

Revealing the similarity between urban transportation networks and optimal transport-based infrastructures

Daniela Leite^{1,*} and Caterina De Bacco^{1,*}

¹Max Planck Institute for Intelligent Systems, Cyber Valley, 72076, Tübingen, Germany

*caterina.debacco@tuebingen.mpg.de

ABSTRACT

Designing and optimizing the structure of urban transportation networks is a challenging task. In this study, we propose a method inspired by optimal transport theory to reproduce the optimal structure of public transportation networks, that uses little information in input. Contrarily to standard approaches, it does not assume any initial backbone network infrastructure, but rather extracts this directly from a continuous space using only a few origin and destination points. Analyzing a set of urban rail, tram and subway networks, we find a high degree of similarity between simulated and real infrastructures. By tuning one parameter, our method can simulate a range of different networks that can be further used to suggest possible improvements in terms of relevant transportation properties. Outputs of our algorithm provide naturally a principled quantitative measure of similarity between two networks that can be used to automatize the selection of similar simulated networks.

INTRODUCTION

Transportation networks are a fundamental part of a city's infrastructure. Their design impacts the efficiency with which the system is operated, hence, they should follow optimal principles while being constrained by limitations like budget or physical obstacles. Existing approaches for studying the quality of network design often rely on the analysis of the topological network properties, and relate them to optimal features like transportation cost, efficiency or robustness. These analyses are usually made *a posteriori*, only once the network has been constructed and thus only resulting properties can be analyzed^{1,2}. A different approach is that of posing *a priori* a principled optimization setup, where one defines a cost function that a network should minimize under a set of constraints, and then searches for optimal solutions in terms of network topologies. Numerous studies have explored this approach in biological networks, transportation networks, etc.^{3,4}. However, most of these approaches rely on an existing backbone of a network infrastructure that can be optimized in terms of traffic distribution^{5,6} but do not consider the possibility of building the network from scratch, starting from a limited set of nodes. Alternatively, as optimizing over all possible topologies is difficult, one can investigate only various simple shapes from a predetermined set of possible geometries⁷⁻⁹ or rely on heuristics^{10,11}. Recently, Kay et al.¹² presented a two-step agent-based model that replicates biologically-grown networks and proposes them as a template for urban design. Further, the lack of a principled metric to measure the similarities between an observed network and a simulated one poses the problem of making this evaluation effectively.

In this work, we show that urban transportation systems exhibit underlying network topologies similar to those that follow optimality principles as defined in optimal transport theory. Specifically, we propose a model to characterize real transportation networks based on a simple optimal transport framework, similar to what is observed in biological systems like the slime mold *Physarum polycephalum*, which adapts its network structure to reach food patches in an optimal way. A previous study by Tero et al.¹³ shows how this mold forms networks with comparable optimal transportation properties, e.g. efficiency and cost, to those of the Tokyo rail system, but provided no rigorous quantitative definition of network similarity beyond measuring these properties. We empirically validate our approach with a systematic characterization of the structure of several urban transportation networks and propose a rigorous definition of similarity in terms of optimal transport theory.

Urban transportation networks often exhibit different network structures based on the goals of network designers. For instance, some networks focus on connecting people living in the outer layers of the city to the city core, while other prefer to develop a robust infrastructure servicing the core¹⁴. Several studies have focused on analyzing network properties like scaling laws and network connectivity¹⁵⁻¹⁷, which are indications of an underlying optimality mechanism that these networks might follow in order to make a city efficient, both in reduced infrastructure costs per capita and in increased productivity. However, our understanding of what optimality principles are captured in real transportation networks is incomplete. In fact, studying

network properties could only partially explain the underlying mechanisms regulating network design, as each property captures a different aspect. Here, we take a different approach and build the network from scratch while comparing with the real ones observed from data, starting with only a few shared nodes in input. Specifically, we model network structures observed in urban transportation networks by adapting a classical optimal transport framework to simulate a network-design problem dependent on realistic travel demand settings and using little information in input. We then compare the resulting networks with those observed from real data and assess their similarity. Importantly, the model can simulate different optimal strategies by tuning a parameter β , which interpolates between minimizing infrastructural and operating costs. On one hand, this allows to simulate optimal networks that resemble those observed in real transportation systems more closely, as we tune β . On the other hand, by comparing the networks resulting for various values of β with those observed from real data, we can also assess the impact of the two types of cost in the design of various urban infrastructures.

We use this model to analyze 17 different transportation networks from 16 cities¹⁸. Despite the complex nature of the mechanisms driving the design of transportation networks, we observe that multiple urban transportation networks follow a surprisingly similar topological pattern, observed as in biological systems. We observed that optimal networks obtained with our approach have similar cost and performance to those observed in real ones.

RESULTS

Modeling network design in transportation networks

Consider an urban area where a set of points of interest are located in certain positions in space. These may correspond to a combination of central and peripheral points where people work and live. The goal is to connect them by building a transportation network in an optimal way. At this point, we do not observe any network, but are rather free to use the whole space where the urban area is located, i.e. a 2D surface. From this, we need to select a set of points and edges connecting them, in other words, a network. In Fig. 1a–b we illustrate the problem setup for the subway network in Rome, where green and red markers denote an example set of such reference points and the lines denote edges in the observed metro network infrastructure.

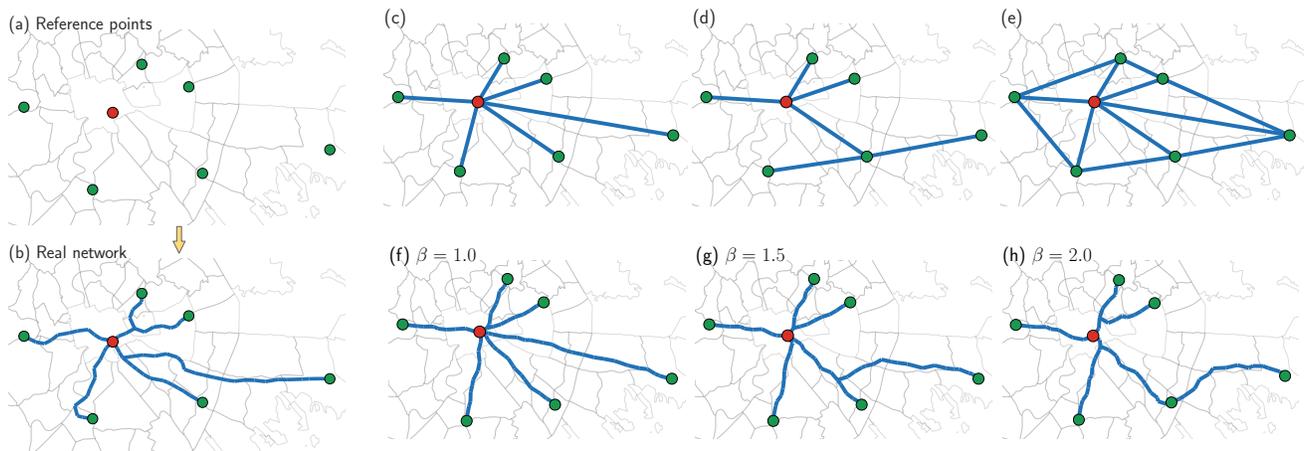


Figure 1. Problem setup for the subway network of Rome. (a) Given a set of real latitude-longitude coordinates denoting origins (green) and destinations (red), we aim to build a network structure that resembles well the observed public transportation network connecting those points, as in (b). (c)–(e) Intuitive ways to build a network structure by connecting origins and destinations, versus networks extracted with our optimal transport-based method in (f)–(h). The only known information is the set of six origins (O) and one destination (D). We capture different optimization mechanisms by tuning the β parameter: in (f), the network is a shortest path-like structure, while in (g) and (h) we show examples of branched transportation schemes.

In general, there are many choices for designing the network. For instance, in Fig. 1c–e we show three examples of intuitive shortest-path-like minimization solutions for the settings shown in Fig. 1a. These are however quite different from the observed network in Fig. 1b. The question we address is what network design principle is producing simulated networks that are more similar to those observed in real urban networks. Optimality could be defined in various ways depending on the designers' goal, but generally, it is not known what principles the designers used when building the network. Instead, we want to assess this by observing real data of transportation networks and fitting them with a flexible and computationally efficient optimization setup guided by optimal transport theory. This allows to select, among the many possible choices, the networks that are optimal in a simple setup and compare them with the observed transportation networks. For this, we adopt the formalism recently developed by Facca et al.^{19–21} that generalizes to a continuous space the original idea of Tero et al.²². In particular, this allows

starting with only a set of relatively few origin and destination nodes in input and then designing a network by exploring the 2D surface, i.e. without the need of an initial existing backbone. The idea is inspired by the behavior of the slime mold *Physarum polycephalum*, which dynamically builds a network-like body shape when foraging. One can thus consider a dynamics for the two main quantities involved, flows and conductivities, that implements this mechanism at any point in space. The stationary solution of this dynamics corresponds to the minimizer of a Lyapunov cost in a standard optimization setup, which has a nice interpretation in terms of infrastructure and operating transportation costs. From these solutions, one can then automatically extract optimal network structures using the approach presented in²³.

