3

RAILROADS AND POPULATION DISTRIBUTION

HGIS data and indicators for spatial analysis

Eduard J. Alvarez-Palau and Jordi Martí-Henneberg

Introduction

This work looks at the impact that the railroad network has had on the distribution of population and urbanization. It fits into what constitutes quite a wide space for reflection about how transport infrastructure has shaped territorial change throughout history.¹ The analysis of the railroad network is particularly relevant in this respect given its rapid expansion throughout Europe from the mid-nineteenth century onwards. The transport network established was completely different from any previous (road, navigable waterways and ports) infrastructure and its efficiency for transporting passengers and goods over long distances left it without any direct competition. In fact, this remained the case until the appearance of the automobile, whose hegemony was not definitively established until the 1950s. There was therefore a significant period (1840–1940) during which the railroad served as the main means of transport wherever it had been introduced.

It is generally accepted that the railroad is essential for understanding the current geography of population and production. Indeed, a number of classical contributions have convincingly shown this interrelationship. In general, academic works coincide in the belief that the provision of rail services facilitated commerce, helped to specialize production, and promoted phenomena such as the integration of markets, movement of people, urbanization and tourism, all of which are elements that have been important in the configuration of our modern society.²

Along these lines, we explain what GIS can add to traditional approaches to the analysis of transport networks. In this chapter, we show that new georeferenced databases can help consolidate the geographical scope of historical studies of transport networks, due to their capacity to integrate data relating to various themes, and provide relevant analyses. This territorial dimension can be used to complement the temporal dimension as the datasets used often cover the last 150 years. In this way, it is possible to propose new perspectives from which to study history and to develop a field that is central to the digital humanities.
In this chapter, we present the indicators that we have developed to analyse both the network itself and its interrelationship with population. In this way, we seek to open new lines of work that complement, but do not seek to replace, traditional approaches to these subjects. It is still, however, necessary to carry out archive work, undertake local monographic studies and make global reflections. Only in this way will it be possible to study interaction between the different dimensions of the global railroad network and other disciplines such as architecture, literature, engineering, and art. Within this context, it will be possible to give meaning to the quantitative results provided by Historical-GIS (HGIS).

The aim of our research is to improve our understanding of the relationship between railroads and the distribution of population in Europe. We have done this by analysing case studies referring to specific geographical areas and periods and by using a number of different methodologies. To achieve this, it has been necessary to enter the field of HGIS and to incorporate new databases that permit quantitative and comparative studies conducted at different scales. This has made it possible to show the complex and diverse interrelationship between increases in population and access to rail services.

The following section includes the state of the art and the most important results obtained from other studies. This is the starting point for obtaining a better understanding of what the use of HGIS databases can contribute to the analysis of phenomena that have a long tradition in the study of the humanities and social sciences.

**State of the art**

We discuss two different questions in this section. First, we examine databases that provide information about railroads and population in Europe from a national and historical perspective. Secondly, we refer to the most significant work that has been undertaken in order to try to quantify the impact of the railroad network on population and urbanization. Both approaches focus on the use of HGIS for spatial analysis; the bibliography cited is therefore relatively recent. Before evaluating the potential contribution of this new approach, we should perhaps first consider some of the earlier contributions to this line of research. The work of Paul Bairoch is particularly relevant as it considers improvements in supplies to major cities using the railroad as a main factor explaining their growth since the nineteenth century. Although economic historians have tried to measure the impact of the railroad in terms of the social savings that it offered the societies into which it was introduced, they have often come to different conclusions about its relevance. It has only been since the emergence of the new economic geography (initiated by Krugman in 1991) that the influence of transport has assumed a clearer role within the field of econometrics. In this field, the growth and distribution of population (the market) and the capacity to access transport are seen as two key explanatory factors.

