

Network science reveals the structure and lasting impact of the Roman road system

Received: 15 November 2025

Accepted: 19 June 2026

Published online: 15 September 2026



Adam Pažout ^{1,2,3,4}, Tom Brughmans ^{1,2,3} ✉, Pau de Soto ⁴,
Matteo Mazzamurro ^{1,2,3,5}, Maria Coto-Sarmiento ^{1,2,3,6}, Clara Filet ⁷,
Eduardo Herrera Malatesta ⁸, Magnus L. Nielsen^{1,3}, Gustav Emil Ølgaard^{1,3} &
Peter B. Vahlstrup ⁹

Ancient infrastructure can leave lasting imprints on modern settlement patterns and mobility. The Roman road network connected vast regions of Europe, North Africa, and the Middle East, and is often mentioned as an example of such long-term impact. Yet its structural characteristics remain largely unquantified. Here we present a high spatial resolution quantitative exploratory analysis of the Roman Empire's road system, enabled by our creation of a spatially detailed Empire-wide dataset. We address long-standing questions concerning the straightness and persistence of these roads, and the centrality of cities like Rome. We show that Roman roads tend to be straight where topography allows but less so than modern roads; provincial capitals were more centrally located than other cities for mediating terrestrial mobility; segments hypothesised to be main roads were more important intermediaries than secondary roads; and Roman road density correlates with major modern roads.

The Roman Empire was established in the late first century BCE, and existed as an integrated political unit until the fifth century CE in the west and in the east until the 15th century CE. At its height in 117 CE it covered an area of almost 5,500,000 km² around the Mediterranean Sea, occupying much of Europe, the Middle East and North Africa. These regions were united for the first time into a single geopolitical unit which enabled mass mobility on a continental scale via maritime, riverine and terrestrial means. The Romans used and extended pre-existing roads. Terrestrial routes of some sort already connected all landlocked places where humans lived and there was a significant degree of continuity of their occupation, and vast areas were already densely urbanised with well-established road infrastructure by the time the Romans incorporated them. Importantly, however, the nature and degree of formalisation of these pre-Roman terrestrial routes

varied markedly across the Empire's later extent. In the eastern Mediterranean, where complex political organisation and urban development long predated Roman expansion, Roman authorities often encountered established long-distance networks of engineered roads linking cities and major economic and administrative centres¹. In contrast, in many parts of the western Mediterranean, pre-Roman routes appear to have consisted largely of tracks and pathways with limited built long-distance infrastructure. Consequently, the role of Roman surveyors and engineers in creating terrestrial routes varied across the Empire, depending on those regional conditions, the density and organisation of pre-Roman communities, and the politico-military context. In some areas their work primarily consisted of upgrading and integrating existing corridors, whereas in others it required the planning and construction of new roads to connect

¹Department of History and Classical Studies, Aarhus University, Aarhus C, Denmark. ²Social Resilience Lab, Aarhus University, Aarhus C, Denmark. ³Centre for Urban Network Evolutions (UrbNet), Aarhus University, Højbjerg, Denmark. ⁴Grup de Recerca en Arqueologia Clàssica (GRAC_UAB), Universitat Autònoma de Barcelona, Barcelona, Spain. ⁵School of Physical and Chemical Sciences, Queen Mary University of London, London, UK. ⁶Department of Archaeology and Anthropology, HUMANE, Human Ecology and Archaeology Research Group, Milà i Fontanals Institution for Research and Humanities, IMF-CSIC, Barcelona, Spain. ⁷Institut Ausonius UMR 5607, University of Bordeaux-Montaigne, Pessac, France. ⁸Faculty of Archaeology, Leiden University, Leiden, The Netherlands. ⁹Center for Humanities Computing, Aarhus University, Aarhus C, Denmark. ✉e-mail: t.b@cas.au.dk

administrative centres, military sites, and newly incorporated territories.

Significant investments were made into terrestrial road infrastructure during the Roman period, in part to support its military conquest and pacification of new regions between the 3rd century BCE and 3rd century CE, but equally to move goods and people between countryside, cities and ports. The gromatici and engineers who designed the road layouts normally chose those routes that were favourable, avoiding steep slopes and promoting roads with conditions enabling wheeled transport^{2,3}. Pliny the Elder argued that Roman engineering distinguished itself from that of earlier civilisations by its practical utility, unlike, for example, ‘the pyramids, so many idle and frivolous pieces of ostentation of their resources’ (*Plin. NH* 36.75). Those engineering qualities were also identified by Vitruvius (*Vitr. De arch.* 1.3.2), ‘durability, convenience, and beauty’ became so embedded in the imperial landscape that parts of this infrastructure appear to have persisted in place for millennia.

Roman roads represent a medium that structured the flow of people, animals, disease, goods and ideas^{4,5}. A detailed understanding of their structure is key to studies of major phenomena in this continental scale region, including the spread and impact of past epidemics, trade in foodstuffs and food security, the spread of technological innovations and beliefs, mass migration, and the millennia-long development of terrestrial transport. But the overall structuring principles of Roman roads are the subject of ongoing debate: their relationship to pre-Roman roads⁶, the preference for optimal paths through the landscape⁷, the degree to which a Roman interventionist government planned an efficient empire-wide road network from a top-down perspective^{8,9}, and the role of local and provincial needs in road development¹⁰.

A precise quantitative description of the structure of Roman roads across the entire empire remains absent, despite Roman roads having been the subject of scholarly debate for over a century^{11,12}. Our current understanding of its structure is largely based on qualitative studies of the available evidence^{6,9,13}, which includes ancient itineraries and physical remains including preserved roads, milestones and bridges. There are several regional studies that have analyzed these engineering tasks applied to mountainous or irregular regions^{14,15}. Previous quantitative studies tend to focus on a single structuring feature such as topography or network structure, or were limited to particular provinces or regions, or used an incomplete empire-wide dataset of low spatial resolution^{16–28}. Previous studies have emphasised a number of long-held assumptions^{25,27,29–32} about Roman roads, including the high density of roads during Roman occupation, a tendency towards straight and paved roads, the centrality of Rome in the road network, and a high degree of correlation between present-day and Roman roads. Yet a quantitative and empirically-informed evaluation of these assumptions at an Empire-wide scale has never been performed. We perform a detailed descriptive approach to the roads’ topographic and network structure, which is complementary to previous work in economic history and historical economics that focused on identifying causal inference. This descriptive approach is crucial since we do not currently know how human interaction potential was shaped by road infrastructure during the Roman period, which in previous qualitative work is often inferred by abstract knowledge of ancient connections between places rather than by concrete spatially explicit paths crossing real-world topography.

Here we present a spatially detailed quantitative exploratory study of the structure of roads of the entire Roman Empire, made possible thanks to our work in creating the most comprehensive and detailed digital model of 299,171.31 km of roads of the Roman Empire³³. We focus in turn on evaluating long-standing questions: (1) Were Roman roads predominantly straight?^{9,29} (2) What was their continuity up to the present day?^{17,18,30} (3) How central were major cities and Rome itself in the network?

Results

The straightness and topographic structure of roads

We analysed four properties that describe basic characteristics of Roman roads in their local topography: average and maximum slope, mean TPI (Topographic Position Index), orientation, and sinuosity (Fig. 1 and Supplementary Datasets 1–2). TPI is a terrain attribute defined as the difference between the elevation of a cell and the average elevation of cells in a given neighbourhood³⁴. Sinuosity can be used as a measure of ‘straightness’ of the road, calculated as a ratio of actual road length to straight line distance between start and end point of 1 km long road segments. These metrics together provide us with valuable insight into the process of planning and construction of Roman roads, concerning constraints of ancient technology and transportation. Results were derived for 1 km stretches of roads to enable comparability. These results do not change much when taking road certainty categories or road types into account (see ‘Topographic structure by certainty’ and Supplementary Fig. 1 for details).

How straight were Roman roads? Sinuosity is very low in general (mean 1.048; median 1.0045) with most low sinuosity roads located in lowlands (mainly in Western Europe, the Pannonian basin, North Africa, the Po valley, Puglia, central Anatolia and Syria), where their course is not heavily dictated by the topography and straight-line connections can be more easily maintained (Fig. 1c). High sinuosity roads are found in mountainous areas and they predominate in Southeastern Europe. However, these results also reveal the variability in data collection efforts and resolution for different parts of the empire, as represented in our representativity map³³.

The straightness of Roman roads is in part a function of topography. In flat terrain connections between two points will tend to be straight, radially extending from the settlements, as can be best seen in Northwest Europe. We observe that sinuosity is slightly positively correlated with maximum slope (Fig. 1f), i.e. growing maximum slope correlates with higher sinuosity of the roads. On the other hand, sinuosity is slightly negatively correlated with TPI, but their relationship is more complex. To a certain point an increasing TPI is correlated with growing sinuosity, but very high TPI values are rather correlated with lower sinuosity. The explanation might be that high TPI values are more associated with ridges which can support more straight roads than lower TPI values which rather indicate more complicated topography (upper/middle slopes etc.). In addition to topography, the Romans brought technological improvements which allowed planning and building (straight roads) more effectively, such as surveying instruments^{29,31} and construction methods³. Nevertheless, major modern roads (Fig. 1d) tend to be straighter than ancient roads, with lower mean and median sinuosity values (1.024 and 1.0026 respectively). The simple explanation is that modern road building avoids sharp turns and aims to build as straight roads as possible by strongly modifying the terrain to facilitate faster travel by car, which is made possible by modern surveying and construction methods.

The slope value (Fig. 1a) tells us about potential limits of each road segment regarding a suitable mode of transport, since animal-drawn carts and wagons are limited to gentle slopes. The upper slope limit for using vehicular traffic is thought to be anywhere between 5 and 16%, or between 2.86° and 9.09°¹⁴. The vast majority of roads (270,955.85 km of roads or 90.57%) are within the limits of possible vehicular traffic below 16% rise (maximum slope less than 9.09°). High slope values are located in the mountainous areas. Interestingly, the Alpine roads show lower slope values (the mean of average slope values is 2.68°; mean of maximum values is 6.79°) compared to roads in other mountain regions e.g., Taurus (3.33° and 8.11°), Pyrenees (3.14° and 7.85°), Cantabrian Mountains (4.98° and 11.26°), Atlas (2.92° and 7.06°), and Greece (2.86° and 7.11°). Most values fall between 0° and 10° (0–17.63% slope). The global median slope of the whole dataset is 3.88% (2.22°). This pattern is likely explained by two complementary factors. First, although the Alps constitute the highest mountain system in Europe,

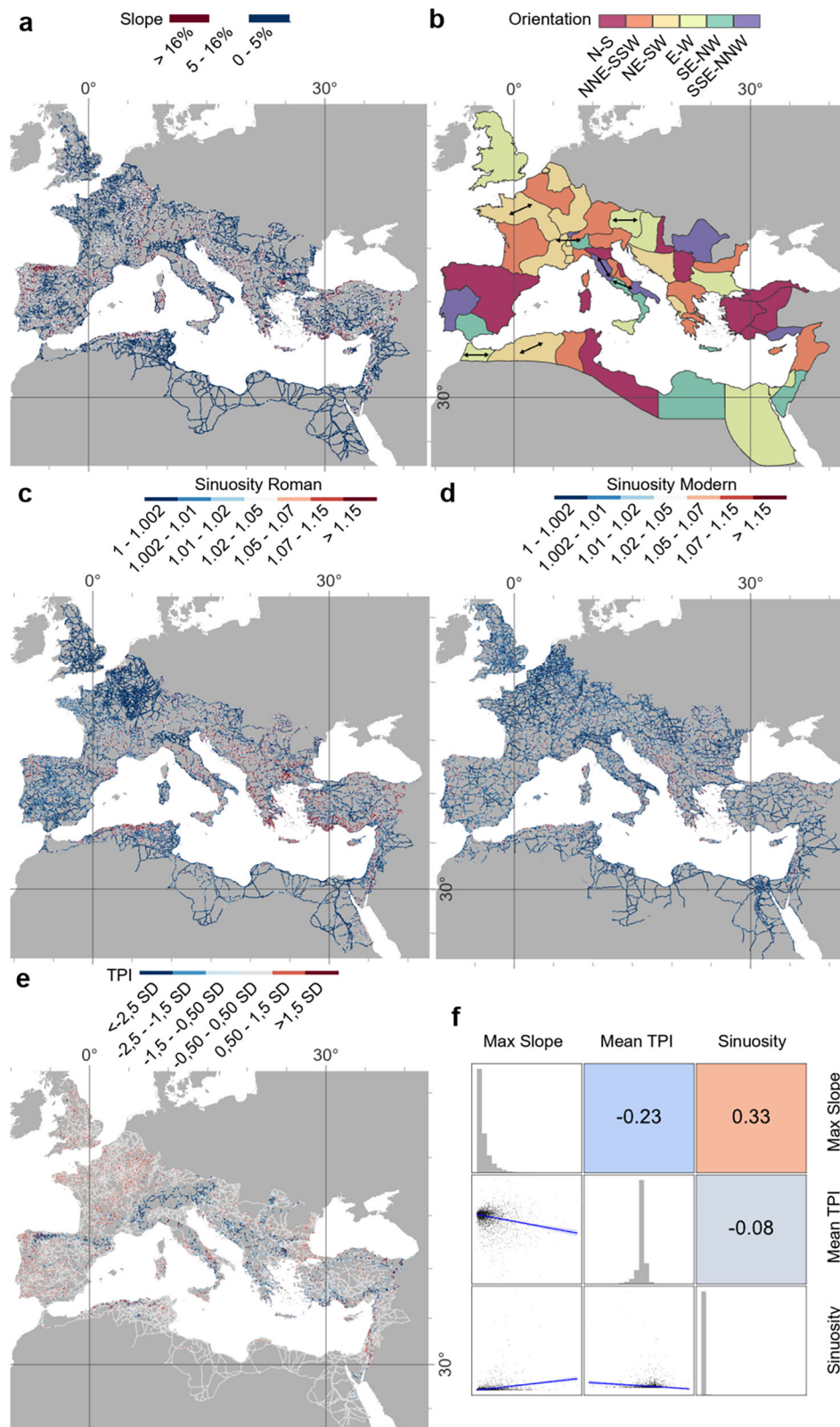


Fig. 1 | Topographic variables. **a** Maximum slope. **b** Dominant orientation of roads by province. Arrows indicate provinces where the dominant orientation of roads follows the orientation of mountains. **c** Roman Road sinuosity. **d** Sinuosity of modern roads. **e** Topographic position index (TPI) of the Roman roads, negative

values indicate valleys, high positive values indicate ridges. **f** Matrix summarising the relationship between the topographic variables, lower left: scatterplots with linear trend, diagonal: histograms, upper right: Pearson's ρ values. All variables were measured for 1000 m long road segments.