Having introduced the main problem and ideas, we now briefly describe the model. Consider a surface in 2D and a set of points on it. Specifically, we denote a set of origins and destinations as f^+ and f^- , respectively. These contain the reference points where people enter and exit the transportation network. By defining $f = f^+ - f^-$, mass conservation can be enforced with the constraint $\int f dx = 0$. The two main quantities of interest are denoted with $\mu(x, t)$, the transport density (or conductivity), and $u(x, t)$ the transport potential. The former can be seen as a quantity proportional to the size of a network edge, while the latter determines the fluxes traveling along them. The dynamical equations in this continuous setting are

$$-\nabla \cdot (\mu(t, x) \nabla u(t, x)) = f, \quad (1)$$

$$\frac{\partial \mu(t, x)}{\partial t} = (\mu(t, x) \nabla u(t, x))^\beta - \mu(t, x), \quad (2)$$

$$\mu(0, x) = \mu_0(x) > 0. \quad (3)$$

Equation (1) determines the spatial balance of the flux, defined as $q = \mu \nabla u$; Eq. (2) enforces optimal solutions; Eq. (3) is the initial condition. The parameter β captures different optimization mechanisms: $\beta < 1$ enforces congested transportation, $\beta = 1$ is the shortest path-like and $\beta > 1$ is branched transportation. In Fig. 1f–h we show examples of different optimal configurations, with $\beta = 1$, $\beta = 1.5$ and $\beta = 2.0$. Here, we consider the cases $1 < \beta \leq 2$, where the approximate support of the conductivity μ displays a network-like structure. A numerical solution for this system is obtained by a discretization in space and time, and the equilibrium solution is a pair (μ^*, u^*) , that minimizes the functional

$$\mathcal{L}(\mu, u) = \frac{1}{2} \int \mu |\nabla u|^2 dx + \int \frac{\beta}{2 - \beta} \mu^{\frac{2 - \beta}{\beta}}. \quad (4)$$

This can be interpreted as the network transportation cost, where the first term is a network operating cost, while the second is the cost to build the infrastructure. By changing β , one can tune their relative contribution to the total transportation cost. Once the optimal (μ^*, u^*) are obtained, one can then use the model described in²³ to extract a final network structure, i.e. a set of nodes, a set of edges connecting them and their weights proportional to the conductivities. This can then be compared with the one observed from real data and repeated for various values of β . An example is shown in Fig. 4a–d.

As we aim at extracting a network, our problem starts by defining a set of origins and destinations (O–D), points in the space with coordinates (x, y) where passengers enter and exit the network. We identify a set of nodes with smallest and higher network centrality as the origin and destination nodes, respectively. Depending on the criteria defining centrality, the coordinates for this set of nodes might change, and so the final networks obtained with our approach may result in different topologies. In this work, we consider degree and betweenness centrality²⁴. Intuitively, both consider peripheral nodes as origins, while differing in how destinations are selected. Degree centrality captures hubs, stations with many connections, whereas betweenness centrality captures nodes that many passengers use in their travel trajectories. The destination nodes are typically located in the city center, while origin nodes are along the periphery. This setting is often encountered in urban scenarios, e.g. in subway networks²⁵. With all these problem inputs at hand, we then use the algorithmic implementation of Nexttrout²³ to extract our simulated networks efficiently.

Revealing the similarity of optimal simulated networks and the observed transportation systems

We apply the proposed dynamics to empirical data collected from 16 different cities in multiple geographical regions around the world. For each city, we selected various available types of public transportation systems, such as rail, subway and tram, keeping the largest connected component. The networks considered in this manuscript have a few loops, as the dynamics can only retrieve loopless structures in the regime where network extraction is meaningful²³. Specifically, we measure the loop ratio $L_{ratio} = n_L/E$ as the number of loops divided by the number of edges and select networks with $L_{ratio} < 0.2$ (see Methods for more details). One could in principle recover loopy structures by employing numerical schemes, e.g. superposition of different outputs²⁶, but we do not explore this here.

The applied OT dynamics successfully describe the transportation network structure observed in different cities at a macroscopic

level. While the selected transportation networks have different topology and include multiple transportation modes, the networks reconstructed by Nexttrout with only little information in input show a high degree of similarity with the real ones (see Fig. 2a–c). This suggests the existence of simple universal optimality rules captured by the DMK dynamics for the modeling of urban transportation rail networks, similar to what has been observed for the behavior of the slime mold *Physarum polycephalum*. Notice that in principle one can increase this similarity further, by simply adding more information in input in terms of origin and destinations (see Fig. 4d for an example with multiple origins and destinations). However, here we are interested in recovering macroscopic structure in a more challenging scenario where input information is strictly limited to one central destination and few peripheral origins.

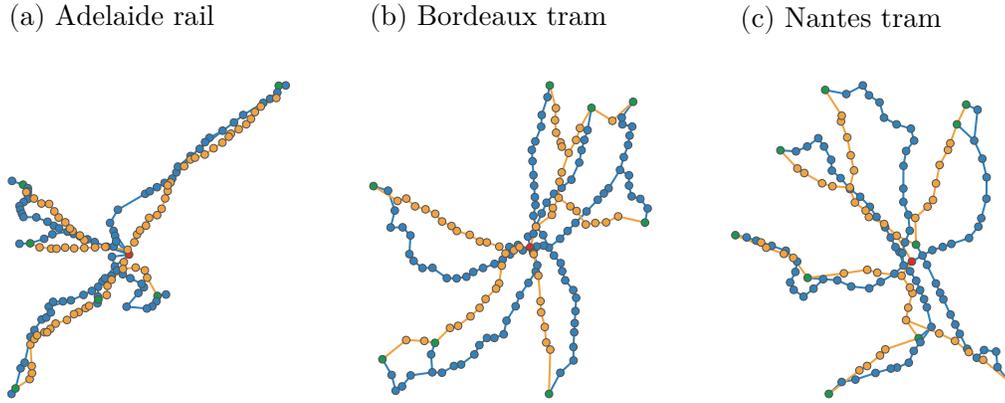


Figure 2. Example of different network topologies generated by Nexttrout (yellow), plotted against the corresponding real network (blue). Green nodes represent those chosen as origins (O), whilst red nodes correspond to the destinations (D). (a) Adelaide rail network, with $N = 87$ nodes, $O = 6$ origins and $D = 1$ destination. (b) Bordeaux tram network, with $N = 108$ nodes, $O = 8$ origins and $D = 1$ destinations. (c) Nantes tram network with $N = 97$ nodes, $O = 10$ origins and $D = 1$ destination.

Beyond a qualitative visual comparison, we explore how our simulated networks score in terms of core network properties relevant for transportation compared to the real networks. For this, we consider the cost, the total path length l , the distribution of traffic and the density of branching points (or bifurcations; here we use the two terms interchangeably), for both extracted and original networks. Similar to Tero¹³, we define the cost as the total length of the network (TL), i.e. the total number of edges. Passengers may not always take the shortest path, but may rather consolidate on fewer main arteries (e.g. to minimize the number of stops or connections)²⁷, a behavior that can be captured by a DMK discrete dynamics (built-in the filtering step of Nexttrout) by varying β and extracting the flows u_e on edges, quantities proportional to the number of passengers using an edge. Hence, we consider an alternative measure of total path length as $l := \sum_{e \in E_i} l_e |u_e|$, where l_e is the euclidean distance. This takes into account u_e , the flow of passengers on an edge e , and its absolute value $|u_e|$ is proportional to the amount of passengers traveling on an edge e , i.e. how traffic is distributed, assuming that passengers follow optimality principles to consolidate paths. This is a reasonable assumption in rail networks as the ones studied here, where the cost to build the infrastructure can be high (and thus should be minimized) and minimizing traffic congestion is not as relevant as in, e.g., road networks. In our experiments, we extract optimal flows u_e by running the discrete DMK dynamics on the extracted and real networks, using the same sets of origins and destinations as used in the original network extraction problem, setting $\beta = 1.5$. This information can also be used to measure the macroscopic behavior of traffic on edges, which can be measured using the Gini coefficient²⁸ $Gini(T_e)$ on the traffic $T_e = |u_e|$. This coefficient ranges from $[0, 1]$, where the closer to 1, the more unequal is the traffic distribution on the network. Finally, we calculate the percentage of bifurcations (D_{BP}) as the fraction of nodes with degree equal to 3. The simulated networks display similar properties as those observe on the real ones, as shown in Fig. 3. While similarity differs depending on the property and datasets vary in their range values, we notice that most of the datasets have at least a pair of properties that superimpose well between their simulated and real property values. For instance, in Fig. 2a we notice that Adelaide rail has an intuitively similar path length, which is confirmed in Fig. 3a. Furthermore, the same network show comparable results for traffic and cost. As for the tram networks of Bordeaux and Nantes in Fig. 2b-c, we observe similar behavior for cost and traffic. In some cases, e.g. the network of Grenoble tram shown in Fig. 4, we see a strong similarity in all the four metrics.

