Infrastructures connecting people: A mechanistic model for terrestrial transportation networks

Luce Prignano
Institute of Complex Systems (UBICS), Condensed Matter Physics, Universitat de Barcelona, Barcelona, Spain; Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain

Lluc Font-Pomarol
Departament d’Enginyeria Química, Universitat Rovira i Virgili, Tarragona, Spain

Ignacio Morer
Institute of Complex Systems (UBICS), Condensed Matter Physics, Universitat de Barcelona, Barcelona, Spain; Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain

Sergi Lozano
Institute of Complex Systems (UBICS), Condensed Matter Physics, Universitat de Barcelona, Barcelona, Spain; Departament d’Història Econòmica, Institucions, Política i Economia Mundial, Universitat de Barcelona, Barcelona, Spain

Abstract
Terrestrial Transportation Infrastructures (TTIs) are shaped by both socio-political and geographical factors, hence encoding crucial information about how resources and power are distributed through a territory. Therefore, analysing the structure of pathway, railway or road networks allows us to gain a better understanding of the political and social organization of the communities that created and maintained them. Network science can provide extremely useful tools to address quantitatively this issue. Here, focussing on passengers transport, we propose a methodology to shed light on the processes and forces that moulded transportation infrastructures into their current configuration, without having to rely on any additional information besides the topology of the network and the distribution of the population. Our approach is based on a simple mechanistic model that implements a wide spectrum of decision-making mechanisms (representing different power distributions) which could have driven the growth of a TTI. Thus, by adjusting a few model parameters, it is possible to generate several synthetic transportation networks, and compare across them and against the empirical system under study. An illustrative case study (i.e. the railway system in Catalonia, a region in Spain) is also provided to showcase the application of the
proposed methodology. Our preliminary results highlight the potential of our approach, thus calling for further research.

Introduction

Historically, transportation infrastructures have had a huge impact over the development of territories. They allow the movement of people and goods and support the access to key resources, thus affecting their long-term capability to cope with changing socio-economic scenarios. Terrestrial transportation infrastructures (hereafter referred to as TTIs) differ from other transportation systems mainly in that their connections are physical.

This particular feature opens the door to the exploitation of TTIs as a fundamental source of information about the societies that created and maintained them. Indeed, only in the case of terrestrial infrastructures, the decision of connecting two previously disconnected places is a crucial, not easily reversible, one. Consequently, the balance between the cost of building a new connection and the provided benefit to the system can be regarded as the result of multiple conflicting interests. As a result, the structure of TTIs reflects not only the needs they are supposed to satisfy, but also the distribution of resources and power across a territory (Herranz-Loncán & Fourie, 2017; Pablo-Martí et al., 2021).

Here, we present an analytical framework allowing for a better comprehension of power interactions involved in the unfolding of a regional TTI, from the study of its structural features. While we agree with the prevailing hypothesis that the core mechanism behind the evolution of TTIs involves balancing costs and benefits, our approach diverges from others that have been formulated to reveal universal traits shared by all TTIs on a global scale (see (Louf et al., 2013) for railways, (Louf et al., 2014) for railways and subways and (Black, 1993) for a more general perspective). Contrary to other approaches, we do not assume that all actors involved are rational. Instead, our focus lies in examining the outcomes resulting from these actors having varied priorities and inconsistent opportunities to achieve satisfaction.

To this aim, we generalized to present-day TTIs a methodology previously proposed for an archaeological case study. In the context of archaeological research, the importance of TTIs for the understanding of the political and social organization of the societies that created and maintained them was initially assessed in relation to the Roman Empire, and more recently in a number of studies in the New World and in Pre-Roman Europe (see, for instance, (Raymond, 1976; Taylor, 1979; Trombold, 1991; Jenkins, 2001; Smith, 2005; Groenhuijzen and Verhagen, 2016; Verhagen et al., 2019)).

Building on this literature, a methodology was proposed to infer the relationship among political entities from the structure of transportation networks in some proto-historical case studies (Prignano et al., 2019; Fulminante et al., 2017, 2021). In contrast with usual Spatial Interacting modelling (Fischer and Reggiani, 2004; Oshan, 2021), that approach was not based on the identification of optimal transportation structures based, to some extent, on entropy maximization (Wilson, 1971). On the contrary, it developed an experimental setting to test different power balances and compare TTIs structures hypothetically generated by each of them. This way, it provided a tool to exploit the relationship between regional power interactions and transportation networks’ complexity (to be understood as defined in (Reggiani, 2021)).

Here, we present a generalized version of such a methodology focusing on passengers transport (PTTIs). In particular, this generalized approach allows to address questions like to what extent an infrastructure was conceived as an independent system rather than as an auxiliary network that complements another; or whether a particular PTTI was shaped by the population spatial distribution.
The rest of the article is organized in five sections. We begin by summarizing the original methodology, which serves as the foundation for our research. This is followed by two sections that detail how this methodology was adapted to fit current-time scenarios. The first of these two sections focuses the processing of empirical data of present PTTIs, discussing their differences with ‘archaeological PTTIs’ in terms of the richness of available details and the necessity to filter relevant information. In the second one afterwards, we introduce the generalized version of our modelling approach that incorporates the impact of pre-existing conditions and population distribution, elements that were absent in the original methodology. In particular how current case studies are richer in details and the various ways this wealth of information could be processed.

In the subsequent section, we apply our methodology to an illustrative case study: the Catalan railway system, a regional PTTI with several decades of history that coexists with a road system. Special attention is paid to the assessment of models’ performance and to drawing conclusions properly from the comparison of empirical and synthetic data. Finally, the article closes with some concluding remarks. Our preliminary results highlight the potential of our approach, thus calling for further research.

**Original methodology**

Our starting point is a collection of works tackling the issue of quantitatively inferring aspects of the political organization of a region from the structure of transportation networks (Prignano et al., 2019; Fulminante et al., 2017, 2021). Their goal was to identify the nature of the interplay between different human communities, going beyond tasks usually addressed by archaeological quantitative analysis, such as establishing the existence of a certain degree of regional organization. To this aim, the authors addressed the analysis of the decision-making processes prioritizing some paths over others by envisioning a methodology consisting of two fundamental ingredients:

(i) a procedure for extracting relevant quantitative data from road maps; (ii) formal models implementing alternative mechanisms for generating synthetic TTIs to be compared against the empirical ones.