they also include a set of major longitudinal valleys and relatively favourable passes that can be exploited to maintain gentler gradients along long-distance corridors, compared to other mountain ranges where routes may be forced into steeper, shorter ascents and descents. Second, and probably more importantly, the comparatively low slopes of Alpine roads reflect a strong Roman engineering and strategic commitment to ensuring reliable overland connectivity from Italy to the rest of Europe. In this context, Roman surveyors and builders did not merely accept the constraints of the terrain, but actively intervened to reshape it where necessary through large-scale cuttings, terracing, retaining structures, and, in exceptional cases, tunnels, as exemplified by the rock-cut sections associated with the road system of the Aosta Valley³⁵.

Most of the road segments fall into low TPI values (mean -17.8 , median -4.98), which suggests that the preferred locations for the roads were flat areas or valleys, whereas fewer roads are located on slopes or prominent ridges (Fig. 1e). This pattern might be partly caused by higher representativity of the dataset for the low-lying regions, especially north-west Europe (see representativity map in de Soto et al.³³). What stands out is the prevalence of high negative values indicating roads in valleys, especially in mountainous areas, south-eastern Europe and Asia Minor. On the other hand in most of north-western Europe, the Iberian peninsula, and North Africa the majority of roads are located in moderate topography, in level areas, on gentle slopes or small elevations.

In mountainous areas the road orientation is strongly determined by topography, it follows the orientation of mountain ranges and valleys, and cuts through mountain passes. This is most striking in a few provinces where the overall road orientation follows the overall orientation of mountain ranges (Fig. 1b): Alpes Graiae, Etruria (Regio VII), Latium et Campania (Regio I), Lugdunensis, Mauretania Caesariensis, Mauretania Tingitana, Noricum. However, in other provinces roads are overall perpendicular to mountain ranges: Africa Proconsularis, Asia, Bithynia et Pontus, Galatia, Hispania Citerior, Iudaea, Moesia Superior, Umbria (Regio VI). In non-mountainous regions other factors determined the orientation, including the settlement structure and river valleys.

Roman and major modern road density correlate

What is the relationship between Roman roads and major modern roads? The difference in kernel density between modern (Fig. 2b) and Roman roads highlights areas exhibiting significant changes in the road density between the Roman and modern period (Fig. 2c). We observe a conspicuous increase in road density in the core urban and industrial areas of Western Europe (northern Italy, Rhine-Meuse basin, Britain) and highly concentrated change around large urban centres (e.g., Paris, Marseille, Madrid). Areas with a lower density of modern roads are located mainly in Greece, Asia Minor and Northern Africa. The patterns observed in other regions are results of more varied processes. The differences in road density in the Iberian Peninsula can be related to Modern era centralisation tendencies focused on Madrid²². The case of Egypt shows increased urbanisation and rising importance of vehicular transport in the Nile Valley compared to the past when the country was reliant mainly on riverine transport³⁶. However, the lower representativity of the dataset for certain regions (central Europe, western Balkans, Tuscany, Marche, central Anatolia, see de Soto et al.) can skew the results so that the difference between Roman and modern road density will be larger.

We subsequently explored the relationship between the location of Roman and major modern roads, and several key topographic and anthropic variables (elevation, TPI, density of Roman sites, and modern population, see 'Linear regression'), using a 0.5° cell grid as the analytical unit (using the provinces as analytical units provides comparable results, see Supplementary Fig. 3). The results (Fig. 3d, Supplementary Table 3) indicate a positive correlation between locations

of Roman and modern roads (Pearson's ρ 0.46). Modern roads are positively correlated with Roman site density (0.2 for 0.5° cells and 0.37 for provinces), but less strongly than Roman roads are (0.25 for 0.5° cells and 0.45 for provinces). However, correlation with the modern population is lower (Pearson's ρ 0.17). The topographic variables show very low negative correlation to Roman roads (Pearson's ρ -0.11 for elevation and -0.02 for TPI) which suggests that topography has a slightly negative impact on the location of Roman roads: slightly fewer roads with growing slope or TPI. Modern roads are slightly more negatively correlated with elevation and TPI than Roman roads are (Pearson's ρ -0.28 and -0.03).

We further explored the spatial relationship of these variables using Geographically Weighted Regression (see 'Spatial regression') in three scenarios: (1) whether the relationship between locations of modern roads and locations of Roman roads is consistent throughout the region, (2) whether the relationship between the density of the modern population and locations of Roman roads is consistent throughout the region, and (3) what is the explanatory power of selected environmental and anthropic variables (density of Roman settlements) on the location of the Roman roads. The results (Fig. 3a–c; Supplementary Table 4) show considerable spatial heterogeneity in scenario 1 with contrasting over- and under-predictions which are not organised into any large contiguous regions. The model explains a large degree of the variance (adjusted $R^2 = 0.61$), but the results suggest that the relationship between Roman and modern roads is mediated by local (micro- to meso-level) circumstances. The model in scenario 2 does not explain much variance (adjusted $R^2 = 0.25$). The relationship between modern population and Roman roads is weak, but spatially consistent. The positive and negative residuals are concentrated in the main population and urban centres across the whole research area. In this case, our evidence is likely biased towards modern population centres where archaeological investigations are concentrated due to modern development, and recovery rates of archaeological remains are larger there than in the countryside. Model 3, the combination of Roman site density, elevation and TPI values, explains only a small part of the variance (adjusted $R^2 = 0.35$). We again observe a highly locally varied and spatially heterogeneous pattern, including systemic overprediction in North Africa, the southern Levant, southern and central Italy, southern Greece, Bulgaria and the Danube frontier, central and northern France, Britain, and central Spain. This suggests that the included variables capture only a limited and spatially uneven share of the processes shaping the location of Roman roads which are mediated at a micro- to meso-level.

The robustness of these results was explored by changing cell size, the neighbourhood size, and comparing with data confidence (see 'Robustness of spatial regression'). Comparable results are obtained when using smaller 0.25° cells (Supplementary Fig. 5; Supplementary Table 6), although the resulting models for scenarios 1–2 perform worse. We equally observed worse performance when using a 1000 km neighbourhood for models of scenarios 1–2 (with different substantive results for scenario 1) and using an optimal number of neighbours for scenario 3 (Supplementary Fig. 5; Supplementary Table 5). Comparison with data reliability and representativity reveals results to be more robust in north-western Europe, Spain, Italy, Greece, the eastern Balkans, most of Asia Minor, the Near East, the Eastern Desert of Egypt, coastal Cyrenaica and Tripolitania, and North Africa. Finally, comparable results are obtained for scenario 1 when excluding cells with low representativity or reliability, although the model explains slightly less variance (adjusted R^2 0.53 as compared to 0.61; Supplementary Fig. 6 and Supplementary Table 7).

There are significant differences in the density of Roman roads across the Empire. The dataset contains 299,171.31 km of roads in a total study area of 5,271,920.243 km², giving a global road density in the Roman Empire of 56.748 m/km² (Fig. 2a). High road density is observed in most of the densely urbanised regions of the Roman

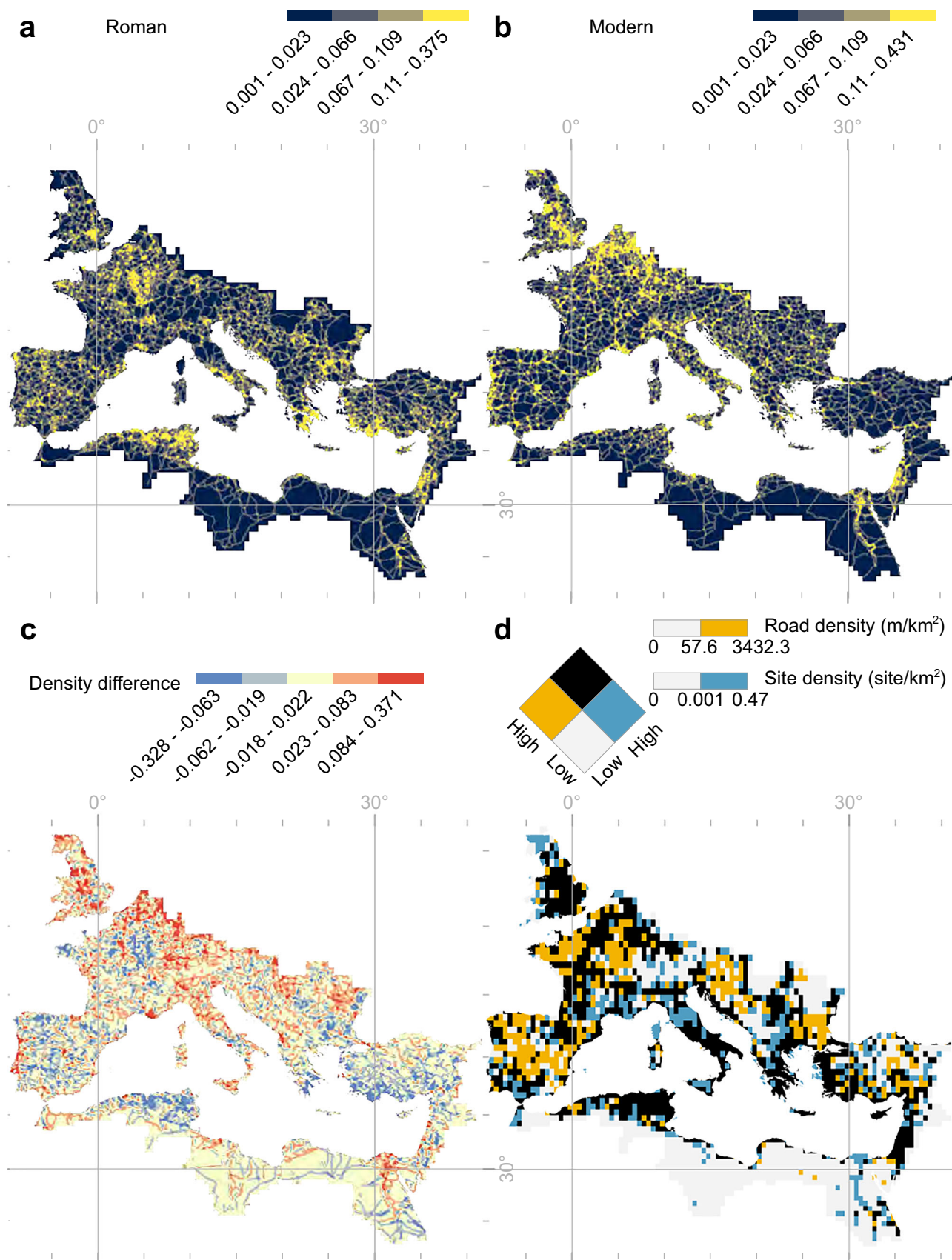


Fig. 2 | Density of roads. a kernel density of Roman roads. **b** kernel density of modern roads. **c** Difference of kernel density between Roman and modern roads, negative values indicate a decrease in road density from the Roman period,

positive values indicate an increase in road density since the Roman period. **d** the relationship between the density of Roman roads and the density of Roman sites.