It should be underlined that many countries have recently made important investments in the creation of HGIS databases at the national scale that are of great relevance. From a demographic perspective, these databases tend to contain information derived from national censuses referring to the nineteenth and twentieth centuries. It is, however, preferable to georeference these data with respect to the
administrative boundaries corresponding to each historical period as these have undergone numerous changes over the past two centuries. It should be underlined that this is quite a difficult task and one that, to date, has only really been accomplished at a detailed scale for England and Wales. For the moment, the resulting datasets only tend to show the reality during each specific historical period. However, undertaking long-term comparative studies is much more challenging. For this reason, significant effort has been dedicated to designing approaches that allow the use of existing data as part of a prior approach for exploring long-term change.

In this article, we have resisted the temptation to make a detailed analysis of the monographic studies of railroads that have been carried out to date, as this has already been done in several earlier publications. These publications provide explanations of such aspects as: the initial cartographic databases; how the different elements and their attributes were georeferenced; the digital layers contained in the homogenized information; and the descriptive analyses of the main results obtained relating to the evolution of the railroad network from 1830 until 2015. Other networks, such as that of navigable waterways, have already been identified and georeferenced for specific periods. In the future, it would be relevant to add further information relating to main roads, ports, and airports in order to maximize the explanatory potential of these models.

In the academic arena, the influence of railroads on modern society and economic development has mainly been studied by economists, geographers, and historians and from a predominantly national and regional perspective. In these studies, discussion has mainly centred on determining whether the influence of railroads on territory has been more or less beneficial and on finding new models for obtaining variables that would make it possible to quantify these interrelationships.

In contrast, the influence of railroads on the geography of population has remained relatively forgotten. Although several interesting initiatives have already been undertaken, more analytical studies will be required if we are to obtain significant results. Georeferenced GIS databases and modelling instruments are now available and it will soon be possible to exploit the potential that they offer.

Studies like these require us to combine two different GIS elements: the railroad network and the distribution of population at the municipal level. As we shall see, the results obtained clearly show the impact that railroads had during the period immediately prior to the large-scale development of private, motorized forms of transport. In territorial terms, the results obtained also reveal the positive influence that railroads had on areas that were well-served by the rail network in comparison with those with deficient levels of accessibility.

The acquisition of new databases has facilitated the appearance of many new lines of research and presented different approximations and results. To date, the majority of such studies have focused on European countries: Finland, Sweden, France, Wales, and the UK.

In Spain, comparable population databases exist for the municipal level and allow very accurate HGIS and econometric analyses. Franch et al. used these data to compare and contrast accumulated population growth during the period 1900–2001 within areas served by rail transport. The results obtained show that urban areas
directly influenced by railroads either grew more, or declined less, than others and that rural areas suffered important losses of population. Barquín et al. also studied the influence of railroads on population during the period 1860–1910, taking into account a number of other variables that included industry, mining, and distance from the coast. They concluded that areas of Spain with rail connections grew by up to 60 per cent more than those without.

More recently, new work has begun which focuses on Europe as a whole. Caruana and Martí-Henneberg carried out an econometric analysis related to the density of the railroad network at the regional level, substituting GDP for population as an indicator of development. This study also demonstrated a positive correlation between the variables analysed.

Along the same lines, but this time outside Europe, this approach has contributed to methodological discussions that have made it possible to explore existing knowledge in greater depth. Works referring to the USA and South America have compared the availability of rail services to the growth of GDP, following the approach established by the economist G. Solow. They again highlight how railroads made a positive contribution to sustained growth in GDP.

Historical databases on railroads and population

The next section presents the most innovative aspects of compiling historical databases. This is a key consideration because the research currently underway and its potential depend on their scope and precision.

The data described refer to the European railroad network and to population trends in different European countries. All of this information was compiled within the framework of a wider HGISe project. It includes a compendium of different research projects that have been carried out in recent years in collaboration with various European universities.

