET C., BEAUUGUITTE L. (2014) Network science and spatial science: review and outcomes of a complex relationship. *Networks and Spatial Economics*, 14(3-4): 297.

DUCRUET, C. and BERLI, J. (2018). Measuring land-sea interactions at ports and cities : insights from geomatics and network analysis. In *Hal SHS*.

DUCRUET C., CUYALA S., EL HOSNI A. (2018) Maritime networks as systems of cities: The long- term interdependencies between global shipping flows and urban development (1890-2010). *Journal of Transport Geography*, 66: 340-355.

DUCRUET, C.; BERLI, J. SPILIOPOULOS, G. ; and ZISSIS, D. (2021). Maritime Network Analysis: Connectivity and Spatial Distribution. In *Guide to Maritime Informatics*, Springer, pp. 299-317.

DUCRUET, C., LUGO I. (2013) Cities and transport networks in shipping and logistics research. *Asian Journal of Shipping and Logistics*, 29(2): 149-170.

DUPUY, G. (1987). Les réseaux techniques sont- ils des réseaux territoriaux, *L'espace Géographique*, 3, París.

DURANTON G., and PUGA, D. (2014). The growth of cities. In *Handbook of economic growth* (Vol. 2, pp. 781-853). Elsevier.

GIMENEZ Y CAPDEVILA, R. (1986): La Geografía de los transportes, en busca de su identidad, *Geocrítica* 62, p. 63, Barcelona.

GINSBURG, N. (1961) *Atlas of Economic Development*. The University of Chicago, Department of Geography, Research Paper No. 68. Chicago

GIPOULOUX F. (2011) *The Asian Mediterranean: Port Cities and Trading Networks in China, Japan and Southeast Asia, 13th-21st Century*. Cheltenham and Northampton, MA, Edward Elgar Publishing. HAGGET, P (1965): *Análisis locacional en la Geografía Humana*, Gustavo Gili, Barcelona

KALUZA P., KOLZSCH A., GASTNER M.T., BLASIUS B. (2010) The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48): 1093-1103.

KANSKY, K.J. (1963). *Structure of transport networks: relationships between network geometry and regional characteristics*, Research papers, 84, Department of Geography, University of Chicago.

KIM, S. (2005). Industrialization and urbanization: Did the steam engine contribute to the growth of cities in the United States?, in *Explorations in Economic History*, Volume 42, Issue 4, pp. 586-598.

KOOPMANS, C., RIETVELD, P., & HUIJG, A. (2012). An accessibility approach to railways and municipal population growth, 1840–1930. *Journal of Transport Geography*, 25, 98–104.

KRINGS G., CALABRESE F., RATTI C., BLONDEL V.D. (2009) Urban gravity: a model for intercity telecommunication flows. *Journal of Statistical Mechanics: Theory and Experiment*, 7: L07003.

LAMBERT, A., BOURQUI, R. and AUBER, D. (2013). Graph visualization for geography. In C. Rozenblat and G. Melançon (Eds.), *Methods for multilevel analysis and visualisation of geographical networks* (pp. 81-102). Springer.

LO, J., and WEI, Y. H. D. (2009). Modeling spatial variations of urban growth patterns in Chinese cities: The case of Nanjing. *Landscape and Urban Planning*, 91(2), 51–64.

MARNOT, B. (2015). *Les villes portuaires maritimes en France. XIX-XXIème siècle*. Armand Collin.

MOJICA, L., and MARTI-HENNEBERG, J. (2011). Railways and population distribution: France, Spain, and Portugal, 1870–2000. *Journal of Interdisciplinary History*, 42(1), 15–28.

NEAL, Z.P. and ROZENBLAT, C. (2022) *Handbook of Cities and Networks*. Research Handbooks in Urban Studies series. Elgar.

NELSON, A. (2008). Estimated travel time to the nearest city of 50,000 or more people in year 2000. Global Environment Monitoring Unit - Joint Research Centre of the European Commission, Ispra Italy.

ODLYZKO, A. (2010) *Collective hallucinations and inefficient markets: The British Railway Mania of the 1840s*. University of Minnesota.