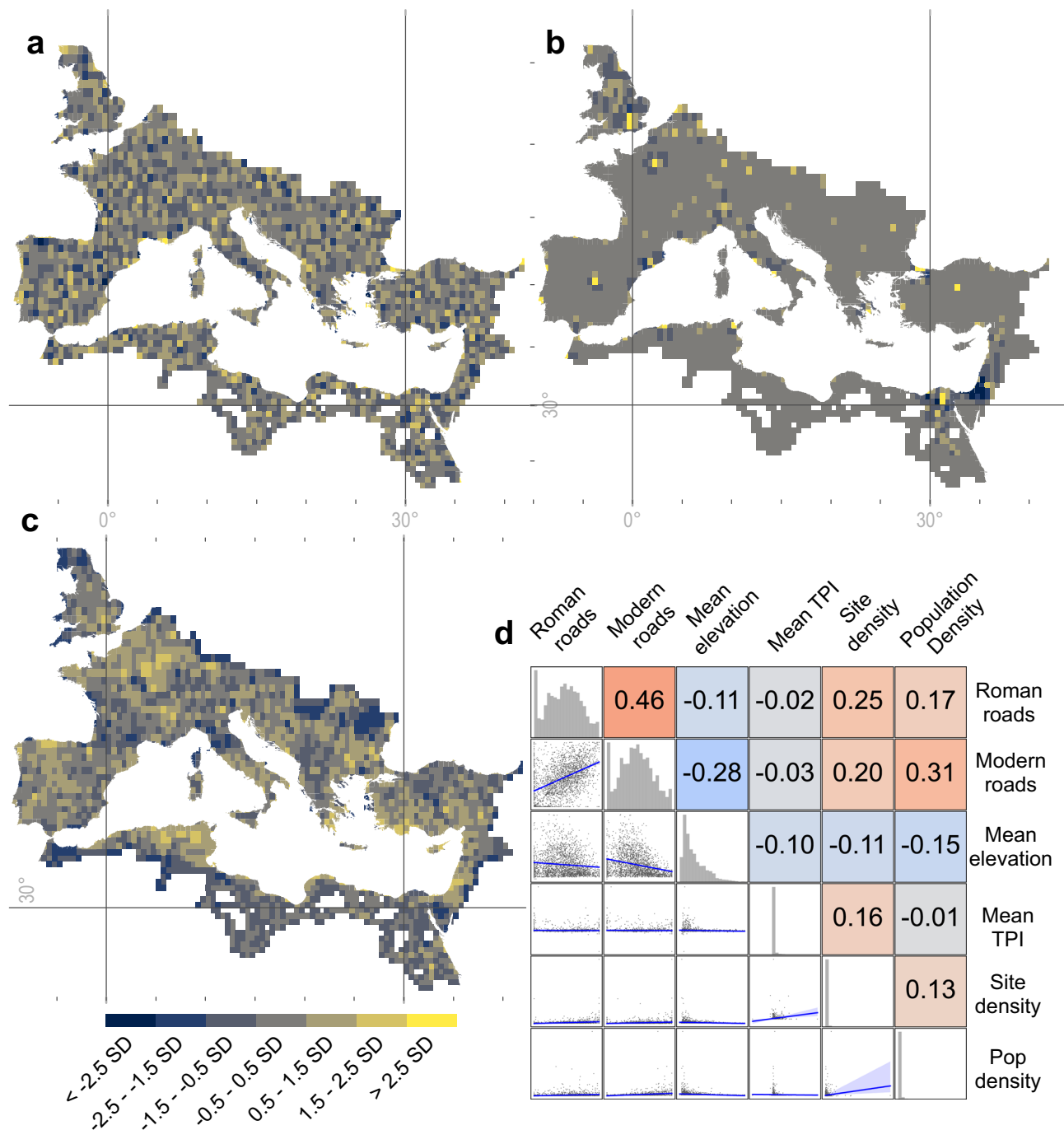


Fig. 3 | Geographically weighted regression analysis, showing standardised residuals of the model. a scenario 1: relationship between Roman and modern roads. **b** scenario 2: relationship between Roman roads and modern population. **c** scenario 3: relationship between Roman roads and topographic and anthropic variables (mean elevation, mean Topographic Position Index, site density). **d** linear

regression of all dependent and explanatory variables (continuous ordinary least squares model), lower left: scatterplots showing linear trend and 95% confidence intervals of the fitted regression line, diagonal: histograms, upper right: Pearson's p values. Asterisk shows a statistically significant relationship ($p < 0.05$): Modern roads $p < 0.0001$, Site density $p = 0.029$.

Empire. These include areas with a long history of urbanism and intensive trade activity (Italy, Greece, Asia Minor, Tunisia). Other high density areas correspond to the militarised border regions with long-term concentrations of army units creating conditions for urbanisation and development of (road) infrastructure (lower Rhine, middle and lower Danube)³⁷. Low density areas include the marginal mountain and desert regions where habitability and roads are restricted, such as the Alps, Pyrenees, the southern Balkans, the Sahara and the Syrian desert. Some highly urbanised areas have lower road densities (such as Baetica, Hispania Tarraconensis or the Northern Levant), since their roads

are concentrated to valleys or coastal plains strongly bounded by mountains. These densities also highlight the long-term effects of differential research intensities and data availability, where the road structure of intensely researched parts of the Roman Empire such as Latium, the Southern Levant, or Northern France, are particularly well known (compare confidence map in de Soto et al.)³³. Density of ancient sites is positively correlated with Roman roads at a provincial scale (Pearson's ρ 0.45, Supplementary Fig. 3b), but the correlation is weaker (Pearson's ρ 0.25, Fig. 3d) when using 0.5° cells as analytical units (see 'Linear regression'), which likely reflects lower representativity of the

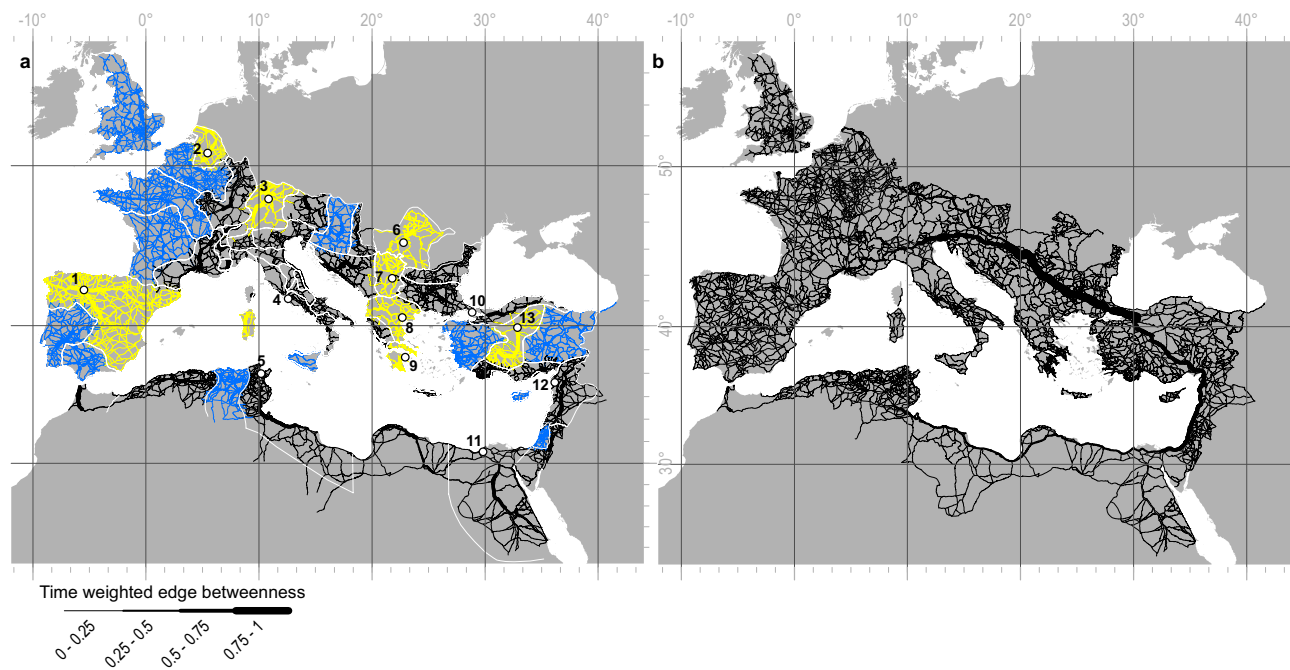


Fig. 4 | Network representation of the roads of the Roman Empire, line width represents time-weighted edge betweenness. a Results per Roman province (marked by white borders) reveal differences in the structure of key intermediary roads to provincial terrestrial movement: linear (black), distributed (blue), or centralised (yellow). **b** Results for the entire empire, revealing the importance of

the Mediterranean coastal roads and the intercontinental crossing at Byzantium. Selected major cities marked for reference (including focal cities in centralised provinces): 1 Legio-León, 2 Atuatuca-Tongeren, 3 Augusta Vindelicorum-Augsburg, 4 Rome, 5 Carthage, 6 Sarmizegetusa, 7 Naissus-Niš, 8 Thessalonica, 9 Corinth, 10 Byzantium, 11 Alexandria, 12 Antioch, 13 Ancyra.

road dataset in certain regions covered (see representativity map in de Soto et al.³³). Some areas are revealed to have a high road density and low settlement density, suggesting an underrepresentation of ancient places in the dataset used (Fig. 2d, see ‘Ancient sites’) such as in the central Iberian peninsula, northern and central France, western Balkans, and Bulgaria. Other areas demonstrate the opposite case of low road density and high site density, identifying underrepresentation in the roads dataset (northern Britain, Tuscany and Marche in Italy, Corsica, south-west Germany; representativity and reliability in the road data is discussed in de Soto et al. 2025 and in ‘Robustness of spatial regression’³³).

The structural location of major cities on the Empire-wide network

How central were major cities and Rome itself on the Roman road network? This question directly relates to the broader issue of transport-network design and to the role of cities in it. We answer this question by first exploring the location of cities on the Empire-wide network, we subsequently explore the most central roads per province, and finally combine this information with the centrality of cities per province to reveal the diverse ways in which roads structured the opportunities of city populations including Rome’s.

The Empire’s road network data representation comprises 10,516 nodes (representing intersections and dead ends) and 14,901 edges (representing roads), forming 13 connected components. The network’s Gamma index, Alpha index and average degree are strongly correlated (Supplementary Dataset 4, see ‘Network analysis’), as expected, since in the case of planar graphs they are all indices of density. To establish the significance of these values, we compare the results of some of the above measures with those of major modern roads (Supplementary Dataset 5). The Gamma Index is consistently higher for Roman roads than for major modern roads, reflecting great network density. The Alpha Index and global clustering coefficients are also generally higher. Although this result compares all Roman roads

with only the major modern roads, it still indicates a remarkable level of terrestrial interconnectedness in Roman times.

We computed a travel time weighted edge betweenness centrality³⁸. This metric quantifies the segment’s importance as an intermediary for terrestrial connectivity, by measuring the proportion of shortest paths passing through each road segment based on travel time using Tobler’s hiking function (see ‘Travel time weighted edge betweenness’). Figure 4b demonstrates the importance of the Mediterranean coastal roads for mediating terrestrial movement throughout the entire empire, in particular in North Africa, the Levant, southern France, and northwestern Spain. It further highlights the structuring effects of natural features like rivers and valleys (Po valley, Rhine, Danube, Drava, Maritsa, Tundzha) and cultural features like the Roman frontier (Rhine, and Danube to some extent) and important urban settlements: the high betweenness path passes Lepcis Magna, Alexandria, Caesarea Maritima, Antioch (Antakya), Nicomedia (İzmit), Byzantium (Istanbul), Philippopolis (Plovdiv), Serdica (Sofia), Aquileia, and Narbo (Narbonne). Indeed, 24 out of 45 provincial capitals are located in the top two deciles for empire-wide edge weighted betweenness centrality (see Supplementary Dataset 7 for empire-wide centrality scores of provincial capitals). Byzantium is highly centrally located in the continental road network as a whole, since it connects the European and Asian components of the Roman road network (via a ferry crossing the Bosphorus, another ferry connecting the European and Asian components is included at the Dardanelles³⁹, see ‘Network analysis’). Interesting deviations include the high betweenness values for roads going south of the Danube, following the Drava instead and crossing the eastern Alps near Graz until it reaches the Rhine near Mannheim. The general decrease of values from the centre to the periphery is to be expected for a betweenness metric, but it equally highlights peripheral areas in the continental Roman Empire’s road system, including Tunisia, the Sahara, most of Egypt, the Syrian desert, most of northeastern Turkey, Greece, Dacia, Italy, and the entire Atlantic facade. The connectivity of many of these regions should

instead be understood as mediating contacts within the province, across the frontier (Dacia, Syria, Sahara), via rivers (the Nile valley in Egypt), or maritime (Greece, Italy, the Atlantic facade, Tunisia, the Red Sea coast, north Turkey). Such terrestrial mobility results are hard to divorce from the reliance on river- and seaborne transport for crossing large distances. Roads were more essential for connecting inland areas and enabling effective short- to mid-distance travel such as travel within provinces.

