

Heinold, Arne; Makowski, Connor

Article — Published Version

Driving the Extra Mile: Comprehensive Analysis of Road and Rail Transportation Distances in Europe

Networks and Spatial Economics

Suggested Citation: Heinold, Arne; Makowski, Connor (2026) : Driving the Extra Mile: Comprehensive Analysis of Road and Rail Transportation Distances in Europe, Networks and Spatial Economics, ISSN 1572-9427, Springer US, New York, Vol. 26, Iss. 2, pp. 633-662, <https://doi.org/10.1007/s11067-026-09729-y>

This Version is available at:

<https://hdl.handle.net/10419/342915>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by/4.0/>



Driving the Extra Mile: Comprehensive Analysis of Road and Rail Transportation Distances in Europe

Arne Heinold¹ · Connor Makowski²

Received: 12 June 2025 / Accepted: 30 January 2026 / Published online: 19 February 2026
© The Author(s) 2026

Abstract

This paper provides two empirical analyses of road and rail transportation distances in Europe. The first analysis examines the proximity to major roads and rail stations, addressing connectivity within regions. The proximity of countries to both networks ranges from 3 to 12 kilometers, with an average of 8 kilometers, and shows a strong correlation between the two. We formulate and test hypotheses demonstrating that population density and historical factors largely explain these connectivity differences. For instance, locations in Eastern European countries are farther from major roads but closer to rail stations than those in other European countries. However, there are exceptions, and we then discuss implications for transportation policies by comparing the performance of European countries against a regression-based accessibility prediction. The second analysis examines circuitry, the ratio between transportation and great-circle distance, to address connectivity between regions. Circuitry between countries ranges from 1.15 to 1.76 (median 1.25) for road transportation and 1.13 to 2.07 (median 1.42) for rail transportation. Hypotheses are formulated and tested to describe the regional and mode-specific variations. For instance, peripheral countries like Italy and Greece have higher circuitry than other countries, and circuitry of road and rail networks is positively correlated. While circuitry decreases for longer road distances, no such relationship is found for rail. However, circuitry reliability improves with distance in both networks, with average estimation errors falling below 5% for distances exceeding 2,500 kilometers. The combined insights from both analyses offer a comprehensive overview of Europe's road and rail transportation infrastructure.

Keywords European transportation network · Proximity analysis · Circuitry · Road and rail transportation · Transportation network analysis

1 Introduction

Distance plays a central role in a variety of academic disciplines and real-world analytical challenges, including the projection of transportation costs, travel behavior, and global trade patterns. The term “distance” refers to the covered mileages or kilometers when moving between two geographical locations, the origin and the destination, and can be measured in several ways. In this context, the great-circle distance, referred to as straight-line distance, Haversine distance, or Euclidean distance, depicts the airline distance between two locations (“as the crow flies”), and the transportation distance refers to the shortest path between two points via the actual transportation network. This paper systematically analyzes distance-based connectivity differences in road and rail transportation for the entire European continent. We do so by providing two analyses: a proximity analysis and a circuitry analysis. Figure 1 illustrates the study’s approach.

In a first step, our proximity analysis calculates the great-circle distance between geographical locations and major roads and rail stations. This proximity has often been used to assess transportation accessibility and spatial effects. Yet, only a few studies have systematically analyzed and reported differences for the entire European continent, considering both the road and rail network infrastructure. This paper closes the gap and formulates and tests hypotheses to describe the varying levels of regional and mode-specific connectivity. Based on this, we present results from comparing the performance of European countries against a regression-based accessibility prediction. Our findings can support the identification of regional performance disparities, which in turn can support designing effective policies, such as decisions on infrastructural investments. In a second step, our circuitry analysis determines the

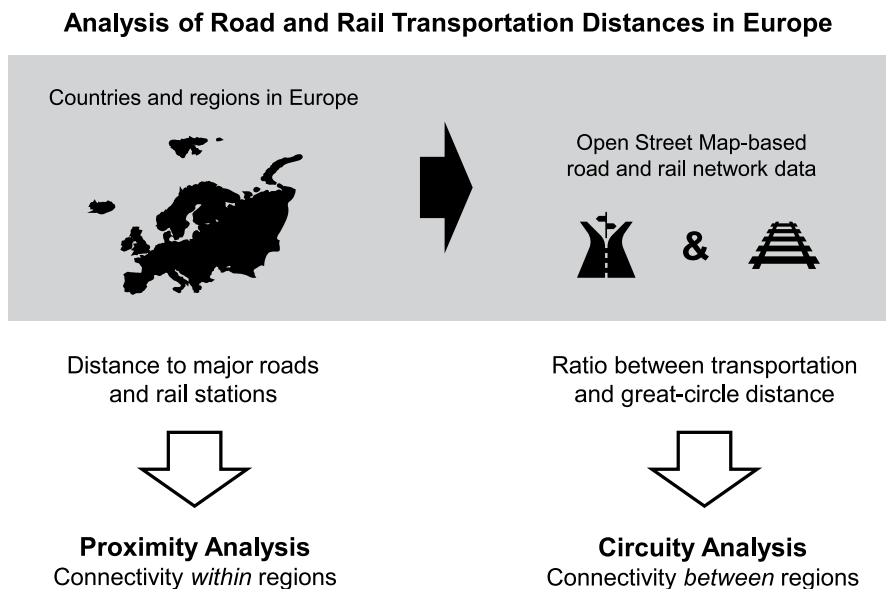


Fig. 1 Scope of this study: Comprehensive analysis of European road and rail transportation distances

ratio between the transportation and great-circle distances, referred to as circuitry or detour index. The transportation distance is by definition higher than the great-circle distance due to factors like detours for transshipment activities or geographical characteristics such as mountains or lakes. However, regardless of the exact cause, circuitry has been used to measure a network's distance-based efficiency, complementing other efficiency measures such as processed trade volume and road congestion (e.g., Wiegmans et al. 2018, Reggiani et al. 2022). A circuitry closer to one generally indicates more efficient networks (e.g., fewer detours) compared to values that are farther from one. In this context, we note that we strictly refer to distance-based efficiency. While closing roads cannot improve distance-based efficiency, it may improve time-based efficiency due to factors such as improved traffic flow. Circuitry values are also used to approximate the actual distance via the great-circle distance, especially when it is (computationally) challenging to compute exact distances for large and dynamic datasets (e.g., Ulmer and Thomas 2018). Despite circuitry's relevance, only a limited number of studies on the European transportation network provide the necessary scale of information. To close this gap, we present circuitry results for the road *and* rail transportation networks on Europe's country *and* regional level. Our analysis quantifies the significant circuitry differences across the continent and between the two modes. Based on this, we formulate and test hypotheses to explain the observed variations.

To sum up, our paper describes results from a proximity *and* circuitry analysis for the road *and* rail network. Each analysis offers unique and novel insights for the entire European continent: the proximity analysis reveals the connectivity *within* regions, while the circuitry highlights the connectivity *between* regions. To the best of our knowledge, none of the previous works have presented results at a comparable scale. We then discuss our findings from both analyses to provide a comprehensive picture of the European road and rail network, identifying regional similarities and differences between the two metrics, which constitute the basis of informed policy-making. The rest of the paper is structured as follows. Section 2 provides a literature review and outlines this paper's contribution. Based on this, Section 3 presents the paper's methodology and data sources. Section 4 describes numerical and visual results from the proximity and circuitry analyses and formulates hypotheses. Section 5 tests the hypotheses and discusses the results and potential policy implications. Section 6 concludes the paper.

2 Literature Review and Contribution

This section reviews relevant literature across two domains. Section 2.1 examines studies on distance-based accessibility, highlighting the contribution of our proximity analysis. Section 2.2 discusses approximation methods for transportation distance, identifying works related to our circuitry analysis. Finally, Section 2.3 synthesizes the research gap and outlines the main contributions of this paper.

2.1 Accessibility to Transportation Networks

Accessibility to transportation networks is a key metric for describing users' opportunities or possibilities to use a specific mode of transport within a network. The topic has a long-standing history of academic research and has been covered from different, often interdisciplinary, perspectives. Macroeconomic analyses have extensively discussed the role of accessibility to transportation for the overall economic development (e.g., Cain 1997, Weisbrod 2008, Caragliu 2008), while microeconomic analyses have focused on specific aspects of an economy or region. For instance, early work by Hansen (1959) analyzes the impact of accessibility on land use, highlighting the predominant importance of considering distance when analyzing accessibility. In fact, distance has been recognized as a key metric for accessibility ever since (e.g., Kwan 1998, Geurs and van Wee 2004, Miller 2018, Siddiq et al. 2021, Pourramazani and Miralles-Garcia 2023, Jang and Park 2025) and has been used to measure accessibility in many different transportation settings. For example, distance-based spatial accessibility is considered in Guagliardo (2004) for primary care coverage and in Mercier et al. (2021) for postal service coverage. Both studies, however, critically reflect on using distance as the sole measure of accessibility, emphasizing the need for other dimensions. Otsuka et al. (2019) demonstrate this in the context of European transport networks. The authors evaluated the accessibility to urban rail stations along the TEN-T (Trans-European Transport Network) corridors. The European Union defines these corridors to describe and improve key cross-country and multi-modal transportation infrastructure connections (e.g., north-south). In their study, the authors used a radius of 800 meters around an urban rail station to define the generally accessible region. However, in a second step, they identified other important aspects, such as walkability and the availability of other transport modes (e.g., bike sharing). The proposed methodology has been applied to six railway nodes along one of the TEN-T corridors. This indicates that, while numerous factors influence outcomes at a fine-grained spatial scale (e.g., below 800 meters), distance remains a dominant accessibility metric when viewed from a broader, global perspective.

In addition to research focused on measuring accessibility ("How?"), a complementary stream has emerged to investigate accessibility differences and the underlying factors contributing to it ("Why?"). The direction of relevant factors and accessibility, however, has been shown to be complex. For example, Levinson (2007) investigate whether land use patterns preceded the development of rail infrastructure or vice versa. Drawing on historical panel data between 1871 and 2001, Levinson (2007) reveal a nuanced, bidirectional relationship between the two. What is clear, though, is that multiple factors explain the different levels of accessibility in transportation networks, including economic, environmental, and social aspects (e.g., Geurs et al. 2009). Consequently, Tartaglia et al. (2025) propose extending traditional cost-benefit analysis frameworks to incorporate broader social impacts, thereby making it inclusive of more dimensions, in transportation planning decisions. Kim et al. (2024) show the importance of this by examining the spatial distribution of bus shelters across 20 U.S. cities, revealing that areas with higher proportions of minority populations tend to have more bus stations. Moreover, they identify a strong correlation with population density, where denser areas usually have more shelters. Furthermore,

Barthélemy and Flammini (2009) analyze the co-evolution between population density and the road network in 20 international cities. More generally, several studies have demonstrated a positive correlation between network infrastructure accessibility (e.g., spatial proximity) and demand indicators (e.g., population density). For example, Lopresti and Tartaglia (2024) show that high-speed rail connections in Italy are correlated to tourism demand and Quinn (2013) even demonstrates that road density can serve as a proxy for population density. This relationship has also been the focus of studies analyzing the historical relationship between population and transportation accessibility in Europe. Alvarez et al. (2013) show the positive impact of the railway network on population growth in England and Wales by analyzing data between 1871 and 1931. Similarly, Kotavaara et al. (1970) analyze the population change between 1970 and 2007 in Finland and relate it to the accessibility to the road and railway network. The authors demonstrate that both road and rail accessibility played a significant role in population change, as the population generally increased in areas with greater accessibility. Finally, several studies have demonstrated the persistent influence of historical infrastructure developments, which remain evident many years after their construction and, in some cases, regardless of whether the infrastructure is still in use (e.g., Brooks and Lutz 2019, Flückiger and Ludwig 2019, Caragliu and Del Bo 2025, Balboni 2025).