The underlying idea is that some models reproduce relevant structural features of the empirical TTI with higher accuracy than others. In other words, they provide a higher quality explanation of the empirical evidence and, therefore, we assume them to be more likely to resemble the actual mechanisms of regional organization (Prignano et al., 2019).

Network science was adopted as a natural framework to address the interplay between connectivity and functionality of TTIs. Indeed, network science provides both with tools to identify and measure structural characteristics of empirical TTIs and with a conceptual framework for formal model building (Boccaletti et al., 2006; Newman, 2003; Costa et al., 2011).

The task of translating road maps into networks is not straightforward and can be performed in many alternative, not equivalent ways. Since they were studying inter-settlement interactions, they needed our nodes to represent the human communities connected through the regional TTI. Then, as the simplest possible option, they established a bidirectional link between any two sites that were directly connected by a terrestrial route, with no other settlement in between. To include the geographical factor in a simple way, they represented sites as geo-localized nodes and assigned weights to the links according to the geodesic distance between the nodes they connected.

They designed a minimalist set up in which each node, at each step, expressed a preference concerning the new link to be established, according to a variably well informed assessment of costs and benefits. Then, they could either compete against each other or reach an agreement about which connection was going to be built next.
The goal of those works was not to understand why in a certain region there were more or less settlements, or more or less roads. On the contrary, they were addressing the question of why and how the settlements that existed in the region built those roads instead of others. Consequently, the models took the set of settlements with their corresponding geographic locations and the amount of available resources – here quantified as the total link length $L_{tot}$ – as inputs, not as parameters to be fitted.

The process ended when the total length of the connections added was equal to the total link length of the corresponding empirical network. Consequently, any synthetic graph generated by a network model replicated the following characteristics of the corresponding empirical network: (1) the total number of nodes $N$, (2) its geographic density $\delta = L_{tot}/\sum_{i,j=1}^{N} d_{ij}$ and, (3) the average node strength $s_i = 1/N \sum_{j=1}^{N} \sum_{j \in V(i)} l_{ij}$, where $V(i)$ is the set of neighbors of $i$ and $l_{ij}$ is the length of the link between $i$ and its neighbor $j$.

Any other metric is, in principle, suitable to be used for comparing different synthetically generated graphs against their empirical counterpart.

**The equitable efficiency model (EE model)**

Out of all the models devised in those works, one presented a higher explanatory power when applied to the proto-historical case studies (i.e. it generated the synthetic networks more similar to the corresponding empirical ones). It equipped nodes with global information about their connectivity, but also with the ability to make coordinated decisions. More concretely, it assumed that each settlement knew the (weighted) length of each one of the existing path joining it with any other settlement. Then, links were prioritized globally according to their normalized closeness $R$ that was calculated as follows:

$$R_i(j) = R_j(i) = R_{ij} = \frac{d_{ij}}{L_{ij}},$$

where $d_{ij}$ is the geodesic distance between node $i$ and node $j$, and $L_{ij}$ is the length of the shortest existing path between them. The normalized closeness $R$ can also be seen as the inverse of what is called the route factor or detour index (Barthélemy, 2011). If $i$ and $j$ are disconnected, that is, belong to different connected component, this means that there exist no finite length path between them. Hence

$$R_{ij} = \lim_{L_{ij} \to \infty} \frac{d_{ij}}{L_{ij}} = 0$$

and the comparison between disconnected node pairs is performed by determining

$$\text{sign} \left( R_{ij} - R_{lm} \right) = \text{sign} \left( \lim_{L_{ij} \to \infty} \frac{d_{ij} - d_{lm}}{L_{ij}} \right) = \text{sign} (d_{ij} - d_{lm})$$

which is obtained assuming that the path length $L_{ij}$ is the same for all the disconnected nodes.

The function $R_i(j)$ balances costs and benefits, prioritizing those links that shorten long paths (large $L_{ij}$) while wasting little resources (short $d_{ij}$). Notice that here, following most of the literature on the topic (Gastner and Newman, 2006a; Gastner and Newman, 2006b; Louf et al., 2013), we are interpreting link costs in an economic sense (building and maintenance costs), thus assuming that they are directly proportional to their length (i.e. distances between connected nodes). Benefits, however, have a pure topological nature: Each node-agent wants to have as many direct connections as possible to avoid long detoured routes.

More concretely, the EE model follows a three-step procedure:
1. For each node $i$, all the $R_i(j)$ values are calculated.
2. Each node $i$ proposes the creation of a link between itself and a node $j^*$ such that the $R_i(j^*)$ was the minimum value among all the $R_i(j)$ (local interest expressed by node $i$).
3. All the proposals are ranked according to their $R$ value and a link is created between the pair corresponding to the global minimum (coordinated decision-making).

Step 1, 2 and 3 are repeated until the summed lengths of all created links reaches that of the empirical system. Nonetheless, the metaphor of the individual priorities that have to be sorted is useful for devising meaningful generalization of the present baseline model.

**EE model with preferential attachment**

The EE model considered all settlements to be on the same ground and the links to be built were selected among the individual preferences according to a fair criterion. In order to explore a slightly different scenario, where preferences of nodes with more and longer links were entitled to a higher priority level, a variant of the EE model was devised. Mathematically, such a bias is obtained by weighting the ratio $R$ with a (negative) power of the strength (or weighted degree, that is, the total length of its adjacent links) of the proposing node, thus introducing preferential attachment among the nodes with the greatest strength. The trade-off between the two ingredients in determining the priority of each link is tuned by the exponent $a$ of such power. Hence, the new value of the biased ratio $R'$ for a connection between node $i$ and $j$ proposed by node $i$ is

$$R'_{ij} = \frac{R_{ij}s_i^{-a}}{L_{ij}s_i^{-a}}$$  \hspace{1cm} (4)

where $s_i$ is the strength of node $i$. Therefore, when $a$ is equal to zero, we recover the EE Model.

It is worth noting that in all the scenarios considered, including the last one, we assumed that all node-settlements were intrinsically equally important. When prioritising new links to be established, the node-settlements based their choice on geographical (distances) and topological (already existing links) information, but on node-settlement attributes (such as power, richness or attractiveness).

