HAL
open science

Measuring one century of railway accessibility and population change in France. A historical GIS approach

Thomas Thévenin, Christophe Mimeur, Robert Schwartz, Loic Sapet

To cite this version:

Thomas Thévenin, Christophe Mimeur, Robert Schwartz, Loic Sapet. Measuring one century of railway accessibility and population change in France. A historical GIS approach. Journal of Transport Geography, 2016, 56, pp.62-76. 10.1016/j.jtrangeo.2016.08.017 . hal-01361067

HAL Id: hal-01361067

https://hal.science/hal-01361067v1

Submitted on 4 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MEASURING ONE CENTURY OF RAILWAY ACCESSIBILITY AND POPULATION CHANGE IN FRANCE: AN HISTORICAL GIS APPROACH

Thomas Thevenin[1], Christophe Mimeur[1], Robert Schwartz[2], Loïc Sapet[1]
[1] ThéMA, UMR6049 CNRS, Université Bourgogne Franche-Comté, France
[2] Department of History, Mount Holyoke College, United States

This is a postprint version, the definitive version of this paper is:

Thévenin, T., Mimeur, C., Schwartz, R., & Sapet, L. (2016). Measuring one century of railway accessibility and population change in France. A historical GIS approach. Journal of Transport Geography, 56, 62-76. https://doi.org/10.1016/j.jtrangeo.2016.08.017

Introduction

This article presents an innovative study of the relationship between railway accessibility and population change in France from 1860 to 1930. Using a large, geo-historical database and spatial analysis, it identifies key advantages of historical GIS in transport history as well as its limitations. It illustrates the challenges of creating reliable data on population change for the 36,000 communes of France over eight decades and integrating them with the main and local lines of the evolving French railways system. In so doing, it reexamines the structuring effect of railway expansion over the long period in which rail transport was the predominant means of moving goods and passengers over long distances. The increasing speed of rapid transport--part of the "speed revolution" according to Studeny (1995) -- reshaped territorial spaces and spatial relationships, in effect, shrinking the space-time connectivity of communities with accessible rail service, and augmenting that of communities not well so served. The "conquest of space" by railways was thus not homogeneous, even, or continuous, but fractured, uneven, and discontinuous.

In France, the "speed revolution" witnessed a six-fold increase in locomotion. There, more notably, the railway system was planned firstly on a national scale as a radial system connecting Paris to major cities by the lines of state sponsored companies, each with its own designated territory. Gradually regional development extended the system as the companies constructed branch lines across their territories. (Caron, 1997). Then, beginning in the 1880s, narrow-gauge tertiary "lines of local interest" were added through a collaborative effort of the central government and local authorities. By the eve of the first world war, the combined system permitted passenger and goods traffic to move from remote rural communities to distant cities and the capital itself. To understand this evolving interconnectivity it is important, we think, to go beyond questions focused on cities and urban hierarchies (Pumain, 1997) to consider the changing spatial relations among national, regional, and local scales by using the national mosaic of French *communes* as the basic geographic units of analysis.

1. Putting the Census into a National H-GIS

The last ten years has witnessed the creation of National H-GIS in a number of countries around the world. They are covering large countries like China (Bol and Ge, 2005) and Great Britain (Gregory, 2005). Others were developed in the United States (McMaster et al., 2005), in Belgium (Vanhaute, 2005). These projects typically followed several common steps in development.

1. Identifying historically accurate boundaries of administrative units for which census and other attribute information was recorded.
2. Encoding the spatial boundaries of the units, with real-world geographic coordinates, in the H-GIS.
3. Assembling historical information recorded for the units: in our case, information on population, railway lines and stations, topography, agriculture, and other characteristics that become available.
4. Linking the historical attributes to the appropriate geo-referenced units.

The development of our H-GIS proceeded through these four steps as follows.

2.1. The E-Geopolis database

The E-Geopolis project is a reference database designed to capture demographic CHANGE in urban areas around the world from 1800 to the present. In Europe, demographic data have been collected at the finest scale of resolution for France (*communes*), Spain (*municipios*) and Great Britain (parishes). The French data were collected in work sponsored by a national program research, first for cities larger than 10,000 inhabitants and then extended to the entire country at the level of the commune.

For the French case, indispensable reference work was the book by Claude Motte and others, entitled *"Communes d'hier, communes d'aujourd'hui"* (Motte et al., 2003). Under Motte's direction, this historical administrative dictionary describes each of some 36,000 communes in which the census enumerations were made in the nineteenth and twentieth centuries. The work also offers informative commentaries on the history of the census, including the serious deficiencies of the first census of 1831, the substitution of nine-year census intervals owing to THREE wars (1870-71, 1914-18, 1939-1945), and the creation in 1946 of a new national bureau of statistics that has administered the census since then.

Despite of the wars and other administrative changes, the boundaries of the French communes were generally stable over the years. As exceptions to the rule, Motte and her team identified 4,800 communes whose boundaries changed from 1801 to 2010, sometimes resulting from the fusion of several communes, sometimes, from the creation of new ones, and sometimes, from the suppression of others. Consequently, the number of communes declined from 40,609 in 1801 to 36565 in 2010. Nevertheless, the French administrative boundaries of communes are relatively stable, each serving as a functional administrative unit, spaced roughly at 2.3 kilometers apart over the entire country. It is the base unit for which census information is collected and can be analyzed.

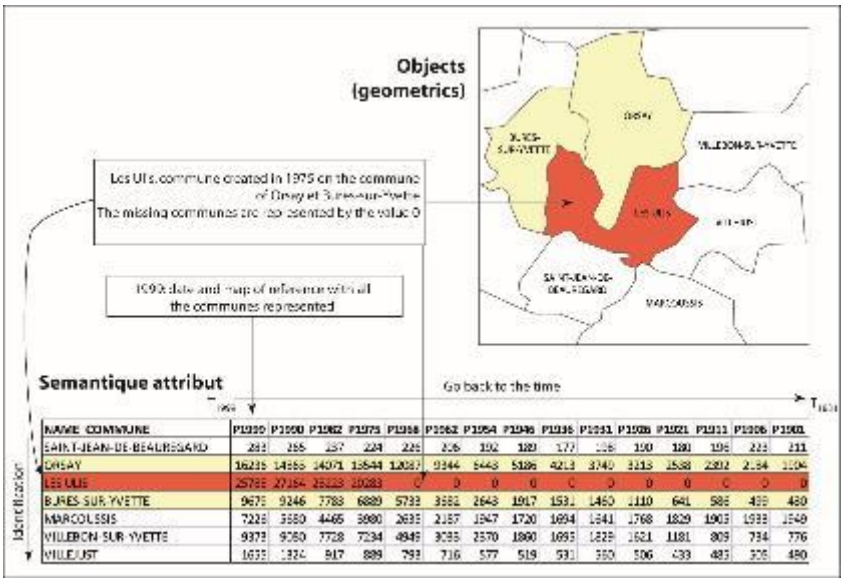
2.2. The space-time composite structure of French communes

To deal with changing boundaries of spatial units, two solutions can be considered. The first one is an aerial interpolation method that was developed by Ian Gregory in the National Great Britain H-GIS project. It consists of digitizing the geometry of each spatial unit over the studied period. Then, areal interpolation is applied to create a standardized set of administrative boundaries onto which can be mapped corresponding estimates of population and other attributes (Gregory and Ell, 2007). The standardized units make it possible to eliminate the confounding effect of boundary modifications when studying evolving patterns across space and time.

Unfortunately, the information required to reconstruct boundary changes is not available for the French communes. Given the historical stability of communal boundaries in France, we use current communal boundaries on which to attach data. In the relatively small number of communes whose boundaries did change, the changes can be incorporated by applying the space-time composite model of Langran and Christman (1992) in GIS software. Based on the principle of decomposition/association, the usual process of foreword projection has been reversed in favor of backward projection from 1999 to 1831. The base map T1999 (figure 1) assembles all the communes of the current period, a total 36, 565 spatial units. The example of “Les Ulis” in the region of Paris is presented in Figure 1. This new commune was created in 1975 from parts of the communes of Orsay and Bueres-sur-Yvettee. In its year of creation, the new commune possessed a population of 20,283, drawn from those of its neighbors. Years prior to 1975 carry zero values, although this encoding is not without problems. From the semantic point of view, the 0 is ambiguous because it represents *the absence* of data. From the geometric point of view, the boundaries of Les Ulis are those that existed in 1999; hence, they do not capture whatever boundary changes that may have occurred since 1975.

According to Motte, difficulties of this kind affect 823 communes, i.e., 2 % of the universe of 36,565 units. Such difficulties—sometimes the result of varied orthography, sometimes, the variability in administrative boundaries—require careful and cautious estimations and verifications to identify probable errors in the sources and the data. In short, the data need to be “clean” to yield reliable results, and “cleaning” requires detecting likely errors and correcting them. One can then meet one of the most important challenges facing the geo-historical community: improving the reliability of the database.

Figure 1: The space-time composite structure



NAME COMMUNE	P1999	P1990	P1982	P1975	P1968	P1962	P1954	P1946	P1936	P1931	P1926	P1921	P1911	P1906	P1901
SAINT-JEAN-DE-BEAUREGARD	283	265	237	224	226	206	192	189	177	106	190	180	196	223	211
ORSAY	16235	14665	14071	13544	12087	9344	6443	5186	4213	3740	3213	2538	2392	2154	1104
LES ULIS	25785	27164	25223	20283	0	0	0	0	0	0	0	0	0	0	0
BURES-SUR-YVETTE	9675	9246	7783	6889	5733	3682	2643	1917	1531	1460	1110	641	586	499	480
MARCOUSSIS	7226	5680	4465	3980	2635	2157	1947	1720	1694	1641	1768	1829	1903	1933	1949
VILLEBON-SUR-YVETTE	9373	9060	7728	7234	4949	3035	2570	1860	1693	1829	1621	1181	809	734	776
VILLEJUST	1633	1324	917	889	793	716	577	519	531	360	506	433	483	305	480

2.3. Detecting potential errors

One method of error detection in historical population data is that used by Gregory and Ell (2006) in the Great Britain HGIS; it proceeds in three steps, as illustrated in (Figure 2, from a to d).