We contribute to this stream of literature by providing a comprehensive proximity analysis. This is, we present detailed results for distance-based accessibility across both road and rail networks for the entire European continent. Following this, we also investigate the relationship to population density data. To the best of our knowledge, no prior research has been done at a comparable scale.

2.2 Approximation Methods for Transportation Distance

This section reviews papers that propose or utilize (analytical) approximation methods for transportation distance. Despite its simplicity, the great-circle distance has been used in various contexts to assess the proximity to the transportation infrastructure. For instance, Gutiérrez and Urbano (1996) use it to evaluate the proximity of locations to the TEN-T network corridors, using a corridor's width of 40 kilometers. It has also been used in studies as a proxy for transportation distance or travel time. In particular, due to its high correlation of up to 98.7% with actual travel time (Phibbs and Luft 1995), the great-circle distance has been used in spatial analyses within the healthcare system (e.g., Bliss et al. 2012, Moynihan et al. 2012) or to assess the accessibility of public service facilities (Stroppe 2023). In general, the great-circle distance has been used when the exact distance was not critical or when it was (computationally) challenging to compute exact distances for large datasets.

A similarly computationally lean approach, however, is to multiply the great-circle distance by some index to achieve more realistic values. For instance, Gibbons and Wu (2020) use the so-called Manhattan distance ($\sqrt{2}$ · great-circle distance), a grid-like approach, to approximate distances between countries and airports. Based on this, the circuitry is a more context-specific index, such as using circuitry for urban areas or circuitry for railway transportation. This can lead to more accurate values, which is important when distance is a critical input, such as when measuring costs,

fuel consumption, or emissions (e.g., Demir et al. 2011, Heinold 2020). For instance, circuitry-based distances are used in the dynamic decision making of same-day deliveries (Ulmer and Thomas 2018; Chen et al. 2023), in the calculation of recreational trips around a river (Keiser and Shapiro 2019), and in a study analyzing geographic access to cancer screening facilities (Tailor et al. 2019). Studies have also considered circuitry as a network efficiency metric. For instance, Mert Cubukcu (2020) use circuitry to assess the efficiency of the transportation network in Paris, Giacomini and Levinson (2015) analyze the circuitry of random points across 51 metropolitan areas in the United States over a period of 20 years, Huang and Levinson (2015) use circuitry as one measure to assess the efficiency of public transportation networks in metropolitan areas in the United States, and Merchán et al. (2020) explain circuitry differences in urban road networks as a result of dimensional and topological properties. Finally, Choudhry and Qian (2024) use circuitry, employing actual distances tracked via GPS as transportation distance, to measure a truck's route efficiency in urban last-mile deliveries.

Due to this widespread interest, a large body of research has emerged on the disparity between great-circle and actual distances, providing empirical circuitry values. The following provides an overview of those studies. Ballou et al. (2002) present results from inter-city distances for selected countries. For Europe, the study reports an average circuitry factor of 1.46, which lies between the lowest (1.12 for Belarus) and highest (2.1 for Egypt) value within the data set. Boscoe et al. (2012) provide insights from analyzing population-based travel paths for over 66,000 census locations in the United States, the District of Columbia, and Puerto Rico. A circuitry factor of 1.42 was found for the total data set. The authors also identify a correlation between the great-circle distance and actual travel distance, with accuracy levels of 90% and 96% if errors up to 5 km and 10 km are tolerated, respectively. There are also several studies focusing on cities. Morita et al. (2014) report a circuitry for 112 cities in Japan, ranging from 1.13 to 1.92, with an average of 1.30. Yang et al. (2018) look at the urban road systems using taxi data from 10 cities in China. Here, the authors identify a correlation between circuitry and distance, stating that the data fits to a model in which “the mean of the detour ratios is inversely proportional to the Euclidean distance with an intercept.” Chen and Chen (2021) support this finding when computing the circuitry for 25 cities in China. While the indices range from 1.24 to 1.48, with an average index of 1.32, the authors note that “short-distance travels in a city usually require more detour than long-distance ones.” Similarly, Mennicken et al. (2024) analyze the road distances to the main center in 300 European cities and reports an average circuitry of 1.34, which generally decreases with increasing size of a city. Finally, Jorge et al. (2024) calculate circuitry values ranging from 1.39 to 2.09 for 25 regions within the London metropolitan area. These values were subsequently utilized in an optimization model, which improved the accuracy of assigning couriers to orders in an instant delivery service.

Based on this, our paper contributes by providing comprehensive circuitry results for road and rail transportation on the entire European continent. Thereby, we also present empirical findings on the relationship between distance and circuitry. To the best of our knowledge, no prior research has addressed this topic at a comparable scale.

2.3 Overall Contribution

This study offers new insights by presenting results from *two* analyses: a proximity analysis, which examines the great-circle distance between geographical locations and major roads and rail stations, and a circuitry analysis, which investigates the ratio between the transportation distance and the great-circle distance. Numerical and visual results are presented at the first (country) *and* second (regional) NUTS level (Nomenclature des Unités Territoriales Statistiques (French); nomenclature of territorial units for statistics (English)), which is the primary nomenclature of territorial units for statistics in Europe. This perspective enables the comparison of political and economic areas, such as countries, and the application to a wide range of applications, such as transportation modeling. While previous studies mainly concentrated on either the road or the rail network, this study examines both networks to reveal proximity and circuitry differences between the two modes. Additionally, it extends findings from previous studies, which have often focused on specific regions or contexts such as urban areas. With this, our results provide a comprehensive picture of the current European road and rail transportation network infrastructure. Based on this, we combine the insights from the two analyses and formulate and test hypotheses to explain the differences in the data. Finally, we note that we describe the efficient processing of the data, such as the use of a shortest path algorithm to compute distances within milliseconds in large transportation networks. In particular, our empirical findings are derived from Open Street Map data, which offers high-quality open-source information on the European road and rail transportation network. The description of the data processing could facilitate further research in this area.

3 Data and Methodology

This section describes the data on regions, the grid system, and the infrastructure as well as the efficient processing of this data when analyzing proximity and circuitry. Specifically, Section 3.1 describes the regions and the creation of the grid, while Sections 3.2 and 3.3 explain the methodologies for the proximity and the circuitry analyses, respectively.

3.1 Regions and Grid

This paper's analyses cover the region commonly referred to as mainland or continental Europe. Additionally, the proximity analysis includes Cyprus, Finland, Iceland, Ireland, Great Britain, Malta, Norway, and Sweden, essentially covering the entire European continent. For this area, artificial geographical locations (e.g., pairs of longitude and latitude values) are created in a grid-like pattern with a cell size of 1 square kilometers, resulting in about 6 million unique grids. Each grid's mid-point serves as a potential location for the analyses. The grid size of 1 square kilometer is considered reasonable, as studies focusing solely on urban areas use only slightly smaller values. For instance, Salonen and Toivonen (2013) rely on a grid size of 250 meters for modeling travel times in urban networks. Results from the

analyses are generally aggregated for different levels of the so-called NUTS regions. Within each NUTS region, we select a representative set of grid locations to ensure the convergence of values, as detailed in the following sections. NUTS regions are hierarchically clustered areas within the European Union (EU), described by EU's official statistical office (European Union 2024), and widely used in academic and non-academic settings (e.g., Leuthold et al. 2012, Condeço-Melhorado and Christidis 2018). Although defined by the EU, the NUTS-classification encompasses additional countries, such as potential EU candidates on the Balkan Peninsula, thereby essentially expanding Europe's continental scope. Results are reported either on the highest (NUTS-0: countries) or second-highest (NUTS-1: regions within a country) NUTS level. The black solid lines in Figs. 3 and 4 show the borders of the NUTS-1 regions.

3.2 Proximity Analysis

In the proximity analysis, the great-circle distance to the nearest major road and rail station is computed for each grid using infrastructural data from the open-source OpenStreetMap project (Osm n.d.). OSM data is of high quality and widely used in geographic analyses (e.g., Sun et al. 2020, Ninivaggi and Cutrini 2025). Europe's complete OSM data set was downloaded from Geofabrik (2024) and then processed country-by-country to extract the major roads and rail stations. In particular, in the context of the OSM methodology and for the purpose of this paper's analysis, ways with the tags [motorway, trunk, primary] were classified as major roads, and nodes with the tags [station, stop, halt] were classified as rail stations. Regarding rail stations, tags often overlap with what is commonly referred to as a "station". For example, a rail station can be tagged with any combination of the three tags, such as "stop" alone or all three tags together. We consider this by identifying geographical clusters of related tags. Results are presented only for grids within a 20-kilometer radius, as grids beyond this range are typically located in remote areas, such as small islands or non-populated regions.

3.3 Circuity

The circuity describes the ratio between the transportation distance and the great-circle distance, computed as follows (Rodrigue 2024):

$$\text{Circuity} = \frac{\text{Transportation Distance}}{\text{Great-Circle Distance}}. \quad (1)$$

For example, a region with a circuity of 1.5 indicates that a great-circle distance of 100 km results in a transportation distance of 150 kilometers. Considering the Earth's sphere, the great-circle distance is computed with the Haversine Formula (Robusto 1957). Approximating the transportation distance is more challenging due to complex route choice behaviors, depending on factors such as road quality, tolls, safety, demand, congestion, and distance (Prato 2009; Jayasinghe et al. 2016). For example, Huang et al. (2024) identify six distinct user clusters based on 290,000 trajectories

from 10,000 users in China, with only one cluster consistently choosing the shortest route. As a nuanced modeling of route choice behavior is not within the scope of this study, we treat transportation distance strictly as the shortest path between two nodes. Especially over long distances, travel distance is a key factor in route choice behavior, due to its strong correlation with transport time (Phibbs and Luft 1995). Despite this simplification, computing the shortest path remains challenging due to the network's size, which has many nodes and links. Luxen and Vetter (2011) describe a real-time shortest-path routing with OSM data, as implemented in the Open Source Routing Machine (OSRM) for the road network. The general idea is to efficiently preprocess the large dataset to allow for fast, real-time routing. We follow a similar approach and compute the transportation distance with the OSM-based Python project *SCGraph* (Makowski et al. 2025). *SCGraph* includes an efficient algorithm for calculating the shortest path through the main elements of the world's infrastructural networks, such as highways and major rail tracks. In particular, *SCGraph* adds the origin/destination into the road/rail network by creating connections to the nearest vertex in that network (connection point) in each direction (North-East, North-West, South-East, South-West). The assigned distance for each connection is the great-circle distance. The way origins and destinations are added to the network is necessary because a specific geographical location is only by coincidence on an existing road or railway track. *SCGraph* then solves for the shortest network transportation distance between the origin and destination points. It is worth noting that the connection points for adding the origin/destination to the network are not necessarily access points to that road or rail system, but rather the nearest vertex in the network data. Otherwise, we would require specific knowledge about the vehicle or train type, e.g., passenger vs. freight trains or electric vs. diesel locomotives, which is beyond the scope of this paper.