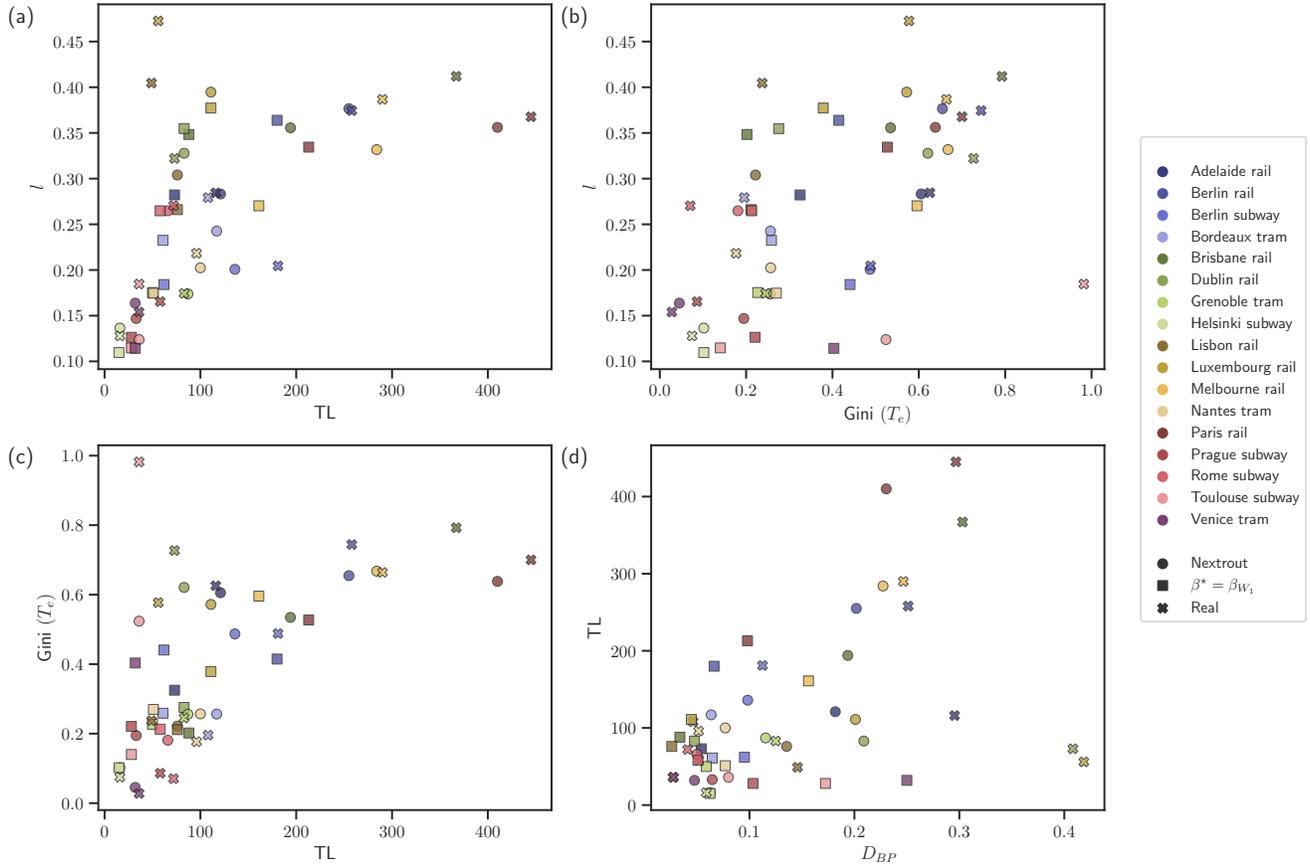


Figure 3. Performance measures for the 17 real networks. Each dataset is assigned to a different color, market shapes distinguish real and simulated networks. Simulated networks are further distinguished based on the one generated via Nexttrout that gives the closest point in terms of the metrics plotted in the figure (circle) or the one corresponding to the best Wasserstein measure (square). (a) Cost (TL) measured in both simulated and real networks, plotted against the total path length. (b) Gini coefficient as a measure of traffic distribution, versus the total path length. (c) Traffic distribution in terms of the Cost. (d) Density of bifurcations plotted against the cost.

Automatic selection of similar simulated networks

Our method allows extracting various simulated networks by varying the parameter β . One can select the one that more closely resembles the observed one in terms of a particular metric of interest, as shown in the previous section. However, different metrics may lead to different most similar simulated networks (i.e. different β), which may not be ideal for a practitioner willing to consider an individual simulated network that resembles well the observed one in terms of all metrics. Hence, the need for a principle automatic selection criteria for choosing the value of β .

The formalism introduced in the previous section suggests a natural way to tackle this problem by considering the Wasserstein similarity measure, a main quantity in optimal transport theory²⁹. Given two graphs that need to be compared, intuitively, this measure captures the minimum "effort" required to move a certain distribution of mass from one to the other. Similar ideas based on optimal transport to measure similarity between graphs have also been proposed in recent works^{30,31}. Here, we describe our proposal for a similarity measure and thus automatic selection of β in detail. Denote the observed network with $G_1(V_1, E_1)$ and the one obtained from the model introduced in previous sections with $G_2(V_2, E_2)$, where V_i, E_i denote the set of nodes and edges, respectively, $i = 1, 2$. We consider the union graph $G_U(V_U, E_U)$, with sets of nodes $V_U = V_1 \cup V_2$ and edges $E_U = E_1 \cup E_2$. One can further assign weights $w_e \in W_U$ to the edges $e \in E$, for instance using the euclidean distance ℓ_e between the nodes $i, j \in V_U$, where $e = (i, j)$, or simply binary values $\{0, 1\}$. Notice that the observed network G_1 may contain nodes that are not corresponding exactly to nodes in G_2 , because in this continuous setting the model uses all the 2D space where the original network is embedded. Only the input origin and destination nodes are guaranteed to be present in both graphs, as they are given in input to the model.

Working on this union network, we then exploit a similar setting as the one already introduced with the model to obtain a

Wasserstein-based similarity measure between G_1 and G_2 . Specifically, denote with B the unsigned incidence matrix of G_U with entries $B_{ie} = +1$ if node i is a start or end point of the edge e , and 0 otherwise. Defining q_i as an indicator vector for the edges in G_U that are also in G_i , i.e. $q_{ie} = 1$ if $e \in E_i$, and $q_{ie} = 0$ otherwise, $\forall e \in E_U$ and $i = 1, 2$, we can set the origin and destination vectors $f = f^+ - f^-$, such that $f^+ = Bq_1$ and $f^- = Bq_2$,

so that G_1 contributes to f^+ and G_2 to f^- . By running a discrete dynamics analogous to the continuous one described in Eqs. (1) to (3), which can be done using Nexttrout²³, one naturally obtains our Wasserstein similarity measure defined as:

$$W_1(G_1, G_2) = \sum_{e \in E_U} w_e \mu_e, \quad (5)$$

where μ_e are the optimal solutions for the conductivities on G_U and w_e is the weight of edge $e = (i, j)$. Here we fix this to be the euclidean distance between nodes i and j . Examples of how the Wasserstein measure changes depending on the different output networks are shown in Fig. 4i.

To evaluate how the simulated networks selected with our Wasserstein-based metric perform in terms of the main topological properties described above, we define a similarity ratio r_p for each of them. Given a particular property p , we compute the ratio between the value measured on the simulated and observed networks. Thus, when $r_p = 1$, the simulated network extracted by Nexttrout matches perfectly with the observed one in terms of property p . For instance, when p is the cost, then $r_{TL} = TL_{\text{Nexttrout}}/TL_{\text{real}}$. In Fig. 5 we compare the ratio values for two of the main transport properties above introduced, comparing them across different automatic selections of β (more details are in the Supplementary Fig.1). Specifically, we obtain for each property p the value of β_p that leads to a simulated network closer to the observed one in terms of property p , i.e. corresponding to a r_p closer to 1. By definition, the simulated graph obtained with β_p has the best r_p for property p , but this may not be true for other properties measured on that same graph. For instance, the graph selected with β_{TL} (second row of Fig. 5) as $r_{TL} \approx 1$, thus reproducing well the number of edges, but it tends to largely overestimate both $Gini(T_e)$ and the number of bifurcations, with $r_p > 2$ for these two properties (see Supplementary Fig.1). Instead, we observe that the simulated graph selected with the Wasserstein measure (first row of Fig. 5) has on average r_p closer to 1 across various properties. In other words, it shows transportation properties that are consistently more aligned to those behold by the observed network.

To validate the consistency of the Wasserstein measure, we compute how each data point in Fig. 5 differs from the perfect match $r_p = 1$. Specifically, we define the minimum mean displacement $\mathcal{D}_d = \frac{1}{|p|} \sum_p (1/n \sum_{i=1}^n |r_{p_i} - 1|^d)^{\frac{1}{d}}$, where p refers to each considered metric ratio ($|p| = 4$), $n = 17$ is the number of data points and $d = 1, 2$. The closer to 0, the less the considered measure deviates from the perfect score. We found that the Wasserstein measure deviates with $\mathcal{D}_{d=1} = 0.5$, whilst the cost has $\mathcal{D}_{d=1} = 0.61$, the total length $\mathcal{D}_{d=1} = 0.69$, the traffic $\mathcal{D}_{d=1} = 0.35$ and the density of bifurcation points has $\mathcal{D}_{d=1} = 0.32$. Similar results are found with the square displacement $d = 2$. The traffic and number of bifurcations have the lowest displacements, with the Wasserstein following. As the Wasserstein tends to select networks with higher β (not shown here), this measure encourages smaller cost (TL). The fact that we see a higher displacement than the one of traffic and density of bifurcations is a sign that real networks could be further optimized in terms of TL. Nevertheless, the gap is small and beyond problem-specific properties (as the traffic or density of bifurcations). These results confirm the benefit of using our Wasserstein-based similarity measure to automatically select a simulated network across different values of β in our OT-based setting. More generally, it provides a robust and meaningful measure to compare network structures that can be used in applications beyond the one considered in this work.

Improving network properties

Simulating networks that follow optimality principles and resemble well those observed in real datasets can be used to assess how urban transportation networks perform in terms of main transportation properties. This can guide network managers towards potential measures directed at improving certain properties. This possibility is conveniently enabled by our approach, as by continuously tuning the parameter β we can simulate various transportation scenarios, thus assessing how a network can increase or decrease a certain property. In Fig. 6 we show the main properties in all the networks simulated with our model and compare with those observed in real data, aiming to compare their trade-offs between various performance metrics. In general, we notice how simulated networks cover a wider range of values than the real ones for the four transportation properties investigated in this work. This allows obtaining, for instance, networks that have shorter total path length ℓ with a comparable cost, as shown in Fig. 6a where many simulated networks are located in the regime $0 < TL < 200$ with ℓ sharply dropping towards 0.1, while many real networks have $\ell > 0.1$. A similar behavior is observed also for the traffic against TL in Fig. 6c, where simulated networks cover areas of the plot where traffic is less congested (smaller $Gini(T_e)$), in contrast to several real networks. Among the simulated networks, those selected according to the best Wasserstein measure tend to have lower cost and a smaller percentage of bifurcations D_{BL} , indicating that this measure encourages not only the usage of a lower amount of edges, but also nodes with low degree.