Railroad data at the European scale

Of the different types of transport infrastructure available, that relating to railroads is probably the easiest to document as it is relatively easy to identify the routes that trains follow and have followed in the past. The fact that stations act as nodes of connection with settlements means that it is possible to locate the exact points at which passengers and goods enter and leave the railroad system. Furthermore, the need for powerful locomotives to move carriages and wagons means that the railroad can only operate as a collective form of transport. We can therefore identify two types of infrastructure that characterize the railroad from the perspective of it offering a transport service: the routes followed by the tracks and the locations of the stations.

The quality of railroad databases will depend on their capacity to include both of these variables. The more precise the data we possess, the more powerful will be the calculation models that we work with and the more reliable the indicators that constitute the study variables.
A first approximation to railroad databases at the European scale can be found within the framework of the HGISe project. This presents an exhaustive compilation of all the wide gauge railroads that have been operational in Europe from the 1830s through to the present day (Figure 3.1). Each line has been digitalized and georeferenced in order to facilitate its visualization and also to obtain certain key attributes via simple spatial calculations. A number of other attributes have also been incorporated, including the number of years that each line has been operative, the type of transport provided (for passengers, freight, or a mixture of the two), and other variables, such as the gauge of the track. This has facilitated the subsequent task of making spatial calculations.

Figure 3.1 (a and b) The wide gauge railroad network in Europe in the years 1850 and 1950.
The population variable from a historico-spatial perspective

Population is the second main subject studied here and considerable time and effort have been invested in obtaining new, updated datasets. Total population and population density are the only two variables that are perfectly comparable over time, with basic population data being available from official censuses. The majority of European countries began to collect population data at ten-year intervals in the middle of the nineteenth century. Although these databases have often been based on different systems of collection and classification, they have a common denominator: they include information about total population at the municipal level. This concept varies slightly from country to country, but always refers to the smallest local unit. It constitutes a single indicator of human activity as it includes geographically and
Table 3.1 Sources and availability of population data at the municipal level for the countries of Western Europe.

<table>
<thead>
<tr>
<th>Country</th>
<th>Name of units</th>
<th>Number of units</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria*</td>
<td>Gemeinde</td>
<td>2,357</td>
<td>1869–2011</td>
</tr>
<tr>
<td>Belgium*</td>
<td>Commune / Gemeente</td>
<td>589</td>
<td>1831–2011</td>
</tr>
<tr>
<td>Denmark (2)</td>
<td>Byer</td>
<td>1,451</td>
<td>1896–2011</td>
</tr>
<tr>
<td>England/ Wales (1)</td>
<td>Civil parish</td>
<td>11,847</td>
<td>1801–1861</td>
</tr>
<tr>
<td>England/ Wales**</td>
<td>Civil parish</td>
<td>11,847</td>
<td>1871–2001</td>
</tr>
<tr>
<td>Finland*</td>
<td>Kunta</td>
<td>336</td>
<td>1880–2011</td>
</tr>
<tr>
<td>France*</td>
<td>Commune</td>
<td>36,570</td>
<td>1801–2011</td>
</tr>
<tr>
<td>Germany (3)</td>
<td>Stadt / Gemeinde</td>
<td>11,710</td>
<td>1870–2011</td>
</tr>
<tr>
<td>Italy*</td>
<td>Comune</td>
<td>8,100</td>
<td>1920–2011</td>
</tr>
<tr>
<td>Luxembourg*</td>
<td>Commune</td>
<td>116</td>
<td>1821–2011</td>
</tr>
<tr>
<td>Netherlands(1)</td>
<td>Gemeente</td>
<td>441</td>
<td>1830–2011</td>
</tr>
<tr>
<td>Norway*</td>
<td>Kommune</td>
<td>430</td>
<td>1769–2011</td>
</tr>
<tr>
<td>Portugal*</td>
<td>Cidades / Freguesias</td>
<td>4,042</td>
<td>1864–2011</td>
</tr>
<tr>
<td>Spain*</td>
<td>Municipio</td>
<td>8,114</td>
<td>1877–1911</td>
</tr>
<tr>
<td>Sweden*</td>
<td>Stad / Församling</td>
<td>1,646</td>
<td>1810–1911</td>
</tr>
<tr>
<td>Switzerland (1)</td>
<td>Commune</td>
<td>2,624</td>
<td>1870–2011</td>
</tr>
<tr>
<td>Scotland (2)</td>
<td>Civil parish</td>
<td>870</td>
<td>1841–2011</td>
</tr>
</tbody>
</table>

Source: Own research based on primary sources available via municipal institutes of statistics.

historically comparable data for all of the different states considered. It is therefore possible to transfer this information to an HGIS, using georeference systems, and to assign attributes corresponding to each municipality.