In theory, circuitry can be computed between all pairs of grids. However, as mentioned before, we are working with approximately six million unique grids, which would result in over 36 trillion possible connections. Even with multithreading and access to high-performance computing resources, such a scale is computationally infeasible. Instead, the circuitry results are presented using the mean values for a random selection of 200 connections between two regions. For example, the circuitry between Spain and Greece is based on 200 random connections, which can go in either direction. Figure 2 demonstrates that the sample size of 200 is sufficient for the mean circuitry values to converge. The figure illustrates the convergence of the cumulative average circuitry for four selected country connections: two within the road network and two within the rail network. All four connections converge after approximately 100 iterations, with some connections, such as road transportation between Poland and Lithuania, converging earlier, and others, such as rail transportation between Italy and Croatia, converging later. The reasoning behind the different convergence behaviors is discussed later, but it is clear that there is a strong overall convergence.

Finally, we note three distinct cases that we observe in the circuitry values. First, extremely narrow locations result in a circuitry value of 1.0. This occurs because the shortest path in *SCGraph* is then computed solely via the great-circle distance, without involving network connection points at either the origin or the destination. Second, very remote areas within a country tend to inflate the circuitry. For instance,

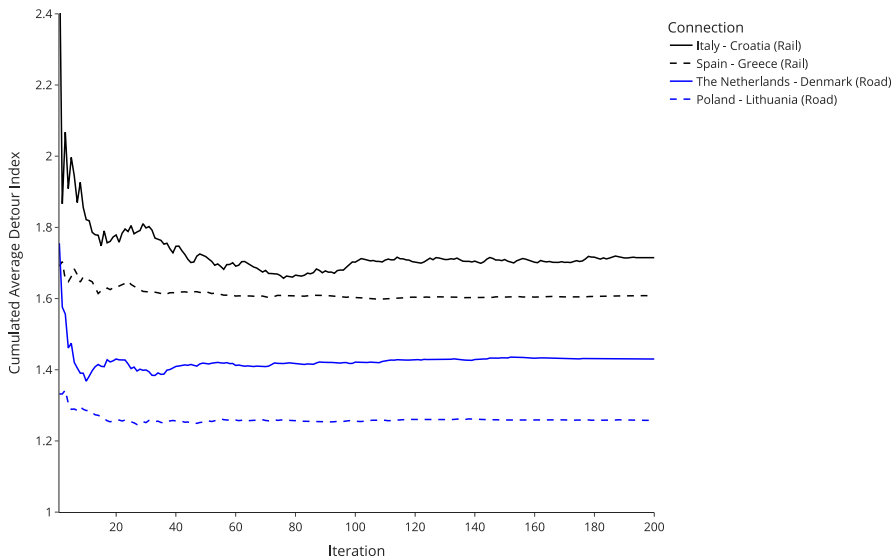


Fig. 2 Convergence of the cumulated average circuitry for four selected country connections

Portuguese islands located far from mainland Portugal yield unusually high average circuitry values for the country. Third, data quality issues in the open-source dataset may affect results. For example, a “missing” rail link could introduce artificial detours within the network. These issues are addressed by excluding outliers. For this, we compute the lower and upper bounds for each connection by subtracting and adding 1.5 times the interquartile range ($Q3 - Q1$) from the first quartile ($Q1$) and to the third quartile ($Q3$), respectively. Then, we clip the bounds to lie within the interval $[1, 3]$, ensuring that the bounds are not extreme and remain meaningful. We then only consider circuitry values within the lower and upper bounds, thereby excluding an average number of outliers of 7.8 (3.9%) and 5.9 (2.9%) on the country and regional level, respectively.

4 Numerical and Visual Findings

This section presents numerical and visual findings from the proximity analysis and the circuitry analysis in Sections 4.1 and 4.2, respectively. Each section offers a comprehensive analysis of the respective metrics, which serves as the basis for our hypothesis formulation. With this, our hypotheses are grounded in empirical observations from the dataset and usually reinforce established findings described in the literature for other contexts (e.g., other geographical regions).

4.1 Proximity Analysis

This section presents results for the proximity of the grid locations to the road and rail transportation network.

Figures 3 and 4 illustrate the distance of grid locations to the nearest major road and rail station, respectively. The figures use a color scale, with darker colors indicating shorter distances and lighter colors representing greater distances. The borders of the NUTS-1 regions are shown as solid black lines. It can be observed that differences exist between road and rail networks and between regions in Europe. For the road network, Western Europe appears to provide better connectivity than Eastern Europe. This can be seen as grids have a lighter color in the Czech Republic and Romania, as well as in countries in the Balkans, indicating that these regions are less connected

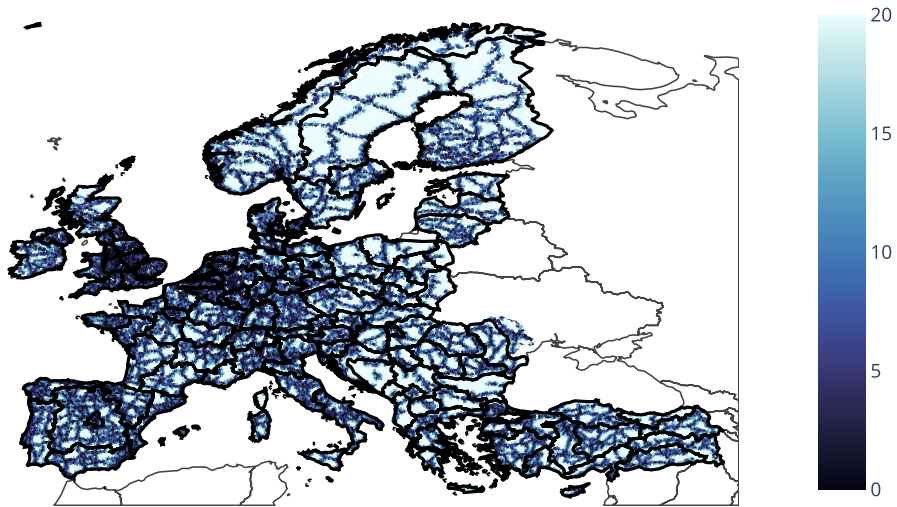


Fig. 3 Color coded distance to nearest major roads (in km)

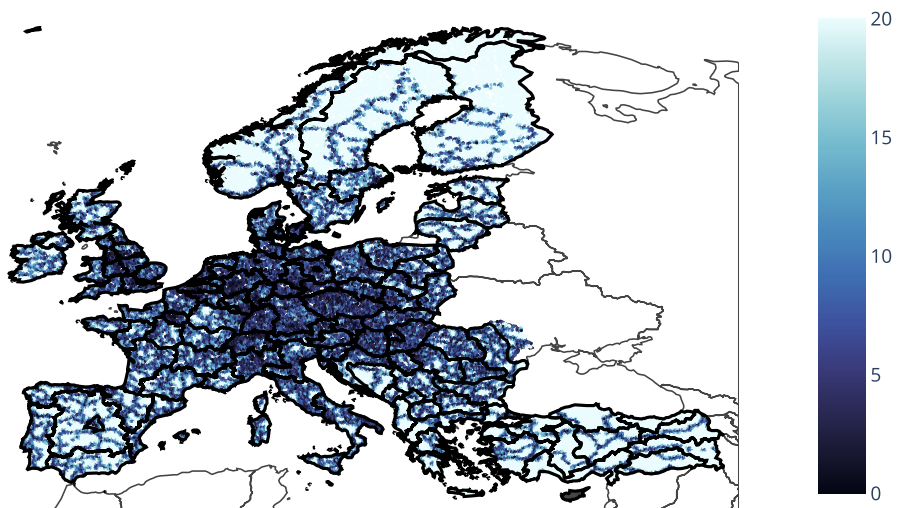


Fig. 4 Color coded distance to nearest rail stations (in km)

when compared to other European regions. The situation is somewhat different for the rail network. Western and Eastern European countries exhibit very good coverage of rail stations, as indicated by the darker color on the maps. In contrast, countries in Southern Europe, such as Spain and Portugal, as well as countries like Greece and Turkey, seem to have a less dense coverage of rail stations. Notable exceptions are Norway, Sweden, and Finland, which appear less connected to major roads and rail stations. This appears to be driven by low population densities in the north. Thus, a low proximity does not necessarily imply that a network is not serving its purpose, as there is less need for extensive connectivity in areas with low population density. It is also interesting to see that major roads *and* railway tracks are visually identifiable on the maps in less populated areas, such as Northern Europe. Likewise, metropolitan areas in Europe, such as the well-connected cities of Madrid in Spain and Paris in France, can be identified. Dark-blue areas on the road *and* rail maps surround both of these major cities. Both of these observations suggest a generally high correlation between proximity to major roads and proximity to rail stations.

To support the visual findings, Table 1 presents the mean, median, and standard deviation of the grid locations' proximity aggregated on a country level as well as for Northern, Eastern, Southern, and Western Europe at the bottom of the table. While the values in the table are not weighted by population density, results from a corresponding analysis are described later in Section 5.1 for countries within the European Union. The table clearly shows the significant differences among the countries. For instance, grid locations in western European countries, such as Belgium, the Netherlands, and Luxembourg, have some of the lowest values for both transportation networks. In contrast, the highest values are observed in Balkan countries, such as Montenegro, Bosnia-Herzegovina, and Albania.