**From maps to accessibility networks**

In the case studies previously considered, all of them from the Iron Age, the region under study was provided of a single PTTI that embedded the footprint of the relationships between the human communities living in the area. When human societies started building roads (Earle, 1991; Lay, 1992), they created, in each territory, a network made of a combination of artificially adapted natural paths and manufactured ways that served for displacements at all the scales – from local to supra-regional – and for all means of transport (pedestrian, by wheeled vehicles, with animals). Then, each settlement had to be connected to others by means of this single PTTI, or it would be otherwise isolated. This is not the case for later scenarios. Already in the Roman Empire, the most important cities were connected through primary roads, while less important towns and villages were reachable thanks to a dense net of secondary roads and less manufactured pathways (see, for instance, (Carreras and De Soto, 2013)). Such a difference plays a crucial role both in the way the empirical network is constructed and in the modelling approach.

Considering their high construction costs, some PTTIs, such as railway systems or highways, are not designed to directly reach each single town and village in a territory. Instead, many human settlements benefit indirectly from the presence of a station or exit nearby.
Since one of our objectives is to capture the effect of population distribution over the territory on PTTI structural patterns, the option to discard some groups of inhabitants based on whether their residence place is reached by the infrastructure under study or not must be ruled out.

We had therefore to disregard some of the most usual network representations, such as the so-called space-of-stations for railway systems, in which nodes are stations and links represent physical connections (Kurant & Thiran, Sep 2006), as well as the option – widely used for urban networks – of building a graph whose nodes represent road intersections and endpoints (Barthélemy & Flammini, 2008).

As an alternative, we sought for a method to integrate information about both population distribution and PTTI’s connectivity in a single empirically based graph (i.e. a PTTI accessibility network). Specifically, we developed a procedure to assign fractions of the population to each train station, highway exit, etc. (hereafter, referred as station for the sake of simplicity) by ‘merging’ neighbouring localities, while redefining both their coordinates and connectivity.

Our merging procedure grasps information about population distribution from the most fundamental units (i.e. in general and hereafter, the municipalities) and their geographical positions. Initially, each municipality (with or without direct train connection) is represented as a node. The process takes into account population sizes and distance between municipalities to, iteratively, joining two nodes into one. In that way, low populated nodes are likely to be merged with closer, larger ones, in a location somewhere in the middle of both. Along the process, newly merged nodes preserve any connection existing in the previous step.

Formally speaking, the process is represented by an algorithm that gets as input information about the initial set of nodes $i$ (geographic coordinates and population $P_i$) and the links among them, and executes the following steps:

1. It creates a list with all possible pairs of nodes $i, j$ computing its inter-distances $d_{ij}$. The list is ranked in ascendant order according to the distances.
2. For the first element in the list, it calculates $\Gamma_{ij} = P_{\min}^* d_{ij}(10^6 \text{inhabitants} \cdot 1 \text{km})$, where $P_{\min} = \min(P_i, P_j)$. This quantity expresses the balance between the importance of the smallest node and the distance that separates it from the largest one.
3. If the condition $\Gamma_{ij} < \Gamma^*$ is fulfilled, both nodes merge into one. The new node will have a population $P' = P_i + P_j$ and a new position $(\text{Lat}'$, $\text{Lon}')$ at the centre of mass of $i$ and $j$ according to their former populations:

   \[
   \text{Lon}' = \frac{P_i \text{Lon}_i + P_j \text{Lon}_j}{P_i + P_j}, \quad \text{Lat}' = \frac{P_i \text{Lat}_i + P_j \text{Lat}_j}{P_i + P_j}
   \]

4. If there existed a link between the two merged nodes, the link disappears. If there were links adjacent to $i$ or $j$, they are rewired correspondingly to the new node. Notice that those links would also change their length since the position of the merged node is now somewhere between former $i$ and $j$ positions.
5. The four previous steps are repeated until no more pairs of nodes are able to merge.

The parameter $\Gamma^*$ fixes how lowly populated a node has to be and how close to another one more populated to merge them. It acts as a restrictive merging condition, being more permissive to nodes’ merging as it increases its value. By means of the parameter $\Gamma^*$, we are considering the finite capacity of a station, allowing for a higher node density in highly populated areas, a necessary feature of almost any PTTI (see Figure 1 for a schematic example).
This method presents several advantages. It is independent on the territory and it is able to capture potential heterogeneity across a region (e.g. rural and urban). More importantly, this method is able to define the scale of the system that suits our goals. In this sense, it is more flexible and less arbitrary than allocating nodes based, for instance, on administrative divisions of the territory (e.g. provinces, counties…) and allows for an easier comparison between case studies.

Although different solutions are possible, in most of the cases it is sensible to set the value of $\Gamma^*$ so that in the output the number of links matches, or slightly exceeds, the number of nodes. In principle, this condition allows to obtain a single connected component. Hence, isolates – if any – can be interpreted as representative of the population and geographical position of less favoured communities in terms of accessibility to the PTTI under study. Selecting a lower $\Gamma^*$ would make the number of nodes to be greater than the number of links, thus forcing some nodes to stay disconnected and making this feature less meaningful. On the contrary, a considerably higher $\Gamma^*$ might hide the existence of poorly connected sub-regions by forcing further node merging. In this sense, the maximum distance between any merged municipalities should not be larger than the distance that one can reasonably assume, in a given context, that people from a smaller village would cover to reach a near town connected to the TTI. Hence, the most appropriate choice for the value of $\Gamma^*$ is intrinsically related to the specificity of the system under study.

**Generalized modelling approach**

The main difference between proto-historical scenarios and more recent case studies is the relevance of the context pre-existing the construction of the infrastructure under study. The first steps of the algorithms of our baseline models consist in building links between isolated nodes that are
completely equal, except for their geographical coordinates. The underlying hypotheses is that differences in power and importance did not preexisted the creation of the very first transportation infrastructure, before which there were just a number of (almost) disconnected settlements. However, if we want to model a PTTI that is not the first ever built communication network in the region, this supposition does not hold. In particular, we must take into account (1) the coexistence of more PTTIs, which implies that there exist alternative ways to reach each location and (2) the uneven level of agency throughout the different communities in the considered territory. Hence, nodes in the initial state of the models are neither equal, nor disconnected. Here, we address such issues keeping the algorithms as well as the input data as simple as possible.