Table 3.1 presents the primary sources for the countries of Western Europe that can be used for this type of analysis. These data have been collected and digitalized within the frameworks of different research projects.

The primary sources used in our HGIS were homogenized decade by decade, based on years ending in ‘0’. The total number of municipalities has constantly changed over time. Municipal data have been assigned to georeferenced points on maps based on their spatial coordinates. (*) The data presented have been both georeferenced and homogenized. (**) Only georeferenced data or only homogenized data. (1) Data in dataset format. (2) Data only available from original censuses. (3) Data in publishable format, but only for urban areas.

The contents of Table 3.1 will serve as the basis for future work requiring the use of HGIS. Total population figures serve as a comparable reference over time. Significant changes can be interpreted as indicators of economic and social transformations; these can then be examined in detail at a later stage. As can be seen from the section referring to our case studies, it is necessary to have detailed data (at least at the municipal level) in order to obtain relevant results. The problem is that this generally implies a great effort in terms of data collection and exploitation and constitutes an important methodological challenge.

The main indicators used to refer to population are its territorial distribution, whether expressed in absolute terms or as density per unit area, and its percentage
increases, which are always contemplated at ten-year intervals. These can then be used to create new variables that can be correlated with railroad indicators of different types. One variable that is intimately related to population and which assumes special relevance in this type of analysis is urbanization. The growth of cities, both in terms of their populations and urban areas, is something that can be shown and measured spatially. This is of particular interest when it comes to providing transport infrastructure to serve a city and its different neighbourhoods.

Based on all of these data and the different possibilities for their analysis, we propose moving on from merely describing the evolution of the geography of population in order to try to determine the factors that explain it. We have sought to do this based mainly on the railroad variable but we have also tried to integrate other indicators at the municipal level that are capable of capturing complementary information. In terms of geographical information, it is possible, for example, to incorporate variables that provide estimates of the altitude and ruggedness of the terrain and of distances from the coast. In terms of political data, we can also incorporate such variables as distances from national borders and/or whether settlements are located near to local or national capitals. Some of these factors have already been used as control variables in previous statistical analysis.

**Spatial interactions between railroads and population: case studies**

The main objective of this research is to identify population patterns in areas served by railroad networks. This requires the creation of detailed databases and the use of quantitative methods to demonstrate the relationships between the two variables.

In the present study, we identify four major groups of indicators. The first three are based on railroad infrastructure and the fourth on urbanization. First, we detail the classical indicators: length, density, and per capita stock of the railroad network. These can all be calculated quite easily, without the need for GIS, although georeferenced databases greatly facilitate both the calculation and presentation of the results obtained. Secondly, we characterize the indicators of railroad coverage. This dichotomous variable indicates whether a territory is covered by rail services. Thirdly, we describe the indicators of accessibility, which measure the time and/or cost required to travel between each pair of nodes within a given network. These indicators can be presented in an aggregated form, at the state level, or in a disaggregated way, relating to each individual city. However, this is quite complicated to calculate and may require the development of complex spatial models. Finally, there is another group of indicators that can be used in the opposite direction. They consider railroad infrastructure as an invariant and try to quantify urban growth in order to see whether it has been conditioned by either the route taken by a railroad or the presence of a particular station.

**Simple indicators: the length, density, and per capita stock of the railroad system**

The simplest and most commonly used indicator for determining the availability of railroad infrastructure is the total length of the network. This indicator is calculated by summing the total length of track in a given territory. It also makes it possible to observe the evolution of this infrastructure over time. At the European level, the
previously cited HGISe database has made it possible to calculate this indicator immediately and then to use it as a variable in later analyses.