Figure 5 extends this analysis by showing the proximity distances as a box plot for the road and rail networks at the country, regional, and grid location levels. The median values across all three levels and for both networks are remarkably similar. In particular, the distance to the nearest major road or railway station in Europe is, on average, approximately 8 kilometers. This indicates that the connectivity of the two transportation modes is comparable. However, the levels have a significant spread, with higher spreads observed in more granular data. For example, the minimum and maximum values at the country level are approximately 3 and 12 kilometers, respectively, while the full range of values between 0 and 20 is seen at lower levels. This can be explained by aggregated data averaging out extreme values. In conclusion, based on the results described above, we propose the following hypotheses:

Hypothesis 1 *Locations in Eastern European countries are farther away from major roads compared to those in other European countries.*

Hypothesis 2 *Locations in Eastern and Western European countries are closer to rail stations compared to those in other European countries.*

Hypothesis 3 *In Europe, the connectivity to major roads and rail stations increases with rising population density.*

Table 1 Distance to nearest major road and rail station for countries in Europe (in km)

Country	Region ¹	Road			Rail		
		Mean	Median	Std. Dev.	Mean	Median	Std. Dev.
Albania (AL)	South	9.67	9.43	5.63	10.10	9.79	5.23
Austria (AT)	West	8.05	7.24	5.39	6.66	5.55	4.68
Belgium (BE)	West	4.55	3.64	3.70	5.64	4.61	4.10
Bosnia-Herzegovina (BA)	South	10.62	10.91	5.68	9.94	9.70	5.45
Bulgaria (BG)	East	10.12	10.09	5.65	8.67	7.89	5.32
Croatia (HR)	South	8.34	7.50	5.50	7.84	6.79	5.17
Cyprus (CY)	South	7.08	5.77	5.43	—	—	—
Czech-Republic (CZ)	East	9.01	8.56	5.62	4.07	3.28	3.05
Denmark (DK)	North	7.71	6.86	5.24	7.30	6.34	4.74
Estonia (EE)	North	9.10	8.61	5.62	10.01	9.72	5.44
Finland (FI)	North	8.71	8.01	5.51	11.58	11.86	5.08
France (FR)	West	8.09	7.22	5.42	8.66	7.96	5.04
Germany (DE)	West	7.08	5.97	5.23	5.25	4.38	3.71
Greece (EL)	South	8.44	7.72	5.64	9.92	9.60	5.43
Hungary (HU)	East	9.37	9.03	5.64	5.34	4.55	3.57
Iceland (IS)	North	9.61	9.24	5.58	—	—	—
Ireland (IE)	North	7.94	7.01	5.39	10.53	10.39	5.01
Italy (IT)	South	7.29	6.22	5.28	7.28	6.17	4.87
Kosovo (XK)	South	9.44	9.10	5.61	9.67	9.44	5.36
Latvia (LV)	North	8.85	8.23	5.50	9.69	9.24	5.24
Liechtenstein (LI)	West	—	—	—	6.45	6.23	3.91
Lithuania (LT)	North	8.55	7.84	5.52	10.11	9.88	5.32
Luxembourg (LU)	West	6.37	5.11	5.00	5.42	4.24	4.22
North Macedonia (MK)	South	9.11	8.71	5.60	9.35	8.88	5.39
Malta (MT)	South	3.31	1.88	3.70	—	—	—
Moldova (MD)	East	8.52	7.63	5.47	8.76	7.99	5.33
Montenegro (ME)	South	10.95	11.43	5.41	9.61	9.25	5.50
Netherlands (NL)	West	4.16	3.21	3.47	6.47	5.46	4.35
Norway (NO)	North	8.71	8.08	5.56	9.91	9.58	5.35
Poland (PL)	East	9.50	9.22	5.69	7.50	6.42	4.98
Portugal (PT)	South	7.35	6.08	5.47	9.15	8.55	5.42
Romania (RO)	East	8.79	8.17	5.57	7.92	6.96	5.07
Serbia (RS)	South	9.35	9.00	5.68	8.86	8.15	5.34
Slovakia (SK)	East	9.43	9.16	5.56	5.73	4.79	4.04
Slovenia (SI)	South	8.43	7.77	5.53	7.02	6.03	4.73
Spain (ES)	South	7.22	6.12	5.27	9.98	9.72	5.35
Sweden (SE)	North	9.06	8.53	5.59	10.13	9.88	5.18
Switzerland (CH)	West	6.87	5.58	5.22	4.63	3.39	3.94
Turkey (TR)	South	8.64	7.91	5.49	10.44	10.28	5.30
United Kingdom (UK)	North	5.04	3.60	4.50	7.24	6.24	4.85
Europe	All	8.03	7.09	5.51	8.18	7.22	5.23
	North	7.99	6.99	5.55	9.40	8.91	5.31
	East	9.23	8.82	5.64	7.09	5.90	4.91
	South	7.99	7.06	5.47	9.18	8.56	5.37
	West	7.41	6.32	5.32	6.98	5.87	4.77

¹: According to the United Nations geoscheme (<https://unstats.un.org>)

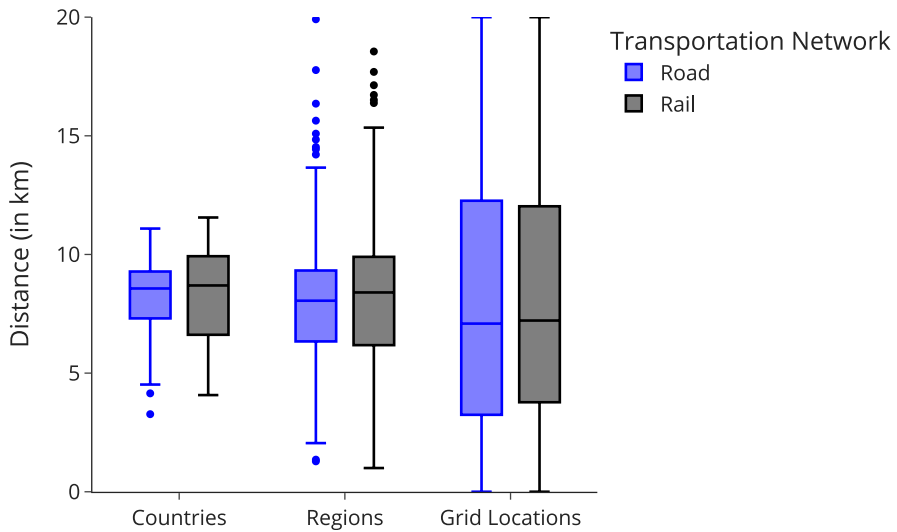


Fig. 5 Distance to nearest major road or rail station on a country, regional, and grid location level (in km)

For testing hypotheses relating to cardinal directions, each country is assigned to a single region based on the United Nations geoscheme (see second column of Table 1). For Hypothesis 2, we perform separate tests for countries from Eastern (Hypothesis 2.1) and Western Europe (Hypothesis 2.2). Likewise, for Hypothesis 3, we separately test the impact on the road (Hypothesis 3.1) and rail (Hypothesis 3.2) networks.

4.2 Circuity Analysis

This section presents circuity results aggregated at the country or regional level. The scope of the results is limited to countries within continental Europe, as the circuity is less meaningful for islands or remote locations. The complete data set, including the average circuity and the corresponding standard deviation, is provided in an electronic appendix for each country pair. Figure 6 shows the distribution of the circuity for each combination of countries (left) and regions (right) for both the road (results colored in blue) and rail (results colored in black) transportation modes. The corresponding box plots are displayed at the top of the figure. The circuity is notably higher for rail compared to road. Across all country connections, the median circuity is 1.25 for road and 1.42 for rail. Similarly, across all regions, the median circuity is 1.23 for road and 1.38 for rail. This indicates that, on average, rail transportation involves significantly longer detours than road transportation. However, it is also evident that there is significant variability in the values within each mode at the country and regional levels. Specifically, the circuity between countries ranges from 1.15 to 1.76 for the road network, and from 1.13 to 2.07 for the rail network, with even greater variability observed when comparing regions. This suggests that it is essential to consider connection-specific circuity for accurate results.

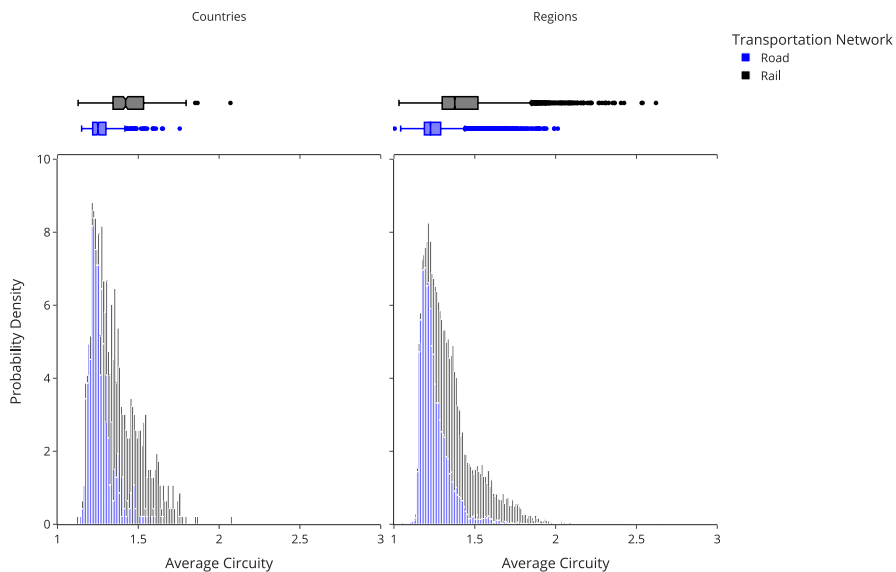


Fig. 6 Probability distribution of the circuitry on the country (left) and regional (right) level

To further analyze the spread within Europe, Figs. 7a and 7b show the average circuitry between country pairs. A blue color scale is used, where lighter shades represent lower circuitry and darker shades represent higher circuitry. Comparing the two figures reveals distinct patterns in the circuitry values, with geography being the most significant determinant. First, centrally located countries like Germany, France, and Belgium exhibit some of the lowest circuitry, while countries located at the periphery of Europe, such as Greece and Italy, have relatively high circuitry. Second, certain connections have exceptionally high circuitry. For instance, the high circuitry between

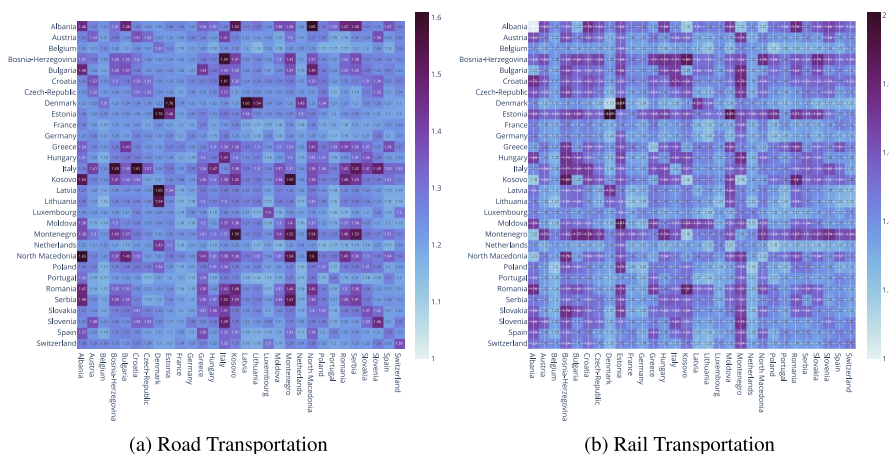


Fig. 7 Circuitry of traveling via road and rail between European countries

Estonia and Denmark is primarily due to the necessity of traveling around the Baltic Sea. Similarly, the Adriatic Sea influences the high circuitry between Croatia and Italy, resulting in significant detours for the relatively short great-circle distance between these countries. Third, at least for road transportation, connections with longer distances tend to have a lower circuitry. For instance, the circuitry is 1.27 between Portugal and Spain (average great-circle distance of ≈ 500 km), whereas it is 1.17 between Portugal and Poland (average great-circle distance of ≈ 2.500 km). This suggests that detours are less significant for longer distances. Overall, the findings presented in this section provide valuable insights into the circuitry for continental Europe, useful in decision-making scenarios where precise location information may not be available. However, there are significant differences, and we propose the following circuitry-related hypotheses:

Hypothesis 4 *In Europe, rail transportation has higher circuitry than road transportation.*

Hypothesis 5 *In Europe, peripheral countries have higher circuitry than other countries.*

Hypothesis 6 *In Europe, connections with longer distances have a lower circuitry.*

To test Hypotheses 5 and 6, we conduct separate analyses for the road and rail networks.