1. Based on the population census for Great Britain parishes (figure 2.a), the normalized rate of population change in *each* sample unit of the decennial time series is calculated for each census year in order to detect suspicious disparities in consecutive values. Significant disparities are commonly known as "spiking values."
2. For any given spiked point in the series, the differences between the spiked value at census year t_0 and the values at the previous and subsequent census years—t_{-1} and t_{+1} respectively—are measured (figure 2.b and 2.c). An example of intermediate spike is shown in figure 2.f.
3. The larger of the differences, expressed in Z-scores (figure 2.d), is then tested by Monte-Carlo simulation to determine the likelihood that the spiked value occurred by chance or by some real process of mistaken recording, mistaking enumeration, or other "real" error.

This method takes into account each demographic series, one by one. For the French case, we used another method better suited to our *quinquennial* data. As shown in Figure 2e, the normalized rate of population change (1831-1999) for a sample of communes exhibits a more complex series than that of Great Britain, owing to a greater number of spiked values and greater variation in the amplitude of those spikes. The objective is to identify spikes as suspect according to times series and longitudinal series. For that and to address such complexity, we applied statistical tests usually employed in climatology to better identify significant fluctuations and extreme values in a time series. Our next step applied a more recent method that Cornelius and Reynolds developed—SMWDA (Split Moving Windows for Dissimilarity Analysis)—to more accurately identify break points in a temporal series in terms of the date, duration, and nature of the fluctuations (Cornelius and Reynolds, 1991).Used successfully by

soil scientists and climatologists, this method of testing worked well in addressing the very numerous fluctuations in our demographic time series. Essentially our analysis was a five-stage process (Figure 2.g).

1. The analysis starts by placing at the beginning of the demographic time series a window/interval of a specified sequence of years, such as twelve years in the figure below. As we have quinquennial census data, twelve years correspond to two intercensal periods.
2. The window is then divided in half to classify values that appear “normal” and those that so extreme as to be suspect.
3. Using the data for the Côte d’Or department as a sample and example, the mean rate of population change is computed for each of the two windows or categories.
4. The difference between “normal” and suspect mean rates of change represent the degree of dissimilarity in the given sample.
5. The process is repeated for each successive 12-year interval until the complete demographic series (1830-1999) is corrected or verified.

One characteristic of SMWDA is that it may point to unusual fluctuations even in a randomly generated series. To avoid “false positives” Monte Carlo tests (Figure 2.h) make it possible to identify fluctuations of statistical significance at conventional levels of confidence (99, 95, and 90%) In the example presented Figure 2h, 3 spikes were judged significant at the 90 % level.

Figure 2: Two methods for identifying significant dissimilarities

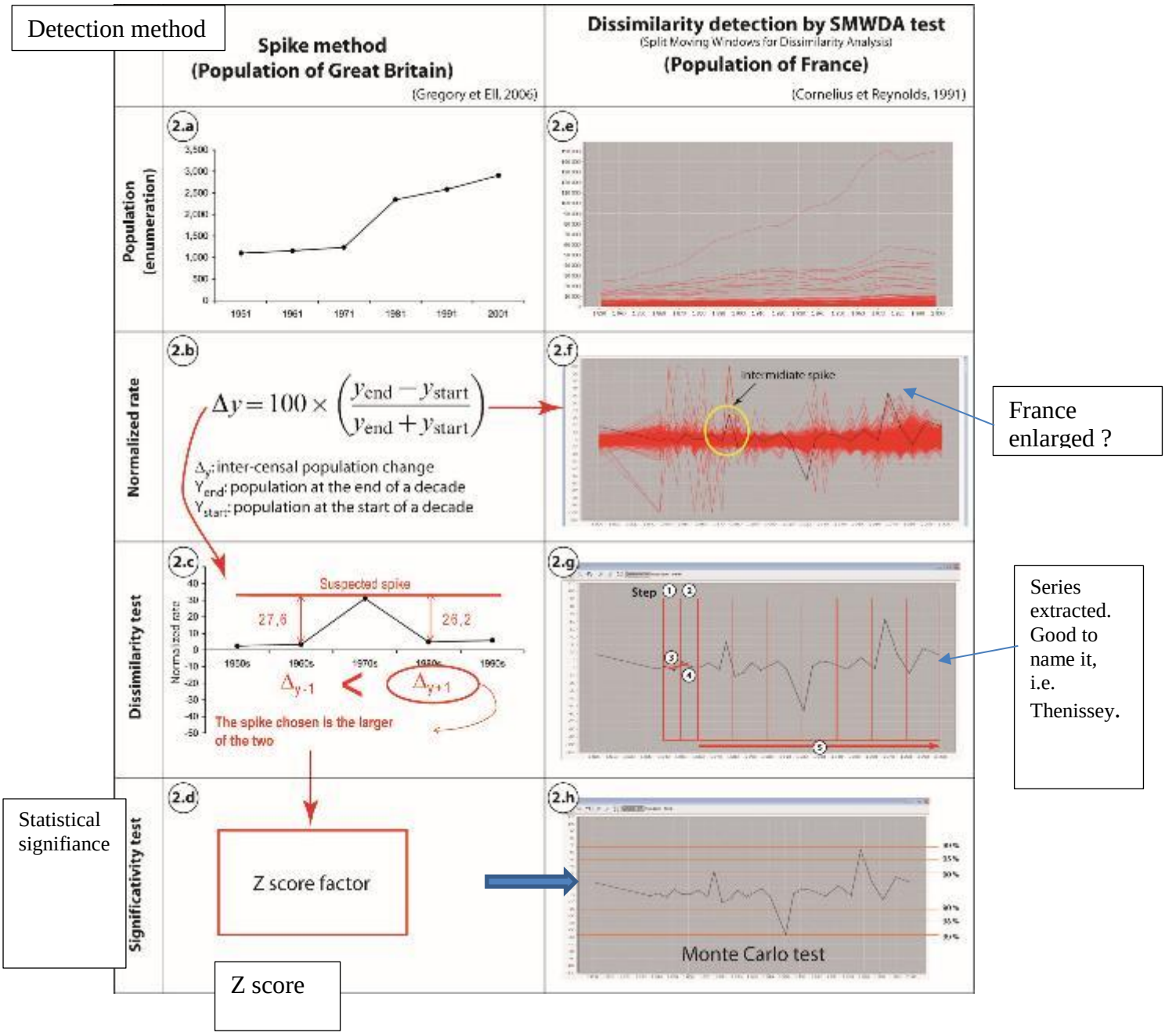
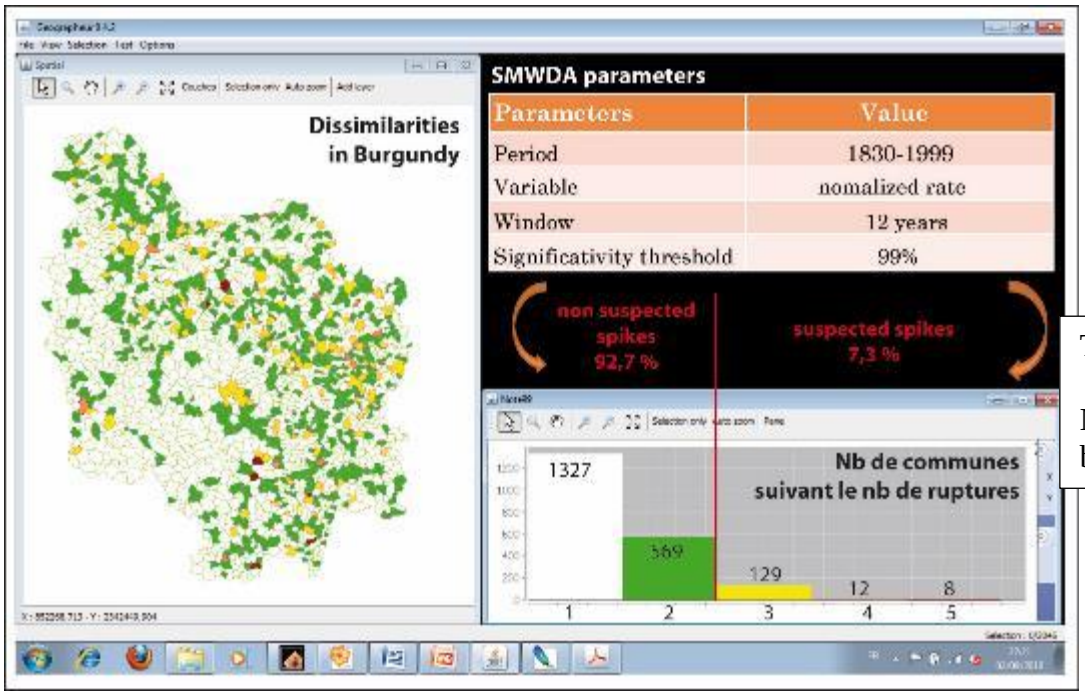