However, if the aim is to compare results for different territories, it is evident that the length indicator is not sufficient on its own. European regions are very different in terms of their surface areas, populations, and economies and this complicates any form of direct comparison. As a result, there is a tendency to normalize the length of the network with respect to the total surface area of the region being studied (railroad density) and its population (per capita stock). This could also be done using economic indicators such as GDP (stock in railroad in relation to GDP). Adopting this approach has enabled us to make more precise and focused comparisons. 21

Figure 3.2 shows the considerable divergence between the two indicators. Railroad density reveals very different provisions of infrastructure between one country and another. The United Kingdom exhibits very high densities in all of the periods studied; this differs notably from what can be observed in all the other countries. France is the next country in terms of performance with respect to this indicator, despite the difference in surface area. Finally, more peripherally located countries, such as Spain and Finland, exhibit relatively low densities. What is interesting here is the fact that, in terms of density, we observe few variations in the rank position of each country over time; this therefore appears to be quite a stable indicator.

The second indicator shows a very different situation. Although the United Kingdom presents high per capita stock during the initial decades of railroad construction, by the end of the twentieth century it has been overtaken by both France and Finland. This could be explained by Britain’s strong demographic growth and by a more intense use of its network. We must also consider the physical dimensions of
Finland and its low density of population; this largely explains the major growth observed. Finally, Spain exhibits much lower and more constant levels of growth.

**Indicators of railroad coverage**

The coverage models seek to offer an index that permits a qualitative and quantitative classification of the distance between a given point within a territory and a particular piece of transport infrastructure. First, we can visualize whether this point is covered by the mode of transport in question. Then we can estimate the corresponding access time and the resulting utility for users. In recent literature, there has been a tendency to calculate coverage by drawing concentric circles around elements of infrastructure, such as railroad stations. The surface area, or population, located within this buffer is then considered to be covered by the network. The sum of all of these surfaces constitutes the total area with coverage within a given country or region and this is contrasted with that without coverage.

A railroad network is generally considered to provide coverage to a territory when it provides it with rail services. This occurs when it satisfies the following mathematical relationship:

\[ P_i \in C_{pf} \text{ if } \exists E_j \text{ so that it serves } P_i \]

Where:
- \( P_i \) represents the population of node \( i \).
- \( C_{pf} \) is a group that includes the total population receiving railroad coverage.
- \( E_j \) represents station \( j \).

The authors have developed an alternative way to approach this calculation. They did this by testing a coverage model for civil parishes for the whole of England and Wales from 1871 to 1931 (Figures 3.3a and 3.3b). In this case, five different levels of coverage were identified: 1) no railroad station within a two-hour walk from the parish centroid, 2) a nearby station, between one and two hours away, 3) a station less than one hour away, 4) one available station, 5) two or more stations available to the civil parish.

The interesting aspect of this case is the correlation between the indicator of coverage shown here and the variable relating to the increase in population (Table 3.2). In England and Wales, our analysis of parishes that did not experience any changes in coverage during the four decades between 1891 and 1931 shows that those with initially low levels of coverage (1, 2, and 3) tended to lose population (with average values of -0.62, -0.38, and -0.01 per cent, respectively). On the other hand, those with intermediate levels of initial coverage (4) tended to gradually gain population (0.34 per cent), while those initially covered by two or more stations gained population at considerably higher rates (1.07 per cent).

The main problem with this approach is that, to date, very few European countries have HGIS databases providing information about the opening and closing of stations. In the case of Spain, for example, there are only registers of the opening and closing of lines, but not of stations. It has therefore been necessary to introduce a number of simplifications when using this indicator. Franch et al. therefore used...
Figure 3.3  (a and b) Spatial distribution of the adapted indicator of railroad coverage for the English county of Norfolk between 1871 and 1931. 
Source: Alvarez-Palau et al. (2013).
The physical distance from the railroad as the base from which to estimate coverage.24 This approach permits an initial approximation when few data are available. However, in this case, it is necessary to carefully define the distance that is considered to denote coverage.