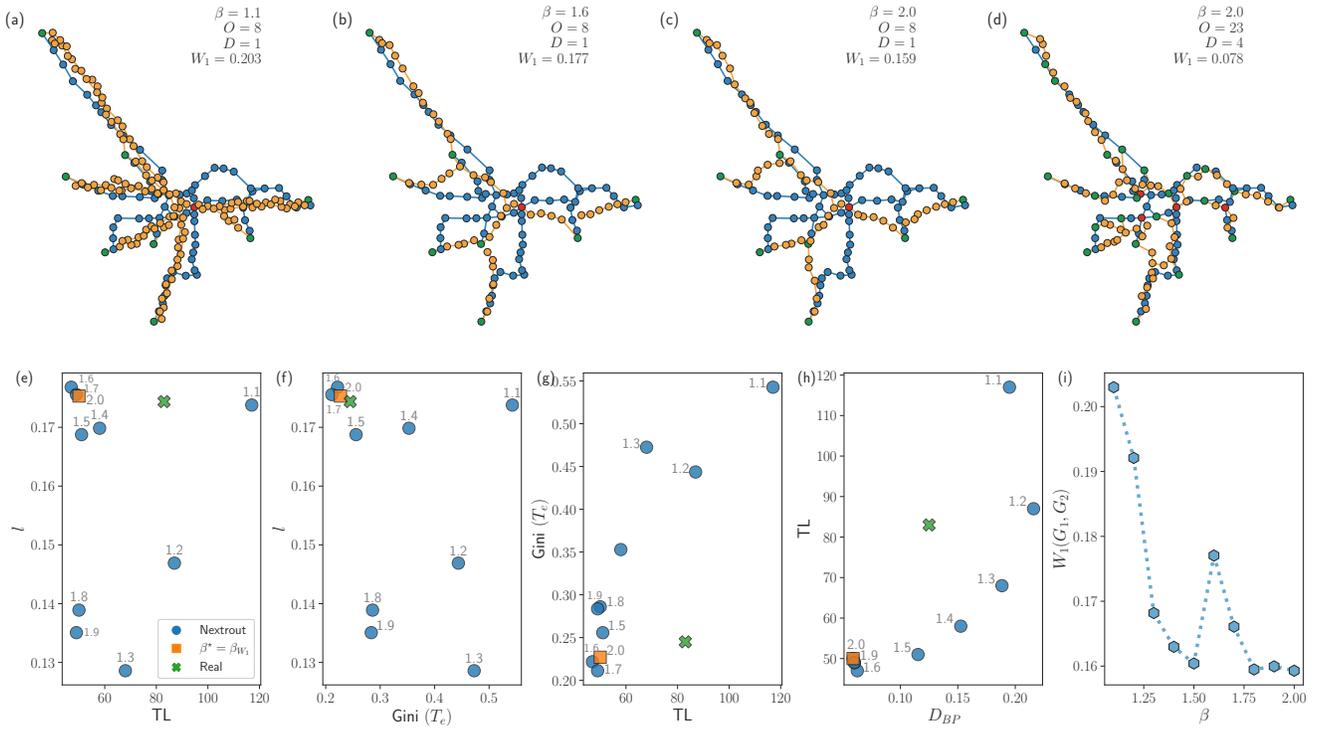


Figure 4. Wasserstein similarity measure between graphs for automatic selection of β . We show examples (a-d) of simulated networks and report their Wasserstein measure from the observed network of the Grenoble tram ($N = 80$ nodes). (a-c): we select the origin nodes based on those with the smallest degree, and a unique destination as the one with the highest degree. Notice that there are some branches in the original network not being captured by our algorithm. This is due to the little information we give as initial input. The more nodes we add corresponding to these branches, the closest to the observed network we would get. In (d) we show how the same network changes as we set more stations as initial input ($O = 23$ origins and $D = 4$ destinations), resulting in smaller Wasserstein, at the cost of a higher amount of information given in input. (e-h): we compare several network properties as measured in the observed and simulated networks. (e) The cost (TL) against the total path length (l), highlighting the different β for each obtained network. If similarity is chosen to be defined in terms of the cost, the closest network to the real one would be that with $\beta = 1.2$. (f) TL against the Gini coefficient of traffic on edges ($Gini(T_e)$). In this case, the optimal network is equivalent to the one with minimal Wasserstein, i.e. $\beta = 2.0$. (g) $Gini(T_e)$ against l . The optimal network is equivalent to the one with minimal Wasserstein, i.e. $\beta = 2.0$. (h) TL against the of branching nodes. The closest network in terms of the number of bifurcations for $\beta = 1.5$). (i) Wasserstein similarity measure for the simulated Grenoble tram networks as a function of β , in the setting of eight origins and one destination. The most similar network in terms of this measure is at $\beta = 2$, when $W(G_1, G_2)$ is minimum. The peak at $\beta = 1.6$ is due to the absence of few edges in the rightmost part of the network that results into disconnecting a small branch, thus causing the distance to increase.

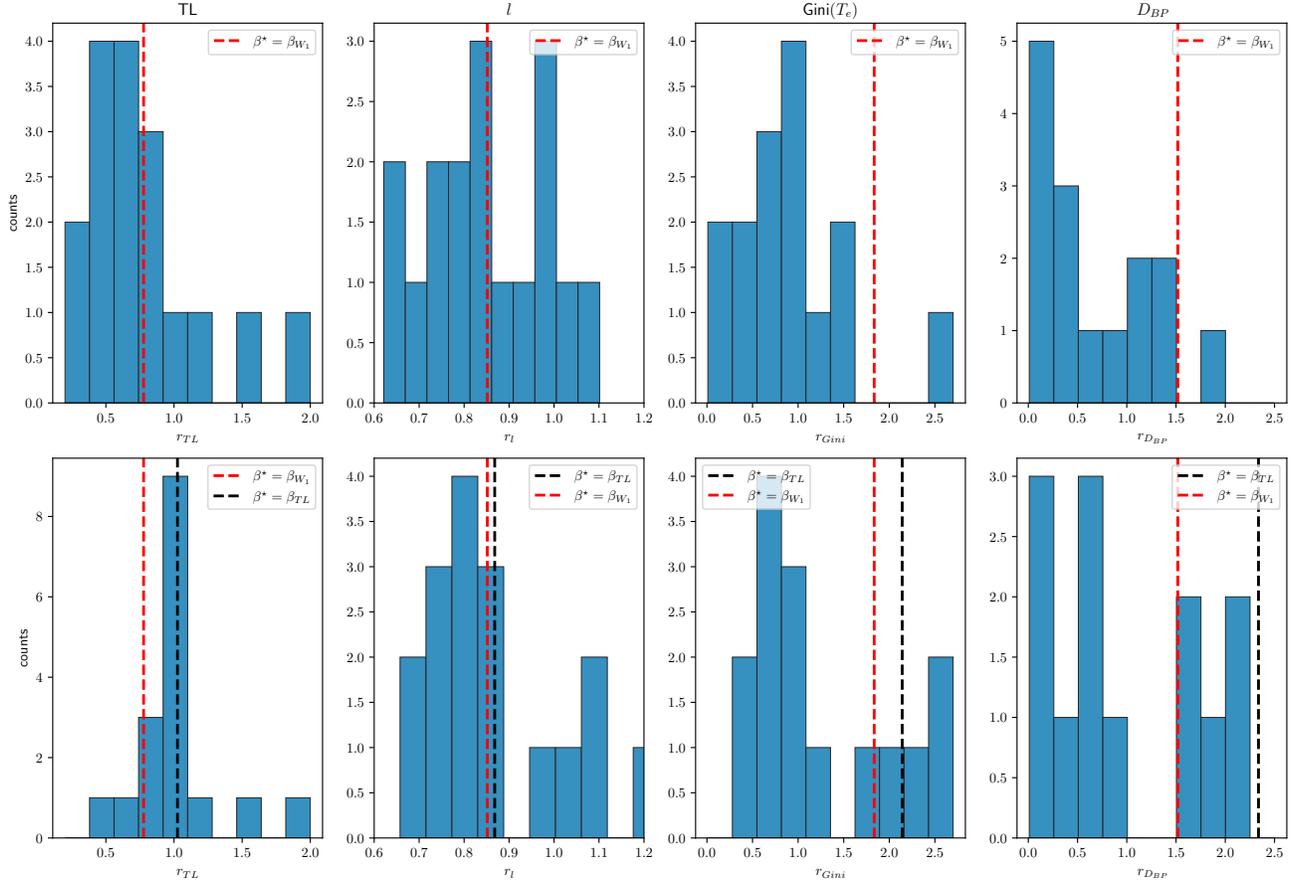


Figure 5. Counts for the metric ratios. Given a particular metric (e.g. Wasserstein, first row), we select the optimal β for each data point and measure the ratio across the multiple considered metrics. The red lines highlight the average for the optimal β given by the Wasserstein measure. We then repeat this procedure for the TL and highlight how the average of $\beta^* = \beta_{Wass}$ changes given other optimal β (the ratios for the remaining metrics can be found in the Supplementary Information).

Discussion

Planning transportation systems in a city is a challenging task. This study has shown that urban transportation systems can be simulated by simple principles based on optimal transport theory. Using empirical data from 16 cities and multiple transportation types, our model provides simulated networks obtained with little information in input and no a priori backbone network structure, exhibiting properties that closely resemble those observed in real transportation systems. Our model interpolates between various transportation regimes by tuning a single parameter, while allowing for a natural definition of a similarity measure to compare the simulated networks with those observed in real systems. We observed how the selected networks with this criterion exhibit transportation properties that on average resemble the corresponding real networks well.