**Generalized equitable efficiency model**

Taking equations (4) and (2) as a starting point, we introduce two modifications and define a new, generalized version of the EE model: The *generalized Equitable Efficiency Model* (hereafter referred as gEEM). We are interested in including the preferential attachment mechanism, since it proved to be useful when handling unbalanced settlements in terms of power (Fulminante et al., 2017). However, instead of measuring the relative relevance across nodes using the node strength (as in Sec. 2.2), gEEM uses an attribute of nodes that is now assumed to be a feature previous to the construction of the PTTI. Since our interest in this paper is focused on population distribution, here we are taking nodes’ population. Notice, however, that this modelling approach could be applied using other attributes (such as nodes’ wealth, for instance) depending on the requirements of each particular application. Thus, a new definition for the normalized closeness is proposed:

\[
R'_{ij} = \frac{d_{ij}(p_i)}{L_{ij}(p_i)} = \frac{P_i}{P_{min}},
\]

where \(P_i\) is the population (or any other desired attribute) of node \(i\) and \(P_{min}\) is the smallest population of all nodes.

On the other hand, we redefined the limit taken to evaluate the shortest path between two existing nodes, which the previous models took as \(L_{ij} \to \infty\). We propose to make it finite defining it as

\[
L_{ij}^{[d]} = e^{-\eta L_{CG}}
\]

where \(i, j\) are nodes in different connected components and \(L_{CG}\) is the total link length of a complete graph built on the same set of geolocated nodes. By redefining this limit, our purpose is to reproduce the overall effect that other infrastructures have in shaping the railway network. Considering that between two disconnected nodes there is a finite existing path, we make them reachable. Thus, this trait allows to consider them not so primordial for the network and to open the possibility to construct other links beforehand (between already connected nodes, for instance). This simple addition to the baseline model is suitable to effectively take into account the existence of denser but slower infrastructures (e.g. secondary roads) when studying a sparser but faster one (e.g. railways or highways). Conversely, it would not be useful if we wanted to account for high-speed shortcuts on an otherwise slow TTI.

**Network characterization by structural metrics**

In order to better characterize the empirical network resulting from the merging process and compare it with synthetic counterparts later on, we calculated several network properties that
provide information about both the structure of the network and its influence on communication
dynamics:

**Clustering coefficient.** It describes the tendency of nodes to form dense groups (clusters). Given nodes
$i$, $C_i$ indicate the proportion of all potential links between the neighbours of $i$ that actually exist:

$$C_i = \frac{m_i}{k_i(k_i - 1)/2}$$

where $m_i$ is the number of links connecting two neighbours of $i$, thus closing a triangle containing $i$. Averaging this ratio over the set of nodes, the average clustering coefficient $\langle C \rangle$ constitute an indicator of the density of closed triangles in the network:

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^{N} C_i$$

**Average link length.** It is a useful metric for a basic characterization of the links in a weighted spatial
network:

$$\langle \ell_i \rangle = \frac{1}{N} \sum_{i=1}^{N} \langle \ell_i \rangle_i$$

where $M$ is the number of edges.

Moreover, in order to get information about link length variability, we computed the standard
development of this metric $\sigma_{\ell_e}$.

**Standard deviation of node strength.** The average value of the node strength $\langle s \rangle$ is, by construction, the
same for the empirical network as for any synthetically generated counterpart, but its standard
development $\sigma_s$ can be used as an indicator of how balanced the node connectivity is.

On the other hand, we found reasonable to assess the efficiency of the empirical network in terms of its main functionality, namely, the interchange of goods, information and people. Several ways to evaluate such property have been proposed (Latora & Marchiori, 2001; Vragović et al., 2005). From the different definitions in this literature, we took one proposed specifically for weighted networks in (Vragović et al., 2005). This definition compares the existing shortest path between two nodes, $L_{ij}$, with the geodesic distance that separates them, $d_{ij}$, assuming that the information flows better when the first value approximates the second. For a single connection, this definition coincides with our normalized closeness $R$ and, as previously mentioned, can be seen as the inverse of what is called the route factor or detour index (Barthélemy, 2011). According to this approach, we calculated the global and local efficiencies of the network as follows:

**Global efficiency.** For each pair of nodes $i$ and $j$, it compares the path connecting them through the
network ($L_{ij}$) with the ideal case corresponding to the straight line ($d_{ij}$). Then, the Global efficiency is
obtained by averaging this ratio over all pairs of nodes.

$$E_{\text{glob}} = \frac{1}{N(N - 1)} \sum_{i \neq j}^{N} \frac{d_{ij}}{L_{ij}}$$
Local efficiency. For each node $i$, it evaluates how well information is exchanged between its first neighbours when $i$ is missing. In other words, it assesses the robustness of the network and how it is able to deal with failures (for instance, when the communication through a specific node is not possible).

$$E_{loc} = \frac{1}{N} \sum_{i=0}^{N} k_i (k_i - 1) \sum_{j \neq k \in \Gamma_i} \frac{d_{jk}}{L_{jk/i}}$$

where $j, k$ belong to the subgraph of first neighbors of $i$, $\Gamma_i$, and $L_{jk/i}$ is the shortest path connecting $j$ and $k$ when the node $i$ is removed.

Finally, we defined a generalized version of the Global efficiency that takes into account the number of inhabitants associated to each node.

Passenger efficiency. For each pair of passengers $p$ and $q$ that do not belong to the same node, it compares the path connecting them through the network ($L_{pq}$) with the ideal case corresponding to the straight line ($d_{pq}$). Then, the Passenger efficiency is obtained by averaging this ratio over all pairs of passenger.

$$E_{pass} = \frac{1}{N_F} \sum_{p \neq q} \frac{d_{pq}}{L_{pq}} = \frac{1}{N_F} \sum_{i \neq j} P_i P_j \frac{d_{ij}}{L_{ij}}$$

where $N_F = P_{tot}^2 - \sum_i P_i^2 = N(NP^2 - \bar{P}^2)$ is the total number of pairs of individuals discarding those associated to the same node, being $P_{tot}$, $P$ and $\bar{P}$ the total population of the system, the mean node population and the mean square node population, respectively. The Passenger efficiency is equal to the Global efficiency when all the nodes have the same population. More in general, it is larger than the Global efficiency when the population is concentrated on nodes whose average normalized closeness $R_i = (1/N)\sum_{j \neq i} \frac{d_{ij}}{L_{ij}}$ is larger compared to the rest of the nodes in the system and, vice versa, it is smaller when the most populated nodes are poorly connected. As an illustrative example, consider a system a with total population $P$ distributed on three nodes connected by two links, such that the central node, which is connected to the other two, has $P_c = cP$ inhabitants while the others both have $P_p = \frac{1-c}{2} P$. If we denoted as $R_p < 1$ the normalized closeness of the two peripheral nodes that are not connected to each other, then the Global efficiency of the system can be expressed as $E_{glob} = (2 + R_p)/3$. The Passenger efficiency depends on the parameter $c$ and is given by $E_{pass} = [R_p(4 - 4c^2 + (4 - 2R_p)c + R_p)]/(1 + 2c - 3c^2)$. Hence, it can be easily shown that for any value of $R_p < 1$ we have that $E_{pass} > E_{glob}$ whenever $c > 1/3$.