**Indicators of railroad accessibility**

Models of accessibility seek to generate indicators that incorporate the travel time required for trips undertaken within the railroad network. These are calculated by summing the total time required to make the complete journey. This includes the time required to access the railroad network, any possible waiting time, and the time required to make the journey. The mathematical formulation used to obtain this indicator is presented here:

\[ I_{GA_j} = \frac{\sum_{i=1}^{n} C_{ji}}{n-1} \]

Where:

- \( I_{GA_j} \) is the Index of General Accessibility calculated for node \( j \).
- \( n \) is the sum of the number of locations included in the study (universe).
- \( C_{ji} \) is the cost matrix of the cheapest way of travelling from node \( j \) to node \( i \).

Where:

\[ C_{ji} = t_{\text{origin}} + t_{\text{waiting}} + t_{\text{navel}} + t_{\text{destination}} \]

\[ t_{\text{navel}} = t_{\text{trip}_1} + t_{\text{transfer}_1} + \ldots + t_{\text{transfer}_{k-1}} + t_{\text{trip}_k} \]

The result of the calculation for each pair of nodes is saved in a cost matrix \( C_{ji} \) with \( n \) origins and \( n \) destinations (general accessibility). The average value of each row is represented by a discrete value that determines the average cost, or time, needed to access a particular node from the rest of the territory (specific accessibility).

---

**Table 3.2** Evolution of the average cumulative annual growth rate in E&W parishes, according to their respective levels of coverage, 1871–1931.

<table>
<thead>
<tr>
<th>Year</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
<th>Level 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1871–81</td>
<td>-0.69%</td>
<td>-0.48%</td>
<td>0.01%</td>
<td>0.52%</td>
<td>1.79%</td>
</tr>
<tr>
<td>1881–91</td>
<td>-0.68%</td>
<td>-0.53%</td>
<td>-0.16%</td>
<td>0.44%</td>
<td>1.23%</td>
</tr>
<tr>
<td>1891–1901</td>
<td>-0.94%</td>
<td>-0.76%</td>
<td>-0.53%</td>
<td>-0.06%</td>
<td>0.75%</td>
</tr>
<tr>
<td>1901–11</td>
<td>-0.25%</td>
<td>0.09%</td>
<td>0.47%</td>
<td>0.74%</td>
<td>1.26%</td>
</tr>
<tr>
<td>1911–21</td>
<td>-0.54%</td>
<td>-0.32%</td>
<td>0.00%</td>
<td>0.23%</td>
<td>0.70%</td>
</tr>
<tr>
<td>1921–31</td>
<td>-0.64%</td>
<td>-0.26%</td>
<td>0.14%</td>
<td>0.14%</td>
<td>0.66%</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td><strong>-0.62%</strong></td>
<td><strong>-0.38%</strong></td>
<td><strong>-0.01%</strong></td>
<td><strong>0.34%</strong></td>
<td><strong>1.07%</strong></td>
</tr>
</tbody>
</table>

Source: Alvarez-Palau et al. (2013).
Figure 3.4 (a and b) Index of specific accessibility calculated for the urban centre of Cambridge (for trips taking up to two hours) in 1931. On the right, it is possible to see the distribution of population in the same civil parishes.

Source: Own work. See HGIS for railways and Gregory et al. (2011).
Figure 3.4 (Continued).
Figure 3.5 (a and b) Correlation between distance from the centre of Cambridge and density of population in 1931. In (b), it is possible to observe the correlation between the index of specific accessibility and the density of population.