A limitation of this study is that our OT-based model does not capture infrastructures with loops, thus limiting its applicability to rail networks, or subway and rail networks with a very small density of loops. Possible extensions of the formalism presented in this work to account for loops are an interesting direction for future work. Beyond simple heuristics, there are interesting modeling choices that one can make to effectively tackle this problem. For instance, one could generalize our approach to situations where travel demands are treated stochastically^{32–34} or change in time³⁵, scenarios where in certain regimes an OT-based approach can naturally lead to the formation of loops. Similarly, loops could emerge in multicommodity settings where fluxes of passengers are distinguished by their origin and destination stations, using an OT-based multicommodity framework as the one proposed in^{36–38}. Both directions, provided they could be generalized to a continuous setting as the one studied here, can potentially result in optimal simulated network infrastructures capturing properties that differ from the ones analyzed here, e.g. robustness to disruptions. Another limitation is the assumption that networks are static, i.e. do not change in time. It would be interesting to compare how differences between simulated and observed networks may arise because

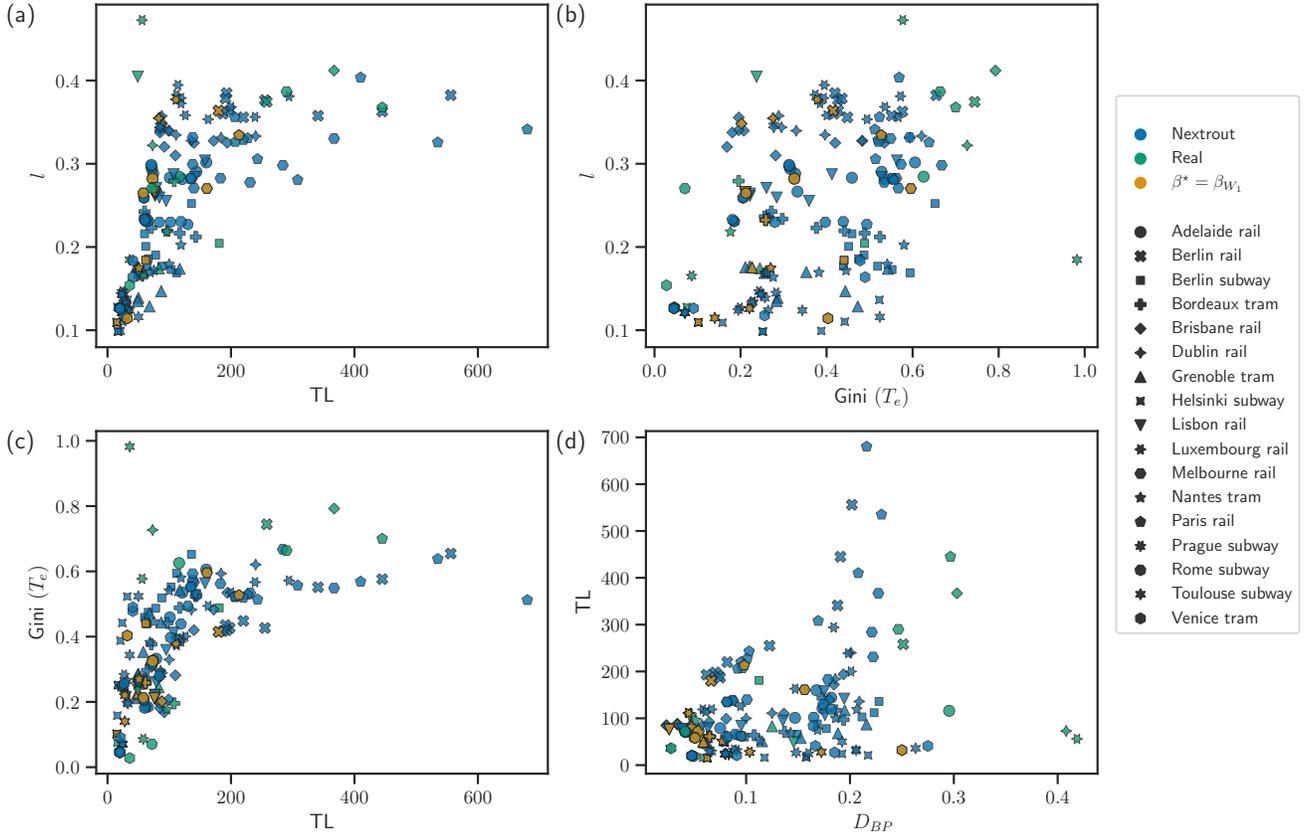


Figure 6. Comparison of simulated and observed networks. We show the values of the main transportation properties investigated in this work for real and simulated networks. Simulated networks cover a wider range of properties’ values, thus allowing in particular to select network that have lower or comparable values of these properties than those observed in the corresponding real networks.

different network branches may have been developed in different time periods. This could be used to study the evolution of transportation properties in time, as done in^{39,40}. Similarly, the networks studied in this work are often one layer of a multimode network. Integrating other transportation modes into a multilayer formalism and suitably adapting our OT-based approach, e.g. borrowing ideas from⁴¹, could give us a deeper understanding of optimal network design in interconnected urban systems.

In summary, there are many factors contributing to the development of urban transportation networks. However, our simple optimization scheme provides a principled benchmark for comparison with real-world networks. By interpolating between different transportation regimes, we find that the networks simulated by optimal transport principles have similar properties as those observed in real systems in many cases. In addition, comparing simulations resulting from different parameters’ values can give us indications on how to improve transportation performance. Our approach has the additional benefit of naturally providing a Wasserstein similarity measure that can be used as a graph comparison tool beyond the application to transportation networks studied here.

Methods

Data collection and analysis

We collected network data from various public transportation networks from 16 different cities¹⁸. Network statistics are detailed in Table 1. Each city had one or multiple transportation modes available. Node id’s were associated with longitude and latitude coordinates of real stations for multiple means of transportation, as well as possible connections between them. Our main goal was to analyze each network individually, therefore we did not address the multilayer case where connections among the different means of transportation exists. For instance, if rail and subway stops have the same coordinates, they are treated as distinct in each network.

To avoid possible redundancies or inconsistencies in the data, such as duplicated nodes or edges that looked too long,

we performed a preprocessing step. Specifically, we considered a threshold τ that corresponds to the minimum distance in kilometers between the pairs of nodes that had the same node id's and no connections between them, matching features stored in the node metadata such as names of the real stations. If the euclidean distance $d(i, j)$ between nodes i and j was smaller than this threshold, we collapsed the two nodes into one, i.e. $i = j$. This was used to avoid that entries and exits of each station would be counted as two distinct nodes in the same network, and possibly affecting the selection of origins and destinations.

To match the latitude and longitude coordinates of the datasets with those in the 2-dimensional plane that Nexttrout uses to solve the continuous problem, we re-scaled every pair (lon, lat) to a (0, 1) system of coordinates. Starting with a total of 62 data points (networks), we extracted the individual disconnected components and number of loops for each of them. Network extraction was performed on the biggest components only.

City	Transport mode	N	E	L_{ratio}	O	$D_{\{d_i, B_i\}}$
Berlin	subway	169	181	0.072	14	1
Bordeaux	tram	110	110	0.009	8	1
Grenoble	tram	80	83	0.048	8	1
Helsinki	subway	17	16	0.0	4	1
Lisbon	rail	48	49	0.041	9	1
Nantes	tram	97	96	0.0	10	1
Prague	subway	24	23	0.0	6	1
Rome	subway	73	72	0.0	6	1
Toulouse	subway	37	36	0.0	4	1
Venice	tram	37	38	0.053	4	1

Table 1. Description of real networks. We report the main network statistics as number of nodes N , number of edges E and loops ratio L_{ratio} defined as the number of loops divided by the number of edges, and selecting networks with $L_{ratio} < 0.2$, as higher values would require the recovery of loops in the extracted networks (see Supplementary Notes for the full table).

We extract networks using Nexttrout²³, selecting $1 < \beta \leq 2$ such that for every pair (*origins, destinations*) we simulate 10 different networks. Since the extracted networks may contain redundancies, we remove them using the graph filtering step from Nexttrout. Outputs of this step have less redundant structures and are closer to the optimal topologies.

We select origins and destinations based on the degree and betweenness centrality measures. The degree centrality d_i of a given node is defined as the number of edges connected to it. The betweenness centrality is defined as the frequency with which a node is on the shortest path between all other nodes,

$$B_i = \sum_{i \neq j \neq k} \frac{\sigma_{ik}(j)}{\sigma_{ik}},$$

where σ_{ik} is the total number of shortest paths from node i to node k and $\sigma_{ik}(j)$ those shortest paths passing through j .

Terminals were chosen as follows: nodes with $d_i \leq 1$ are assigned as origins, while those with $d_i = \max_n \{d_n\}$ or $B_i = \max_n \{B_n\}$ as destinations. We selected this set {origins, destinations} to be small, in order to use the least amount of information in input. In multiple cases the set of origins was equivalent in both centralities, with the difference being on the location of the destinations, thus the output networks were different. In terms of final networks properties, the results are comparable for both studies centralities (see SI for more details).

References

- [1] Zhang, J., Zhao, M., Liu, H. & Xu, X. Networked characteristics of the urban rail transit networks. *Phys. A: Stat. Mech. its Appl.* **392**, 1538–1546 (2013).
- [2] Ding, R. *et al.* Application of complex networks theory in urban traffic network researches. *Networks Spatial Econ.* **19**, 1281–1317 (2019).
- [3] Navlakha, S. & Bar-Joseph, Z. Algorithms in nature: the convergence of systems biology and computational thinking. *Mol. systems biology* **7**, 546 (2011).
- [4] Easley, D., Kleinberg, J. *et al.* *Networks, crowds, and markets*, vol. 8 (Cambridge university press Cambridge, 2010).
- [5] Watanabe, S., Tero, A., Takamatsu, A. & Nakagaki, T. Traffic optimization in railroad networks using an algorithm mimicking an amoeba-like organism, physarum plasmodium. *Biosystems* **105**, 225–232 (2011).

- [6] Liu, L., Song, Y., Zhang, H., Ma, H. & Vasilakos, A. V. Physarum optimization: A biology-inspired algorithm for the steiner tree problem in networks. *IEEE Transactions on Comput.* **64**, 818–831 (2013).
- [7] Aldous, D. & Barthelemy, M. Optimal geometry of transportation networks. *Phys. Rev. E* **99**, 052303 (2019).
- [8] Mc Gettrick, M. The role of city geometry in determining the utility of a small urban light rail/tram system. *Public Transp.* **12**, 233–259 (2020).
- [9] Barthelemy, M. Optimal transportation networks and network design. In *Spatial Networks*, 373–405 (Springer, 2022).
- [10] Cantarella, G. E., Pavone, G. & Vitetta, A. Heuristics for urban road network design: lane layout and signal settings. *Eur. J. Oper. Res.* **175**, 1682–1695 (2006).
- [11] Laporte, G., Mesa, J., Ortega, F. & Perea, F. Planning rapid transit networks. *Socio-Economic Plan. Sci.* **45**, 95–104 (2011).
- [12] Kay, R., Mattacchione, A., Katrycz, C. & Hatton, B. D. Stepwise slime mould growth as a template for urban design. *Sci. reports* **12**, 1–15 (2022).
- [13] Tero, A. *et al.* Rules for biologically inspired adaptive network design. *Science* **327**, 439–442 (2010).
- [14] Derrible, S. & Kennedy, C. Characterizing metro networks: state, form, and structure. *Transportation* **37**, 275–297 (2010).
- [15] Levinson, D. Network structure and city size. *PloS one* **7**, e29721 (2012).
- [16] Derrible, S. & Kennedy, C. The complexity and robustness of metro networks. *Phys. A: Stat. Mech. its Appl.* **389**, 3678–3691 (2010).
- [17] Lin, J. & Ban, Y. Complex network topology of transportation systems. *Transp. reviews* **33**, 658–685 (2013).
- [18] Kujala, R., Weckström, C., Darst, R. K., Mladenović, M. N. & Saramäki, J. A collection of public transport network data sets for 25 cities. *Sci. data* **5**, 1–14 (2018).
- [19] Facca, E., Cardin, F. & Putti, M. Towards a stationary Monge–Kantorovich dynamics: The physarum polycephalum experience. *SIAM J. on Appl. Math.* **78**, 651–676 (2018).
- [20] Facca, E., Daneri, S., Cardin, F. & Putti, M. Numerical solution of Monge–Kantorovich equations via a dynamic formulation. *J Sci Comput.* **82**, 1–26 (2020).
- [21] Facca, E., Cardin, F. & Putti, M. Branching structures emerging from a continuous optimal transport model. *J. Comput. Phys.* **447**, 110700 (2021).
- [22] Tero, A., Kobayashi, R. & Nakagaki, T. A mathematical model for adaptive transport network in path finding by true slime mold. *J. theoretical biology* **244**, 553–564 (2007).
- [23] Baptista, D., Leite, D., Facca, E., Putti, M. & De Bacco, C. Network extraction by routing optimization. *Sci. reports* **10**, 1–13 (2020).
- [24] Newman, M. *Networks* (Oxford university press, 2018).
- [25] Roth, C., Kang, S. M., Batty, M. & Barthelemy, M. A long-time limit for world subway networks. *J. The Royal Soc. Interface* **9**, 2540–2550 (2012).
- [26] Baptista, D. & De Bacco, C. Principled network extraction from images. *Royal Soc. Open Sci.* **8**, 210025 (2021).
- [27] Ibrahim, A. A., Leite, D. & De Bacco, C. Sustainable optimal transport in multilayer networks. *Phys. Rev. E* **105**, 064302 (2022).
- [28] Dixon, P. M., Weiner, J., Mitchell-Olds, T. & Woodley, R. Bootstrapping the gini coefficient of inequality. *Ecology* **68**, 1548–1551 (1987).
- [29] Santambrogio, F. Optimal transport for applied mathematicians. *Birkäuser, NY* **55**, 94 (2015).
- [30] Petric Maretic, H., El Gheche, M., Chierchia, G. & Frossard, P. Got: An optimal transport framework for graph comparison. In *Advances in Neural Information Processing Systems* **32**.
- [31] Xu, H., Luo, D., Zha, H. & Duke, L. C. Gromov-wasserstein learning for graph matching and node embedding. In *International conference on machine learning*, 6932–6941 (PMLR, 2019).
- [32] Corson, F. Fluctuations and redundancy in optimal transport networks. *Phys. Rev. Lett.* **104**, 048703 (2010).
- [33] Hu, D. & Cai, D. Adaptation and optimization of biological transport networks. *Phys. review letters* **111**, 138701 (2013).
- [34] Katifori, E., Szöllösi, G. J. & Magnasco, M. O. Damage and fluctuations induce loops in optimal transport networks. *Phys. review letters* **104**, 048704 (2010).

- [35] Lonardi, A., Facca, E., Putti, M. & De Bacco, C. Infrastructure adaptation and emergence of loops in network routing with time-dependent loads. *arXiv preprint arXiv:2112.10620* (2021).
- [36] Lonardi, A., Facca, E., Putti, M. & De Bacco, C. Designing optimal networks for multicommodity transport problem. *Phys. Rev. Res.* **3**, 043010 (2021).
- [37] Lonardi, A., Putti, M. & De Bacco, C. Multicommodity routing optimization for engineering networks. *Sci. reports* **12**, 1–11 (2022).
- [38] Bonifaci, V. *et al.* Physarum-inspired multi-commodity flow dynamics. *Theor. Comput. Sci.* (2022).
- [39] Pei, A., Xiao, F., Yu, S. & Li, L. Efficiency in the evolution of metro networks. *Sci. Reports* **12**, 1–10 (2022).
- [40] Cats, O. Topological evolution of a metropolitan rail transport network: The case of stockholm. *J. Transp. Geogr.* **62**, 172–183 (2017).
- [41] Ibrahim, A. A., Lonardi, A. & Bacco, C. D. Optimal transport in multilayer networks for traffic flow optimization. *Algorithms* **14**, 189 (2021).

Acknowledgements

The authors thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Daniela Leite.

Data Availability

The synthetic data can be obtained from the corresponding author upon request.

Additional information

Accession codes: open source codes and executables are available at <https://github.com/Danielaleite/opt-urban-nexttrout>.

Competing interests. The authors declare no competing interests.

Supplementary Notes

Supplementary note 1: Metric ratios

We have already shown that the final optimal network changes depending on specific selected transportation properties. But how distinct are these networks and how do such properties favor different optimal β ? In Fig. S1 we show an example for the subway network of Berlin to illustrate this difference. While $\beta^* = \beta_{Gini}$ favors a network with more branches and wider coverage, for $\beta^* = \beta_{D_{BP}}$ the optimal network is that with fewer branches.

To evaluate this difference across multiple properties, we use the metric ratios to compute the distance between observed and real networks. We have already shown the ratios for Wasserstein and cost (TL) in the main text. Fig S2 contain the counts for the other three studied metrics. We notice that the Wasserstein encourages smaller path length, but higher traffic and density of bifurcation points (which is confirmed by the minimum mean displacement).

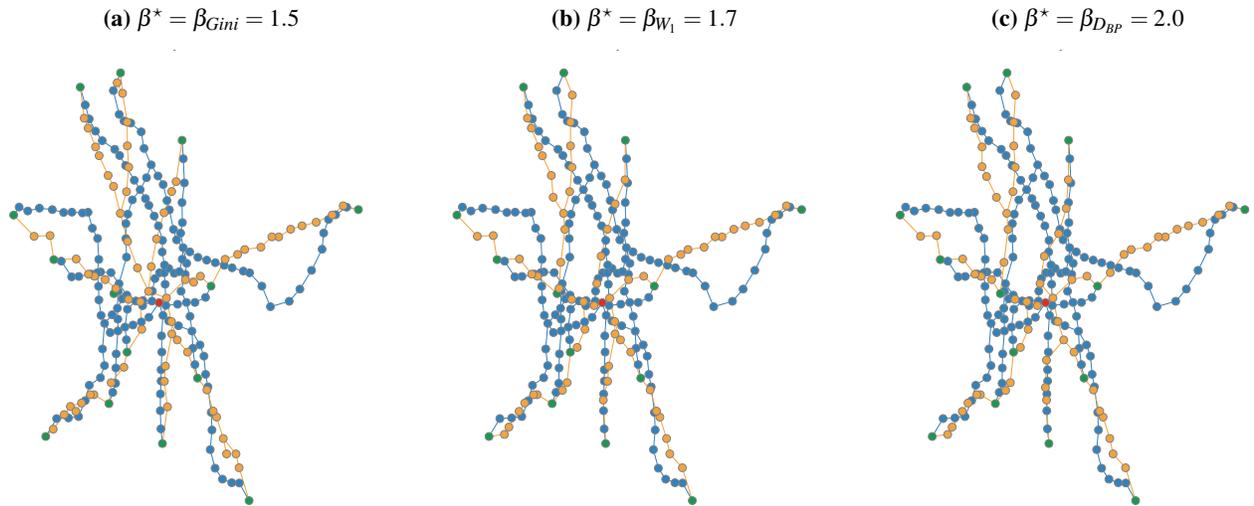


Figure S1. Optimal networks for the subway network of Berlin. In (a) the optimal β is that given by the *Gini*, which tends to favor the networks with more branches and wider coverage to distribute the traffic; in (b) we show the optimal network as selected by the Wasserstein measure, whilst in (c) we show the one given by the density of branching points, which tends to favor networks with fewer branches, so higher values of β .

Supplementary note 2: Centrality criteria for origins and destinations selection

The first step of our pipeline consists in defining the set of Origin-Destination pairs (OD) taken from a real network. We perform a preprocessing step to remove possible redundancies found in the original spatial data collected from¹⁸.

We map each stop, which consists in a pair longitude-latitude, to a node in a $[0, 1]$ system of coordinates as a starting point. We then perform the preprocessing to obtain the original network mapping, and compute different network centralities to define the set of OD points: nodes with lowest values of degree, betweenness and closeness centrality are set to be origins and those with the highest values of these properties are set as destinations. In nearly all networks we noticed that the selected nodes for betweenness and closeness were equivalent, therefore the Nexttrout generated networks would have the same set of nodes and edges (and hence the measured properties would be the equivalent), so we find it enough to show the results for networks generated using only degree and betweenness centrality.

As we have already shown results obtained for the networks with OD points selected based on the degree, we now present the network properties obtained based on the betweenness centrality. Notice that despite the absolute number of destinations being the same, they still differ in terms of location, which will impact the final networks generated using our pipeline. The final betweenness properties are shown in Fig. S3 and Fig S4. Supplementary Table 1 shows the full list of studied networks and its corresponding network properties (number of nodes and edges, loops ratio, number of origins and destinations).

Supplementary note 3: Weights for the Wasserstein measure

The Wasserstein similarity measure (W_1) is designed to capture the amount of ‘effort’ necessary to move information between the original network and the one extracted using our pipeline. Since it is defined as the weighted sum of the optimal fluxes, the choice of such weights might influence the final optimal network. We compared the differences of selecting unitary or euclidean weights in Fig. S3.

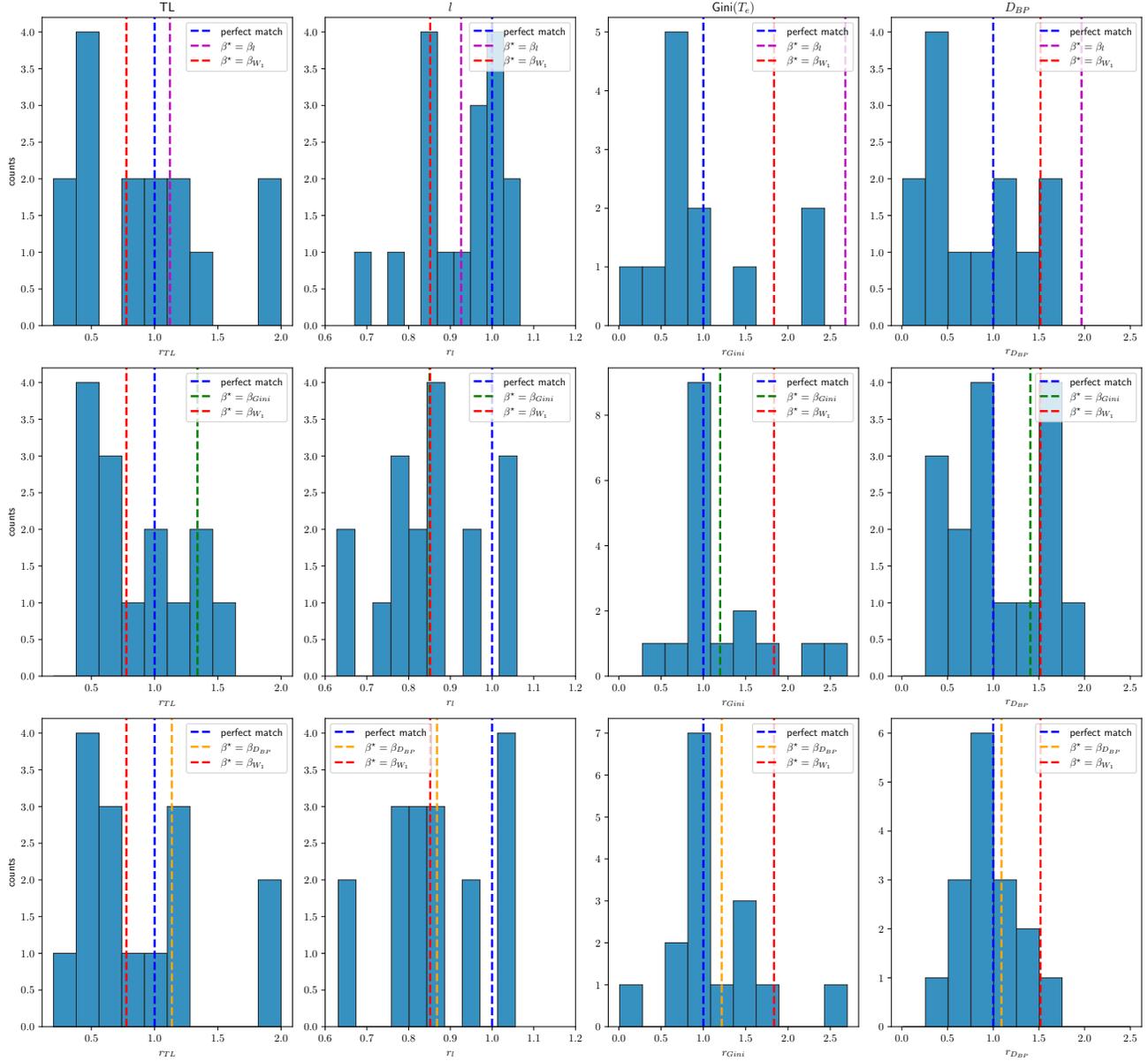


Figure S2. Metric ratios for the path length l , Gini and density of branching points. We measure the ratio r_p across the three metrics, highlighting how the average of $\beta^* = \beta_{Wass}$ - given by the red lines - changes given other optimal β , and how it deviates from the ‘perfect match’, i.e. when $r_p = 1$.

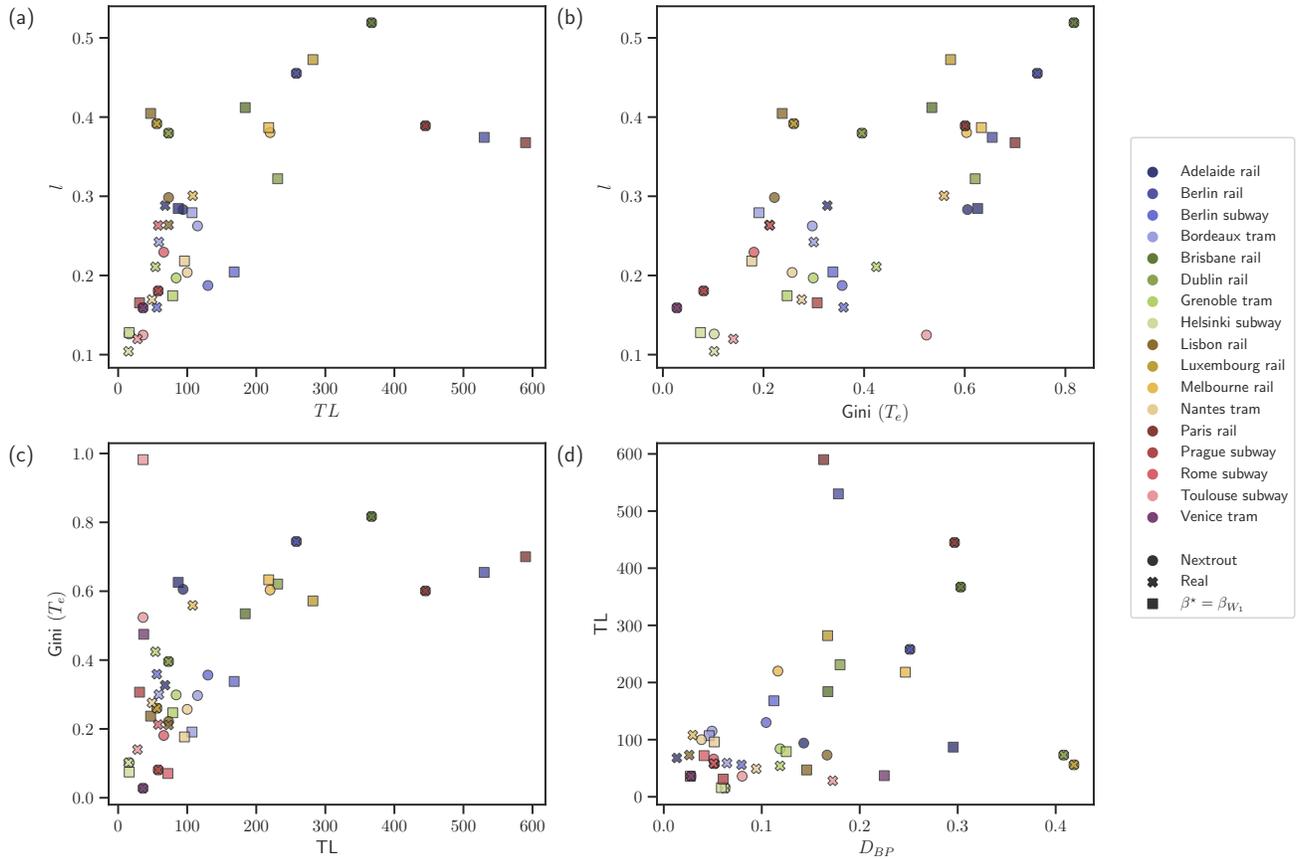


Figure S3. Measures for the 17 studied networks. Each dataset is assigned to a different color, market shapes distinguish real and simulated networks. The ladder are further distinguished based on the one generated via Nexttrout (having betweenness as the destinations selection criteria) that gives the closest point in terms of the metrics plotted in the figure (circle) or the one corresponding to the best Wasserstein measure (square). (a) We measure Cost (TL) for both simulated and real networks, plotted against the total path length l . (b) Gini coefficient as a measure of traffic distribution, versus the total path length. (c) Traffic distribution in terms of the Cost. (d) Density of bifurcations plotted against the cost.

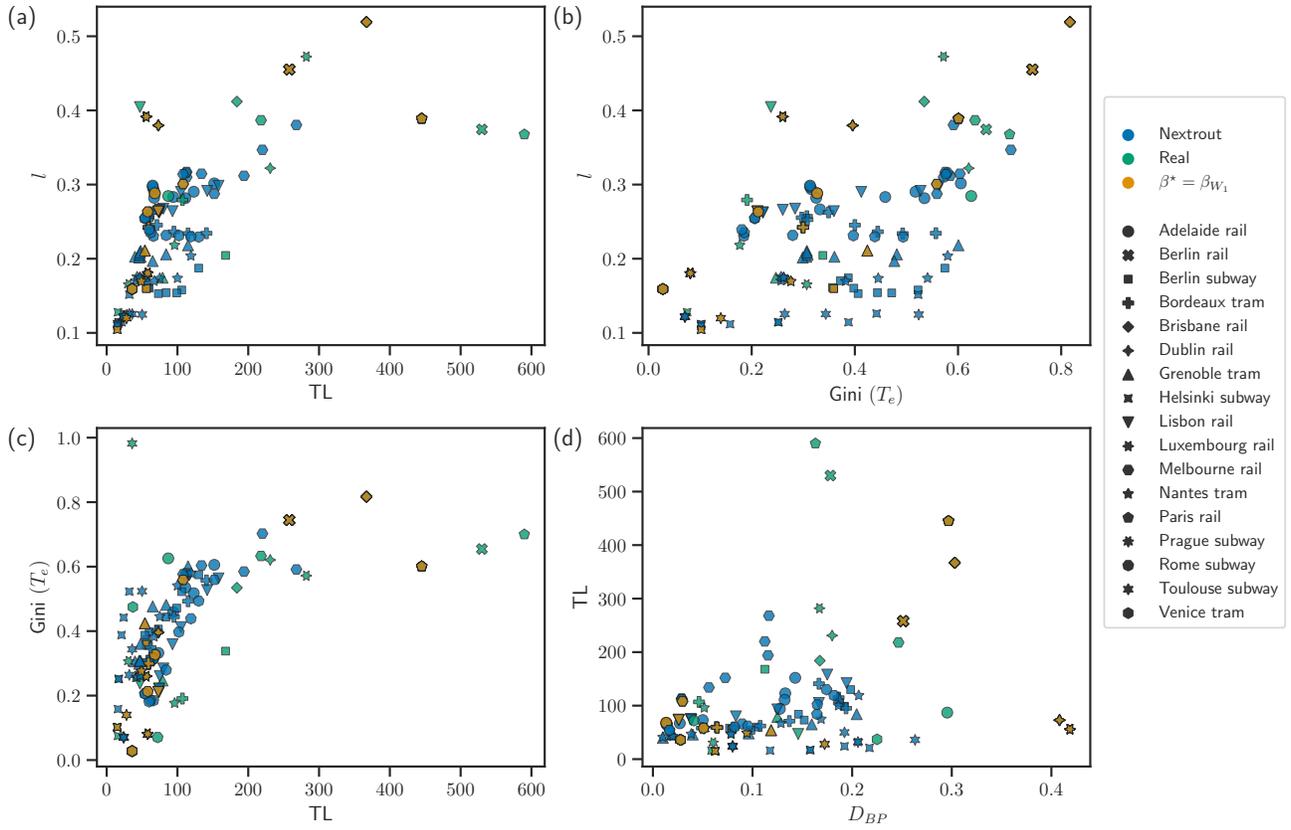


Figure S4. Comparison of simulated networks using the betweenness criteria and the one from observed data. We show the values of the main transportation properties investigated in this work for real and simulated networks. Simulated networks cover a wider range of properties' values, thus allowing in particular to select network that have lower or comparable values of these properties than those observed in the corresponding real networks.

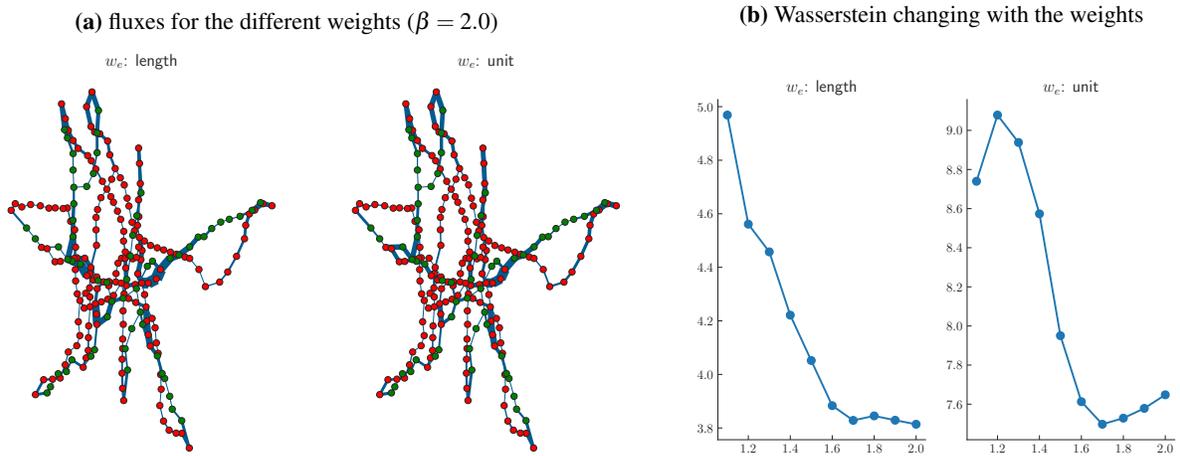


Figure S5. Selecting weights for the Wasserstein similarity measure. (a) Given the source graph (red nodes, real network) and the sink graph (green nodes, extracted network), we observe how the optimal flux distributes along the union graph. The euclidean length (left) distributes the flux along the more central nodes, thus capturing more realistic scenarios, whilst unitary weights distribute more equally along the network. We then compare in (b) how the Wasserstein measure changes across different values of β . We design the Wasserstein measure to capture more realistic scenarios, thus for smaller values of β , the networks are expected to have more redundant nodes, so we expect this measure to be higher, and the opposite as β increases. This pattern seems to be better captured by the euclidean weights.

City	Transport mode	N	E	L_{ratio}	O	$D_{\{d_i, B_i\}}$
Adelaide	rail	88	116	0.25	6	1
Berlin	rail	203	258	0.21	19	1
Berlin	subway	169	181	0.072	14	1
Bordeaux	tram	110	110	0.009	8	1
Brisbane	rail	297	367	0.193	10	1
Dublin	rail	59	73	0.03	10	1
Grenoble	tram	80	83	0.048	8	1
Helsinki	subway	17	16	0.0	4	1
Lisbon	rail	48	49	0.041	9	1
Luxembourg	rail	43	56	0.025	11	1
Melbourne	rail	219	290	0.248	22	1
Nantes	tram	97	96	0.0	10	1
Paris	rail	337	445	0.244	24	1
Prague	subway	24	23	0.0	6	1
Rome	subway	73	72	0.0	6	1
Toulouse	subway	37	36	0.0	4	1
Venice	tram	37	38	0.053	4	1

Supplementary Table 1. Full description of the 17 studied real networks. We report main network statistics as number of nodes N , number of edges E and loops ratio L_{ratio} defined as the number of loops divided by the number of edges, and selecting networks with $L_{ratio} < 0.27$, as higher values would require the recovery of loops in the extracted networks. We also report the selected number of origins and destinations used.

Supplementary note 4: Traffic distribution

We also show how traffic is distributed along the network. We measure traffic on edges (T_e) by setting the same origins and destination nodes used on the first step of our extraction pipeline, and running the discrete DMK-dynamics with $\beta = 1.5$. Fig. S6 shows that with this setting, both extracted and real networks tend to distribute traffic towards the more central edges, which is usually observed in real transportation systems, where stations with more connections register higher flow of passengers.

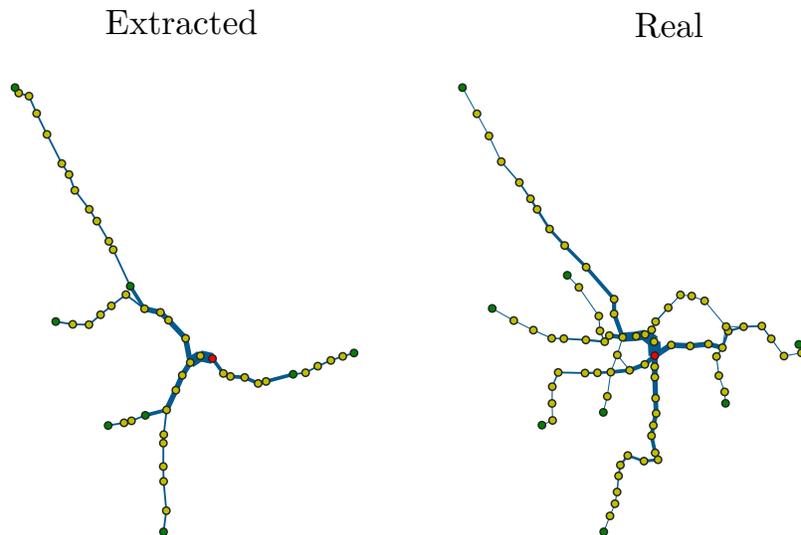


Figure S6. Traffic in extracted and real networks. Using the same set of origins and destinations we observe a higher distribution of traffic towards more central edges in both cases.