Figure 3 : Errors depicted with SMWDA method



Translate to English

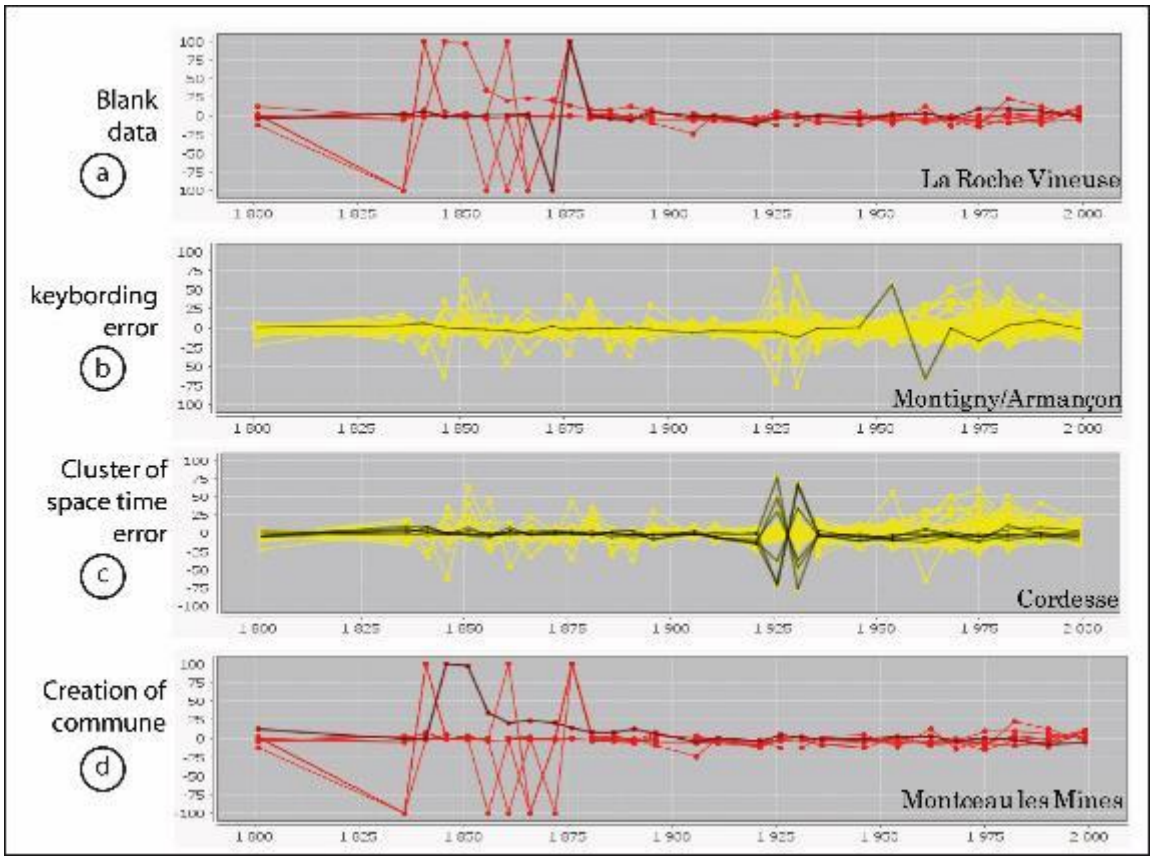
Number of communes by number of spikes

Implemented in a spatio-temporal exploratory software, we applied SMWDA in a pilot study (Figure 3). This method is here applied to population data for the region of Burgundy. Using the figures for each quinquennial census from 1831 to 1930, we used a window of 12 years that typically included *two* census enumerations. To identify only the major spikes, the significance level was set to 99%. In interpreting the results we assume that reliability of the successive population totals varies *inversely* with the number of breaks revealed by the SMWDA: the more numerous the statistically significant breaks, the less reliable the population totals. In our case, the threshold of reliability ends at 2 breaks, i.e., a series with more than 1 break in a twelve- year period is a likely skewed by error. Using this threshold, the testing indicated that 7.3 % (149) of the communes had highly suspect values, and mapping the affected communes showed that they were more or less equally distributed across the map of Burgundy. In other words, about 92 percent of all the 2,045 number series pertaining to region's communes are very likely composed of accurate figures. Moreover, there was very little if any spatial clustering among the suspect communes, indicating randomness and the absence of an event or process that would produce spatial autocorrelation.

In the same exploratory software, the SMWDA gives us the opportunity to establish typology of potential errors in the census data. Indeed, the use of parallel plots helps classify the nature of the statically significant breaks. Figure 4 shows the plots pertaining to four kinds of significant breaks resulting from either a) missing data b) errors in data entry c) clusters of mistaken data assignments to geographic and temporal locations, or d) creation of new communes. The interpretation of these patterns requires caution and, in some cases, further research. It is possible, for example, that any given commune could experience a large increase or decrease in population because of demographic events or the effects of substantial economic change in years between census enumerations. In the commune of Thenissey (Côte d'Or department), for instance, the arrival of more than 100 temporary

workers engaged to build a tunnel for a new railway line in 1849 meant that the population recorded in the 1851 census (450) was one-third larger than that recorded in 1846 (334). In the census of 1856, Thenissey's population had returned to a "normal" if smaller level (292)[1]. In this and other ways, small communes could see their populations increase temporarily from 100 to 200 inhabitants, producing in our study notably suspect spikes on plots that were *not* the result of mistaken counts, mistaken data entries, or other errors of this kind.

Figure 4: Four examples of errors



All this means that further validations are needed to make good use of the available demographic data that was drawn from the E-Geopolis database. One important validation compares the E-Geopolis figures with those published elsewhere, noting carefully both absolute and relative differences in data values to reach conclusions about the origins and character of dissimilarities. In our case, consulting the web site of Cassini [2] helped verify the accuracy of our demographic data, for there was a strong convergence between Cassini and the E-Geopolis databases.

With these verifications completed, the pilot study of Burgundy served to calibrate the testing process and prepare the way for applying SMWDA to the whole of French territory. The SMWDA deals both with potential error in a time series and also with error between the time series. The results for the whole of the country indicated that 953 communes (2.6 % of all)

[1] Archives Départementales de la Côte-d'Or, 10 M 627, Census enumerations of 1846, 1851, 1851.
[2] http://cassini.ehess.fr/cassini/fr/html/index.htm

possessed population totals that were very likely erroneous in one twelve-year period or another. This small proportion of error strengthened our confidence in the ability to investigate the relationships between demographic change and the expansion of the railway network.

3. Measuring railway accessibility over a long period

The concept and measurement of accessibility offer a revealing way of studying the performance of a transportation network in terms of the population served at any given time. Firstly defined by an economist (Hansen, 1959), the concept of accessibility has received considerable attention, producing many varied definitions and results for roads and railways in a recent period (Gutiérrez and Urbano, 1996; Hackett et al., 2005; Holl, 2007). In transportation geography, Mei Po Kwan (2003) has defined the concept broadly as the study of both potential and actual travel across space to a given place and its commercial amenities, and community facilities. Complementing this abstract notion are more specific, functional definitions of accessibility such as that proposed by Guers: "*Measure of transportation system used for traveler or supply to reach an activity or a destination with one or several modes"* (Geurs and Ritsema von Eck, 2001).

The analysis of accessibility by time duration in the long run shows how graph theory has the best potential to investigate paths along the network. It requires firstly to build a database that produces a cartographic representation of the french railway network over time. This database has been developed for several projects, inclucing its integration in an international gis for the history of European Integration (Martí-Henneberg, 2013). As shown in primary results of the program, the French network is the LARGEST in Europe in 1930 (37,137 km), larger thanGermany (33,106 km) and United Kingdom (30,939 km). (FOOTNOTE a study)

In this article, we combine these two definitions in building the foundations of a geo-historical study of the relationships between population dynamics and the railway expansion from 1830 to 1930. Geographic studies of transport usually focus on the city scale to describe and understand the urban hierarchy (Pumain, 1982). Our approach differs by focusing on a local scale composed of some 36,000 French communes by measuring accessibility over an historical period of one century. In this way we are able to map the changing contours of accessibility across national space *and* historical time.

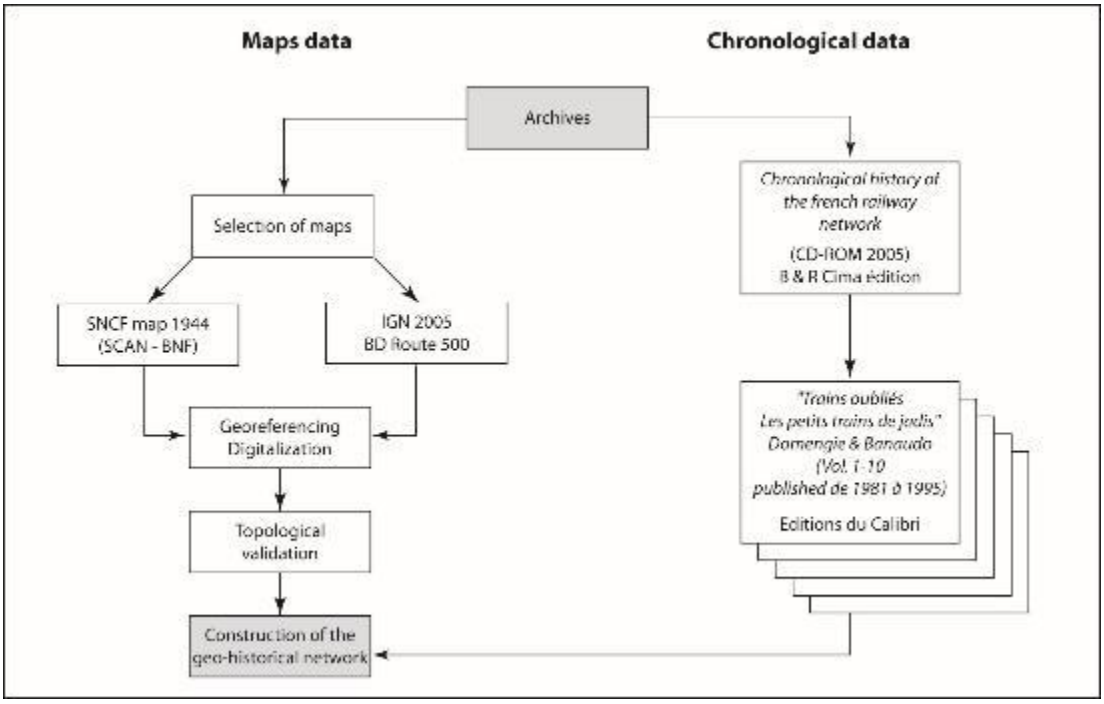
Our study is similar to a recent article on accessibility and *population growth* in England and Wales over the sixty-year period from 1871 to 1931 (Alvarez, Franch and Henneberg, 2013), but there are differences to note. In the case of France, we give due attention both to *population growth* and *population decline*. In England and Wales, rural depopulation increased in the 19th century and dramatically so in the period 1860 to 1900—a situation we treated elsewhere (Schwartz, 2010; Schwartz and Thevenin, 2012). By 1851 the majority of the English population lived in towns and cities, and population growth was very much an urban phenomenon. In France rural depopulation was less extensive and relatively gradual compared to Great Britain. Indeed, the rural population in France remained the majority until 1931. Further, France was more than four times larger than England and Wales so the expansion of the rail network was a territorial challenge, especially in the attempt to bring rail service to rural communities. This was complicated by extensive areas of rugged terrain in

uplands and mountainous regions. The constraints of French topography thus require particular attention in estimating accessibility.
Since the quality of the data determines, to a large degree, the reliability of accessibility measures and estimates, it is important to describe our sources and their limits and how they influenced specific aspects of our methods.