Application of our generalized methodology to a present-day case study

This section illustrates the applicability of our generalized methodology (including the merging procedure to obtain the empirical network and gEMM as the model generating synthetic counterparts) to present-day scenarios.

Empirical accessibility network from node merging

The specific territory used in this study to illustrate the applicability of our generalized methodology was Catalonia, a Mediterranean region in northeast Spain. This region is densely populated and presents a heavy unbalance between rural and urban areas, with more than 70% of its population living in cities. It also has a relevant industrial sector and diverse and developed transportation infrastructures, among them, the railway network. Such infrastructure can be divided into high-speed railroads and regional...
services. The former is actually part of a larger-scale structure, which has been framed according to a
country-wide perspective (i.e. the Spanish high-speed rail service AVE). Therefore, here we consider
only the regional service, which is, composed of two structures managed by two different operators
(ADIF and FGC) owned by the Spanish and Catalan regional governments, respectively.

In Figure 2 (left), we see Catalonia’s regional railway networks, together with the municipalities
served by them. As already pointed out in the section regarding the processing of empirical data of
present PTTIs, many localities in Catalonia (and their corresponding share of the population) are not
directly served by the railway networks. This justifies obtaining a train accessibility network as our
empirical network instead of a physical railway one (i.e. the space-of-stations representation).

In this particular case, we chose to set the threshold parameter $\Gamma^*$ so that it is potentially possible
for the resulting total link length to connect all the nodes in a single connected component. Starting
from an initial state in which nodes outnumber links 945 to 205, the merging process progressively
reduces the number of nodes, with some sporadic link losses as a collateral consequence.
Meanwhile, the total link length of a hypothetical minimum spanning tree built on the same set of
geolocated nodes also decreases, and does it faster than the total link length. Depending on the value
of $\Gamma$, the resulting final network has different number of nodes $N(\Gamma)$ and links $M(\Gamma)$, and different
total link length $L_{\text{tot}}(\Gamma)$ and minimum spanning tree length $L_{\text{MST}}(\Gamma)$ (see Figure 3, left). In general,
the larger the value of $\Gamma$, the smaller the ratios $Q_{NM}(\Gamma) = N(\Gamma)/M(\Gamma)$ and $Q_{LL}(\Gamma) = L_{\text{MST}}(\Gamma)/L_{\text{tot}}(\Gamma)$,
until they both reach a plateau. The metrics assessing for the connectivity of the network change
accordingly, following a generally increasing trend that gets more and more irregular when $N$ is of
the order of few tens of nodes or less (see Figure 3, right).

Once the condition $Q_{LL} > 1$ is satisfied, the EE model always builds a connected network, and
actually if the ratio is equal to 1 it produces the MST. However, in order to allow the present
generalized version of the model to generate connected networks different from the MST, we chose
a value $\Gamma^*$ such that the additional condition $Q_{NM} > 1$ (i.e. the number of links in the resulting
empirical network exceeds the number of nodes) is also satisfied.

Figure 2. Empirical network. Left: Municipalities ($N = 945$) and train connections ($M = 205$) before the
merging procedure. The initial total link length is $L^0_{\text{tot}} = 1208.51\,(\text{km})$ and the total length of the
 corresponding MST is $L^0_{\text{MST}} = 3737.26\,(\text{km})$. Colours according to merging results. Right: Empirical network
after merging procedure for $\Gamma^* = 0.82$. Labels assigned according to the most populated municipality. A node
size is proportional to the total population of the municipalities inside it.
This is just a heuristic criterion, but has the clear advantage of not forcing the graph to have isolates while permitting their existence. In this way, the ‘surviving isolates’ in the empirical network represent a meaningful trait of the system under study, not just a mere consequences of a wrong scale choice. We can thus apply our methodology to investigate the reason behind network features such as isolated notes, multiple connected components and cycles (closed loops).

The empirical network is showed in Figure 2, before (left) and after (right) applying the merging procedure setting $\Gamma^* = 0.82$. The merged version has $N = 39$ nodes and $M = 44$ links, while the length of the corresponding minimum spanning tree is almost a 20% shorter than that of the actual connectivity pattern, $L_{tot} = 965.3 \text{ (km)}$.

To better grasp the role that the distribution of the population across the territory plays in shaping the merged empirical network (hereinafter, referred to as ‘empirical network’), we reshuffled the number of inhabitants of the municipalities and repeated the merging processes several times. By tuning the $\Gamma$ parameter, we generated a set of networks with $N = 39$ nodes, but their locations, the municipalities composing them, as well as their connections are, in general, different. In particular, we observed that all the randomized merged networks have almost the same total link length as the empirical network, but significantly less links and more isolates. The (individual) link length is, on average, larger and less heterogeneously distribute among connections. For what concerns the efficiency, both global and Passenger efficiency decrease drastically, as expected for topologies with so many more disconnected nodes. The Local efficiency also reduces its value in a similar proportion, although the difference is less significant, while other magnitudes, such as the average clustering coefficient or the standard deviation of node strength, do not vary significantly. In Table 1, we summarized the most significant differences between the empirical network and the networks obtained from randomized node populations.

![Figure 3](image-url)

**Figure 3.** Characterization of the output of the merging process. Left: The number of nodes $N$ and edges $M$ in the empirical network (left axis, orange, circles and diamonds, respectively) and the corresponding total link length $L_{tot}$, together with the length of the minimum spanning tree $L_{MST}$ built on the same set of geolocated nodes (right axis, purple, squares and triangles, respectively), are plotted as functions of the parameter $\Gamma$. Right: Some network metrics are plotted as function of $\Gamma$: the clustering coefficient (green diamonds); the Local efficiency (blue triangles); the Global efficiency (orange circles); the Passenger efficiency (purple squares). The vertical line indicates the value $\Gamma^* = 0.82$. 