5 Discussion and Hypotheses Testing

This section discusses the numerical and visual findings from the previous section and tests the hypotheses. The hypotheses are categorized as either comparative (e.g., x is higher than y) or correlative (e.g., x is positively correlated with y). For the comparative hypotheses, the mean values of each group and the t -test results are presented in Table 2. For the correlative hypotheses, both the Spearman correlation coefficient, which identifies linear and non-linear relationships, and the Pearson correlation coefficient, which identifies linear relationships, are shown in Table 3. The

Table 2 Statistical results for comparative Hypotheses 1, 2, 4, and 5

Hypothesis	Analysis	Network	Group 1	Group 2	t -test ¹	Mean		t -value
						Group 1	Group 2	
1	Proximity	Road	East	Other	greater	9.23	7.84	175.6332***
2.1	Proximity	Rail	East	Other	lower	7.09	8.48	-215.4489***
2.2	Proximity	Rail	West	Other	lower	6.98	8.62	-274.3622***
4	Circuitry	-	Rail	Road	greater	1.42	1.25	81.1585***
5.1	Circuitry	Road	Periphery	Other	greater	1.30	1.24	21.9955***
5.2	Circuitry	Rail	Periphery	Other	greater	1.49	1.41	16.0050***

¹: The mean of group 1 is greater or lower compared to the mean of group 2.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3 Statistical results for correlation Hypotheses 3 and 6

Hypothesis	Analysis	Group 1	Group 2	<i>t</i> -test ¹	Country		Region	
					Spearman ²	Pearson ²	Spearman ²	Pearson ²
3.1	Proximity	Road	Pop. Density	lower	-0.6236***	-0.7563***	-0.6059***	-0.66**
3.2	Proximity	Rail	Pop. Density	lower	-0.6667***	-0.5610***	-0.6836***	-0.67**
6.1	Circuitry	Road	Distance	lower	-0.4767***	-0.3994***	-0.3955***	-0.38**
6.2	Circuitry	Rail	Distance	lower	-0.0347	-0.0198	-0.1412***	-0.12**

¹: Values in group 1 are positively (> 0, greater) or negatively (< 0, lower) correlated with values in group 2.

²: -1 (negative correlation), 0 (no correlation), 1 (positive correlation)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

table also includes *t*-test results to assess the significance of the correlations. With this, Section 5.1 provides reasons for differences in the proximity results (Hypotheses 1 to 3) and, based on this, Section 5.2 discussed policy implications. Section 5.3 explores the accuracy of circuitry values (Hypotheses 4 to 5), and Section 5.4 investigates the relationship between distance and circuitry (Hypothesis 6).

5.1 Proximity Differences

We start by examining Hypotheses 1 and 2, which address the variations among countries classified as Northern, Eastern, Southern, and Western Europe. The assignment of countries to regions is shown in Table 1. We test the hypotheses by comparing the mean values across all locations within a region to the mean values across all locations outside that region. For Hypothesis 1, the *t*-test results determine whether locations in Eastern Europe have a greater mean (e.g., are farther from major roads) and for Hypothesis 2, they determine whether locations in Eastern (Hypothesis 2.1) and Western (Hypothesis 2.2) Europe have a lower mean (e.g., are closer to rail stations). The results presented in Table 2 indicate that the hypotheses are significant at the 1% level and indicate differences that may result from historical developments. We refer to Section 2.1 for a review of studies focusing on these effects.

Based on this, Hypothesis 3 investigates the relationship between proximity to transportation infrastructure (roads and rail stations) and population density. We utilize census data to determine the population density of countries within the European Union (EU), excluding Liechtenstein as well as island nations of Malta and Cyprus. The census dataset provides population density information for an EU-wide 1 km² grid (GEOSTAT 2024), which can be aligned with the grid system used in our study. Figure 8 illustrates the results. Here, the proximity to major roads and rail stations is shown on the horizontal and vertical axes, respectively. The population density is depicted using a continuous color scale (capped at 250 persons per km²), with darker colors indicating higher densities. The results are further differentiated at the country and regional levels in Figs. 8a and 8b, respectively. The two-letter country code

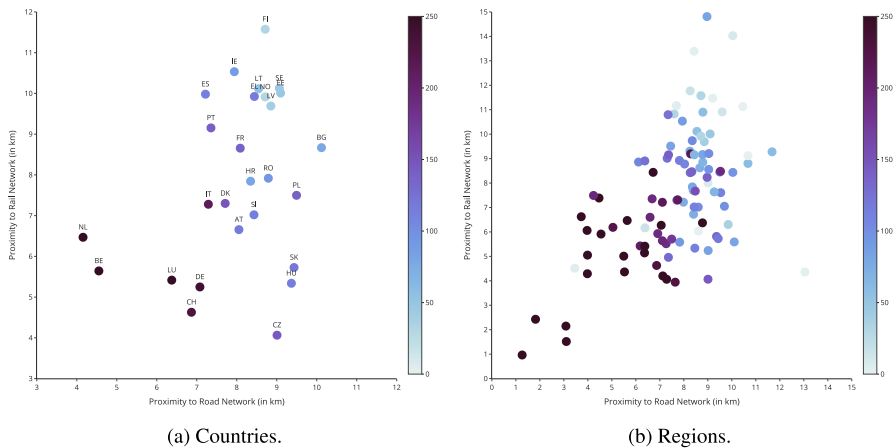


Fig. 8 Relationship between proximity to the road and rail network. The color scale depicts the population density (in persons/km²)

for each of the considered countries is provided in Table 1. Two key insights can be drawn from these figures:

1. **Correlation between population density and network proximity:** The proximity is generally positively correlated with the population density (Hypothesis 3). This is evident as the color of the countries/regions fades towards the top right corner, indicating that regions less connected to the road and rail network tend to have lower population densities. Likewise, countries and regions with high population densities, such as the Netherlands or Belgium, are located in the lower left corner of the matrix, indicating high proximities to both road and rail networks. This observation is also supported in the correlation statistics for Hypothesis 3 (see Table 3) and is in line with the long-standing research in this field, see Section 2.1. We further examine the relationship and discuss its potential policy implications in the following section.
2. **Correlation between the road and rail network proximity:** The figures show that the better a grid is connected to the road system, the better it is also connected to the rail system, vice versa. This correlation is particularly evident in Fig. 8b, where a nearly linear relationship between the two measures can be observed for the regional level. Compared to the country level, the stronger trend at the regional level may be due to significant regional variations within countries, particularly in large nations like France, Spain, and Germany. Thus, the road and rail networks are more similar at the smaller regional level, resulting in a stronger correlation. This general observation of a positive correlation between the public road and rail networks aligns with findings from the literature. It is primarily driven by the previously discussed correlation between population density and spatial proximity, observed in both road and rail networks (e.g., Kotavaara et al. 1970). However, such relationships are not universally observed across all modes, as demonstrated, for instance, by Blumenfeld-Lieberthal (2009) in the

case of railway and air transport networks. We refer to Mishra et al. (2012) for an overview of public transit connectivity in multi-modal transportation networks.

Overall, being closer or farther from the rail or road network alone does not necessarily indicate the quality of the network. In particular, a grid's population density has proven to be an essential factor to consider. At the same time, it is important to recognize other relevant factors, such as historical differences between Western and Eastern European countries.

5.2 Policy Implications

Building on the findings from the previous proximity analysis, this section presents an extended examination aimed at informing policymakers. The purpose is to identify potential best practices and highlight areas where improvements may be needed. In particular, we take a closer look at the long-standing relationship between accessibility and population density. As already illustrated in Fig. 8a, there are several notable exceptions to expected patterns. For instance, the Czech Republic (CZ) and Portugal (PT) have similar population densities (146 vs. 140 persons per km²). Still, the former has, on average, a higher proximity to rail stations than the latter (4.1 vs. 7.5 km), while their proximity to major roads is similar (9.0 vs. 9.5 km). This could indicate a mismatch in infrastructural development between the two countries.

To provide further guidance, Fig. 9 illustrates the relationship between population density (horizontal axis) and proximity to the transportation network (vertical axis). We tested several ordinary least squares (OLS) regression models (linear, exponential, and logarithmic functions) to describe the functional relationship between proximity (y) and population density (x). The best-fitting functions are shown as red lines in Figs. 9a and 9b for the road and rail networks, respectively. To visualize deviations from the estimated values, we applied a color scale: darker brown tones

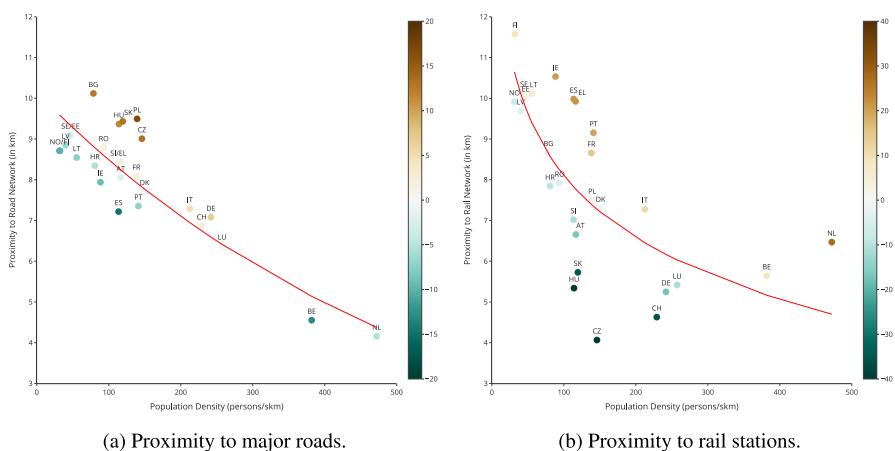


Fig. 9 Relationship between population density (in persons/km²) and proximity to the road and rail network (in km). The color scale depicts the residuals relative to the regression (in %)

indicate positive deviations (i.e., proximity is higher than predicted), while darker blue tones indicate negative deviations (i.e., proximity is lower than predicted). Light colors represent values closely aligned with the regression estimates. We note that the general negative relationship is in line with other empirical studies, as described in Section 2.1: more densely populated areas are usually more accessible. Moreover, there is a diminishing relationship between population density and proximity in both network types, which also aligns with the literature (e.g., Wu et al. 2021). Thus, if population size doubles, accessibility does not double. This, however, is more pronounced in the rail network, as becomes clear when looking at the following specific observations:

1. **For the road network**, we observe the following relationship that best fits our data: $\log_{10} y = -0.000774 \cdot x + 1.006255$. This indicates a mildly diminishing relationship. The coefficient of determination (R^2) is 81.80%, indicating the significance of the relationship. Nevertheless, it is noteworthy that the linear model also exhibits a relatively high coefficient of determination ($R^2 = 76.82\%$). Anyhow, it is now straightforward to spot countries performing “better” or “worse” than predicted by the regression. For instance, countries such as Spain (ES) and Belgium (BE) are below the regression line, indicating that, based on the European regression function, road network accessibility is better than expected. The opposite is true for several Eastern European countries, such as Bulgaria (BG) and Poland (PL), which are above the regression line and have a dark brown coloring.
2. **For the rail network**, the following relationship best fits our data: $y = -5.053962 \cdot \log_{10} x + 18.214097$. We note a relatively high R^2 value of 54.09%, which is, however, lower than the R^2 value found for the road network. Again, with the help of the regression function, it is now easy to identify those countries that perform “better” (e.g., the Czech Republic (CZ) and Hungary (HU)) or “worse” (e.g., the Netherlands (NL) and Portugal (PT)) than predicted by the regression.