3.1. Geo-historical reconstruction of the French railway network.

The construction of historical geographic information system of railway networks is a growing initiative in the fields of geohumanities and economic geography (Atack, 2013; Martí-Henneberg, 2013). The first step in this part of our research was devoted to finding historical maps of railway lines and other evidence that would provide the opening and closing dates of lines and stations—to name the essentials. In the end, the best maps were found in the Department of Maps and Plans at the Bibliothèque Nationale, while supplementary information on tertiary, local lines, opening and closing dates, and other matters were collected from a digital atlas and reference works. This corpus constituted the evidence used in the digitizing process. (See Figure 5.)

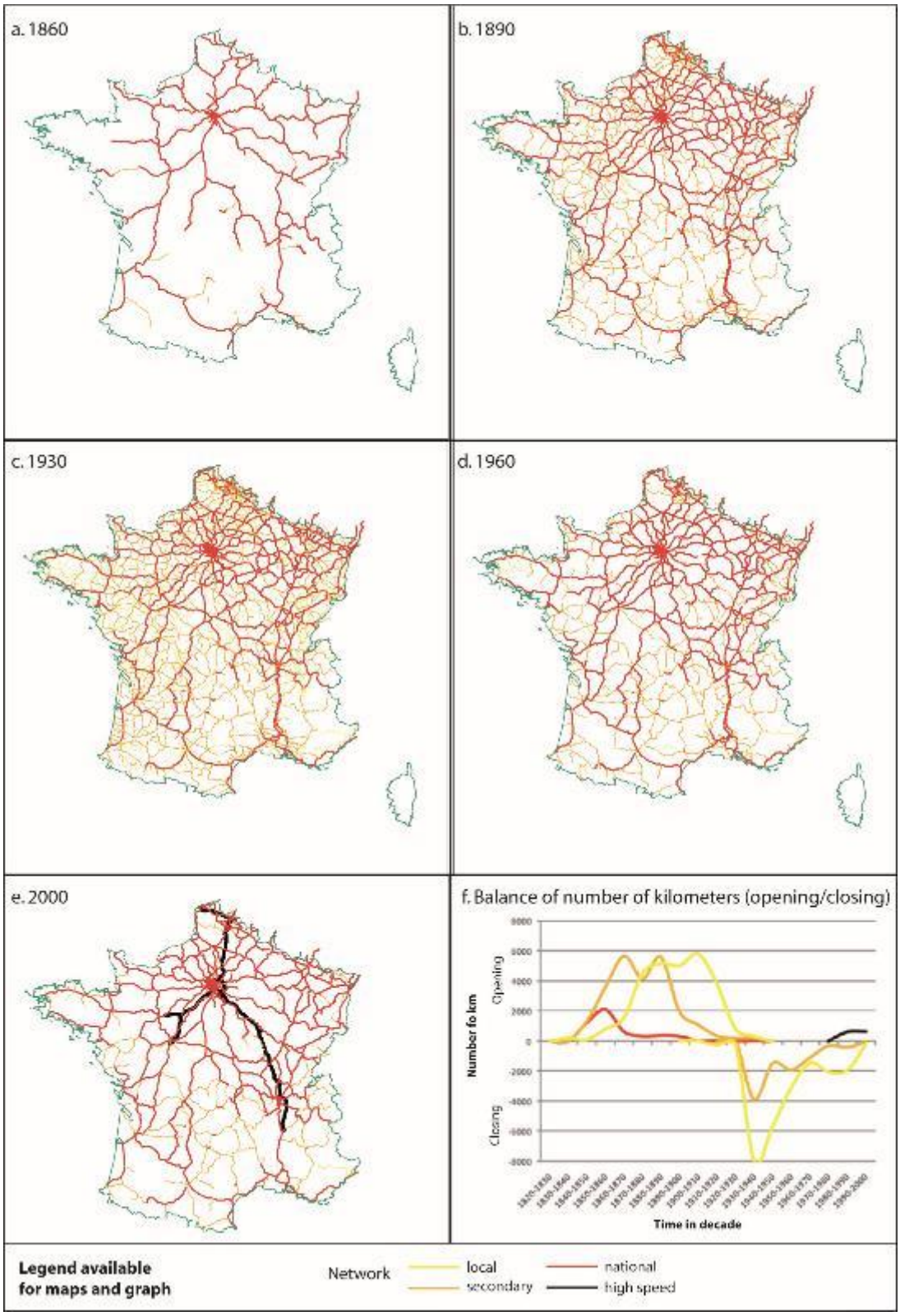
Figure 5 : Operating procedure for data collection



The cartographic data in our H-GIS comes from a map published by the French National Railways (Société Nationale des Chemins de Fer) in 1944, the date at which the railway network reached its maximum extent and would diminish gradually thereafter. With a scale of 1/800 000, the map offers a highly detailed geography of the French network. After receiving a high resolution scan of the map from the Bibliothèque Nationale, we geo-referenced the document, using as a base map the railway layer of the current Route 500 database produced by the Institut Geographique National. To the current network, we added

all the lines present on the 1944 map, yielding a total of some 28,000 kilometers of railway. Next were added the lines belonging to a tertiary network of what were called “lines of local interest”. To improve the geographic accuracy of the digitized rail lines, we verified the topology of the lines and nodes using appropriate methods of graph theory.

Figure 6: Expansion and compression of the french railway network (1860-2000)



Secondary sources supplied the dates needed to establish the chronology of railway construction with opening and closing dates of lines and stations, the nature of a given line—main line with two sets of tracks, local line with narrow gauge rails [1 meter width], ownership [Paris-Lyon-Mediterranean Company, the French state, a French department] and so forth in order to date the construction of the railway network[3].

The results of the data collection and its geo-processing are illustrated in Figure 6. These maps depict 3 main periods of the French network growth. After the construction of the national lines reinforced the centrality of Paris by connecting the capital to important cities in the provinces before 1860, the next expansion reinforced the density of lines to Paris by connecting each departmental seat of administration (*chef lieu)* to the capital city between 1860 and 1890. Importantly, in the same period, the fist lateral lines were opened, connecting, for example, Dijon in the east to Tours in the west. The last phase of the railway building, from 1890 to 1930, was largely devoted to develop "lines of local interest" in keeping with the Freycinet plan of the central government. Drawn up in 1878 by the Minister of Public Works, Charles Freycinet, one major aim of the program was to extend rail service to communities without it, including all *chef lieux* (seats) of *cantons* and as many other small country towns for which state subsidies and local funds were sufficient to bear the cost. By diminishing the great disparities in service that existed between well-served and under-served regions, the program was designed to affirm the Third Republic's commitment to equality while facilitating the modernization of the vast rural and agricultural domain of the country. In this way, some 20,000 kilometers of local lines came into existence by 1930, bringing the total French network to between 56,000 to 61,000 kilometers at that date. From this maximum point, a phase of contraction began, gathering pace after the Second World War to the 1960s. (See Figure 6.) The contraction of the system all but eliminated the local lines, as improved roads, automobiles, buses, and trucks replaced them. At the same time, the modernization of the equipment and lines belonging to the remaining network went forward, beginning with the electrification of locomotives in the 1960's and moving in the 1980s to the construction of high-speed railways. All these changes are incorporated into our H-GIS database.

3.2. Building a geo-historical graph of the railway network

In the study of transportation, concepts of graph theory are central to network modeling and the calculation of accessibility indices (Haggett and Fréchou, 1973). Many algorithms exist to determine the optimal path across a network, according to such selection criteria as physical distance, travel time and other components of generalized cost. To determine the optimal path requires estimating the various constraints of speed, direction, terrain, and so forth. The archival and printed sources for our estimates are given in the bibliography.

For estimating speed, there are two possibilities. One is to derive estimates from the time tables published by rail companies for each date selected for study. T. Luenig has employed