Main roads central in provincial road networks

What were the most central roads in each province? Much has been written on the typological diversity of Roman roads and on how they should be classified. When referring to these routes using ancient terminology, one may invoke *viae publicae*, *viae vicinales*, and *viae privatae*⁹. For several decades, however, scholarship has tended to adopt a simpler dichotomy between primary and secondary roads. The former seems to comprise the major arterial corridors that structured overland connectivity across the Empire and are commonly recorded in the *Itineraria*, or marked by milestones⁶, yet quantitative evidence-driven confirmation that they could indeed have been more important (in highly specified ways) for structuring mobility are absent.

Across nearly all provinces with both main and secondary roads (40/43), main roads are more likely than secondary roads to have higher time-weighted edge betweenness (median probability = 0.65), indicating that they occupy more strategically important positions within the network (Fig. 5a). This shows the potential of betweenness to identify main roads in the Roman road network, and supports the interpretation that those sections classified as main roads were key intermediaries in Roman terrestrial transport. The degree and edge betweenness results are robust to random removal of uncertain road segments with a 20% probability (see 'Robustness of network analysis results').

The betweenness results further highlight roads that were important intermediaries for terrestrial mobility in each province, revealing a linear, centralised or distributed structuring of the roads in Roman provinces (Fig. 4a). Key intermediary roads are often structured linearly by a river (the Nile: Aegyptus; the Rhône: Gallia Narbonensis) which also marks the Roman frontier (the Danube: Pannonia Inferior, Moesia Inferior; the Rhine: Germania Superior), or along the coastline and the coastal mountain ranges around the Mediterranean (Italia, Epirus, Dalmatia, Lycia et Pamphylia, Cilicia, Bithynia et Pontus, eastern part of Africa Proconsularis, Mauretania Caesarensis, Cyrenaica). Interesting deviations include the frontier provinces Moesia Superior and Pannonia Superior, where the central roads are not along the Danube and through the provincial capitals, but rather centre on inland supply routes. Mountain ranges and passes structure key intermediary roads in mountainous parts of inland provinces (Noricum, Thracia, Alpes). In some provinces the high betweenness roads are revealed as being centred on focal settlements, from which the rest of the province can be effectively reached (Naissus in Moesia Superior, Sarmizegetusa in Dacia, Augusta Vindelicum in Raetia, Ancyra and Iconium in Galatia, Atuatuca in Germania Inferior, Corinth in Achaia, Thessalonica in Macedonia). Other provinces instead demonstrate a wider distributed coverage with multiple high betweenness roads (Pannonia Superior, Cappadocia, Asia, Britannia, Iudaea, western part of Africa Proconsularis, Numidia, Gallia Aquitania, Gallia Lugdunensis, Gallia Belgica, Hispania Tarraconensis, Lusitania, Hispania Baetica).

Capitals central in provincial road networks

Were major Roman cities and provincial capitals characterised by central locations on the provincial road system? Nodes with high centrality values play an important role in the road network, serving as major local crossroads (degree centrality), strategic points along shortest paths (time-weighted betweenness), or locations that are quickly reachable from across the system (time-weighted closeness). The time-weighted betweenness and closeness histograms (Fig. 5c–f)

reveal that capitals tend to be particularly centrally located as reachable places or intermediaries in a province when compared to other cities. Provincial capitals tend to be in central locations (35 capitals have at least one centrality score in the top decile; 14 out of 45 capitals have all three scores in either top or second deciles) and particularly on major local crossroads (degree centrality: 25 capitals in top decile, plus 8 in second decile, Fig. 5h; although all cities tend to have high degree scores since roads connect cities, Fig. 5g). Capitals with no centrality score in the first decile are important riverine (Lugdunum, Colonia Agrippina) or coastal ports (Carthage, Alexandria, Tingis, Caesarea Maritima, Perinthus, Caralis). These results quantitatively confirm Roman provincial capitals are beneficially located for either maritime or provincial terrestrial mobility, or both (Rome, Antioch, Corinth, Corduba, Alalia, Nicopolis, Iol Caesarea, Viminacium; Fig. 5b). The estimated population of major Roman cities is not correlated with centrality on the road network, with the exception of a weak but significant Spearman correlation between population and degree centrality in Italia ($p = 0.22$, $p < 0.001$), and Britannia ($p = 0.58$, $p < 0.01$); a significant correlation between population and closeness centrality in Achaia ($p = 0.54$, $p < 0.001$); and in Germania Superior both closeness ($p = 0.69$, $p < 0.01$) and betweenness ($p = 0.73$, $p < 0.01$).

Discussion

Our exploratory analysis of the topographic and network structure of the roads of the Roman Empire, presents spatially detailed answers at an Empire-wide scale to long-standing questions: were Roman roads straight, what was their relationship to the location of major modern roads, and how central were cities to structuring flows in the road network? Our findings are complementary to those of previous research in the economics literature on long-term persistence which, instead of the exploratory and descriptive spatial methods used here, uses econometric methods to identify causal relationships^{19,21,24–27}. These studies have been performed on smaller parts of the Empire using lower resolution road data, and we argue our work can facilitate more comprehensive analyses in economic geography, urban studies, and transport geography, by integrating the spatially-detailed empire-wide dataset in their models. Our work further goes beyond previous comparable qualitative and exploratory research through a much higher degree of specification: we specify under which conditions Roman roads are straight, how straightness and correlation with major modern roads varies precisely throughout the Empire, which cities are central, how central they are, and what we mean by centrality.

How straight were Roman roads? Our results can inform ongoing debates to reassess what has traditionally been interpreted as the straightest road system of antiquity, by highlighting the variability of how road construction was shaped by the physical environment and where shorter but more sinuous routes were preferred. We conclude that Rome created a road system where straight alignments were consistently favoured when the terrain permitted, but at the scale of the whole empire the common assumption that Roman roads were uniformly rectilinear is not accurate. This pattern suggests planning aimed at improving directness and efficiency within technological and topographical constraints, rather than pursuing straightness at any cost.

The general 'straightness' of Roman roads is explained by them crossing such areas where topography enabled the construction of rather straight roads. Main roads, both split into 1000 m segments and unsplit, exhibit slightly lower sinuosity values than secondary roads, closer to modern roads (Supplementary Table 1, median sinuosity split: main roads 1.0035, secondary roads 1.0051, modern roads: 1.0026). Therefore, a case for straight Roman roads might be made with regard to the main roads, which likely represent public and military roads (*viae publicae*, *viae militares*), or axes of land divisions (*limites*) laid out and engineered or adapted by specialists (*mensores*, *gromatici*).

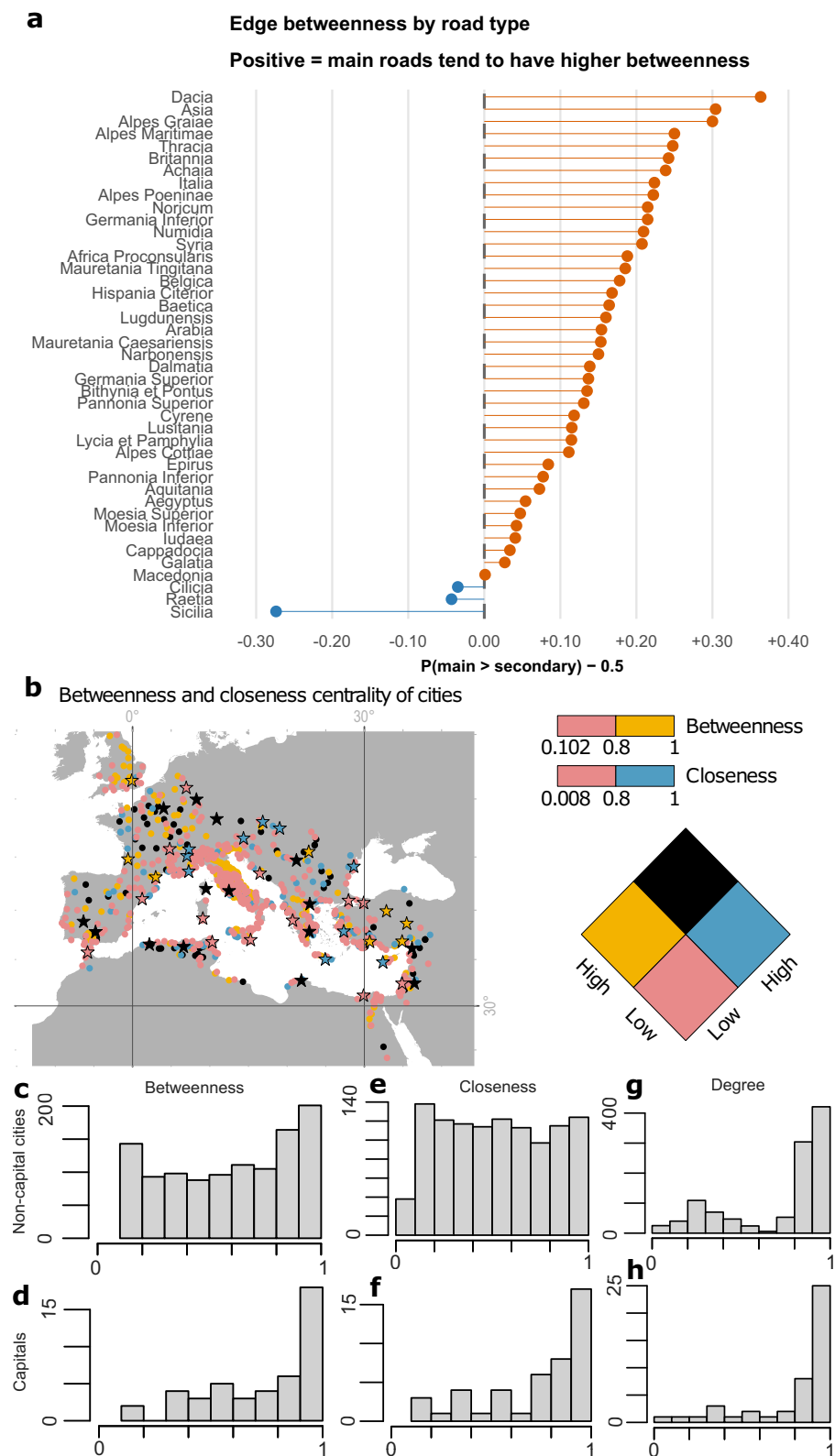


Fig. 5 | Centrality of roads and cities. **a** The lollipop chart shows, for each province, the probability that a randomly selected main road has higher betweenness than a randomly selected secondary road, providing a distribution-based measure of the relative strategic importance of the two road types. **b** betweenness and closeness centrality of Roman cities, provincial capitals are represented by a star

and tend to have high betweenness or closeness or both, and those who do not are major maritime or riverine ports (in white, e.g. Alexandria, Lugdunum).

c–h histograms of centrality deciles for closest nodes to cities. Capitals are skewed to the top 2 deciles of betweenness and closeness.

Our result that Roman roads are less straight than major modern roads is to be expected. While the Romans had highly accomplished engineers, capable of constructing bridges and tunnels, and at times cutting through or levelling terrain to maintain viable alignments^{40,41}, present day technologies allow for far more extensive and systematic landscape modification. Modern surveying and large-scale earth-moving enable straighter, high-speed road corridors that are less constrained by local topography, whereas ancient roads more often negotiated the surrounding topography.

Mountain ranges, river valleys and settlement distribution are important structuring features for the location of Roman roads. The vast majority of roads (270,955.85 km of roads or 90.57%) are within the limits of possible vehicular traffic below 16% rise, and are built in areas with low topographic variability. Mountains provide the most formidable barriers to wheeled traffic in the Roman Empire, but the relatively low slope values of Alpine roads as compared to other mountain ranges reflects its major valleys and passes as well as significant resources in adapting the terrain to facilitate terrestrial transport out of and into the Italian peninsula, as demonstrated e.g., by the works on the Aosta Valley road³⁵. The mountain roads are those where the difference between Roman and modern road sinuosity is most pronounced, as climbing strategies for walking (including pack animals) on steep slopes involve considerable use of switchback paths (zigzagging)⁴².

As a result of these calculations, our conclusion that Rome did seek to build straight roads wherever terrain conditions allowed confirms and expands on previous scholarship on Roman road planning³¹. At the scale of the whole Roman Empire, however, the common assumption that Roman roads were uniformly rectilinear is not entirely accurate, even if straight alignments were consistently favoured when feasible. As expected, slope and local topography appear to be the main factors shaping the structure of Roman roads. Thus, although Roman engineers could modify the landscape to a considerable degree, they could not avoid a minimum level of sinuosity. In this sense, geographic conditions directly structured the network. This is particularly clear in provinces of southern Europe and the Mediterranean basin, where more pronounced relief is associated with fewer straight roads, in opposition to north-western Europe, the Po valley, and the Pannonian basin with fewer obstacles for engineering straight roads.