Synthetic networks generated by the gEMM

Given a node layout \( \mathcal{L} \), that is, set of geolocated nodes, and a total link length \( L_{tot} \) such that \( L_{tot} > L_{MST}(\mathcal{L}) \), by varying the two parameter \( a \) and \( m \), the gEMM is able to produce a great variety of network topologies, ranging from star-like graphs to disconnected cliques (all-to-all sub-graphs) and almost regular lattices, which may include a variable amount of isolates and connected components of different sizes.

To assess how some basic structural properties of the model’s outputs depend on \( a \) and \( m \), we applied the gEMM to the node layout and total link length of the empirical network obtained in Sec. 5.1 varying both parameters.

Since the number of different network topologies that can be constructed on a given node layout for a fixed total link length is finite, the parameter plane \( m/a \) can be ideally partitioned into regions corresponding to an identical, unique output of the model. However, such an exhaustive exploration is beyond the scope of the present study. At a more coarse-grained scale, we observed that whenever \( m \leq 4.5 \), this parameter does not affect the output of the model, that is, the output for any \((a, m)\) is the same as for \((a, 4.5)\) for any \( a \). Something analogous happens for the region \( m \geq 10 \) where the output for any \((a, m)\) is the same as for \((a, 10)\) for any \( a \). Similarly, for \( a \geq 1 \), the model does not produce any topological novelty and any network created for \((a, m)\), with \( a > 1 \), is also generated for at least another pair \((a', m')\), where \( a' < 1 < a \) and \( m' < m \).

In particular, for \( a = 0 \) and \( m \leq 5.7 \), the output of the gEEM is the same as that of the original EE model (Figure 4(a)). The first connections to be created belong to the MST, afterwards the algorithm adds few shortcuts until the total link length equals that of the empirical network. When the value of any of the two parameters is increased, some of the links belonging to the MST start to be rewired and some nodes may be left disconnected from the largest component of the resulting network. While the principal effect of increasing the value of \( m \) (Figure 4(b)–(f)) is the appearance of closed triangles, larger values of \( a \) (Figure 4(b)–(d)) force the links to be redirected towards most populated nodes.

Comparing synthetic and empirical networks

The empirical network presents an almost vanishing clustering coefficient and a considerably high Global efficiency \( E_{glob}^{emp} = 0.732 \), just slightly lower than the EEM network (see Table 2). When comparing with the network topologies generated by the gEEM, we find that these two feature can be observed when \( 0.55 \leq m \leq 0.65 \) and \( a \leq 0.4 \) (see Figure 5(a) and (b)). If we look at other metrics such as the local efficiency \( E_{loc}^{emp} = 0.245 \), the average link length \( \langle l_e \rangle^{emp} = 21.7 \text{ km} \) and the standard deviation of the node strength \( \sigma_s^{emp} = 14.9 \text{ km} \), similar values can also be found in the same region of the parameter space.

Additionally, we computed a network distance between each synthetic topology and the empirical network based on the shortest path length between node pairs:

| \( \Gamma \) | \( M \) | \( I \) | \( Q_{LL} \) | \( \langle l_e \rangle \) | \( \sigma_s \) | \( E_{glob} \) | \( E_{loc} \) | \( E_{pass} \) |
|---|---|---|---|---|---|---|---|---|
| \( \chi \) | 1.09 | 34.2 | 8.8 | 0.96 | 27.8 | 9.4 | 0.435 | 0.162 | 0.489 |
| \( \sigma \) | 0.09 | 2.2 | 1.6 | 0.07 | 1.3 | 1.4 | 0.042 | 0.045 | 0.097 |
| \( P \) | >0.99 | <0.01 | >0.99 | 0.99 | >0.99 | <0.01 | <0.01 | 0.05 | <0.01 |
| \( e \) | 0.82 | 44 | 2 | 0.82 | 21.7 | 14.9 | 0.732 | 0.245 | 0.834 |

Expected value (\( \chi \)), standard deviation (\( \sigma \)) and the corresponding p-value (\( P \)) of some metrics that differ significantly between a set of 100 networks obtained from randomized node populations and the empirical network, whose values are reported in line e. “I” stands for the number of isolates. All metrics have no dimension except for \( \langle l_e \rangle \) and \( \sigma_s \), which are expressed in Km.
Figure 4. Artificial networks for different pairs of values for the parameters $a$ and $m$. Edges’ colour illustrates the order of construction of the links by the algorithm. Darker colours stand for earlier edges while brighter stand for later ones.

Table 2. Metrics values for some examples of synthetic networks.

| $a$ | $m$ | $\langle C \rangle$ | $\langle I \rangle_c$ | $\sigma_L$ | $\sigma_s$ | $I$ | $E_{\text{glob}}$ | $E_{\text{loc}}$ | $E_{\text{pass}}$ |
|-----|-----|---------------------|---------------------|-----------|-----------|-----|----------------|----------------|----------------|
| 0   | ≤5.7| 0                   | 22.1                | 13.5      | 30.5      | 0   | 0.800          | 0.192          | 0.830          |
| 0   | ≥9.9 | 0.295               | 16.0                | 6.45      | 75.3      | 27  | 0.307          | 0.089          | 0.261          |
| 0.02| 5.84 | 0                   | 21.8                | 13.3      | 30.5      | 0   | **0.821**     | 0.254          | 0.850          |
| 0.03| 7.68 | 0.438               | 16.3                | 7.06      | 40.4      | 9   | 0.310          | **0.634**      | 0.509          |
| 0.17| 7.73 | 0.475               | 17.2                | 10.5      | 69.6      | 15  | 0.208          | 0.587          | 0.412          |
| 0.22| 6.42 | 0.032               | 20.1                | 11.2      | 29.4      | 2   | 0.764          | 0.358          | **0.877**      |
| 1.0 | ≤6.5 | 0                   | **37.7**            | **24.6**  | **153.1** | 12  | 0.329          | 0              | 0.559          |
| 1.0 | ≥8.7 | 0.285               | 16.1                | 6.47      | 74.7      | 27  | 0.307          | 0.089          | 0.261          |
| Empirical | 0.039 | 21.7 | 14.9 | 33.1 | 2 | 0.732 | 0.245 | 0.834 |