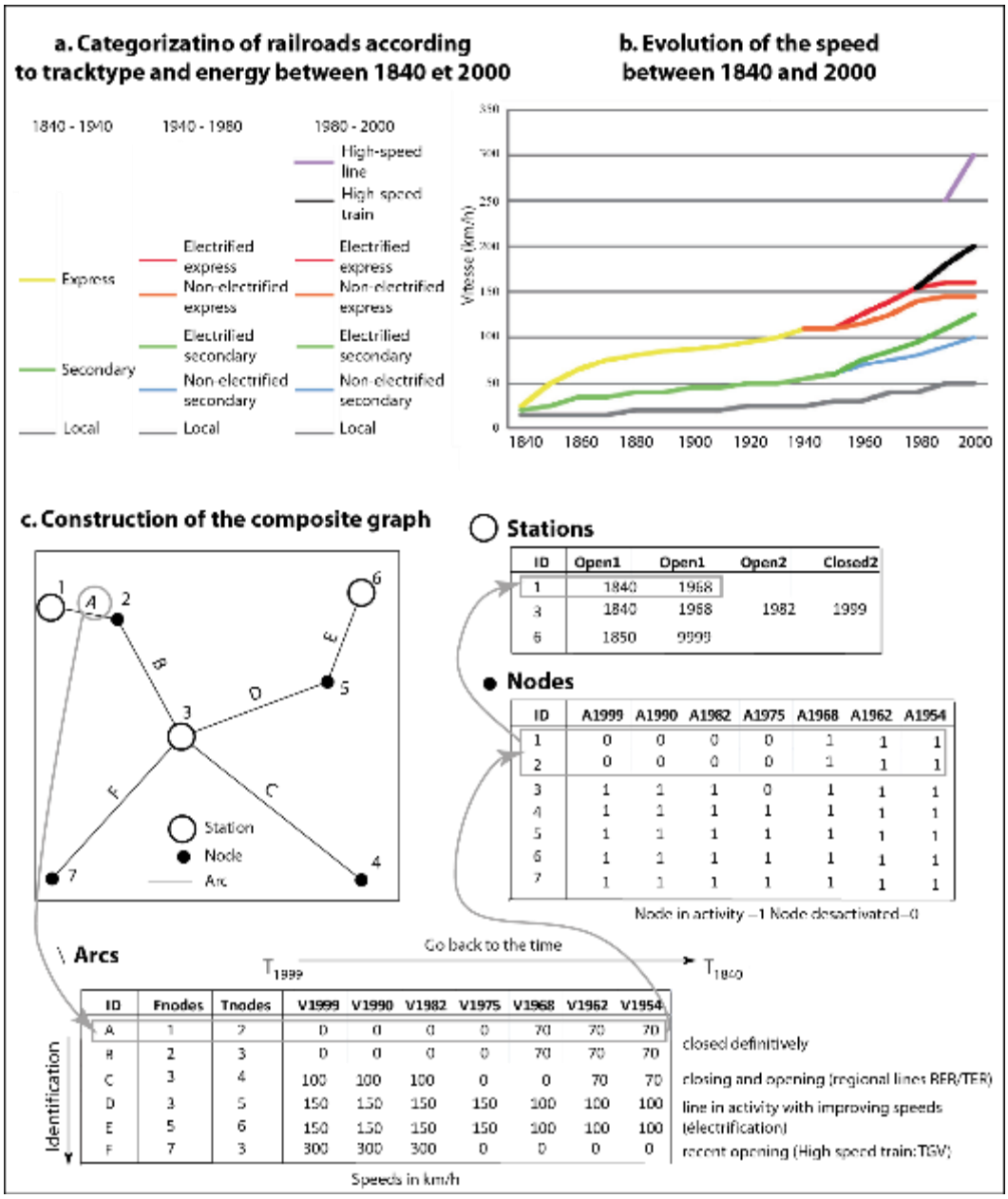
[3] This current limit has been criticized by number of H-GIS authors: 8. Gregory, I.N. and R.G. Healey, *Historical GIS: structuring, mapping and analysing geographies of the past.* Progress in Human Geography, 2007. **31**(5): p. 638, 12. Knowles, A.K., *Emerging trends in historical GIS.* Historical Geography, 2005. **33**: p. 7-13.

this method in his study of the British rail network, drawing from Bradshaw's time tables a sample of the daily travel times over 50 routes in three different chronological periods to estimate average rates of speed over time (Crafts and Leunig, 2005).

Using speed estimates mentioned in other kinds of sources—time tables being largely unavailable in France—our approach uses a different method that estimates over time and geography the accessibility, not of a sample, but of the *entire* railway network. To reach that point was a four-stage process:

1. Identifying the main sources on railway construction (Caron, 1997; Studeny, 1995), and the high speed train development of the contemporary period.
2. Deriving from these sources estimates of speed for each of 9 categories of lines and for each decade of our study period (Figure 7a.).
3. Incorporating travel constraints owing to topography and levels of service on any given line.
4. Calculating the speed estimates for each decade.
5. Calculating and graphing the mean speeds for the national network for each decade according to the service and the energy available along the network (Figure 7b.).

Figure 7: A composite graphe



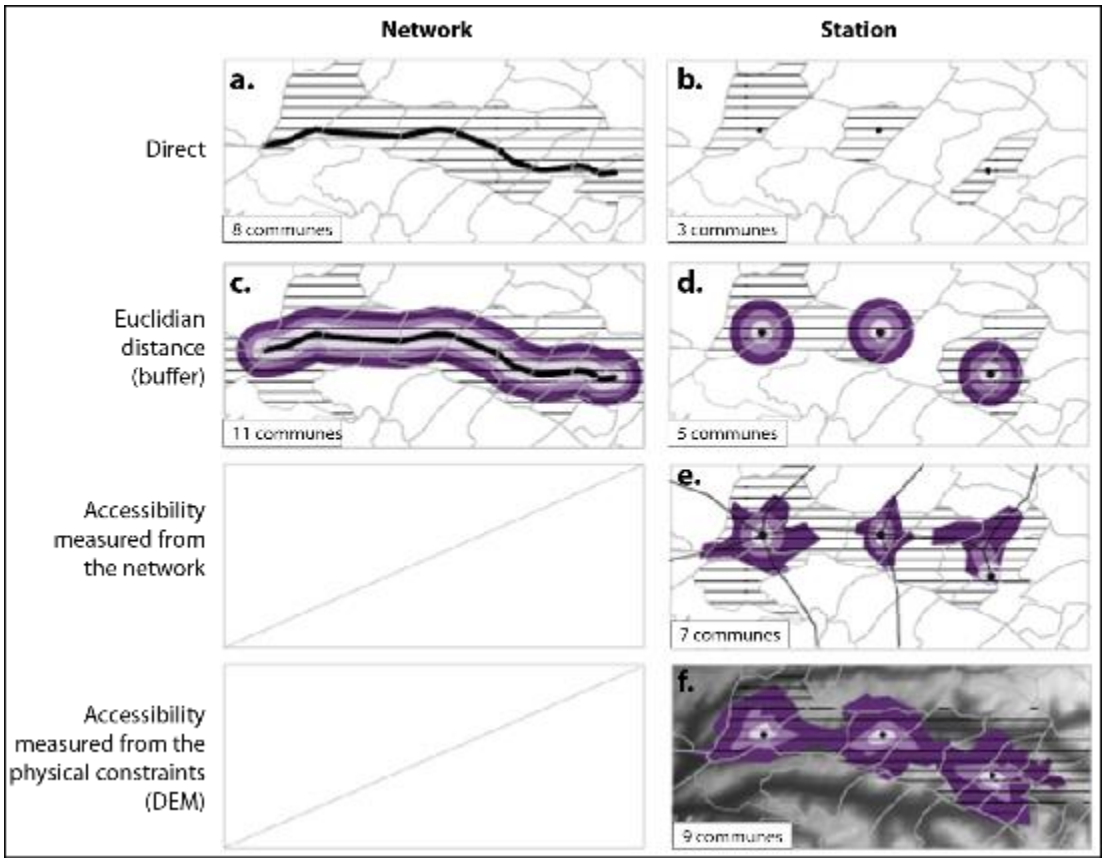
These steps yield a measurable and structured network in the form of a graph of arcs and nodes *without* directional constraints (See Figure 7c.). The transformation of the HGIS into a graph requires a definition of terms. Each section of the network is an edge, whereas each station and intersection is A node. The graph is valued in each edge, where the time to cross the edge is the confrontation of the length and the speed observed in the edge. The constitution of the graph aims to calculate travel time between each station. To this static representation, one has to build an adaptive graph to incorporate the temporal dimension, producing thereby a network of arcs and nodes that evolve and change as lines and stations are created or removed and speeds (typically) increase. This network modeling is less accurate than the sampled results Leunig has produced, but is more comprehensive and more complete than his. If time and sources were available, it would be interesting to apply his method to the French network and, vice versa, to apply our method to the British railway system. A valuable

comparison would certainly result that would provide alternatives to researchers in other countries, according to the sources they have at hand.

3.3 Pedestrian accessibility to the railway network

The various ways one might reach a transport network must be taken into account in estimating accessibility. During the first period of our study, walking to a train station was the principal means of getting aboard the network, to which should be added travel by horse-drawn cart. We treat both as a pedestrian mode of travel and use this mode in analyzing accessibility to reach the closest railway station from 1860 to 1930. The purpose is to estimate the travel time from each functional center of the 36,000 French communes to the nearest station. Just as accessibility varies by mode of transport, it also varies according to the conception of the target object. In our case, we can consider two forms: the linear railways and the punctual stations. Taking account of this difference, employing six alternative methods of calculation provides a range of accessibility estimates. In each of six examples presented in Figure 8, one railway line crosses 8 communes and 3 stations.

Figure 8: Six accessibility measure to the network



Classic approaches Typically study a direct accessibility to the network, as the edge (Figure 8a) or the station (Figure 8b) crosses the boundaries of a commune. The first example serves all 8 communes and the second serves the 3 communes that possessed stations only. According to Gregory either one or the other form of direct accessibility is typically used in historical GIS (Gregory and Ell, 2007).