What was the continuity of Roman roads up to the present day? Because we are comparing two networks separated by almost 2000 years, a wide range of historical, political, and economic processes have necessarily shaped the contemporary system. Traces of the Roman network are therefore unsurprising: many present-day cities either originated in, or already existed during, the Roman period, and connections between them often have deep historical roots. Over the intervening centuries, however, shifting demographic patterns, state formation, and changing economic priorities have also generated new motivations for building infrastructure where none previously existed, fostered the emergence of new communication hubs, and, conversely, reduced the strategic or economic relevance of other regions and corridors. Globally, Roman roads are moderately positively correlated (Pearson's ρ 0.46) with major modern roads, consistent with a previous study on a less detailed dataset on an empire-wide scale²⁵ and for the case of Italy²⁷. Our more spatially detailed results allow for the identification of a conspicuous increase in road density in the core urban and industrial areas of Western Europe (northern Italy, Rhine-Meuse basin, Britain) and highly concentrated change around large urban centres (e.g., Paris, Marseille, Madrid). This increase is most pronounced in small ancient population centres that grew into major metropolitan areas (e.g. Birmingham, Brussels, Paris, Madrid). Areas with a lower density of modern roads are located mainly in Greece, Asia Minor and North Africa (Tunisia and Algeria), which were exceptionally densely urbanised

areas in the Roman Empire. These differences reflect long-term historical shifts in terrestrial infrastructure driven by industrialisation, and urbanisation within nation states.

The high spatial resolution of our results and the wide coverage of our Roman roads dataset reveal enormous local variability in the effect of Roman roads on modern roads, suggesting the importance of local-scale mechanisms, in addition to the regional or continental scale ones focused on in previous studies. Our geographically weighted regression model (GWR) in turn reveals that the relationship between Roman and modern roads is overall strong (adjusted $R^2 = 0.61$), but it is shaped by highly local conditions. This conclusion broadly corresponds to findings by Dalgaard et al., who have identified such persistence in transport infrastructure through correlations between Roman and present-day road density, as well as exploring persistence in economic prosperity through comparisons with present-day nightlights and population density²⁵. In contrast, our spatial regressions did not reveal a strong relationship between Roman roads and modern population densities, suggesting a weaker pattern of long-term persistence along this axis of population density only.

We further notice that there is no geographically continuous region where Roman roads alone have strong local association with modern road infrastructure. Model residuals show highly heterogeneous patterns throughout the research area. Which might suggest different trajectories of historical development across the former Roman Empire, down to the very local level. These results are more complementary to previous research in what concerns the correlation between Roman and modern roads, where high local variability was revealed: the persistence of the Roman road system into the later periods as one of the explanations for long-term persistence of transport networks, economic activities, integration and sustained economic development was demonstrated by Flückiger et al. for Western Europe²⁶, by De Benedictis et al. for Italy²⁷, and by Wahl²¹ for the formerly Roman part of Germany. Bosker and Buring further demonstrated the importance of Roman roads in shaping urban development throughout the following two millennia⁴³. The dataset used in this research can thus be used for enhancing the empirical basis of studies of long-term persistence in economic activities^{19,21,24–27,43–46}, which have identified correlations between pre-industrial and present-day prosperous places.

Local regression slope coefficients and t-scores reveal regions where the location of Roman roads has a positive local effect on the location of modern roads and are statistically significant (Supplementary Fig. 8). They are found mostly in North Africa and the Near East with few locations around the Mediterranean and in Britain. This result puts into question the previous finding by Dalgaard et al.²⁵ where high correlation between Roman and modern roads was postulated for Western Europe and low correlation for North Africa and the Near East (explanatory power of the model not just slope magnitude), explained by changes in modes of transport (abandonment of vehicular traffic) in the Medieval period^{25,47}.

Our findings imply that while we well capture the dominant spatially varying relationship, there are highly localised processes that influence the location of modern roads. The linear regression model hints at some of the possible local processes: (a) local topography (elevation, TPI), to which we may add (b) influence of natural movement corridors (in mountainous areas) suggested by road orientation results, (c) Roman settlement patterns and persistence of the urban form, and finally (d) by historical and institutional development and persistence, which were not considered in this research. Our results should be explored in the future with different econometric methods and assessing the impact of the role of major Roman roads (which make up a much larger proportion in the dataset used in previous work) in driving long-term economic persistence, the different modern population datasets used, as well as the spatially detailed nature of the road dataset.

We demonstrate that road sections classified as main roads were key intermediaries in provincial terrestrial transport, as revealed by consistently higher edge betweenness centrality scores. As such, we have presented a metric that quantitatively describes how main roads structured terrestrial mobility (made possible due to the ability to compare with 195,693.3 km of secondary roads in our dataset). Previous research identified the main roads based on qualitative characteristics (presence of milestones, itineraries, etc.) and in turn assumed their importance. Our results independently confirm these previous assumptions using quantitative metrics. This result might reflect conscious choice by the Romans in planning and developing the road system on a macro-level. These important intermediary roads further revealed differences in the structures of provincial road networks (distributed, linear, or centralised), which indicates differences in how provincial administration and inhabitants were able to use the roads, and how the flow of people, goods and ideas were structured in provinces. A linear or centralised system affords intermediary centres to exercise significant control in trade, information flows or provisioning of a standing army (e.g. Naissus in Moesia Superior), but it is equally more vulnerable to targeted attack, blockades or failure in a central settlement. Provinces with a more distributed road network (e.g. the *Galliae*) are less vulnerable in that sense. These structures demonstrate how Rome used its road network to organise and structure each of the territories it dominated. Through road construction, the framework for mobility was built, and the territories were organised, facilitating transport between the main cities of each province.

How central were major cities in the network? We quantitatively confirm that Roman provincial capitals are beneficially located for provincial terrestrial mobility, more so than non-capital cities, and specify the nature of this centrality using metrics for all Early Imperial capitals. Capitals tend to be particularly centrally located as reachable intermediaries in a province when compared to other cities.

Despite the proverb ‘All roads lead to Rome’, major cities like Antioch, Ancyra, Byzantium and Corinth are better located than Rome as intermediaries for Empire-wide terrestrial mobility. Rome is spatially central in the Italian Peninsula (which is itself an appendix to the highest betweenness path on an Empire-wide scale) but Rome is not on the road with the highest time-weighted betweenness centrality in Italy. The current paper concerns the structure of terrestrial transport infrastructure only, and our results therefore cannot be seen to reflect all mobility through the empire. The importance of riverine and maritime mobility should be highlighted for most transport over long distances, for which Rome is highly centrally located in the empire. Future work should combine this road dataset with navigable rivers and sea connections and alternative modes of transport (e.g. horse courier, ox cart, pack animal), to identify the structure of the multi-modal Roman transport system, which will reveal a much higher centrality of regions central in the Mediterranean (such as Italy, Greece, and Tunisia) and especially of major Mediterranean ports¹⁸. Scheidel has demonstrated how sea-lanes bind together the ‘core’ Mediterranean areas of the Roman Empire¹⁸. We further expand on Scheidel’s analysis by identifying capital cities with exceptional locations for both waterborne and terrestrial mobility (Rome, Antioch, Corinth, Corduba, Thessalonica, Nicopolis ad Istrum, Iol Caesarea, Viminacium, Augusta Vindelicum, Mogontiacum).

High time-weighted edge betweenness of the main roads and high time-weighted betweenness and closeness centrality of many provincial capitals when viewed together might imply system-level decision making on provincial and even supra-provincial level in developing the road system for the needs of the imperial power. This raises contentious issues whether the Roman state was capable of such planning and execution, how this planning could have been organised, what was the role of maps in this process and whether Romans were capable of producing maps that could support such planning^{8,29,48–53}.

We argue that our results (and particularly the high time-weighted betweenness of main roads and capital cities at a provincial and empire-wide level) support the conclusion that such planning and development of the Roman transport network was undertaken on a provincial level (and that empire-wide high betweenness centrality of provincial capitals could reflect empire-wide planning). However, our analyses do not reveal the causes, mechanisms or sequence of planning. Future work should focus on collecting the detailed temporal data about road construction and refurbishment to enable evaluating the explanatory power of factors that structured planning, such as the physical environment, the pre-existing road system, military conquest and pacification, and changes of provincial capitals and administrative divisions.

Caveats to our data analysis should be noted. Chief among these is that the road data used is the best available but not complete due to variation in the reliability and intensity of scholarly work and excavations/surveys across the research area, the heterogeneity across the dataset affecting its representativity, and the varying empirical basis of individual road segments. We performed robustness checks of our main results in light of data uncertainty, and future work could additionally integrate uncertainty into estimations directly. We demonstrated that degree and edge betweenness results are robust to random removal of uncertain edge segments with 20% probability. We provide concrete guidance for targeting future fieldwork and future improvements to the dataset, through our categorisation of road certainty and our confidence map³³, as well as our identification of areas with high site density but low road density (Fig. 2d). At the same time, it is also important to point out that the creation of this absolutely extraordinary network of roads was a gradual process. The Roman road system was a development of earlier routes and changed significantly throughout the imperial period. This long development unfolded in stages, driven by various military, political, economic, and other motivations, all influencing the decisions made. The layout of these roads, therefore, is the result of the evolution of pre-Roman roads, military needs, policies of domination, regional and local interactions, and, above all, economic interests. Our results are derived from an aggregation of the available information and necessarily assume contemporaneity and continuous use of settlements and roads. Like the dataset itself, they are therefore more representative of the 2nd–4th centuries CE structure of the empire, because some settlements and major road infrastructure were not yet in existence in the Republican and early Imperial period, and the system would be less uniformly in use and integrated in late antiquity. For all these reasons, it will be crucial in future research to evaluate the network from a diachronic perspective, assessing the existing network at each historical moment and the modifications made in relation to past and future periods. Comprehensive and comparable empire-wide data on this temporal change does not exist and future work should focus on systematically collecting it to enable the evaluation of the robustness of our findings in light of the dynamic nature of the Roman road system: temporal changes in empire-wide and provincial-level road network structure, development of centrality of cities, and the structural impact of major infrastructure projects including on the economic, administrative and military needs of the time. A further challenge is that the structure of terrestrial transport is most valuable for understanding regional mobility, especially in landlocked places, rather than long-distance movement for which interlocking means of transport should be considered. Future work needs to integrate terrestrial routes with maritime and riverine routes and infrastructure, to create a model of ancient transport at a higher spatial accuracy than is currently available⁵⁴. This would be an invaluable resource for studying the development of Empire-wide flows of people, goods, ideas and disease that were not merely structured or restricted to terrestrial means of transport.

Methods

Empire and province boundaries

The geographical limits of the road data and the analysis region are defined by the extent of the Roman Empire during the Antonine dynasty, ca. around 150 CE, when the Empire reached its largest extent. Provincial boundaries are based on provincial boundaries and Roman Empire extent by the year AD 200 available at the Ancient World Mapping Center⁵⁵ and are only approximate. Provinces east of Euphrates were excluded, due to being occupied by the Roman Empire for short time periods and to being excluded from the road data collection³³. All boundaries were clipped to modern coastlines. Inclusion of Italian *regiones* enhances the granularity of the results in the Italian peninsula. Unfortunately, a similar finer division is not entirely possible for other provinces.

Analysis grid cells

Summary statistics for Roman roads, modern roads, modern population, mean elevation, mean TPI, road density, site density, and comparison between roads and sites were aggregated in 0.5° latitude x 0.5° longitude cells covering the entire research area. These cells were used as analytical units because they provide good commensurability of the data and results whilst still offering a fine-grained resolution reflective of the density of Roman roads. Cells without Roman and modern roads were excluded from the analysis.

Modern roads data

The dataset of modern roads used in the article is a spatial subset of the World Roads dataset available at ArcGIS Online⁵⁶. It is based on a world transportation dataset supplied by Garmin International, Ltd. and distributed by Esri. It contains modern major roads and highways, and selected 'Local roads' in certain areas. All are paved roads. Ferries were removed from the dataset for the analysis. Only roads overlapping with 0.5 degree cells were selected for analysis. This dataset was selected because it is open access, covers the entire research area, provides comparable data coverage across the more than 30 modern countries in the research area, and crucially because it represents a comparable level of road inclusion to the Itiner-e dataset since both focus on major connections between settlements, neither includes city streets, private roads, or local tracks.

Ancient sites

The site density and spatial network models methods use locations of 14,317 ancient sites that intersect with our research area derived from the Pleiades: A Gazetteer of Past Places⁵⁷ dataset with, according to the Pleiades ontology, the type values of 'settlement', 'villa', 'fort', and 'station' (referring to a road station), and time period 'Roman' (featureTyp LIKE '%settlement%' Or featureTyp LIKE '%villa%' Or featureTyp LIKE '%station%' Or featureTyp LIKE '%fort%' And timePeri_1 LIKE '%roman%').