Source: Own research.
When we analyse historical data rather than current data, it is not always possible to appropriately calculate all of the terms in the equation and it is therefore necessary to introduce several simplifications. For example, it is not possible to calculate the access time to the station using surveys. Instead, it is necessary to estimate the average journey time from each location. Waiting time is a completely random variable as it depends on each individual and on the number of train services available. In current studies, authors have tended to use half of the frequency time, but when adopting a historical perspective, there is a tendency to remove this item from the equation. Travel time can be estimated with the help of rail timetables, but these are not always available; as a result, average values must also be used. The commercial speed, type of locomotion, and number of tracks per line also vary and are difficult to estimate. It is therefore necessary to propose approximate methods of calculations until current HGIS databases can be improved.

The indicator for the local scale has been adapted in order to capture the spread of accessibility from any specific node to its periphery. Figures 3.4a and 3.4b present a graphic representation of the index of specific accessibility of the English city of Cambridge. In this case, it is evident that municipalities with rail coverage exhibit lower travel time values than those lacking coverage.

As expected, it can be clearly observed how the nearest parishes concentrate the highest accessibility values. However, if we observe the grading of the different colours around the railroad infrastructure, there is evidence of induced territorial anisotropy. Parishes with railroad coverage have good levels of accessibility to the central node. In contrast, unconnected locations may be geographically closer, but could be considered to lack coverage and to be spatially marginalized.

What is interesting about this case is the cross between the railroad and spatial population indicators. As a central city, Cambridge exhibits much higher levels of population than the settlements on its periphery, where it is possible to detect a clear decline with distance from the centre. This phenomenon is shown in the first graph in Figure 3.5. However, if we use the accessibility variable instead of the distance from the centre, it is possible to observe how the logarithmic correlation increases the level of significance from 0.401 to 0.456. We also begin to detect examples of territorial polynuclearism as the most important locations become better interconnected.

This correlation makes it possible to study the spatial distribution of population within the territory resulting from the provision of railroad services. The resulting anisotropy of accessibility and population decay can both be clearly observed. However, this situation could be reversed by applying appropriate urban planning and transport policies favouring greater territorial homogeneity. Whatever the case, the contrast between accessibility and population remains a vital instrument for understanding such aspects of day-to-day life as urban mobility, the management of public services, and urbanism itself. Although the railroad does not have the same importance today as it did in the past, the same analysis could be applied to consider other transport infrastructure and the results obtained would probably be quite similar.

Indicators that refer to the urbanization process

The previous section identified the main indicators that can be applied when analysing the railroad network. This one focuses on indicators that can be related to the
Figure 3.6 Graphic representation of the indicator showing the direction of accumulated urban growth in the city of Vic.

Source: Solanas et al. (2015).

Urbanization variable in order to analyse how its evolution has been influenced by railroad infrastructure. In other words, we here use the railroad as an invariant to examine the spatial growth of cities. To do this, it is necessary to work with a powerful HGIS containing information relating to the process of urban expansion. Georeferencing historic urban maps has opened up an important line of new research. It is not difficult to document and measure urban growth if enough comparable maps are available to chart urban development over time.
Focusing on urban reconstruction, the HGIS makes it possible to measure urban growth using spatial vectors. By doing this, it is possible to establish the position of the geographic centre of a city prior to the arrival of the railroad and to calculate how the city has subsequently grown. It can show the location of each new area of a city and its total surface area. This indicator is calculated according to the following formula:

$$
\vec{I}_i = \frac{\sum_{i=1}^{n}(\vec{V}_i \cdot S_i)}{\sum_{i=1}^{n} S_i} \quad \text{for } i = 1, 2, \ldots, n
$$

Where:
- $\vec{V}_i$ is the vector between the urban centroid and the area of growth.
- $S_i$ is the surface area of each area of growth.

Applying this indicator to the Spanish city of Vic between 1860 and 2001 gives a set of results covering the period prior to the opening of its urban station and for the subsequent period (Figure 3.6).