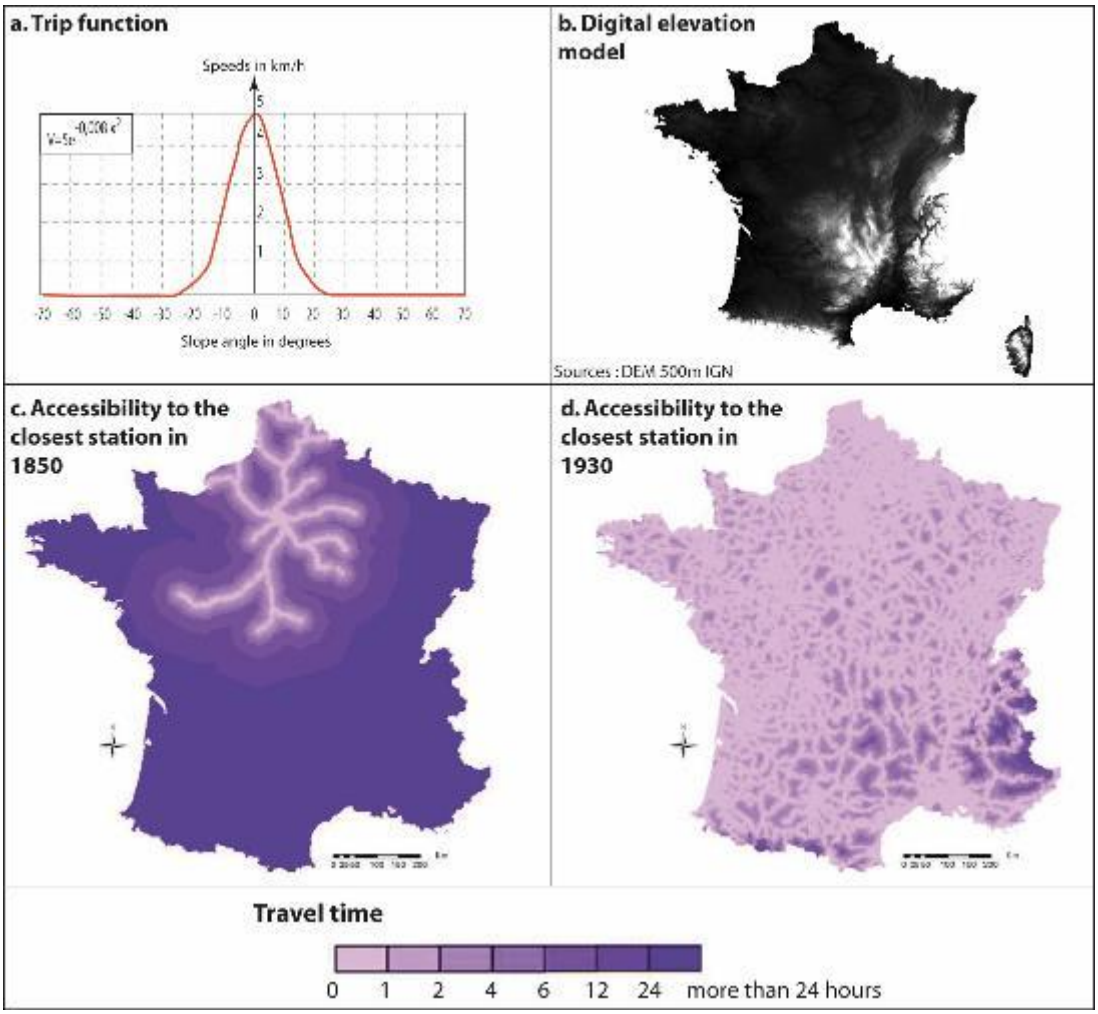
But, if some research into historical accessibility treats only the simple, direct distance it neglects the varied topography of the distance in question. After all, it takes more time to travel through mountains than through plateaus. The second examples (8c and 8d) illustrate the improved measurements that result from placing buffer zones around the line and the stations. At a given distance, the network is accessible by 11 communes in Example 8c, whereas there are 5 communes in Example 7d.

As this approach could be comprehensive in urban areas (O'Sullivan and Morrall, 1996), a simple distance could not express the real effort, which people needed when they had to reach the station on foot. Our purpose here is to take into account two major aspects of geographical space:

- Heterogeneity: human settlements are irregularly dispersed,
- Anisotropy: the pace of travel varies depending upon topographic constraints

The further complexity is represented in example 8e, arguably the best measurement scheme because it links the road and rail networks at station nodes. Lacking the data needed on roadways, we settled on the scheme in Example 8f. This includes the important physical constraint of topography, a factor derived from a digital elevation model (DEM) that covers the whole national space and is suitable for use over the entire period. While the DEM used is inaccurate, it contributes a desirable degree of reality as well as coverage of the entire study area. We adapt here some works about physical accessibility to reach a resource on foot (Kaddouri, 2007) For these reasons, the best method of gauging physical accessibility, it seemed, was to combine the constraints of distance and topography: placing graduated buffer zones around the stations to represent varied distances to the rail network and adding slopes of elevation to represent topographical constraints. [Although a similar approach by Alvarez et al. (2013), the authors' description was brief so a more elaborate discussion here will be helpful.]

Figure 9 : Pedestrian accessibility to the railway station



Using the locations of *functional centers of communes* (as opposed to centroids) and the digital elevation model (Figure 9b), we derived estimates of the influence of heterogeneous settlement patterns and topographic constraints. With these estimates in hand, it was possible to ESTIMATE the pedestrian path from each commune to the closest railway station for each decade in our study period. Taking the conventional rate of 5 km per hour as the average pace of pedestrian travel (Kaddouri, 2007), that rate will vary with terrain elevation: decreasing as the slope of the travel path rises, and increasing as the slope heads downward (Figure 9a). Translated to a map, this estimate of pedestrian speed is the major factor in our modeling of rail accessibility in the 19th century.

According to our estimates, in 1850 an inhabitant of a typical commune had to walk more than 24 hours to reach the railway station closest to his home (Figure 9c). As shown in Figure 6, railway lines at this time radiated from Paris like tentacles reaching into favored cities and regions of the country. To reach Paris by train then usually meant, for the vast majority of the French people, a two or three-day journey from home to the nearest station.

As the rail network grew, so accessibility improved (Figure 9d). At the turn of the 19th century, and with the exception of mountainous regions, the whole of France was more or less closely

connected to the rail network. By then the great majority of communes were within 10 to 15 kilometers (as the crow flies) of a station (Schwartz et al., 2011; Schwartz and Thévenin, 2014). To reach Paris by train in 1900 meant a walk of three or fewer hours to the nearest stations, instead of the two or three days needed four decades before.

4. Accessibility and population served

All railway links were *not* created equal, of course, as the disparities in rail travel to Paris shows clearly. We suggest disparities grew or diminished over time, change well worth pursuing in detail. This can begin by studying the temporal disparities in train travel from any given commune to the nation's capital. Because the requirements for this include a network that is fully connected, our beginning date must be 1860, for it was then when the required interconnectivity came into existence. The year 1930 seems a good endpoint because after that date the closing of local lines began, and the future dominance of the automobile was already visible. To go further than 1930 would require factoring in automobile and bus transport to stations, a complexity we cannot pursue at this point. In sum, from 1860 to 1930, we model and estimate accessibility in terms of walking to a station and then riding a train.

4.1. Accessibility to Paris

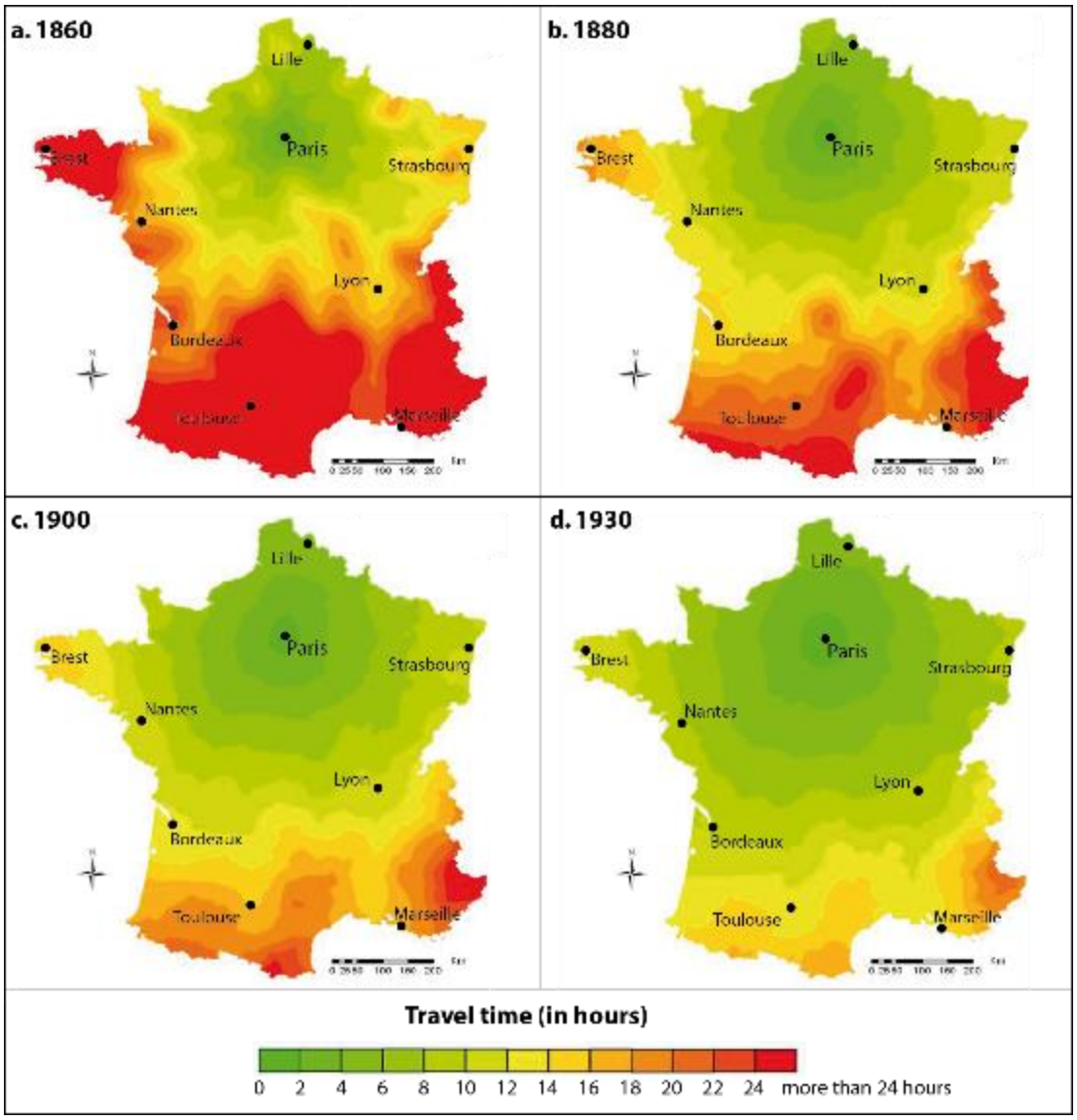
To begin, we assume that a commune is served by rail transport when the distance to the closest station is 5 km or less. Once on the network, speeds of train travel are assumed to vary by type of line—main, secondary, and local lines respectively. The mean speed of each type of line is assigned accordingly to their appropriate routes.