Ancient city population

The city centrality results use Hanson's dataset⁵⁸ of urban settlements in the Roman Empire, with associated population estimates (871 cities).

Modern population

Modern population data are based on the LandScan Global 2022 population dataset⁵⁹. The population values were summarised within 0.5° cells.

Density

For both Roman and modern roads, kernel density was calculated in a 25 km neighbourhood over an area bounded by the extent of the 0.5° cells defining the research area. A smooth density surface is fitted over each road segment, with a density value of 1 where the surface overlaps with the highest values of the segment, and the density decreases away

from the segment, reaching 0 at the specified neighbourhood limit of 25 km. No custom weighting or rescaling was applied. A 25 km neighbourhood was selected as approximately representing the average 1-day walking distance of a traveller. The kernel density was calculated using the 'Kernel Density' tool in ArcGIS Pro 3.2⁶⁰.

Comparison road and settlement density

Digital high-resolution demography data for the entire Roman Empire does not currently exist. Here we use the 14,317 sites described in 'Ancient sites', with approximately equal coverage and representativity across modern-day national states. A site density (sites/km²) was then calculated for each 0.5° cell. Provincial site density was calculated using the 13,638 sites within the provincial boundaries.

Slope, TPI and sinuosity

We present results for three basic topographic properties of roads: slope (average and maximum), mean TPI (topographic position index), and sinuosity. All road segments are split into 1 km sections (resulting in 407,800 road segments), used as an analytical unit to enable comparison of results.

Slope is calculated along each segment of a line over a Digital Elevation Model (DEM). The length of the segment is defined by the resolution of the DEM (see below). Maximum slope is obtained from the segment with the largest value. Average slope is obtained by taking a weighted average of the slope from each line segment. Slope was calculated using the 'Add Surface Information' tool in ArcGIS Pro 3.2⁶¹ and then converted from percentages to degrees.

TPI^{62,63} is a measure of prominence of a raster cell in a DEM in a specified neighbourhood, calculated as a difference between the elevation of the focal cell and the mean elevation of its neighbourhood. Positive values indicate elevated areas, whereas negative values represent valleys. The TPI was calculated using the 'Topography Toolbox' for ArcGIS Pro 3.2⁶⁴.

Sinuosity is a measure of the 'straightness' of the road. It is calculated as a ratio of actual road length to straight line distance between the start and end point of a 1 km road segment. The values closer to 1 indicate straight roads (with 1 being completely straight), higher values indicate roads with a more winding course. Sinuosity was calculated using 'Stream Gradient' and the 'Sinuosity Toolbox' for ArcGIS 10.1⁶⁵.

Both slope and Mean TPI were calculated over a Copernicus GLO-90 3-arcsec resolution DEM covering the whole research area. The Copernicus DEM was chosen because it offers better coverage in the mountainous areas and better vertical accuracy compared to SRTM or Aster GDEM data. Mean TPI was calculated using a 47 cell radius neighbourhood, which corresponds to ca. 5 km (the resolution of the DEM is ca. 107 m at the latitude of the research area). The 5 km neighbourhood was selected as we assume that local topography has a decisive influence on the location of the road.

Topographic structure by certainty

8191.1 km of road data is certain (2.737%), 22,280.9 is hypothetical (7.445%) and 268,801.2 is conjectured (89.818%). Certainty categories and their implications for data confidence are discussed in detail in de Soto et al.³³. 'Certain' indicates roads digitised with high spatial accuracy, 'conjectured' roads are digitised with lower spatial accuracy, and 'hypothetical' roads are those that were identified to exist but were not located or had a less fixed track.

Supplementary Fig. 1a, b captures variability of the average and maximum slope values between the certainty categories, showing that 'Certain' road segments (i.e., those digitised with high spatial accuracy) have slightly lower mean values and lower maximum values compared to 'Conjectured' segments (1.45° vs 1.54°, and 16.1° vs 31.2°). This might reflect the lower spatial accuracy of the conjectured segments since they might be located on unusually steep slopes. Supplementary Fig. 1c demonstrates that the TPI values are skewed towards slightly negative

values among all certainty categories, comparable to the overall results for the roads. Supplementary Fig. 1d shows variability of the road sinuosity, where ‘Certain’ segments have slightly lower values compared to ‘Conjectured’ and ‘Hypothetical’ segments. As with the average slope values, this might reflect lower spatial accuracy of the ‘Conjectured’ and ‘Hypothetical’ segments.

Slope, TPI and sinuosity by road type

For definitions of main and secondary roads, see de Soto et al.³³. Supplementary Datasets 1–2 demonstrate that main roads tend to have lower mean slope and sinuosity values, but slightly higher TPI values. However, the differences are slight. When looking at both road type and segment certainty we see a similar picture, but the Certain road segments tend to have higher average slope values and Hypothetical segments tend to have TPI values closer to zero. This is likely caused by the fact that many certain segments are recorded in upland areas, on slopes and mountains. Hypothetical segments are then mostly found in the desert areas of North Africa and Syria in moderately flat areas (hence TPI values closer to zero).

Orientation

Orientation was determined based on road segments from one intersection to the next intersection. All road segments in the dataset are split at intersections with other road segments. Segment orientation was classified into six direction categories covering 30° of the compass each: N-S, NNE-SSW, NE-SW, E-W, SE-NW, SSE-NNW (see Supplementary Table 8 for orientation classes in degrees). Only the start and end point of each segment is taken into account during the calculation, using the ‘Linear Directional Mean’ tool in ArcGIS Pro 3.2⁶⁶. Predominant orientations of roads per province were calculated by aggregating the total length per orientation. To enable this, all roads going across provincial borders were cut at the borders. Road segments were excluded from the analysis if they go beyond the borders of the Empire (captured by the polygon describing provincial borders). Supplementary Dataset 3 provides a comparison with the predominant orientations of mountain ranges in mountainous provinces (aggregated by total area of the orientations) computed from the GMBA Mountain Inventory v2 dataset⁶⁷.

Network analysis

We represented the road system of the Empire overall and of individual provinces as planar undirected networks, with links representing road segments and nodes representing intersections and dead ends. To enable an analysis of the entire continental part of the Roman Empire, two artificial road connections were added at known ancient ferries³⁹ across the Bosphorus and Dardanelles that connect Europe to Asia Minor. As a result, the largest component spans Europe, the Middle East, and Africa, while smaller components include Great Britain and Mediterranean islands (see Fig. 4).

Supplementary Dataset 4 summarises essential spatial network properties of the Roman road networks and Supplementary Dataset 5 for major modern roads, both per Roman province⁶⁸.

Gamma index⁶⁹: the ratio $\frac{e}{3(n-2)}$ where e is the number of edges and n is the number of nodes. $3(n-2)$ is the maximum possible number of edges in an undirected planar network with n nodes. It is an index of network density.

Alpha index⁷⁰ of the largest connected component: the ratio $\frac{e-n+1}{2n-5}$, measuring the density of bounded faces in the largest connected component. It ranges between 0, for a tree, and 1, for a maximally connected planar network.

Additional network analysis measures were derived to enable comparison with non-spatial networks:

Global clustering coefficient: the number of closed triples over the total number of triplets (both open and closed).

Average degree: the average number of connections of a node.

Travel time weighted edge betweenness

The edge betweenness centrality is the number of shortest paths between any two nodes in a network that include a given edge³⁸. This measure can be interpreted as an indicator of the edge’s role in ensuring efficient connections in the network.

We computed the edge betweenness for the empire as a whole (Fig. 4b) as well as per province for all road segments that fall within a provinces’ borders (Fig. 4a), using the edge_betweenness function from the R package igraph (v.1.4.3)⁷¹. Different definitions of a shortest path can be employed; in our case, we defined it on the basis of the travel time to traverse each segment.

To estimate the travel time, we applied Tobler’s hiking formula Eq. (1)⁷². Let l be the length of a road segment in kilometres and let s be its average slope. The walking speed v in kilometres per hour is given by

$$v = 6 \exp(-3.5[s + 0.05]) \quad (1)$$

and the travel time in hours is calculated as $t = l/v$.

The data set classifies each road segment as either main or secondary (see de Soto et al.³³). For each province, we computed the probability that a randomly selected main road has higher betweenness than a randomly selected secondary road. This provided a measure of the relative strategic importance of the two types of roads (Fig. 5a).

Cities and node centrality

Supplementary Dataset 6 presents for each province’s road network the time-weighted betweenness, time-weighted closeness and the degree of the nearest node to a Roman urban settlement (see ‘Ancient city population’). Capital cities are more strongly skewed towards the top decile for betweenness and closeness than all cities are.

Robustness of network analysis results

Since the road dataset is characterised by a degree of uncertainty and incompleteness, we tested the robustness of node degree and betweenness centrality, and of edge betweenness centrality with respect to random removal of uncertain road segments (i.e., those classified as either conjectured or hypothetical) using methods from refs. 73,74.

Supplementary Fig. 2 illustrates the effect of randomly removing uncertain road segments (i.e. those classified as ‘conjectured’ or ‘hypothetical’) with a 20% probability across 100 iterations, demonstrating robustness of the network results presented here. Supplementary Fig. 2a, b show that nodes with high degree are slightly less frequent in the randomised networks. For instance, the node with the highest degree (10) retains an average degree just above 8, never dropping below 5. The edge betweenness centrality remained larger for main roads than for secondary roads in all sampled random networks (Supplementary Fig. 2c). The betweenness of main roads is on average 1.95 times that of secondary roads, with a minimum ratio of 1.30 and a maximum of 3.01 across the random networks.

Supplementary Fig. 2d–i shows the centrality deciles for the closest nodes to cities for 100 test networks, each generated by randomly removing conjectured or hypothetical road segments with a 20% probability. The skewedness of capitals to the top 2 deciles of betweenness and closeness remarked in Fig. 5c–h persists in the test networks, indicating that the association between the locations of provincial capital cities and nodes with relatively high values of these centrality measures is robust with respect to the removal of uncertain road segments from the provincial road networks.

Linear regression

We use 0.5° cells as analytical units (Fig. 2d), as well as Roman provinces (Supplementary Fig. 3). Within these units we summarised several relevant environmental and anthropic variables (see

Supplementary Table 2). Both Roman and modern roads data were compared based on the area within a 5 km buffer along roads, an approach used in previous studies that enables comparison with previous results²⁵. All polygons containing no Roman and modern roads were excluded from the analysis. A comparison of the percentage of these buffers occupying each analytical unit demonstrates only a small difference between Roman and modern roads. However, in the case of Roman provinces, the variation is more significant and it displays a different pattern with more variability than Roman roads (Supplementary Fig. 3a). We compute Pearson correlation to quantify the strength of a linear relationship between these variables. The results indicate a moderate positive linear correlation between Roman and Modern roads (for 0.5° cells a coefficient of approximately 0.460, $p < 2.2 \times 10^{-16}$, and for provinces approximately 0.539, $p = 1.541 \times 10^{-5}$). These results suggest that as the percentage of buffer area for Roman roads increases, there tends to be a corresponding increase in the percentage of buffer area for modern roads, and vice versa (Fig. 3d and Supplementary Fig. 3b; Supplementary Table 1).

Spatial regression

We use Geographically Weighted Regression (GWR) analysis using 0.5° cells as analytical units, to explore in a spatial context the relationship between the key anthropic and environmental variables and the location of Roman roads, and the association of Roman roads and the location of modern roads. We first identify the level of spatial autocorrelation to evaluate spatial dependencies and clustering of the data using Moran's I index. Given the resolution of the dataset, 100 km distance bands were selected. A field representing the percentage of the 5 km buffer around Roman roads in each polygon cell was used as input (see 'Linear regression'). The results show the highest Moran's I index in the first distance band (100 km) which sharply decreases at larger distances. However, as we evaluate z-values against the null hypothesis that observed phenomena are not spatially correlated, we see the most significant spatial autocorrelation around the 900 km distance band (Supplementary Table 3). Therefore, when selecting a neighbourhood for GWR the neighbourhood threshold was set at 1000 km. Both Moran's I index and GWR analysis were undertaken using ArcGIS Pro 3 tools^{75,76}. Notably high z-scores and therefore very low p values might be caused by the effect of a large sample size which decreases the variance of Moran's I index and inflates the z-scores. Moreover, it implies very strong spatial dependencies in the dataset. However, since Moran's I is a global diagnostic useful for global spatial models, it might not be the best indicator for defining a neighbourhood for a local GWR model. It was decided to compare GWR results using the 1000 km neighbourhood with adaptive neighbourhood using a Golden Search method to identify the optimal number of neighbours in order to find the model that better explains the variance and provides a better fit (using AICc and Adjusted R^2 as a measure).