For different representative pairs of values for $a$, $m$, the values of the metrics with the addition of the number of isolates $I$ are reported, along with the corresponding values for the empirical network. All metrics have no dimension except for $\langle I \rangle_c$, $\sigma_L$, and $\sigma_s$, which are expressed in Km. In bold, the maximum value for each metric.
\[ DSPL = \frac{2}{N(N-1)} \sum_{i \neq j}^{2} \frac{\left| L_{ij}^{\text{emp}} - L_{ij}^{\text{syn}} \right|}{L_{ij}^{\text{emp}} + L_{ij}^{\text{syn}}} \]  

where \( L_{ij} \) is the sum of the lengths of the links in the shortest path between \( i \) and \( j \) if both nodes belong to the same connected component, and \( L_{ij}^{d[l]} = e^{-m}L_{CG} \) otherwise. For disconnected nodes in the empirical network, we assume the same value of \( m \) that generated the synthetic counterpart we are comparing it to.

Once more, it is confirmed that the most similar topologies can be obtained by setting \( m \in [5.6, 6.4] \) while keeping the value of \( a \) small, namely, \( a \leq 0.16 \) or \( a \in [0.25, 0.3] \) (see Figure 6 (Right)).

For such values of \( m \), we have \( L_{ij}^{d[l]} \in [103, 229] \) (km), that is, larger than the average (geodesic) distance between node locations \( \langle d \rangle = 83.5 \) (km), and also larger than the largest link in the empirical network \( l_{\text{max}} = 85 \) (km), but considerably smaller than the maximum distance between nodes \( d_{\text{max}} = 269.5 \) (km).

In the system, for the selected value of merging parameter \( \Gamma^* \), there are 8 nodes (20.5%) whose average geodesic distance to other locations is larger than 103 (km). For these nodes, the underlying assumption of the gEEM that the overall pre-existing connectivity allowed them to reach any other place in the system through a path of length \( L_{ij}^{d[l]} \) is, in this range of the parameter \( m \), incorrect and, from the viewpoint of the infrastructure design, unfair.

In particular, the presence of isolates (more than one) is a trait that can be observed in synthetic topologies only if \( a > 0.28 \), unless \( m > 6 \) and hence \( L_{ij}^{d[l]} < 150 \) (km) (see Figure 6 (Left)), that is, under the hypothesis of a quite unrealistic connectivity provided by the presence of alternative or pre-existing TTIs in the same territory.

Therefore, we can conclude that for the model to be able to reproduce this particular feature of the empirical network, a non-negligible share of unfairness towards the periphery of the system is required, either underestimating geographical distances \( (m) \), or explicitly favouring the most populated locations \( (a) \), or both. However, the model is able to produce synthetic
network topologies with none or just one isolate that closely resemble the empirical one regarding almost any other feature for smaller values of $a$ and $m$, namely, $a \leq 0.2$ and $m \leq 6$ ($L_{ij}^d \leq 150$ km).

At a more local scale, the first node to get disconnected in the topologies generated by the gEEM is La Seu d’Urgell (see Figure 2), which has the smallest merged population associated to it ($p = 53154$) and the furthest distance to the nearest neighbour ($d = 66.331$ km). The isolates in the empirical networks have quite larger populations ($P(Palafrugell) = 113773$ km and $P(Berga) = 84238$) and nearer neighbours (at 24.607 km and 35.379 km, respectively). Although they both are far from the top ranking nodes, the features considered by the present version of the model do not make them plausible candidates to be the sole disconnected nodes of an otherwise connected topology. Possibly, the explanation for their specific condition could be found in the details of the historical process that shaped the TTI under study in its current state. For instance, until 1973, there existed a railway connection Berga-Manresa (Llorens, 1981), a link that the gEEM builds in most of the topologies that are similar to the empirical network.

Conclusions

We have presented a methodology to shed light on the mechanisms, power balances and competing interests that shaped a given TTI into its current configuration. Relying on the analytical and theoretical toolbox provided by network science and adopting an inverse engineering approach, we propose a two-step procedure that allows us to address a broad variety of systems.

The first step consists in mapping the infrastructure under study into a geographic network in a flexible way, adjusting the spatial scale of the representation to the specific situation of the system and its function by tuning a single parameter.

The second step enables us to investigate how different the considered system is from an ideal model of TTIs that has proven to be able to capture most of the relevant features of ancient proto-historical networks of pathways, that is, the EEM. The main characteristics of such a model are (i) the equitable treatment of the necessities and interests of each node-place in the system, regardless of its real power or importance; (ii) the assumption that if in the network there is no path connecting two nodes, then it is impossible to reach one place from the other, that is, the considered network represent the only TTI existing in the territory.
We have devised a mechanistic network model, gEEM, that using EEM as a starting point, progressively and independently relaxes both its main assumptions by means of two parameters.

Finally, we have shown how to effectively infer information about a real case study by applying the proposed methodology to an illustrative example, that is, the regional railway connections in Catalonia. Comparing the output of the gEEM for different values of the parameters to the empirical network of the Catalan regional train service, we evinced that network topologies similar to the empirical one can be generated if the most geographically central and demographically important places are slightly favoured when designing the TTI.

One may argue that these satisfactory results might not be achievable for case studies strongly conditioned by geographic constraints or temporal path-dependence. In this sense, notice that the proposed methodology can accommodate usual constraints of TTI's construction. The effect of natural barriers (e.g. mountains) could be introduced by removing the affected links from the list of potential connections, hence never be allowed to be built. Moreover, path-dependence on previously constructed TTI sections could be considered by establishing such links in the original scenario (before starting the model dynamics).

In any case, the case study presented here should be seen as a benchmark or reference. Knowing that the model can work fine under certain circumstances, allows us to inquire what makes it fail when some of these conditions are not met.

These conclusions are in line with common knowledge but drawn into more precise quantitative terms. This confirms the potential of simple mechanistic models as powerful explanatory instruments for better understanding the configuration of complex transportation systems. In particular, they provide new ways to approach the interaction between two key geographic concepts: Spatial interaction and spatial structure (Oshan, 2021).

Declaration of conflicting interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Agència de Gestió d’Ajuts Universitaris i de Recerca through grant no. 2017 SGR 1466.