The bi-modal accessibility measure is derived from two specific calculations. A measure of time-distance is determined for each paring of commune/station according to the following formula:

$$A_{(C,Go)} = \min[e(C, So)]$$

C : commune
So : Station of origin

Figure 10: Accessibility to Paris (Pedestrian-Rail)



Using the mean rates of speed appropriate for main, secondary, and local lines, the second calculation estimates the travel times from station to station until the destination of Paris is reached. The combination of these two estimated values gives a measure of the unipolar accessibility (UA) from any of 36,565 French communes to the closest Paris railway station.

$$UA_{(C,Sp)} = \min[e(C,So)] + \min[e(C,Sp)]$$

C : commune
So : Station of origin
Sp :Station of Paris

Frequently used in transportation geography, estimates of unipolar accessibility can best be visualized and studied on an isochronic map that displays the spatially smoothed values of (potential) accessibility. Comparing the maps in Figure 10, each representing a different period, reveals striking patterns of historical change in accessibility. Using a line from Bordeaux to Lyon as the division between north and south, the maps shows that great disparities in accessibility between the two regions at least to 1880. By 1900 travelers south

of the line who were going to Paris faced a journey of some 12 to 14 hours, a long haul indeed when compared to the swifter journeys available to their counterparts in the northeast. In 1860, the network became fully connected as the last sections of the "Great Hub and Spokes" were put in place (Caron, 1997). This created a national network with branches radiating from Paris and two transversals between Atlantic Ocean and Mediterranean Sea and between Marseille and Strasbourg. However, main lines were limited to serving paris while branch lines were only emerging on the drawing boards of the private companies. Evident also in 1860 are the great disparities in accessibility between north and south. Only after 1880 did these disparities signficantly diminish. Between 1860 and 1880, the unipolar accessibility MEASURES are correlated with the density of rail lines. Decades later, between 1900 and 1930, ACCESSIBILITY WAS A function of the distance to Paris and of the position of any given locality in the hierarchy of the network.

4.2 Generalized accessibility

Going to Paris was not every French person's dream, of course. In reality, many simply wished to get their goods to the nearest market town or travel to a departmental *chef lieu* to handle a business or personal matter. To take account of these every-day realities, generalized accessibility (GA) is a convenient measure that can be visualized and studied on a map. Generalized accessibility maps yield results similar to those of the unipolar measure because of the star structure of network concentrated to Paris. Nevertheless, the display and meaning of the results are quite different because GA estimates the mean access time from one commune and station of departure to *any* other station in the national network. The formula for this measure thus incorporates the average time from a designated station to any other.

$$GA_{(C,Sd)} = \min[e(C,So)] + \frac{\sum_{i=1}^{n} e(So,Sd)}{n-1}$$

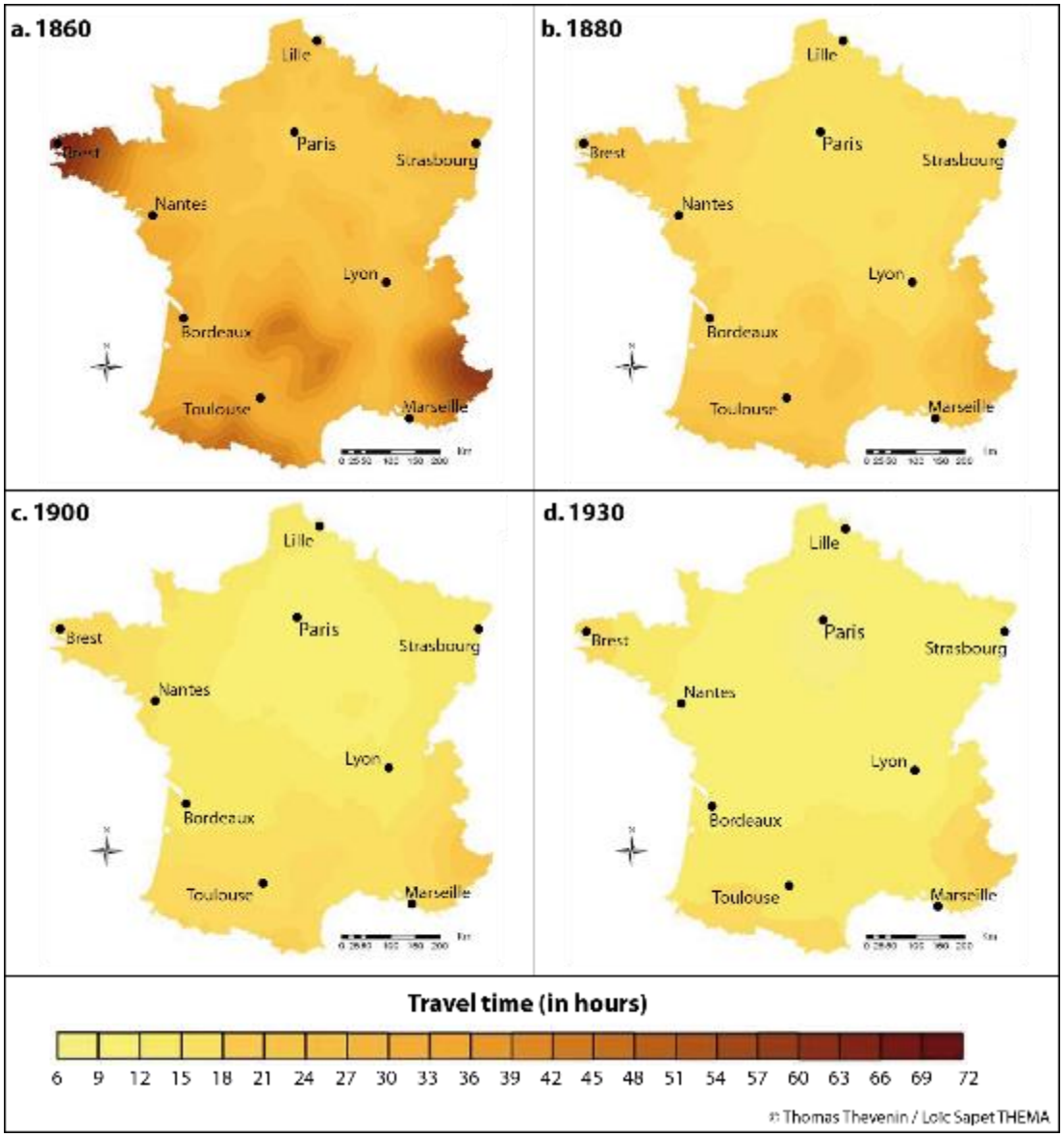
C : commune
So : Station of origin
Sd :Station of destination

To review the results for 1860, the *average* accessibility from Paris to any of the outlying communes in the country was less 20 hours. More important than the national average, the travel time to a given commune varied considerably: travel time to Brest, for example, was about 72 hours in 1860. The time required to reach mountainous zones by train, of course, was much greater in 1860 because few lines or stations were in easy reach of upland communes. The highest values in 1860 are drawing the relief of France: they are concerning Alpes, Pyrenees and Massif Central. They are showing the particular case of the Brittany where the network is only open until Rennes, in the east of the Brittany. Reduced but still considerable travel times affected communes located on the central-eastern regions of Burgundy and Champagne. We assume that the general accessibility in the beginning of our period is also drawing border effects of the network, as it has primarily been developed around Paris. The most change occurred between 1880 and 1900 as the Freycinet project brought the arrival of a station in more than a third of French communes.

From 1900 to 1930, Although the expansion of local lines improved the accessibility in many regions and small communities, it still took about 12 hours of travel from home to stations in the most accessible zones of rail service, and up to 30 hours to reach stations in the most

remote regions. The border effect was also becoming less important as the density of network converged in all regions of the country. However, areas of mountains were still concerned by difficulties of network development. Moreover, Alsace and Lorraine were greatly changed by geopolitical events in the end of the 19th century but in the beginning of the 20th century the density of railroads in what were now German provinces roughly the same as in the other regions of the Eastern France. The greatest disparities remained between North and South of France. One century after the establishment of the "the Great Hub and Spokes", the centrality of Paris in the network remained and persisted.[i]

Figure 11 : Generalised accessibility to the French communes



5. CONCLUSION

Admittedly imprecise, our results and maps represent simplified estimates of highly complex historical processes. Nonetheless, this initial analysis of accessibility in the long run shows how this type of investigations could be extended to others countries and others networks, in keeping with a European initiative that has underscored the importance of infrastructure in investigations about differentiated development (Vickerman, 1995). The measures of accessibility we proposed have the advantage of being applicable to all types of transport networks, depending on the type of data available in each case and each period.

More broadly, the construction of our geo-referenced database regarding population change and railway accessibility illustrates both some of the limits and some of the contributions of combining GIS, spatial analysis, and historical research. Some historians, particularly those unacquainted with GIS technology and suspicious of quantitative methods *per se*, are quick to criticize the collapsing historical reality and the false precision entailed in assigning information of varied origin and nature to a single point, line, or polygon (Gregory and Geddes, 2014; Knowles, 2005; Southall, 2014). Although these are valid concerns, some critics should convince themselves that geography and spatial relationships may well have at least some relevance for their own work, that ignoring these aspects of historical reality sacrifices historical complexity. As researchers and writers we should do better job of making the convincing case.

Turning to methods of geo-history and spatial history, we think it essential to develop specific solutions for validating data. The exploratory data analysis presented in this article represents our effort to estimate the accuracy of census data and railway development. Moreover, when appropriate or necessary, it is important to move beyond the limitations of commercial GIS products to find solutions and methods suitable to very large databases such as ours. As we have discovered, such databases often call for computational power well beyond what even the fastest desktop computer can supply, as well as programing expertise of computer scientists.