- OSHAN T.M., LI Z., KANG W., WOLF L.J., STEWARD FOTHERINGHAM A. MGWR: A python implementation of multiscale geographically weighted regression for investigating process spatial heterogeneity and scale. *ISPRS International Journal of Geo-information*. 2019; 8(6)
- PARSHANI, R. ;ROZENBLAT, C.;IETRI, D.; DUCRUET, C.; and HAVLIN, S. (2010) Inter-similarity between coupled networks, *Physics > Data Analysis, Statistics and Probability*, Cornell University.
- PERIS A., MEIJERS E., van Ham M. (2018) The evolution of the systems of cities literature since 1995: schools of thought and their interaction. *Networks and Spatial Economics*, 18: 533-554.
- PRED, A. (1977) *City Systems in Advanced Economies: Past Growth, Present Processes, and Future Development Options*. Wiley.
- RAIL DELIVER GROUP (2018), 'Rail Freight – Working for Britain', June.
- RIETVELD P., and van NIEROP, J. (1995) Urban growth and the development of transport networks: the case of the Dutch railways in the nineteenth century, in *FLUX Cahiers scientifiques internationaux Réseaux et Territoires*, 19 pp. 31-43
- RODRIGUE, J.P. (2014). " Conceptual Corridor Development. The Geography of Transport Systems." <https://people.hofstra.edu/geotrans/eng/ch2en/conc2en/corridordev.html>
- RODRIGUE J.P., DUCRUET, C. (2020) Graph theory: definition and properties. *The Geography of Transport Systems*, https://transportgeography.org/?page_id=5976
- ROSO, V. and WOXENIUS, J. (2009): The dry port concept: connecting container seaports with the hinterland, *Journal of Transport Geography*, 17(5):338-345
- SADIGOV. R. (2022). Rapid Growth of the World Population and Its Socioeconomic Results, in *Scientific World Journal*.
- SEGUI PONS, J.M., y PETRUS BEY, JM. (1991). *Geografía de redes y sistemas de transportes*, Síntesis, Madrid.
- SHIBASAKI, S.; HIRONORI, K. and DUCRUET, C. (2021). *Global logistics network modelling and policy*. Elsevier, Amsterdam.
- STANILOV, K. (2013). Planning the growth of a metropolis: Factors influencing development patterns in West London, 1875–2005. *Journal of Planning History*, 12(1), 28–48.
- SCHWARTZ, R., GREGORY, I.,and THEVENIN, T. (2011). Spatial history: Railways, uneven development, and population change in France and Great Britain, 1850–1914. *Journal of Interdisciplinary History*, 42(1), 53–88.

TAYLOR P. J., HOYLER M., Verbruggen R. (2010) External urban relational process: introducing central flow theory to complement central place theory. *Urban Studies*, 47(13): 2803-2818.

TAYMAN, J., SMITH, S.K., and RAYER, S. (2011) Evaluating Population Forecast Accuracy: A Regression Approach Using County Data, in *Popul Res Policy Rev.* 2011; 30(2): 235–262.

TOVAR, B. and WALL, A. (2022) The relationship between port-level maritime connectivity and efficiency. *Journal of Transport Geography*, Volume 98.

WATERS, N. (2006). " Network and Nodal Indices. Measures of Complexity and Redundancy: A Review." In *Spatial Dynamics, Networks and Modelling*, edited by A. Reggiani and P. Nijkamp, 13– 33. Northampton, MA: Edward Elgar Publishing

WOLKOWITSCH, M (1992): *Géographie des transports*, Armand Colin, Paris.

XIE, F., and D. LEVINSON. (2009). "Modeling the Growth of Transportation Networks: A Comprehensive Review." *Networks and Spatial Economics*, 9(3): 291– 307.

ZANON MOURA, T.C.; GARCÍA-ALONSO, L. and SALAS-OLMEDO, M. H. (2017). Delimiting the scope of the hinterland of ports: Proposal and case study, *Journal of Transport Geography* Volume 65, pp. 35-43

Declarations

Data availability

The data that support the findings of this study are available from Lloyds but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of César Ducruet, director of Magnetics.

Funding

All sources of funding for the research reported come from the ANR project 'MAGNETICS' 2023-2026

Authors' contributions

"BP performed the historical context and analysed and interpreted the population data regarding PCA and OLS models and made the figures, and was a major contributor in writing the manuscript. CD performed complex network context, made tables and was a major contributor in writing the manuscript. All authors read and approved the final manuscript."