Overall, while the analysis relies on simplified functional relationships, it reveals significant disparities across Europe, offering policymakers valuable insights into areas that may warrant further investigation. This is particularly true when applying the same analysis to fine-grained datasets (e.g., regional or grid-level data), for which the corresponding data structures are available in our GitHub repository. However, it is important to note that the presented analysis relies solely on European data, and it would be interesting to compare the results to functional relationships based on a global context.

5.3 Circuitry Differences and Accuracy

This section starts by analyzing differences between circuitry values, as stated in Hypotheses 4 and 5. Following this, it provides results for the average estimation error to assess the accuracy of using circuitry and straight-circle distance for distance

estimation. The impact of distance (Hypothesis 6), a key factor mentioned in the literature (see Section 2.2), is discussed in the next section.

Hypothesis 4 states that rail transportation has higher circuitry than road transportation. This was clearly visible when comparing the probability density of the two modes in Fig. 6. The results from a *t*-test confirm this by showing that the mean circuitry of rail (1.42) is greater than that of road (1.25) at the 1% significance level, as seen in Table 2. Thus, while the proximity analysis reveals a correlation between the spatial distribution of road and rail networks, it remains crucial to distinguish between the two networks when considering circuitry. Another finding from the numerical and visual analysis was that peripheral countries exhibit higher circuitry than others (Hypothesis 5). There is no distinct definition of what constitutes peripheral countries. To test the hypothesis, we use representatives from various cardinal directions in Europe and define Portugal, Estonia, Denmark, Moldova, Italy, and Greece as peripheral countries. The statistical results presented in Table 2 show higher circuitry values for both the road (1.30 vs. 1.25, Hypothesis 5.1) and rail (1.49 vs. 1.41, Hypothesis 5.2) networks at the 1% significance level.

Following these general findings, the remainder of the section discusses the accuracy of using circuitry values. In particular, we assess the accuracy by investigating the estimation error associated with estimating transportation distance by multiplying the great-circle distance by the circuitry as follows:

$$\text{Estimation Error} = \frac{|\text{Transportation Distance} - \text{Estimated Distance}|}{\text{Transportation Distance}} \quad (2)$$

with

$$\text{Estimated Distance} = \text{Great-Circle Distance} \cdot \text{Circuitry}. \quad (3)$$

This estimation error is computed for all the considered connections. For instance, the estimation error between Germany and France is calculated retrospectively using the circuitry from the 200 random connections. Figure 10 shows the probability distribution of these errors for the road and rail networks. The left part of the figure presents results at the country level, while the right part shows results at the regional level. It can be observed that the circuitry can serve as both a relatively accurate and a relatively inaccurate estimate of the actual transportation distance. For instance, using the average circuitry between two countries can result in an estimation error as low as 0.7% and 1.0% for road and rail transportation, respectively, but also as high as 26% and 30% for road and rail transportation, respectively. The median estimation error is 2.7% for road and 5.2% for rail transportation at the country level and 1.8% for road and 2.8% for rail transportation at the regional level. Thus, the estimation error decreases at the regional level, as smaller areas can better capture region-specific aspects, especially in large countries like France or Germany. Despite this, even at the regional level, it is clear that some pairs have exceptionally high estimation errors, typically resulting from geographic characteristics, such as lakes or mountains. Overall, it can be concluded that the circuitry, combined with the great-circle

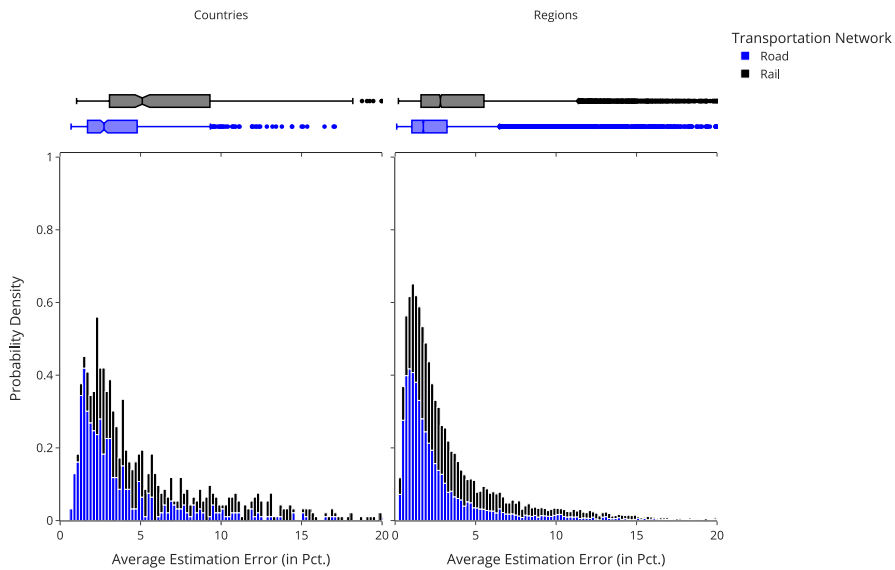


Fig. 10 Probability distribution of the estimation error (in pct.) on the country and regional level

distance, can reasonably estimate the actual transportation distance. However, the accuracy should be verified to avoid unacceptably large errors based on the use case.

5.4 Circuity and Distance

This section discusses the relationship between distance and the circuity. Based on the review in Section 2.2, previous studies have identified a (linear or non-linear) negative relationship between these two metrics. In other words, as stated in Hypothesis 6, shorter distances usually require more detours. To test this, Fig. 11 depicts the circuity alongside the corresponding transportation distance for each considered country pair. Similar results are observed when using Haversine distances. The red line in the figure represents the (best-fitted) OLS regression between transportation distance and circuity. Several insights can be drawn from the figure:

1. **Negative Relationship for Road Network:** A clear negative trend for the road network confirms the literature's finding that the circuity is generally higher for shorter distances. The coefficient of determination (R^2) of the regression model ($y = -0.1256 \cdot \log_{10} x + 1.6535$) is 21.02%, which is relatively high given the significant impact of geographical characteristics previously discussed. For instance, the three country pairs between Denmark and the Baltic states are highlighted with a blue box in the figure and are obvious outliers. Here, the exceptionally high detours are caused by the Baltic Sea, which the regression model cannot accurately capture. The negative relationship is also expressed by both correlation coefficients at the 1% significance level for the country and regional circuity datasets, see Hypothesis 6.1 in Table 3.

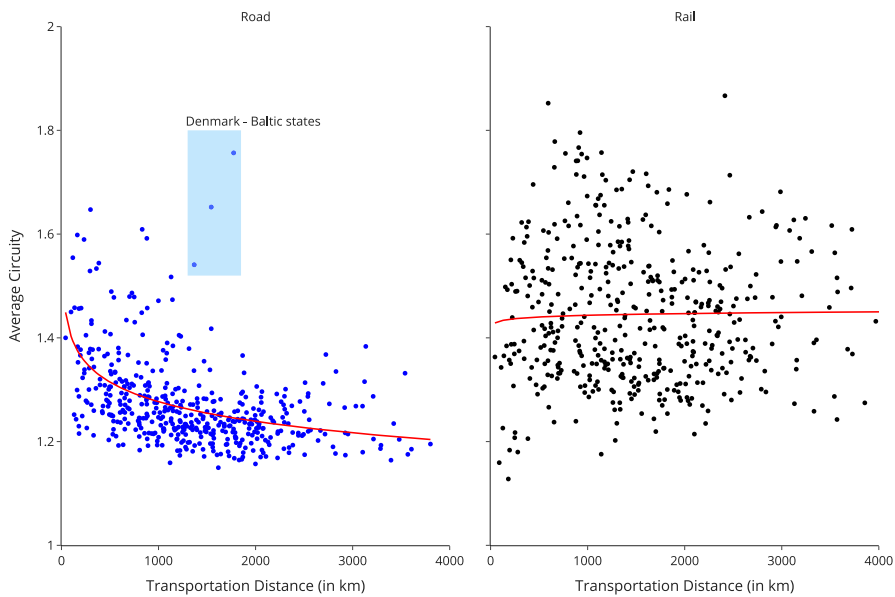


Fig. 11 Relationship between the average circuity and the transportation distance (in km) among considered country pairs

2. **Natural Lower Limit for Road Distance:** The circuity has a natural “lower limit” of about 1.15 for road distances, where even very large distances do not fall below this limit. The reasoning behind this lower limit is that, in Europe, there always seems to be a minimum amount of detour required, regardless of the actual transport relation, which Europe’s distinct geography might cause. This observation is further supported by comparing the above logarithmic regression with a linear regression, the latter of which yields a lower R^2 value of only 15.95%.
3. **No Trend for Rail Network:** There is no clear trend in the data for the rail transportation system, nor is there a converging lower limit ($y = 0.0110 \cdot \log_{10} x + 1.4105$, $R^2=0.0667\%$). This highlights distinct differences between the road and rail network, where the detour required in rail transportation is not only higher on average (Hypothesis 4) but also independent of the actual distance. This observation is also evident from the correlation coefficients for Hypothesis 6.2 as shown in Table 3. Both correlation coefficients are negative at the country level, but only the Spearman coefficient shows significance at the 10% level. While the coefficient’s statistical significance increases at the regional level, the relationship remains weak as both coefficients are relatively low, indicating neither a linear nor a non-linear correlation. One reason for this might be the generally less dense rail network compared to Europe’s relatively dense road network. In other words, there is still a gap in the rail infrastructure between what is geographically possible and what is built.