As for contributions, the construction of a GIS database on population change at the level of the commune over nearly two centuries, built on the prodigious work of Francois Moriconi and Claude Motte, represents a significant contribution to the field. To our knowledge, it is the first GIS containing reliable communal level data covering the entire period of census enumerations for the whole of France. Second, it is the only railway GIS to be comprehensive and complete because it incorporates the tertiary lines of local interst that connected rural areas with the main system. With these data in hand, it will be possible to undertake a detailed and comprehensive re-examination of the rural-exodus of the second half of the nineteenth century, a subject of great concern to contemporaries and one often referred to in modern scholarship but examined only in part. Previous studies of rural out migration in that period have been limited to one or another region or to inter-departmental movements because migration statistics for the whole country exist only at that level (Bretagnolle, 2003). Our H-GIS permits the mapping of population change and estimates of out migration at the communal level, and, more specifically, the changing patterns of population movements in relation to the expanding rail network (Schwartz et al., 2011). In this effort it is essential to examine not only population growth typical of 19th century urbanization but also the varied patterns of population stability and decline in rural areas. As Kasarain et al.(2016) point out in their excellent review of the literature on rail transport and population change over the longue

durée, much about this history in France has been unexamined because reliable data at the local level for the whole country has been lacking. Hence the methods and results used to produce estimates of railway accessibility and population change in the France of 1860 to 1930 contribute significantly to future work in transport history and geography, bringing France out of *terra incognito*.

Finally, in French history, our results attest to the impact of the Freycinet program: the expansion of rail transport through the combination of top-down state planning and subsidies with regional and local initiatives. Carrying out the plan to connect every cantonal chef lieu by rail and to bring modernizing effect of rail transport to the countryside certainly improved the potential advantages of rapid transportation for the bulk of the French population. Economically, the expansion of rail service under the freycinet program was likely an important cause of the narrowing of disparities in agricultural productivity between the north and the south, a subject to be treated in future publications (Schwartz and Thévenin, 2014). As significant as these changes were, we recognize that continuity is part of the story, too. French transportation networks persisted over the *longue durée* by following the same or similar routes. Comparing the routes of Roman roads, of nineteenth-century railways, and of the high-speed trains of today makes clear the prominent lines of persistence. What remains to be done is to carry the story from 1930 to the present, taking into account the competition between railways and roadways.

7. Acknowledgment

We gladly express gratitude for the conscientious review, which permits us to improve our contribution and replace it in the both context of research questioning and European initiatives about HGIS of railway network. We also thank the Collaborative Grants program of the US National Endowment for Humanities and the French National Agency for Research for the indispensable financial support of this research.

References

Alvarez, E., Franch, Xavi and J. Marti-Henneberg (2013). "Evolution of the Territorial Coverage of the Railway Network and its Influence on Population Growth: The Case of England and Wales, 1871-1931." Historical Methods 46(3): 17p.

Atack, J., 2013. On the Use of Geographic Information Systems in Economic History: The American Transportation Revolution Revisited. J. Econ. Hist. 73, 313–338.

Banister, D., Berechman, Y., 2001. Transport investment and the promotion of economic growth. J. Transp. Geogr. 9, 209–218.

Bavoux, J.-J., Beaucire, F., Chapelon, L., Zembri, P., 2005. Géographie des transports. A. Colin, Paris.

Bazin, S., Beckerich, C., Blanquart, C., Delaplace, M., Vandenbossche, L., 2011. Grande vitesse ferroviaire et développement économique local : une revue de la littérature. Rech. Transp. Sécurité 27, 215–238.

Blier, G., 1996. Nouvelle géographie ferroviaire de la France. Vie du rail, Paris.

Boarnet, M.G., 1995. Highways and Economic Productivity: Interpreting Recent Evidence.

Bol, P., Ge, J., 2005. China historical GIS. Hist. Geogr. 33, 150–152.

Bretagnolle, A., 2003. Vitesse des transports et sélection hiérarchique entre les villes françaises, in: Données Urbaines. Paris, p. 433.

Caron, F., 1997. Histoire des chemins de fer en France : 1740-1883. Fayard, Paris.

Chapelon, L., Leclerc, R., 2007. Accessibilité ferroviaire des villes françaises en 2020. La Documentation Française.

Cornelius, J.M., Reynolds, J.F., 1991. On Determining the Statistical Significance of Discontinuities with Ordered Ecological Data. Ecology 72, 2057–2070.

Crafts, N., Leunig, T., 2005. The Historical Significance of Transport ofr Economic Growth and Productivity (No. 1), Research Annex.

De Moor, M., Wiedemann, T., 2001. Reconstructing Belgian territorial units and hierarchies: an example from Belgium. Hist. Comput. 13, 71–97. doi:10.3366/hac.2001.13.1.71

Fishlow, A., 1965. American Railroads and the Transformation of the Ante-bellum Economy. Harvard University Press Cambridge, MA.

Flonneau, M., Guigueno, V., 2009. De l'histoire des transports à l'histoire de la mobilité. Presses universitaires de Rennes, Rennes.

Fogel, R.W., 1962. A Quantitative Approach to the Study of Railroads in American Economic Growth: A Report of Some Preliminary Findings. J. Econ. Hist. 22, 163–197.

Garrison, W.L., 1960. Connectivity of the Interstate Highway System. Pap. Reg. Sci. 6, 121–137.

Garrison, W.L., Levinson, D.M., 2014. The Transportation Experience: Policy, Planning, and Deployment. OUP USA.

Geurs, K.T., Ritsema von Eck, J.R., 2001. Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transportation scenarios, and related social and economic impact (No. RIVM Report 408505 006). National Institute of Public Health and the Environment, Bilthoven.

Gregory, I.N., 2005. The Great Britain historical GIS. Hist. Geogr. 33, 136–138.

Gregory, I.N., Ell, P.S., 2007. Historical GIS: technologies, methodologies, and scholarship. Cambridge University Press.

Gregory, I.N., Ell, P.S., 2006. Error-sensitive historical GIS: Identifying areal interpolation errors in time-series data. Int. J. Geogr. Inf. Sci. 20, 135–152.

Gregory, I.N., Geddes, A., 2014. Toward Spatial Humanities: Historical GIS and Spatial History. Indiana University Press, Bloomington.

Gutiérrez, J., Urbano, P., 1996. Accessibility in the European Union: the impact of the trans-European road network. J. Transp. Geogr. 4, 15–25.

Hackett, S., Parmanto, B., Zeng, X., 2005. A retrospective look at website accessibility over time. Behav. Inf. Technol. 24, 407–417.

Haggett, P., Fréchou, H., 1973. L'analyse spatiale en géographie humaine, Armand Colin. ed. Paris.
Hansen, W.G., 1959. How accessibility shapes land use. J. Am. Inst. Plann. 25, 73–76.
Holl, A., 2007. Twenty years of accessibility improvements. The case of the Spanish motorway building programme. J. Transp. Geogr. 15, 286–297.
Janelle, D.G., 1969. Spatial reorganization: a model and concept. Ann. Assoc. Am. Geogr. 59, 348–364.
Kaddouri, L., 2007. Structures spatiales et mises en réseaux de villes pour la régionalisation des territoires. Université Paul Valéry - Montpellier III, Montpellier.
Knowles, A.K., 2005. Emerging trends in historical GIS. Hist. Geogr. 33, 7–13.
Kwan, M.-P., Janelle, D.G., Goodchild, M.F., 2003. Accessibility in space and time: A theme in spatially integrated social science. J. Geogr. Syst. 5, 1–3.
Langran, G., 1992. Time In Geographic Information Systems. CRC Press.
Martí-Henneberg, J., 2013. Introduction to the Special Issue on the Common Ground of History and Geography. Hist. Methods J. Quant. Interdiscip. Hist. 46, 113–116.
McMaster, R.B., Noble, P., Master, R.B.M., 2005. The U.S. National Historical Geographic Information System. Hist. Geogr. 33.
Motte, C., Séguy, I., Théré, C., 2003. Communes d'hier, communes d'aujourd'hui : les communes de la France métropolitaine, 1801-2001 : dictionnaire d'histoire administrative. INED.
Offner, J.-M., 1993. Les « effets structurants » du transport : mythe politique, mystification scientifique. Espace Géographique 22, 233–242.
Ollivro, J., 2000. L'homme à toutes vitesses : de la lenteur homogène à la rapidité différenciée. Presses universitaires de Rennes, Rennes.
O'Sullivan, S., Morrall, J., 1996. Walking Distances to and from Light-Rail Transit Stations. Transp. Res. Rec. J. Transp. Res. Board 1538, 19–26.
Pumain, D., 1997. Pour une théorie évolutive des villes. Espace Géographique 26, 119–134.
Pumain, D., 1982. Chemin de fer et croissance urbaine en France au XIXe siècle. Ann. Géographie 91, 529–550. doi:10.3406/geo.1982.20140
Schwartz, R., 2010 "Rail Transport, Agrarian Crisis, and the Restructuring of Agriculture: France and Great Britain Confront Globalization, 1860-1900, *Social Science History* 34:2: 229-55.
Schwartz, R., Gregory, I., Thévenin, T., 2011. Spatial History: Railways, Uneven Development, and Population Change in France and Great Britain, 1850–1914. J. Interdiscip. Hist. 42, 53–88.Schwartz, R., Thévenin, T., 2014. Railways and agriculture in France and Great Britain, 1850 to 1914, in: Ian N. Gregory, A.G. (Ed.), Toward Spatial Humanities: Historical GIS and Spatial History, The Spatial Humanities. Indiana University Press, pp. 4–34.
Southall, H.R., 2014. Applying Historical GIS beyond the Academy: Four Use Cases for the Great Britain HGIS, in: Toward Spatial Humanities. Indiana University Press, Bloomington, p. 205.
Studeny, C., 1995. L'invention de la vitesse : France, XVIIIe-XXe siècle. Gallimard, Paris.
Thévenin, Thomas, Robert Schwartz and Loïc Sapet, 2013. "Mapping the Distortions in Time and Space: The French Railway Network 1830-1930." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46, no. 3 (2013): 134-143.
Vanhaute, E., 2005. The Belgium Historical GIS. Hist. Geogr. 33, 136–139.
Vickerman, R.W., 1995. The regional impacts of Trans-European networks. Ann. Reg. Sci. 29, 237–254.

[i] These results show longer travel times than those estimated by C. Studeny, presumably because his figures did not take into account the time to walk to the nearest station (Blier, 1996; Studeny, 1995).