ORCID iD
Luce Prignano https://orcid.org/0000-0003-4092-0699

Notes
1. Such a procedure is equivalent to simply building, at each step, the link to the minimum of the Rij matrix, that is, the link that will provide the maximum increment \( \Delta R_{ij} = 1 - R_{ij} \) once it is created.
2. A graph is said to be complete when it exists a link between each pair of nodes, creating a fully connected structure.

References
Barthélemy M (2011) Spatial networks. Physics Reports 499(1–3): 1–101.
Barthélemy M and Flammini A (2008) Modeling urban street patterns. Physical Review Letters 100: 138702.
Black WR (1993) Black. Transport route location: a theoretical framework. Journal of Transport Geography 1(2): 86–94.
Boccaletti S, Latora V, Moreno Y, et al. (2006) Complex networks: structure and dynamics. *Physics Reports* 424(4): 175–308.

Carreras C and De Soto P (2013) The roman transport network: a precedent for the integration of the european mobility. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46(3): 117–133.

Costa LF, Oliveira ON Jr, Travieso G, et al. (2011) Analyzing and modeling real-world phenomena with complex networks: a survey of applications. *Advances in Physics* 60(3): 329–412.

Earle T (1991) Paths and roads in evolutionarily perspective. In: Trombold CD (ed), *Ancient Road Networks and Settlement Hierarchies in the New World*. Cambridge, MA: CUP.

Fischer MM and Reggiani A (2004) Spatial interaction models: from the gravity to the neural network approach. *Contributions to Economic Analysis* 266: 317–346.

Fulminante F, Prignano L, Morer I, et al. (2017) Coordinated decisions and unbalanced power. How Latin cities shaped their terrestrial transportation network. *Frontiers in Digital Humanities* 4(4).

Fulminante F, Lozano S, Prignano L, et al. (2021) Why rome and not veii? analysing geographical networks in eturia and latium vetus between the early iron age and the archaic era. In: Dawson H and Iacono F (eds), *Bridging Social and Geographical Space through Networks, Proceedings of the Topoi Conference*. Leiden: Sidestone Press.

Gastner MT and Newman MEJ (2006) The spatial structure of networks. *The European Physical Journal B - Condensed Matter and Complex Systems* 49: 247–252.

Gastner MT and Newman MEJ (2006) Shape and efficiency in spatial distribution networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(01):P01015

Groenhuijzen MR and Verhagen P (2016) Testing the robustness of local network metrics in research on archeological local transport networks. *Frontiers in Digital Humanities* 3(6).

Herranz-Loncán A and Fourie J (2017) “For the public benefit”? Railways in the British Cape Colony. *European Review of Economic History* 22(1): 73–100.

Llorens JP (1981) Carrilet manresa-berga. aproximació a les influències socio-econòmiques, el. *Dovella*(3): 9–14.

Jenkins D (2001) A network analysis of inka roads, administrative centers, and storage facilities. *Ethnohistory* 48: 655–687.

Kurant M and Thiran P (2006) Extraction and analysis of traffic and topologies of transportation networks. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics* 74: 036114.

Latora V and Marchiori M (2001) Efficient behavior of small-world networks. *Physical Review Letters* 87: 198701.

Lay M (1992) *Ways of the World: A History of the World’s Roads and of the Vehicles that Used Them*. Piscataway, NJ: Rutgers University Press.

Louf R, Jensen P and Barthelemy M (2013) Emergence of hierarchy in cost-driven growth of spatial networks. *Proceedings of the National Academy of Sciences of the United States of America* 110(22): 8824–8829.

Louf R, Roths C and Barthelemy M (2014) Scaling in transportation networks. *PLoS One* 9(7): e102007.

Newman MEJ (2003) The structure and function of complex networks. *SIAM Review* 45: 167–256.

Oshan TM (2021) The spatial structure debate in spatial interaction modeling: 50 years on. *Progress in Human Geography* 45(5): 925–950.

Pablo-Martí F, Alañón-Pardo Á and Sánchez A (2021) Complex networks to understand the past: the case of roads in Bourbon Spain. *Cliometrìca* 15: 477–534.

Prignano L, Morer I, Fulminante F, et al. (2019) Modelling terrestrial route networks to understand inter-polity interactions (southern eturia, 950-500 bc). *Journal of Archaeological Science* 105: 46–58.

Raymond C (1976) *Roman Roads*. Berkeley and Los Angeles: University of California Press.

Reggiani A (2021) Complexity and spatial networks. In: Fischer MM and Nijkamp P (eds). *Handbook of Regional Science*. Berlin, Germany: Springer-Verlag GmbH.

Smith ML (2005) Networks, territories, and the cartography of Ancient States. *Annals of the Association of American Geographers* 95(4): 832–849.
Taylor C (1979) Roads and Tracks of Britain. London, UK: Orion.
Trombold CD (1991) Ancient Road Networks and Settlement Hierarchies in the New World. Cambridge, MA: CUP.
Verhagen P, Nuninger L and Groenhuijzen MR (2019) Modelling of pathways and movement networks in archaeology: an overview of current approaches. Finding the Limits of the Limes 12: 217–249.
Vragović I, Louis E and Díaz-Guilera A (2005) Efficiency of informational transfer in regular and complex networks. Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics 71: 036122.
Wilson AG (1971) A family of spatial interaction models, and associated developments. Environment and Planning A: Economy and Space 3(1): 1–32.

Luce Prignano is a postdoctoral researcher at the Universitat de Barcelona’s Condensed matter department. Her research focuses on applying complex systems and network science across multiple disciplines, including archaeology, social science, and communication studies.

Lluc Font-Pomarol is a PhD candidate with degrees in Physics and Modelling for Science and Engineering from the Universitat de Barcelona and Universitat Autònoma de Barcelona. His research since 2019 has focused on complex networks in judicial decisions.

Ignacio Morer Zapata currently serves as a data scientist at IAS Artificial Intelligence Innovations. His prior roles include research technician at Fundació Bosch i Gimpera, where he specialized in data analysis using Python for the European EPNET project, and a brief tenure at the Institut Català de Paleocologia Humana i Evolució Social (IPHES).

Sergi Lozano is Associate Professor at University of Barcelona’s Department of Economic History, Institutions and Policy and World Economy, and a member of the University of Barcelona Institute of Complex Systems. His research focuses on the study of long-term cultural and socioeconomic phenomena (usually through the application of computational modeling and network analysis).