In addition to examining the impact of distance on the *value* of the circuitry, it is also interesting to analyze the impact of distance on the *accuracy* of the circuitry. Figure 12 illustrates the average estimation error (in percentage) alongside the corresponding transportation distance. The OLS regression between the estimation error and logarithmic transportation distances is shown as a red line (road: $y = -7.2121 \cdot \log_{10} x + 25.8723$, rail: $y = -7.8194 \cdot \log_{10} x + 30.9449$). A clear negative trend is visible for road and rail transportation, indicating a log-relationship where higher distances converge to estimation errors close to zero. The regression model's coefficient of determination (R^2) is 40% for the road network and 26% for the rail network, demonstrating that the road network data is more robust. This general negative relationship can be explained by the fact that large distances usually balance out detours caused by regional characteristics, such as rivers. However, there are exceptions. For instance, the country pairs between Italy and three Balkan states, highlighted with a blue box in the figure, deviate from this trend due to their very short great-circle distances across the Adriatic Sea. This further underscores our previous discussion on the impact of peripheral countries (Hypothesis 5). Table 4 presents the predicted circuitry and estimation error across Europe for varying transportation distances. The data indicates that longer distances generally yield more accurate results. However, as previously discussed, this relationship is more robust for road transportation, and significant exceptions may still occur. It is important to emphasize that the circuitry and estimation error values presented in the table are independent of each other. Specifically, a transportation distance of 3,000 km corresponds to a predicted circuitry value of 1.2168 and a predicted estimation error of

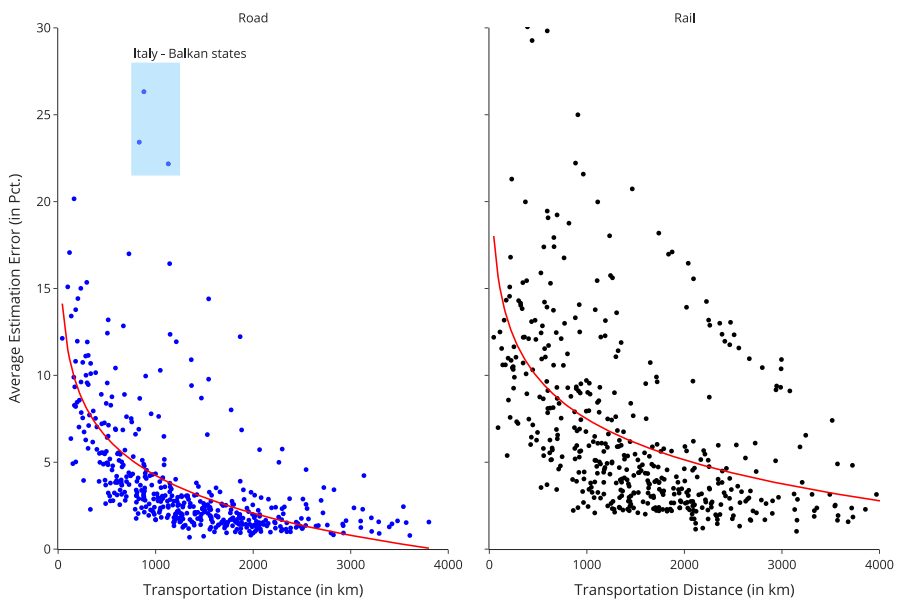


Fig. 12 Relationship between the average estimation error (in pct.) and the transportation distance (in km) among considered country pairs

Table 4 Predicted circuitry and estimation errors across Europe for varying transportation distances

Transportation Distance	Road		Rail	
	Circuitry	Estimation Error (in %)	Circuitry	Estimation Error (in %)
500	1.3145	6.4071	1.4402	9.8406
1,000	1.2767	4.236	1.4435	7.4867
1,500	1.2546	2.966	1.4454	6.1098
2,000	1.2389	2.0649	1.4468	5.1328
2,500	1.2267	1.366	1.4479	4.3750
3,000	1.2168	0.7950	1.4487	3.7559

0.7950% in the road network. However, this does not imply that a circuitry value of 1.2168 can be used for distances of 3,000 km, nor that it is associated with an estimation error of 0.7950%. Rather, using country-pair-specific circuitry values for pairs of countries separated by 3,000 km, we predict a circuitry of 1.2168 and, when using country-pair-specific circuitry values, an estimation error of 0.7950%. We also note that, as discussed above, the functional relationship between distance and average circuitry is relatively weak for the road network and virtually non-existent for the rail network (see Fig. 11).

Overall, the analysis demonstrates the importance of considering distance when applying circuitry. Generally, there are fewer detours for larger distances in the road network, while this relationship is not evident for the rail network. However, for both networks, the accuracy of the circuitry improves with increasing transportation distance.

6 Conclusion

This paper provided a comprehensive analysis of road and rail transportation distances in Europe, presenting results from two types of analyses.

First, the proximity analysis offered insights for major roads and rail stations in Europe. Significant differences were identified at the country level, varying between the road and rail networks. The relationship between proximity and population density was analyzed, revealing that developed countries with dense populations, such as Belgium and the Netherlands, generally have very low proximity to both rail and road networks. Historical differences were also noted, particularly between countries in Eastern and Southern Europe. To support this, we conducted an analysis comparing the performance of European countries against a regression-based accessibility prediction. This analysis can serve as a starting point for identifying regions currently suffering from poor connectivity, which could be the focus of future policies or political initiatives. Future research could also investigate how proximity in Europe has changed over time and how it relates to proximities on other continents and other economic metrics (e.g., Gross Domestic Product).

Second, the circuitry analysis provided insights into the detours when traveling in Europe's road and rail networks. The results showed a wide variation across the continent, with some regions, such as France and Germany, having relatively low detours, while oth-

ers, such as Italy and the Baltic states, require relatively high detours. These differences often result from geographical characteristics. The estimation error was computed for all considered country pairs, showing that the circuitry can be both a very good and a very poor metric for estimating transportation distances based on great-circle distance. Furthermore, the study provided insights into the impact of transportation distance, confirming that longer distances usually decrease the circuitry for road transport, but no such relationship was found for the rail network. However, longer distances improve the accuracy of the circuitry for both network types. Future research could explore additional factors that impact circuitry, such as measures on the infrastructural development of regions and networks, as well as more nuanced route choice behavior, recognizing that users do not always select the shortest path. Furthermore, we see a circuitry-based network, where the entire routing is based on pre-computed circuitry for specific regions and transportation modes, as an interesting field of study. A correctly calibrated circuitry-based network would be extremely fast and accurate while using only a few critical input factors.

6.1 Supplementary Information

This study is accompanied by an electronic appendix presenting the average and standard deviation of the circuitry and estimation error for each country pair and the road and rail networks. It also consists of an analysis of the high correlation between standard deviation and estimation error.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11067-026-09729-y>.

Acknowledgements Map data copyrighted OpenStreetMap contributors and available from <https://www.openstreetmap.org>. We sincerely thank the reviewers for their valuable suggestions that contributed to substantial improvements in the manuscript.

Author Contributions A.H. wrote the main manuscript text and performed the numerical computations. A.H. and C.M. contributed to the presentation of the results, reviewed the manuscript, and reviewed the Python code.

Funding Open Access funding enabled and organized by Projekt DEAL. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data Availability Data is provided within the manuscript or supplementary information files.

Declarations

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alvarez E, Franch X, Martí-Henneberg J (2013) Evolution of the Territorial Coverage of the Railway Network and its Influence on Population Growth: The Case of England and Wales, 1871–1931. *Historical Methods A Journal of Quantitative and Interdisciplinary History* 46:175–191. <https://doi.org/10.1080/001615440.2013.804788>
- Balboni C (2025) In Harm's Way? Infrastructure Investments and the Persistence of Coastal Cities. *American Economic Review* 115:77–116. <https://doi.org/10.1257/aer.20191943>
- Ballou RH, Rahardja H, Sakai N (2002) Selected country circuitry factors for road travel distance estimation. *Transportation Research Part A Policy and Practice* 36:843–848. [https://doi.org/10.1016/S0965-8564\(01\)00044-1](https://doi.org/10.1016/S0965-8564(01)00044-1)
- Barthélemy M, Flammini A (2009) Co-evolution of Density and Topology in a Simple Model of City Formation. *Netw Spat Econ* 9:401–425. <https://doi.org/10.1007/s11067-008-9068-5>
- Bliss RL, Katz JN, Wright EA, Losina E (2012) Estimating Proximity to Care: Are Straight Line and Zip-code Centroid Distances Acceptable Proxy Measures? *Med Care* 50:99–106
- Blumenfeld-Lieberthal E (2009) The Topology of Transportation Networks: A Comparison Between Different Economies. *Netw Spat Econ* 9:427–458. <https://doi.org/10.1007/s11067-008-9067-6>
- Boscoe FP, Henry KA, Zdeb MS (2012) A Nationwide Comparison of Driving Distance Versus Straight-Line Distance to Hospitals. *Prof Geogr* 64:188–196. <https://doi.org/10.1080/00330124.2011.583586>
- Brooks L, Lutz B (2019) Vestiges of Transit: Urban Persistence at a Microscale. *Rev Econ Stat* 101:385–399. https://doi.org/10.1162/rest_a_00817
- Cain LP (1997) Historical perspective on infrastructure and US economic development. *Reg Sci Urban Econ* 27:117–138. [https://doi.org/10.1016/S0166-0462\(96\)02148-5](https://doi.org/10.1016/S0166-0462(96)02148-5)
- Caragliu A (2022) Better together: Untapped potentials in Central Europe. *Pap Reg Sci* 101:1051–1086. <https://doi.org/10.1111/pirs.12690>
- Caragliu A, Del Bo CF (2025) Regional institutions and the urban digital divide. *Pap Reg Sci* 104:100118. <https://doi.org/10.1016/j.pirs.2025.100118>
- Chen X, Chen Y (2021) Quantifying the relationships between network distance and straight-line distance: Applications in spatial bias correction. *Ann GIS* 27:351–369. <https://doi.org/10.1080/19475683.2021.1966503>
- Chen X, Wang T, Thomas BW, Ulmer MW (2023) Same-day delivery with fair customer service. *Eur J Oper Res* 308:738–751. <https://doi.org/10.1016/j.ejor.2022.12.009>
- Choudhry A, Qian S (2024) Quantification of truck accessibility in urban last-mile deliveries using GPS probe data. *Transportation Research Part E Logistics and Transportation Review* 186:103536. <https://doi.org/10.1016/j.tre.2024.103536>
- Condeço-Melhorado A, Christidis P (2018) Road Accessibility in Border Regions: a Joint Approach. *Netw Spat Econ* 18:363–383. <https://doi.org/10.1007/s11067-017-9362-1>
- Demir E, Bektas T, Laporte G (2011) A comparative analysis of several vehicle emission models for road freight transportation. *Transp Res Part D: Transp Environ* 16:347–357. <https://doi.org/10.1016/j.trd.2011.01.011>
- European Union (2024) NUTS - Overview. (Date Accessed: 2025–04-07) <https://ec.europa.eu/eurostat/web/nuts/overview/>
- Flückiger M, Ludwig M (2019) Transport infrastructure, growth and persistence: The rise and demise of the Sui Canal. *Canadian Journal of Economics/Revue canadienne d'économie* 52(624–666):12378. <https://doi.org/10.1111/caje>
- Geofabrik (2024) Geofabrik: It is our ambition to help you to understand OpenStreetMap and to support you using the data. (Date Accessed: 2025–04-07) <https://www.geofabrik.de/>
- GEOSTAT (2024) GEOSTAT: Census Grid 2021. (Date Accessed: 2025–04-07) <https://ec.europa.eu/eurostat/web/gisco/geodata/population-distribution/geostat>
- Geurs KT, Boon W, Van Wee B (2009) Social Impacts of Transport: Literature Review and the State of the Practice of Transport Appraisal in the Netherlands and the United Kingdom. *Transp Rev* 29:69–90. <https://doi.org/10.1080/01441640802130490>
- Geurs KT, van Wee B (2004) Accessibility evaluation of land-use and transport strategies: Review and research directions. *J Transp Geogr* 12:127–140. <https://doi.org/10.1016/j.jtrangeo.2003.10.005>
- Giacomin DJ, Levinson DM (2015) Road network circuitry in metropolitan areas. *Environ Plann B Plann Des* 42:1040–1053. <https://doi.org/10.1068/b130131p>