It can clearly be observed how the accumulated growth of the city is aligned with the vector between the city centre and the station from the creation of the railroad network until the 1930s. In fact, the differences in the angle between the two vectors (accumulated growth and distance from the city to the station) during the first four periods are 10.62, 15.83, 16.80, and 29.5 degrees (over 180 degrees), respectively. From this moment onwards, growth becomes less dependent on the railroad. This can be shown by the emergence of autonomous sectors during the final phase that have more links with the use of the automobile than with other modes of transport. In this last stage, the difference in angle grows by up to 54.10 degrees. This shows the dissociation between urban growth and railroad infrastructure in the city.

In summary, once the railroad had been constructed, it was possible to observe how this infrastructure had a great influence on the city’s urban growth. The majority of the cities analysed by Solanas et al. showed clear growth towards their respective stations. As urban land was used up, due to this growth, there was a tendency for the city to expand around the track until the necessary crossings were provided. Finally, when society adopted individual forms of motorized transport, such as the private automobile, the continuous city model was broken and new areas of growth began to appear in the form of isolated polygons of constructed territory.

**Synthesis and final reflections**

The examples presented here demonstrate the important contribution that HGIS can make to understanding spatial phenomena from a historical perspective. The present contribution focuses on different interrelationships between railroad infrastructure and the distribution of population and demonstrates a strong spatial link between the two variables. To achieve this, we identified a series of indicators that can be applied to railroads in order to obtain explanatory variables and facilitate the understanding of the phenomena studied. Based on these results, accessibility to the railroad
highlights the territorial anisotropy created by transport networks and has consequently been revealed as a predictive variable for population concentration. At the same time, railroad coverage could also explain differential growth between locations related to their level of connection. However, this is not all. An analysis of urban growth in relation to infrastructure also shows how the urban fabric of cities has also been influenced by the routes taken by railroads and the locations of rail stations.

Even so, there still remain many improvements to be made. Associating the locations of stations with attributes relating to the years in which they opened (and possibly closed) would be of great interest and would help us to understand the interaction between the railroad network and the territory. Having more information relating to railroad timetables and the frequency of services would also provide us with knowledge of such aspects as the speed at which commercial services circulate (or once circulated) and the carrying capacities of different lines. Adding narrow gauge28 lines to this study would also represent an important milestone in the development of regional studies, as would the addition of data relating to metropolis tram and subway train networks. It would similarly be interesting to incorporate information referring to the demand for transport within the network, relating to both passengers and freight. Such indicators would also provide the basis for an initial approach to assess the economic profitability of each railroad system.

By the way of some final reflections, we should underline that the development of HGIS and its application has opened up new possibilities for understanding urban and demographic evolution in relation to transport infrastructure. The current work exclusively focuses on railroad and population, considering a number of specific indicators. The analyses presented here could, however, also be extended to other transport networks, such as highways, navigable waterways, ports, and airports. The study of each network requires appropriate databases, specifically designed indicators and the creation of tailor-made spatial models. The level of precision of HGIS therefore depends on the subsequent adaptation of its indicators and the possibility of providing reasonable simplifications. Having said that, ongoing improvements to databases mean that in the short to medium term it should be possible to apply much more sophisticated models. The data obtained from analyses can then be statistically or cartographically crossed with other spatial variables in order to establish correlations, contrasts, and/or econometric models. These should then allow us to establish relationships between different variables in our search for the maximum significance. It is necessary to continue working on this approach and to improve its results iteratively.

Notes


Alvarez-Palau and Martí-Henneberg


7 See: http://vps63872.ovh.net/projects/inland-waterways/.


20 This was based on different research projects led by J. Martí-Henneberg, which included: ‘HGISe: A Platform Analysing Transport, Population and Socioeconomic Data for Europe (1850–2010). The case studies of England & Wales and the Iberian Peninsula’ and ‘The Development of European Waterways, Road and Rail Infrastructures: A Geographical Information System for the History of European Integration’.


24 Ibid.


28 The majority of the railroad network was designed using Standard Gauges (1,435 mm in Europe, except in some countries where is slightly wider). Narrow Gauge lines (1,000 mm or less) tended to be used for regional purposes and were rarely integrated into the main railroad network.