We explored three different scenarios with GWR (Fig. 3a–c; Supplementary Table 4): (1) whether Roman roads are associated with the presence of modern roads, (2) or the density of the modern population; and (3) to what degree mean elevation, the mean TPI calculated in a 5 km neighbourhood, or site density are associated with the location of the Roman roads. With exception of scenario 3, the models using the optimal number of neighbours method (Supplementary Table 4) provided better fit than the 1000 km neighbourhood (Supplementary Table 5). The 1000 km neighbourhood model at this scale tends to oversmooth local relationships and its results are closer to a global linear regression model (Fig. 3c and Supplementary Table 5, note very similar adjusted R^2 values to the linear regression). The differences are most pronounced in scenario 1, where the model using the optimal number of neighbours shows strong spatial non-stationarity (i.e., the relationship between Roman and modern roads is spatially very varied). The following paragraphs describe the best performing models reported in the article (Supplementary Table 4). It shows that for

scenario 1 and 2 the scale of coefficient non-stationarity is smaller than the dominant scale of spatial clustering.

Scenario 1 (Fig. 3a and Supplementary Table 4): the model shows pronounced spatial non-stationarity, the relationship between Roman roads and the location of modern roads varies widely across the study area. There is no contiguous region or area where the model shows good local fit. The model explains a large share of the variance ($R^2 = 0.69$, adjusted 0.61), but the scale at which the relationship operates is small to medium, considering the optimal number of neighbours (31). We may surmise that the relationship between Roman and modern roads is heavily dependent on local influences rather than on transregional or continental scale circumstances e.g., institutional differences between various European regions or between Europe and North Africa and the Near East that shaped persistence and development of transport infrastructure, while local effects of historical and institutional development are more likely. Among other causes influencing the varied effect of Roman roads on modern roads is undoubtedly (a) topographical constraints and natural movement corridors, and (b) varied representativity and reliability of the dataset that influences the confidence in the reported results in certain regions (chiefly central Europe, Balkans, and marginal and (semi-)desert areas) (see 'Robustness of spatial regression').

Scenario 2 (Fig. 3b; Supplementary Table 4): The model explains little of the variance (R^2 values 0.39, adjusted 0.25) but while the modelled relationship is spatially smooth, it fails to capture the key drivers of the variation. It is only in regions with a very high density of population - mainly the major cities (i.e. London, Paris, Madrid, Rome, Milan, Naples, Istanbul, Ankara, etc.) and a few regions such as the Lower Rhine and the Nile Valley, where under- or overpredictions are present. We may suggest that in this scenario, our results may be biased by modern development around contemporary major urban areas, with more investigations and higher recovery rates of archaeological remains providing detailed data on the Roman road network compared to other areas (e.g., rural Spain, western Balkans, etc.). But substantively, Roman roads are not very well associated with modern population density.

Scenario 3 (Fig. 3c and Supplementary Table 4): we explored the spatial relationship between Roman roads and several environmental and anthropic variables, which were selected to help us understand where they had the most decisive influence on the formation of the Roman road network. A global adjusted R^2 value of 0.35 (0.5-degree cells) suggests a rather low fit of the model when using a limited number of chosen variables. This might be explained by poor coverage of the road dataset in certain regions (south-western France, upper Danube, central-south Balkans, central Italy, Sicily, Corsica) or by underrepresentation/overrepresentation of site data in the Pleiades dataset (coast of the Near East and North Africa, see 'Ancient sites'). Additional environmental factors, especially related to climate, precipitations and soil quality might become helpful in marginal desert and mountainous environments.

Robustness of spatial regression

We can assess the effect of the neighbourhood size by comparing with the GWR model using a 1000 km neighbourhood. Scenarios 2 and 3 (Supplementary Table 5) do not substantially differ when using either a 1000 km neighbourhood or optimal number of neighbours. In scenario 1 (Supplementary Fig. 4, Supplementary Table 5) the large neighbourhood tends to oversmooth the results and so the performance of the model is close to the non-weighted linear regression. The pattern of model residuals is locally varied and the presence of Roman roads does not have strong explanatory power on the location of modern roads, as in the case of the GWR model using a smaller neighbourhood. The global R^2 value (0.47, adjusted 0.46) implies that the model explains a substantial share of the variance but the relationship between the two is more complex. The regions of

overprediction (using standardised residuals) show interesting patterns. They cover the former Roman frontier in Europe, along the Rhine and most of the Danube. This may suggest that the sustained presence of the Roman army on the borders of the Roman Empire and investment in developing the road system, might have served as a framework for later development, and may be a process driving road persistency. Among other areas where we can observe high model residuals are regions around large modern agglomerations (London, Paris, Madrid, Marseille, Milan, Istanbul, Athens, Cairo, etc.). Extensive construction and improvements of the modern road network for the growing major cities could bias the model in such cases. In other instances, the high residuals might be caused by geography which was the major factor in defining settlement patterns and movement (the Nile Valley, Cyrenaica and Tripolitania in Libya, Israel, Lebanon, and Palestinian Territories). However, this pattern cannot be observed e.g., in the Alps or the Taurus Mountains, which require different explanations. In yet another instance, high residuals could be explained by both geography and the presence of modern major cities (e.g. Madrid, Ankara). We may observe a particularly telling pattern in the Iberian Peninsula. The highest residuals are found on the coast and in the interior around Madrid. This corresponds well with previous conclusions²² on the historical development of the road network in Spain, where centralisation tendencies focusing on Madrid resulted in a higher concentration of population and the development of a road network radially extending from Madrid to the detriment of the rest of the interior of Spain. Therefore, outside of Madrid, the model shows neutral or slightly negative residuals.

We compared the GWR results obtained for 0.5° cells with results obtained for a finer resolution dataset composed of 0.25° cells (Supplementary Fig. 5, Supplementary Table 6). We used the same Golden Search method to identify the optimal number of neighbours. Surprisingly, the models for scenarios 1 and 2 produced at 0.25° resolution performed worse than models obtained for 0.5° polygons both regarding global and adjusted R² values. However, all 3 scenarios essentially reveal the same patterns (or lack thereof) as the coarser resolution models. In scenario 2, more places of over- and underpredictions are identified, but these still represent major modern urban centres, and so the issue of representativity of the road dataset with regards to recovery of archaeological data in modern cities remain. Only in scenario 3, the global and adjusted R² values slightly improve, thus explaining more variance (0.4 vs. 0.35), implying that elevation, TPI and Roman settlement patterns have stronger influence on the location of Roman roads at a small scale.

We compare the GWR results with the representativity and reliability provided for the Itiner-e dataset³³. Representativity quantifies how the density and spatial detail of the dataset is representative with regard to the global Roman road density average. Reliability expresses reliability of the sources used for creation of the dataset. We overlay these categories over the GWR results for all 0.5° cells that have both Roman and modern roads (Supplementary Fig. 7). The definition of the categories was selected to highlight the differences between low to negative standardised residuals (<0.5) and high standardised residuals (>0.5), and low representativity ('Low' category in de Soto et al. (2025)), and high representativity (categories 'Average', 'Above Average', and 'Exceptional' in de Soto et al. (2025)). The reliability categories ('Low', 'Medium', 'High') were not changed. It shows regions where we have higher confidence in the results (high representativity and medium to high reliability) in north-western Europe, Spain, Italy, Greece, the eastern Balkans, most of Asia Minor, the Near East, the Eastern Desert of Egypt, coastal Cyrenaica and Tripolitania, and North Africa. Problematic areas (low representativity and low reliability) are found mostly in central Europe, the western Balkans, the Nile valley, marginal semi- and desert areas, and Corsica. However, poor results for the western Balkans are caused mainly by low reliability of sources,

while the representativity of the data is high and therefore the residuals shown by GWR are likely robust and comparable to high confidence regions. The most problematic region is then central Europe roughly from the upper Rhine to Pannonia, which consistently exhibits low representativity and low reliability. The same applies to semi-desert and desert areas, but it is unlikely that the Roman road system there was more extensive, and therefore GWR results should hold. Areas of low representativity and medium reliability, covering peripheral regions of Britain, Portugal, Switzerland, Sardinia, parts of the Alps and Italy, central Anatolia, and most of Egypt, are places where the results might be potentially flawed, due to underrepresentation of roads in the dataset. We provide a further robustness check by re-creating scenario 1 using only 0.5° polygons that have either 'Medium' or 'High' reliability and 'Average' to 'Exceptional' representativity. The model (Supplementary Fig. 6 and Supplementary Table 7) leads to comparable results, although it has slightly lower R² values (0.59, adjusted 0.53).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Roman roads dataset: the version of the Itiner-e dataset used to perform this analysis is available as a polyline vector layer in the shapefile format (itinerere.roads.shp), geopackage (itinerere.roads.gpkg) and as a geojson database (itinerere.roads.geojson) on the Zenodo repository⁷⁷. A growing gazetteer version of the dataset can be viewed and queried on the Itiner-e linked open data platform (<https://itiner-e.org>)⁷⁸. The creation of the dataset is described in de Soto et al.³³. The version of the road dataset split into 1000 m long segments with calculated topographic variables, the polygons representing various analytical units (provinces, 0.5° and 0.25° cells) containing topographic and anthropic variables, and additional data generated with analytical tools and used in this article are stored at a Zenodo repository⁷⁹. Modern roads: World Roads dataset available at ArcGIS Online⁵⁶. Provincial boundaries: Roman Empire extent by the year AD 200 available at the Ancient World Mapping Center⁵⁵. Ancient site locations: Pleiades: A Gazetteer of Past Places⁵⁷. City population estimates: Hanson Cities Database (OXREP)^{58,80,81}. Modern population data: LandScan Global 2022 population dataset⁵⁹.

Code availability

The spatial datasets used and the code to reproduce the network analysis results is available at: <https://github.com/MatteoMazzamurro/roman-road-networks>⁸². Code and sources for the linear regression analysis are available here: https://github.com/Mcotsar/Ancient_Roads_Exploratory_Analysis

References

1. Şahin, H. CIL XVII, 5, 3: Neue Meilensteine und Straßen aus der Cilicia Aspera. in *Roman Roads. New Evidences - New Perspectives* 166–190 (De Gruyter, 2019).
2. Grewe, K. Alle Wege führen nach Rom - Römerstraßen im Rheinland und anderswo. 'Alle Wege führen nach Rom....': *Internationales Römerstraßenkolloquium Bonn. [vom 25. bis 27. Juni 2003]* **6**, 9–42 (2004).
3. Quilici, L. Land Transport, Part 1: roads and bridges. in *The Oxford Handbook of Engineering and Technology in the Classical World* (ed. Oleson, J. P.) 551–579 (Oxford University Press, 2008).
4. Witcher, R. Roman roads: phenomenological Perspectives on Roads in the Landscape. *Theor. Rom. Archaeol. J.* **1997**, 60–70 (1998).
5. McCormick, M. Rats, communications, and plague: toward an ecological history. *J. Interdiscip. Hist.* **34**, 1–25 (2003).

6. Kolb, A. *Roman Roads: New Evidence, New Perspectives* [Papers from the International Conference 'Viae Romanae - Roman Roads: New Evidence - New Perspectives' Held at the Historical Seminar of the University of Zurich on the 1st and 2nd of June 2017] (De Gruyter, 2019).
7. Lewis, J. Estimating the scale-dependent influence of natural terrestrial corridors on the positioning of settlements: a multi-scale study of Roman forts in Wales. *J. Archaeol. Sci.* **170**, 106055 (2024).
8. Luttwak, E. *The Grand Strategy of the Roman Empire: From the First Century A. D. to the Third.* (Johns Hopkins Univ. Pr, 1979).
9. Chevallier, R. *Les Voies Romaines* (Picard, 1997).
10. Hitchner, R. B. Roads, integration, connectivity, and economic performance in the roman empire. in *Highways, Byways, and Road Systems in the Pre-Modern World* (eds Alcock, S. E., Bodel, J. & Talbert, R. J. A.) 222–234. <https://doi.org/10.1002/9781118244326.ch11>. (Wiley, 2012).
11. Cuntz, O. *Itineraria romana 1, Itineraria Antonini Augusti et Burdigalense* (Teubner, Leipzig, 1929).
12. Miller, K. *Itineraria Romana. Römische Reisewege an Der Hand Der Tabula Peutingeriana.* (Strecker und Schröder, 1916).
13. Quilici Gigli, S., Quilici, L. *Tecnica stradale romana* (L'Erma di Bretschneider, 1999).
14. Carreras, C., De Soto, P. & Muñoz, A. Land transport in mountainous regions in the Roman Empire: Network analysis in the case of the Alps and Pyrenees. *J. Archaeol. Sci. Rep.* **25**, 280–293 (2019).
15. Bichet, V. et al. Traverser les montagnes du Jura: identification de voies antiques de franchissement de la haute chaîne jurassienne par analyse LiDAR. in *Des routes et des hommes: la construction des échanges par les itinéraires et les transports* (ed. Lemaître, N.). <https://doi.org/10.4000/books.cths.4356>. (Éditions du Comité des travaux historiques et scientifiques, 2019).
16. Isaksen, L. The application of network analysis to ancient transport geography: a case study of Roman Baetica. *Digit. Mediev.* **4**, <https://doi.org/10.16995/dm.20> (2008).
17. Carreras, C. & De Soto, P. The roman transport network: a precedent for the integration of the European mobility. *Hist. Methods* **46**, 117–133 (2013).
18. Scheidel, W. The shape of the Roman world: modelling imperial connectivity. *J. Rom. archaeol.* **27**, 7–32 (2014).
19. Percoco, M. Highways, local economic structure and urban development. *J. Econ. Geogr.* **16**, 1035–1054 (2016).
20. Orengo, H. A. & Livarda, A. The seeds of commerce: A network analysis-based approach to the Romano-British transport system. *J. Archaeol. Sci.* **66**, 21–35 (2016).
21. Wahl, F. Does European development have Roman roots? Evidence from the German Limes. *J. Econ. Growth* **22**, 313–349 (2017).
22. De Soto, P. & Carreras, C. The role of politics in the historical transport networks of the Iberian Peninsula and the Pyrenees from Roman times to the nineteenth century. *Soc. Sci. Hist.* **45**, 233–260 (2021).
23. Licio, V. When history leaves a mark: a new measure of roman roads. *Ital. Econ. J.* **7**, 1–35 (2021).
24. Bottasso, A., Conti, M., Robbiano, S. & Santagata, M. Roads to innovation: evidence from Italy. *J. Reg. Sci.* **62**, 981–1005 (2022).
25. Dalgaard, C.-J., Kaarsen, N., Olsson, O. & Selaya, P. Roman roads to prosperity: persistence and non-persistence of public infrastructure. *J. Comp. Econ.* **50**, 896–916 (2022).
26. Flückiger, M., Hornung, E., Larch, M., Ludwig, M. & Mees, A. Roman transport network connectivity and economic integration. *Rev. Econ. Stud.* **89**, 774–810 (2022).
27. De Benedictis, L., Licio, V. & Pinna, A. M. From the historical Roman road network to modern infrastructure in Italy. *J. Re. Sci.* **63**, 1162–1191 (2023).
28. Lewis, J. Modelling past land routes in archaeology: a 'Roman Roads' theoretical and methodological rethinking. <https://doi.org/10.17863/CAM.109283>. (Apollo - University of Cambridge Repository, 2024).
29. Davies, H. E. H. Designing roman roads. *Britannia* **29**, 1 (1998).
30. Laurence, R. *The Roads of Roman Italy: Mobility and Cultural Change* (Routledge, 1999).
31. Lewis, M. J. T. *Surveying Instruments of Greece and Rome* (Cambridge University Press, 2001).
32. Margary, I. D. *Roman Roads in Britain* (Baker, 1973).
33. De Soto, P. et al. Itiner-e: a high-resolution dataset of roads of the Roman Empire. *Sci. Data* **12**, 1731 (2025).
34. Wilson, J. P., Gallant, J. C. *Terrain Analysis: Principles and Applications* (Wiley, 2000).
35. Mollo Mezzena, R. La strada romana in Valle d'Aosta: procedimenti tecnici e costruttivi. in *La tecnica stradale romana* (eds Quilici Gigli, S. & Quilici, L.) 57–72 (L'Erma di Bretschneider, 1999).
36. Adams, C. E. P. *Land Transport in Roman Egypt: A Study of Economics and Administration in a Roman Province* (Oxford University Press 2007).
37. Strobel, K. Vom Marginalen Grenzraum zum Kernraum Europas. Das römische Heer als Motor der Neustrukturierung historischer Landschaften und Wirtschaftsräume. in *The Impact of the Roman Army (200 B.C. - A. D. 476): Economic, Social, Political, Religious and Cultural Aspects* (eds de Blois, L., Lo Cascio, E., Hekser, O. & de Kleijn, G.) Vol. 6 207–237 (Brill, 2007).
38. Girvan, M. & Newman, M. E. J. Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA* **99**, 7821–7826 (2002).
39. Belke, K. *Tabula Imperii Byzantini 13: Bithynien Und Hellespont* (Verlag der Österreichischen Akademie der Wissenschaften, 2020).
40. Quilici, L. *Le Strade. Viabilità Tra Roma e Lazio* (Quasar, 1991).
41. Busana, M. S. *Via per Montes Excisa. Strade in Galleria e Passaggi Sotterranei Nell'Italia Romana. Il Sottosuolo Nel Mondo Antico* (L'Erma di Bretschneider, 1997).
42. Llobera, M. & Sluckin, T. J. Zigzagging: theoretical insights on climbing strategies. *J. Theor. Biol.* **249**, 206–217 (2007).
43. Bosker, M. & Buringh, E. City seeds: geography and the origins of the European city system. *J. Urban Econ.* **98**, 139–157 (2017).
44. Bosker, M., Brakman, S., Garretsen, H., De Jong, H. & Schramm, M. Ports, plagues and politics: explaining Italian city growth 1300–1861. *Eur. Rev. Econ. Hist.* **12**, 97–131 (2008).
45. Bosker, M., Buringh, E. & van Zanden, J. L. From Baghdad to London: Unraveling urban development in Europe, the Middle East, and North Africa, 800–1800. *Rev. Econ. Stat.* **95**, 1418–1437 (2013).
46. Michaels, G. & Rauch, F. Resetting the Urban Network: 117–2012. *Econ. J.* **128**, 378–412 (2018).
47. Bulliet, R. W. *The Camel and the Wheel* (Harvard University Press, 1975).
48. Dilke, O. A. W. *Greek and Roman Maps* (Cornell University Press, 1985).
49. Talbert, R. J. A., Elliot, T. & Steinman, M. *Rome's World: The Peutinger Map Reconsidered.* (Cambridge University Press, 2010).
50. Isaac, B. *The Limits of Empire: The Roman Army in the East* (Clarendon Press, 2004).
51. Wheeler, E. L. Methodological Limits and the Mirage of Roman Strategy: Part I. *J. Mil. Hist.* **57**, 7 (1993).
52. Wheeler, E. L. Methodological limits and the mirage of Roman Strategy: Part II. *J. Mil. Hist.* **57**, 215 (1993).
53. Cioffi, R. L. Travel in the Roman World. in *Oxford Handbook Topics in Classical Studies* (ed. Oxford Handbooks Editorial Board). <https://doi.org/10.1093/oxfordhpb/9780199935390.013.110>. (Oxford University Press, 2016).
54. Scheidel, W. & Meeks, E. ORBIS: The stanford geospatial network model of the Roman World. <https://orbis.stanford.edu> (2013).

55. Ancient World Mapping Center. Ancient World Mapping Center/ geodata. <https://github.com/AWMC/geodata> (2024).
56. ESRI. World Roads. ArcGIS Online <https://www.arcgis.com/home/item.html?id=83535020ce154bd5a498957c159e3a99> (2011).
57. Bagnall, R., Talbert, R. J. A. & Elliott, T. *Pleiades: a Gazetteer of Past Places*. <https://pleiades.stoa.org/>.
58. Hanson, J. W. & Ortman, S. G. A systematic method for estimating the populations of Greek and Roman settlements. *J. Rom. archaeol.* **30**, 301–324 (2017).
59. Sims, K. et al. LandScan Global 2022. Oak Ridge National Laboratory <https://doi.org/10.48690/1529167> (2023).
60. ESRI. Kernel Density (Spatial Analyst). *ArcGIS Pro Tool Reference* <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/kernel-density.htm> (2024).
61. ESRI. Add Surface Information (Spatial Analyst). *ArcGIS Pro Tool Reference* <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/add-surface-information.htm> (2024).
62. Jenness, J. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v. 1.2. Jenness Enterprises <http://www.jennessent.com/arcview/tpi.htm> (2006).
63. De Reu, J. et al. Application of the topographic position index to heterogeneous landscapes. *Geomorphology* **186**, 39–49 (2013).
64. Dilts, T. Topography Toolbox Pro. <https://hub.arcgis.com/content/247f56c7ff48229c9b1fe132d1b5e9/about> (2023).
65. Dilts, T. & Yang, J. Stream Gradient and Sinuosity Toolbox for ArcGIS 10.1. <https://hub.arcgis.com/content/c8eb4ce1384e45258ccba1b33cd4e3cb/about> (2015).
66. ESRI. Linear Directional Mean (Spatial Statistics). *ArcGIS Pro Tool Reference* <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/linear-directional-mean.htm> (2024).
67. Snethlage, M. A. et al. *GMBA Mountain Inventory v2: a hierarchical inventory of the world's mountains for global comparative mountain science*. 322mb GMBA-EarthEnv <https://doi.org/10.48601/EARTHENV-T9K2-1407> (2021).
68. Barthelémy, M. Simple Measures. in *Spatial Networks* 39–64. https://doi.org/10.1007/978-3-030-94106-2_4. (Springer International Publishing, 2022).
69. Karsky, K. J. Structure of transportation networks: relationships between network geometry and regional characteristics. (University of Chicago/Department of Geography, 1963).
70. Buhl, C. et al. Topological patterns in street networks of self-organized urban settlements. *Eur. Phys. J. B* **49**, 513–522 (2006).
71. Csárdi, G. et al. *igraph for R: R interface of the igraph library for graph theory and network analysis*. Zenodo <https://doi.org/10.5281/ZENODO.7682609> (2025).
72. Tobler, W. Three presentations on geographical analysis and modeling: non- isotropic geographic modeling; speculations on the geometry of geography. *NCGIA Tech. Rep.* **93**, 1–25 (1993).
73. Costenbader, E. & Valente, T. W. The stability of centrality measures when networks are sampled. *Soc. Netw.* **25**, 283–307 (2003).
74. Brughmans, T. & Peeples, M. A. *Network Science in Archaeology*. <https://doi.org/10.1017/9781009170659>. (Cambridge University Press, 2023).
75. ESRI. Spatial Autocorrelation (Global Moran's I) (Spatial Statistics). *ArcGIS Pro Tool Reference* <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/spatial-autocorrelation.htm> (2024).
76. ESRI. Geographically Weighted Regression (GWR) (Spatial Statistics). *ArcGIS Pro Tool Reference* <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/geographicallyweightedregression.htm> (2024).
77. de Soto, P. et al. A high-resolution dataset of roads of the Roman Empire: Itiner-e static version 2024. Zenodo <https://doi.org/10.5281/zenodo.17122148> (2025).
78. Brughmans, T., De Soto, P., Pažout, A. & Bjerregaard Vahlstrup, P. *Itiner-e: the digital atlas of ancient roads*. <https://itiner-e.org/> (2024).
79. Pažout, A. et al. Network Science reveals the structure and lasting impact of the Roman road system: Supplementary datasets. Zenodo <https://doi.org/10.5281/ZENODO.18610056> (2025).
80. Hanson, J. W. *Cities database*. <https://doi.org/10.5287/BODLEIAN:EQAPEVAN8> (University of Oxford, 2017).
81. Hanson, J. W. *An Urban Geography of the Roman World, 100 B.C. to A.D. 300*. (Archaeopress, 2016).
82. Mazzamurro, M. *MatteoMazzamurro/roman-road-networks: R project for Network Science reveals the structure and lasting impact of the Roman road system - version 1.0.0*. Zenodo <https://doi.org/10.5281/ZENODO.20446953> (2026).

Acknowledgements

We thank the authors of de Soto et al. (2025) for contributing to the Itiner-e dataset of roads of the Roman Empire, and all Itiner-e and Viabundus workshop participants for helpful feedback on the data and its analysis. For support and feedback we thank Aarhus University's Past Networks team, Social Resilience Lab, Mie Egelund Lind, CORE, Antikmuseet, AU students in Classical Archaeology, Troels M. Kristensen, Kamil Kopij. This work was supported by Danmarks Frie Forskningsfond (DFF) Sapere Aude research leadership grant (0163-00060B) awarded to T.B. for project MINERVA; The Carlsberg Foundation Semper Ardens: Accelerate Fellowship awarded to T.B. for The Past Social Networks Project (CF21-0382); Danish National Research Foundation (DNRF) Centre of Excellence for Urban Network Evolutions (UrbNet) (DNRF119); and the Viator-e project RTI2018-098905-J-I00 was funded by MCIN/AEI/10.13039/501100011033/ and by FEDER Una manera de hacer Europa awarded to P.S.

Author contributions

A.P., T.B., and P.S. conceived and designed the study; contributed equally. A.P. prepared and documented all datasets used. A.P. and M.C.S. prepared and performed linear and spatial regression. A.P. performed the topographic and density analyses, with help from E.H.M., M.L.N., P.B.V., C.F., and G.E.Ø. M.M. and T.B. conceived and designed the network analysis, and M.M. performed the network analysis. A.P., T.B., P.S., and M.M. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-75453-3>.

Correspondence and requests for materials should be addressed to Tom Brughmans.

Peer review information *Nature Communications* thanks Luca de Benedictis who co-reviewed this paper with Anna Maria Pinna, and Alessio Palmisano, Eric E. Poehler and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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/>.

© The Author(s) 2026