- Gibbons S, Wu W (2020) Airports, access and local economic performance: evidence from China. *Journal of Economic Geography* 20:903–937. <https://doi.org/10.1093/jeg/lbz021>
- Guagliardo MF (2004) Spatial accessibility of primary care: concepts, methods and challenges. *Int J Health Geogr* 3:3. <https://doi.org/10.1186/1476-072X-3-3>
- Gutiérrez J, Urbano P (1996) Accessibility in the European Union: The impact of the trans-European road network. *J Transp Geogr* 4:15–25. [https://doi.org/10.1016/0966-6923\(95\)00042-9](https://doi.org/10.1016/0966-6923(95)00042-9)
- Hansen WG (1959) How Accessibility Shapes Land Use. *J Am Inst Plann* 25:73–76. <https://doi.org/10.1080/01944365908978307>
- Heinold A (2020) Comparing emission estimation models for rail freight transportation. *Transp Res Part D: Transp Environ* 86:102468. <https://doi.org/10.1016/J.TRD.2020.102468>
- Huang J, Levinson DM (2015) Circuitry in urban transit networks. *J Transp Geogr* 48:145–153. <https://doi.org/10.1016/j.jtrangeo.2015.09.004>
- Huang Y, Jin X, Fan M, Yang X, Jiang F (2024) Personalized Route Recommendation Based on User Habits for Vehicle Navigation. <https://doi.org/10.48550/arXiv.2409.14047>
- Jang S, Park H (2025) Spatial equity in public rental housing based on commuting accessibility. *Regional Science Policy & Practice* 17:100196. <https://doi.org/10.1016/j.rsp.2025.100196>
- Jayasinghe A, Sano K, Kasemsri R, Nishiuchi H (2016) Travelers' Route Choice: Comparing Relative Importance of Metric, Topological and Geometric Distance. *Procedia Engineering* 142:18–25. <https://doi.org/10.1016/j.proeng.2016.02.008>
- Jorge D, Rocha T, Ramos TRP (2024) A time-driven simulation-optimization framework for the dynamic heterogeneous order-courier assignment problem for instant deliveries. *Transportation Research Part E Logistics and Transportation Review* 192:103783. <https://doi.org/10.1016/j.tre.2024.103783>
- Keiser DA, Shapiro JS (2019) Consequences of the Clean Water Act and the Demand for Water Quality. *Q J Econ* 134:349–396. <https://doi.org/10.1093/qje/qjy019>
- Kim J, Park J, Lee J, Jang KM (2024) Examining the socio-spatial patterns of bus shelters with deep learning analysis of street-view images: A case study of 20 cities in the U.S. *Cities* 148:104852. <https://doi.org/10.1016/j.cities.2024.104852>
- Kotavaara O, Antikainen H, Rusanen J (2011) Population change and accessibility by road and rail networks: GIS and statistical approach to Finland 1970–2007. *J Transp Geogr* 19:926–935. <https://doi.org/10.1016/j.jtrangeo.2010.10.013>
- Kwan M (1998) Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. *Geogr Anal* 30:191–216. <https://doi.org/10.1111/j.1538-4632.1998.tb00396.x>
- Leuthold FU, Weigt H, von Hirschhausen C (2012) A Large-Scale Spatial Optimization Model of the European Electricity Market. *Netw Spat Econ* 12:75–107. <https://doi.org/10.1007/s11067-010-9148-1>
- Levinson D (2007) Density and dispersion: the co-development of land use and rail in London. *Journal of Economic Geography* 8:55–77. <https://doi.org/10.1093/jeg/lbm038>
- Lopresti I, Tartaglia M (2024) The Relationship Between High-Speed Rail Accessibility and Tourism Demand: The Case Study of Italy. In: Pagliara F (ed), *Socioeconomic Impacts of High-Speed Rail Systems*, pp 473–491. https://doi.org/10.1007/978-3-031-53684-7_22
- Luxen D, Vetter C (2011) Real-time routing with OpenStreetMap data. In: *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*, ACM, New York, NY, USA, pp 513–516. <https://doi.org/10.1145/2093973.2094062>
- Makowski C, Saragih A, Guter W, Russell T, Heinold A, Lekkakos S (2025) SCGraph: A Dependency-Free Python Package for Road, Rail and Maritime Shortest Path Routing Generation and Distance Estimation. <https://doi.org/10.2139/ssrn.5388845>
- Mennicken E, Lemoy R, Caruso G (2024) Road network distances and detours in Europe: Radial profiles and city size effects. *Environment and Planning B Urban Analytics and City Science* 51:174–194. <https://doi.org/10.1177/23998083231168870>
- Merchán D, Winkenbach M, Snoeck A (2020) Quantifying the impact of urban road networks on the efficiency of local trips. *Transportation Research Part A Policy and Practice* 135:38–62. <https://doi.org/10.1016/j.tra.2020.02.015>
- Mercier A, Souche-Le Corvec S, Ovtracht N (2021) Measure of accessibility to postal services in France: A potential spatial accessibility approach applied in an urban region. *Pap Reg Sci* 100:227–250. <https://doi.org/10.1111/pirs.12564>
- Mert Cubukcu K (2020) Using circuitry as a network efficiency measure: the example of Paris. *Spat Inf Res* 29:163–172. <https://doi.org/10.1007/s41324-020-00342-w>

- Miller EJ (2018) Accessibility: measurement and application in transportation planning. *Transp Rev* 38:551–555. <https://doi.org/10.1080/01441647.2018.1492778>
- Mishra S, Welch TF, Jha MK (2012) Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A Policy and Practice* 46:1066–1085. <https://doi.org/10.1016/j.tra.2012.04.006>
- Morita M, Suzuki K, Ki O (2014) An Empirical Study of the Ratio of the Road Distance to the Straight Line Distance in Major Cities in Japan. *Theory and Applications of GIS* 22:1–7. <https://doi.org/10.5638/thagis.22.1>
- Moynihan K, Franc, a UL, Casavant DW, Graham RJ, Mcmanus ML (2023) Hospital Access Patterns of Children With Technology Dependence. *Pediatrics* 151. <https://doi.org/10.1542/peds.2022-059014>
- Ninivaggi F, Cutrini E (2025) Resilience, Proximity and the 15-Minute City: A Case Study for Central Italy. *Netw Spat Econ*. <https://doi.org/10.1007/s11067-025-09669-z>
- OSM (2024) OpenStreetMap is built by a community of mappers that contribute and maintain data about roads, trails, caf  s, railway stations, and much more, all over the world. (Date Accessed: 2025–04–07) <https://www.openstreetmap.org/about/eng>
- Otsuka N, Delmastro T, Wittowsky D, Pensa S, Damerau M (2019) Assessing the accessibility of urban nodes: the case of TEN-T railway stations in Europe. *Applied Mobilities* 4:219–243. <https://doi.org/10.1080/23800127.2019.15737>
- Phibbs CS, Luft HS (1995) Correlation of Travel Time on Roads versus Straight Line Distance. *Med Care Res Rev* 52:532–542. <https://doi.org/10.1177/107755879505200406>
- Pourramazani H, Miralles-Garcia JL (2023) Evaluating Urban Transportation Accessibility: A Systematic Review of Access Dimensions and Indicators. *International Journal of Transport Development and Integration* 7:331–339. <https://doi.org/10.18280/ijtdi.070407>
- Prato CG (2009) Route choice modeling: Past, present and future research directions. *Journal of Choice Modelling* 2:65–100. [https://doi.org/10.1016/S1755-5345\(13\)70005-8](https://doi.org/10.1016/S1755-5345(13)70005-8)
- Quinn P (2013) Road density as a proxy for population density in regional-scale risk modeling. *Nat Hazards* 65:1227–1248. <https://doi.org/10.1007/s11069-012-0379-y>
- Reggiani G, van Oijen T, Hamedmoghadam H, Daamen W, Vu HL, Hoogendoorn S (2022) Understanding bikeability: A methodology to assess urban networks. *Transportation* 49:897–925. <https://doi.org/10.1007/s11116-021-10198-0>
- Robusto CC (1957) The Cosine-Haversine Formula. *Am Math Mon* 64:38. <https://doi.org/10.2307/2309088>
- Rodrigue JP (2024) *The Geography of Transport Systems*. 5 ed., Routledge, London. <https://doi.org/10.4324/9781003343196>
- Salonen M, Toivonen T (2013) Modelling travel time in urban networks: Comparable measures for private car and public transport. *J Transp Geogr* 31(143–153):011. <https://doi.org/10.1016/j.jtrangeo.2013.06>
- Siddiq F, D Taylor B (2021) Tools of the Trade? Assessing the Progress of Accessibility Measures for Planning Practice. *J Am Plann Assoc* 87:497–511. <https://doi.org/10.1080/01944363.2021.1899036>
- Stroppe AK (2023) Left behind in a public services wasteland? On the accessibility of public services and political trust. *Polit Geogr* 105:962–6298. <https://doi.org/10.1016/j.polgeo.2023.102905>
- Sun X, Wandelt S, Hansen M (2020) Airport Road Access at Planet Scale using Population Grid and Openstreetmap. *Netw Spat Econ* 20:273–299. <https://doi.org/10.1007/s11067-019-09480-7>
- Tailor TD, Choudhury KR, Tong BC, Christensen JD, Sosa JA, Rubin GD (2019) Geographic access to CT for lung cancer screening: a census tract-level analysis of cigarette smoking in the United States and driving distance to a CT facility. *J Am Coll Radiol* 16:15–23. <https://doi.org/10.1016/j.jacr.2018.07.007>
- Tartaglia M, Fiduccia A, Cisco G (2025) Wider effects of transportation planning: Focusing on socioeconomic dynamics and issues. *Regional Science Policy & Practice* 17:100186. <https://doi.org/10.1016/j.r spp.2025.100186>
- Ulmer MW, Thomas BW (2018) Same-day delivery with heterogeneous fleets of drones and vehicles. *Networks* 72:475–505. <https://doi.org/10.1002/NET.21855>
- Weisbrod G (2008) Models to predict the economic development impact of transportation projects: historical experience and new applications. *Ann Reg Sci* 42:519–543. <https://doi.org/10.1007/s00168-007-0184-9>
- Wiegman B, Champagne-Gelinas A, Duchesne S, Slack B, Witte P (2018) Rail and road freight transport network efficiency of Canada, member states of the EU, and the USA. *Research in Transportation Business & Management* 28:54–65. <https://doi.org/10.1016/j.rtbm.2018.10.004>

- Wu H, Avner P, Boisjoly G, Braga CKV, El-Geneidy A, Huang J, Kerzhner T, Murphy B, Niedzielski MA, Pereira RHM, Pritchard JP, Stewart A, Wang J, Levinson D (2021) Urban access across the globe: An international comparison of different transport modes. *npj Urban Sustainability* 1:16. <https://doi.org/10.1038/s42949-021-00020-2>
- Yang H, Ke J, Ye J (2018) A universal distribution law of network detour ratios. *Transportation Research Part C Emerging Technologies* 96:22–37. <https://doi.org/10.1016/j.trc.2018.09.012>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Arne Heinold¹ · Connor Makowski²

✉ Arne Heinold
arne.heinold@klu.org

Connor Makowski
conmak@mit.edu

¹ Department of Operations and Technology, Kühne Logistics University (KLU), Hamburg, Germany

² Center for Transportation and Logistics, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA