

# Lawrence Berkeley National Laboratory

## LBL Publications

### Title

Understanding congested travel in urban areas.

### Permalink

<https://escholarship.org/uc/item/906783m4>

### Authors

Çolak, Serdar

Lima, Antonio

González, Marta

### Publication Date

2016-03-15

### DOI

10.1038/ncomms10793

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

ARTICLE

Received 11 Aug 2015 | Accepted 21 Jan 2016 | Published 15 Mar 2016

DOI: 10.1038/ncomms10793

OPEN

# Understanding congested travel in urban areas

Serdar Çolak<sup>1</sup>, Antonio Lima<sup>1,2</sup> & Marta C. González<sup>1,3</sup>

Rapid urbanization and increasing demand for transportation burdens urban road infrastructures. The interplay of number of vehicles and available road capacity on their routes determines the level of congestion. Although approaches to modify demand and capacity exist, the possible limits of congestion alleviation by only modifying route choices have not been systematically studied. Here we couple the road networks of five diverse cities with the travel demand profiles in the morning peak hour obtained from billions of mobile phone traces to comprehensively analyse urban traffic. We present that a dimensionless ratio of the road supply to the travel demand explains the percentage of time lost in congestion. Finally, we examine congestion relief under a centralized routing scheme with varying levels of awareness of social good and quantify the benefits to show that moderate levels are enough to achieve significant collective travel time savings.

<sup>1</sup>Department of Civil and Environmental Engineering, MIT, Cambridge, Massachusetts 02139, USA. <sup>2</sup>School of Computer Science, University of Birmingham, Edgbaston B15 2TT, UK. <sup>3</sup>Engineering Systems Division, MIT, Cambridge, Massachusetts 02139, USA. Correspondence and requests for materials should be addressed to M.C.G. (email: martag@mit.edu).

Cities have a long-standing history cultivating technological innovations that allow citizens to efficiently access goods and opportunities. However, the ease of access has been increasingly difficult to maintain under rapid urbanization<sup>1–7</sup>. As growing population densities create excessive demand for cities' infrastructure, the increasing penetration and advancement of technology generates massive amounts of multidimensional data that can be used to study and mitigate this demand. Specifically, the availability of mobile phone data has led researchers to quantify fundamental spatiotemporal patterns to better understand human mobility in urban areas<sup>8–12</sup>. With the continuous increase in the volume and accuracy of new data sources, new methods that process and distill mobile phone data are consistently refined, and traditional models of mobility such as the gravity-, radiation- or activity-based models are being updated in tandem<sup>13–18</sup>. In the context of travel demand estimation, previous efforts focused on developing models that combine household travel surveys with census and land-use information<sup>19,20</sup>. Despite the robust methodology and meticulous implementation of these models, the high costs associated with obtaining the infrequent and small data have proven to be the bottleneck. To supplement these approaches, traffic simulations and demand estimation models have begun incorporating big data sources into their forecasts, building portable data pipelines to create data-driven decision-making tools for policy makers<sup>21–23</sup>.

Understanding of the complex interplay of road infrastructure and travel patterns to model travel times and congestion in not a single city but many at once has been a particular challenge in this line of research<sup>24–26</sup>. Road networks, the circulatory system sustaining a city's accessibility and cultivating its economic prosperity<sup>27–29</sup> are seized with congestion in most large metropolitan areas. In their 2013 report, TomTom, a leading GPS company, states that in cities such as Moscow, Istanbul, Rio de Janeiro, Mexico City and Beijing, people on average spend >75% extra time travelling due to traffic. The resulting loss of time, money and energy are borne by the city's citizens and travellers. Municipalities continually invest in road infrastructure construction and maintenance to increase supply, although controversies on whether more roads alleviate congestion persist<sup>30</sup>. Other efforts to reduce congestion aim to decrease driving demand by promoting alternative travel modes, high-occupancy driving lanes, carpooling, congestion pricing and, in extreme cases, road space rationing. Even with all these measures, congestion remains inherent and drivers are increasingly leveraging real-time information through GPS devices and online routing tools to move faster. With everyone having easy access to traffic information, drivers make decisions without coordination based on near-perfect information, resulting in suboptimal system configuration. This general trend of using raw real-time information in decision-making has significant implications, as it might be also used as a tool to guide drivers to make choices for the benefit of the city, thus creating a more optimal traffic configuration. The extent of the global inefficiency has been of great interest<sup>31–34</sup> in many contexts, ranging from wireless networks to transportation<sup>35–40</sup>. Theoretical approaches to bring the system to optimality generally converge to marginal cost taxation, which essentially forms the basis of congestion pricing schemes today<sup>41,42</sup>. Despite the abundance of research on optimal flow configurations and their implications in the transportation, urban planning and economics literature, there is a shortage of works that use big data sources to understand the role of travel demand and actual travel times in metropolitan regions when comparing cities. This highlights a need to build a framework that can be replicated to systematically generate meaningful travel times to not only understand cities better but

also test solutions to urban problems such as congestion or pollution.

In this work, we address this issue by coupling travel demand profiles and travel time estimates to analyse how efficiently people move across cities. We begin by modelling the supply by parsing publicly available OpenStreetMap data to obtain road networks. To model travel demand, we mine massive mobile phone data sets, also referred to as call detail records (CDRs)<sup>43</sup>. This procedure requires home and work location detection for millions of users, mining of their location shifts, and the proper sampling procedures to represent accurately the trip tables for the whole city (see Supplementary Notes 1–3). Using this information of the trip distribution within the city, we estimate morning peak vehicular volumes from origins to destinations and compare the inferred travel times based on demand with the estimates of an online map provider in the respective routes and hour of the day. We then explore the relationship between travel distance and travel time across many cities. We show that the time lost due to congestion in each city can be accounted by a dimensionless parameter  $\Gamma$  that measures the ratio between the vehicular travel demand and the road infrastructure supply for the city. To a lesser extent, the differences in congestion levels depend on the population density and the spatial distribution of population. Next, we calculate the detrimental effects of selfish routing by comparing obtained travel times to those that would be observed if the routes were selected to attain the social optimum. We then explore the bounds of the benefits of leveraging information technologies to influence route choices in ways that would help create a more optimal system configuration for vehicular travel. To do so, we implement a generalized selfish routing model that generates expected travel times for varying levels of consideration of overall social good, or  $\lambda$ . We analyse the system gains of socially aware driver behaviour, as well as exploring the distributions of benefits and losses at the individual level. We present our findings for five major cities around the world: Boston and San Francisco Bay Area in the United States, Rio de Janeiro in Brazil, and Lisbon and Porto in Portugal.

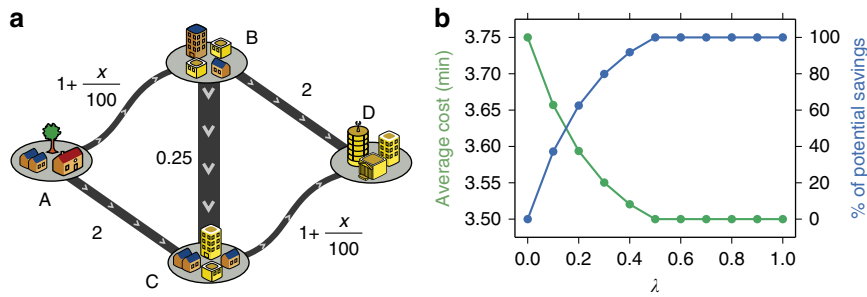
## Results

**Approach.** We formalize the traffic problem by modelling route choice as follows: every driver  $i$  makes a choice of the route  $p$  to their destination. This choice depends on a personal utility  $u_i = \sum_{e \in p} c_e(x_e)$ , expressed as the sum of the costs  $c$  of every road segment  $e$  along the chosen route. For simplicity, we assume that the cost of a road segment for driver  $i$  is equal to the travel time,  $c_e(x_e) = t_e(x_e)$ , where  $t_e(x_e)$  represents the travel time  $t$  observed on road  $e$  for vehicle flow  $x_e$ . We can then define the total cost incurred by all users as  $C = \sum_{e \in E} x_e t_e(x_e)$ . The flow configuration that results in the optimal cost is referred to as the socially optimal flows obtained by a typical minimum cost network flow programme<sup>44</sup>:

$$\begin{aligned} & \text{Minimize}_{x_e \forall e \in E} C \\ & \text{subject to} \quad \sum_p f_p^{st} = f^{st} \\ & \quad x_e = \sum_s \sum_t \sum_p f_p^{st} \delta^{st}(p, e), \\ & \quad x_e \geq 0, f_p^{st} \geq 0. \end{aligned} \quad (1)$$

where  $x_e$  refers to the flow on road  $e$ ,  $f_p^{st}$  is the flow between the source  $s$  and target  $t$  on route  $p$ , and  $\delta^{st}(p, e) = 1$  when road  $e$  lies on route  $p$ .

As drivers make selfish choices, the system settles into a suboptimal state. Although driver  $i$  only experiences and considers his/her own travel time, the cost the whole system incurs also includes the marginal cost driver  $i$  imposes on all



**Figure 1 | Illustration of routing equilibrium.** (a) In this small network, 100 drivers are going from A to D. The road labels represent the costs of travel as a function of vehicle flows. User equilibrium allocates the flows between paths as  $f_{ABD} = f_{ACD} = 25$  and  $f_{ABCD} = 50$ , and the average travel time is 3.75 min for all drivers. Socially optimal flows decrease total travel time to 3.5 by  $f_{ABD} = f_{ACD} = 50$  and  $f_{ABCD} = 0$ , with road BC remaining unused. (b) Achieved percentage of potential savings for increasing values of social good weight  $\lambda$ : 10 and 20% social good weight results in 40 and 60% of potential savings, respectively.

other drivers on the road segments he/she takes. The set of flows that occur when every driver minimizes their own travel time is referred to as the user equilibrium flows. Theoretically, in the resulting system state, no driver can benefit from deviating from their route. This idea, essentially describing a Nash equilibrium in roads, is captured in Wardrop’s principles in transportation<sup>36</sup>: the journey times on all the used routes for an origin–destination (OD) pair are equal and are less than those that would be experienced by a single vehicle on any unused route. This routing game is solved through a potential function  $\phi_e(x_e) = \int_0^{x_e} t_e(x) dx$  such that  $\phi'_e(x_e) = t_e(x_e)$  (ref. 45). The convex programme for the user equilibrium problem has been formulated<sup>46</sup> as follows:

$$\begin{aligned} & \text{minimize} && \sum_{e \in E} \phi_e = \sum_{e \in E} \int_0^{x_e} t_e(x) dx \\ & \text{subject to} && \text{constraints in Eq. 1.} \end{aligned} \tag{2}$$

Figure 1a depicts an example that captures solutions for equilibrium and optimal flows for a widely used toy network. For the demand of  $d_{AD} = 100$ , the user equilibrium flows allocate 50 drivers on path ABCD and 25 drivers on paths ABD and ACD each, resulting in a travel time from A to D of 3.75, regardless of the path chosen. The socially optimal configuration avoids allocating too much flow on the path ABCD, as its marginal cost is higher than those of paths ABD and ACD. By minimizing the marginal cost, path ABCD receives no flow and the average cost is minimized at 3.5.

To assess the benefits of different scenarios based on travel demand information, we make use of the formulation proposed in ref. 47. We reconfigure the utility function of a driver as a linear combination of the cost he/she will incur and the total marginal cost his/her choice imposes on everyone else:

$$\begin{aligned} c_e^\lambda(x_e) &= (1 - \lambda)t_e(x_e) + \lambda \frac{d[x_e t_e(x_e)]}{dx_e} \\ &= t_e(x_e) + \lambda x_e \frac{dt_e(x_e)}{dx_e} \end{aligned} \tag{3}$$

$\lambda$  defines the weight towards social good; it is a parameter ranging between 0 and 1. A driver with  $\lambda = 1$  chooses routes with respect to the marginal costs, thus moving the system closer to the system optimum. Conversely, a user with  $\lambda = 0$  only considers the cost of his route and potentially moves the system away from optimality. The resulting convex programme for the socially aware routing problem is as follows:

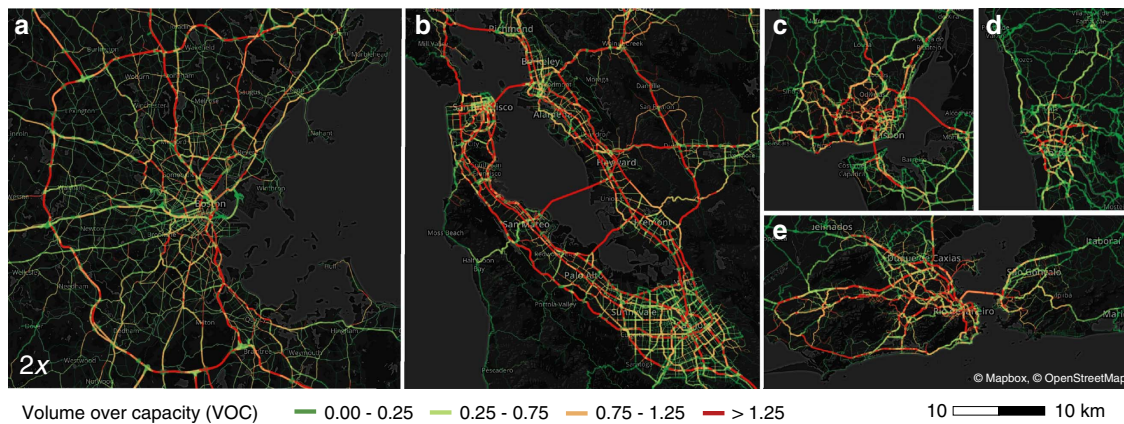
$$\begin{aligned} & \text{minimize} && \sum_{e \in E} \int_0^{x_e} c_e^\lambda(x_e) \cdot x_e dx_e \\ & \text{subject to} && \text{constraints in Eq. 1.} \end{aligned} \tag{4}$$

For the city depicted in Fig. 1a, the user equilibrium configuration results in an average cost of 3.75 min per driver versus 3.5 min the

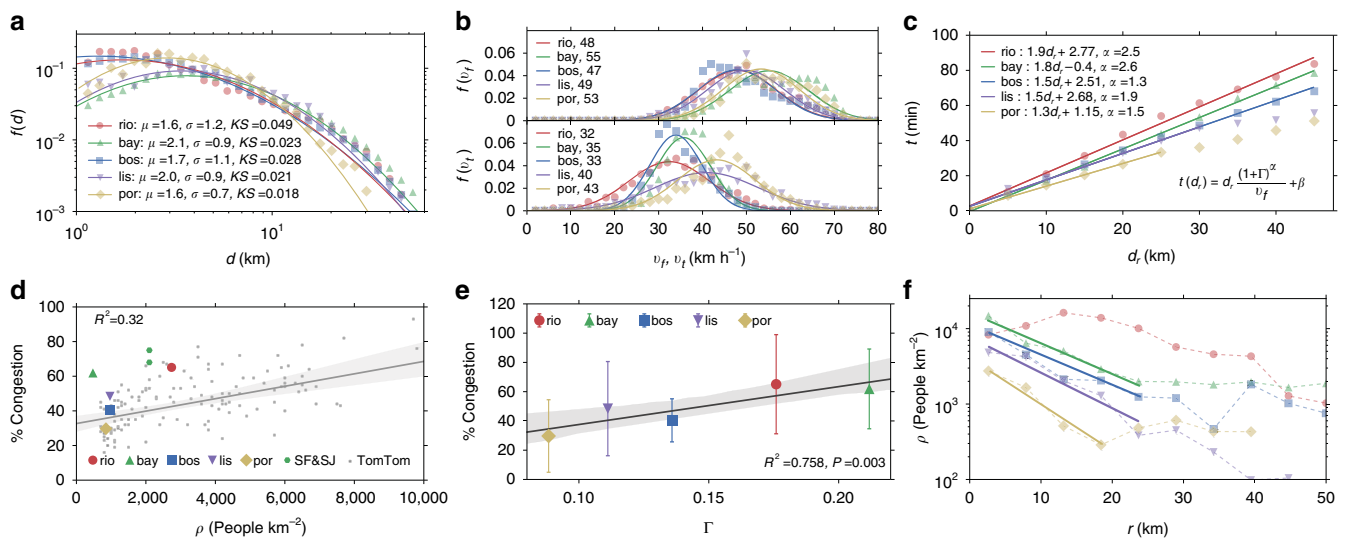
system optimum, meaning solely by adjusting routing behaviour to  $\lambda = 1$ , a benefit of 0.25 min can be achieved per driver. Figure 1b shows that for  $\lambda = 0.1$ , when the drivers begin valuing social good as well, the average cost drops to  $\sim 3.65$  and almost 40% of potential savings are realized. In fact, the social optimum is achieved at  $\lambda = 0.5$ .

**Travel times.** To understand the relationship between travel demand and driving travel times, we begin by comparing our five cities during estimated morning peak period traffic conditions. The areas of analysis are significantly diverse: Rio is very highly populated over its large extensions, whereas Porto’s population density considerably decreases after  $r > 20$  km from the most dense location. Rio de Janeiro, the Bay Area and Lisbon extend across Guanabara Bay, the Bay and Tagus, respectively, and have many inhabitants commuting on few bridges (see Supplementary Fig. 3 and Supplementary Table 1 for more details). As a consequence of their differences, cities demonstrate varying traffic conditions, as shown in Fig. 2. The volume-over-capacity ratio (VOC) measures how successfully a road segment is able to cope with the assigned volume of vehicles, with high VOC values indicating more congestion. High VOCs are generally observed on highways, as they provide faster means of travel due to their wider roads, increased number of lanes and higher speed limits. In addition, bridges and roads that lie central in the network topology are typically congested due to a lack of alternative routes.

We begin by analysing the efficiency of urban mobility for the five regions to understand the mechanisms underlying observed travel times. The main determinant of congestion is travel demand, which is heavily tied to commuting trip distances during weekday peak travel times. In Fig. 3a, we demonstrate that the straight-line (Euclidean) commuting distances,  $d$ , follow a lognormal distribution,  $f(d) = \frac{1}{\sqrt{2\pi}\sigma d} e^{-\frac{(\ln(d) - \mu)^2}{2\sigma^2}}$  with means ranging from 5 to 8 km ( $\mu = 1.6$ – $2.1$ ) and s.d. ranging from 2 to 4 km ( $\sigma = 0.7$ – $1.2$ ) (see Supplementary Fig. 7). It can be observed that majority of trips span relatively short distances and trips over 25 km are uncommon. However, what makes a city more traversable are the speeds at which drivers can span these distances. In Fig. 3b we investigate the effective speeds in both free and congested traffic conditions. It can be observed that cities exhibit similar free travel-speed distributions, normally distributed with  $\mu$  fluctuating around  $50 \text{ km h}^{-1}$  with mean values reported in the legend. The differences in road network supply  $S = \sum_{x_e > 0, e \in E} l_e C_e$  (km vehicles per hour), where  $l_e$  and  $C_e$  are the length (km) and the flow capacity (vehicles per hour) of a road segment  $e$ , explains the slight differences in free flow speeds, as



**Figure 2 | The maps of VOCs (volume over capacity) of the roads in the user equilibrium configuration.** The depicted cities are (a) Boston, USA, (b) San Francisco Bay Area, USA, (c) Lisbon, Portugal, (d) Porto, Portugal, and (e) Rio de Janeiro, Brazil. Higher VOCs are generally observed in highways, as they provide faster means of travel. (Boston is 2x the distance scale.) Maps under © OpenStreetMap contributors BY-SA.



**Figure 3 | Comparisons of cities and their congested travel.** (a) Distributions of commuting trip distances,  $d$ , in the morning peak period with parameters of the fitted lognormal distribution depicted in the legend (see Supplementary Fig. 7 and Supplementary Table 2 for more detail). (b) Distribution of trip free flow speeds,  $v_f$ , and in traffic conditions,  $v_t$ . (c) Commuting travel times versus route distances of commuters,  $d_r$ . (d) Estimates of overall mean % of time lost in congestion versus population density  $\rho$  for TomTom Traffic Index estimates and our analysis. (e) Relationship of overall mean % congestion to the demand to supply ratio,  $\Gamma$ , for the five subject cities, with error bars specifying the s.d. (see Supplementary Fig. 8). (f) Average population density  $\rho$  as a function of distance from the most dense area in the region,  $r$ .

seen in Table 1. These differences are significantly more apparent in speed distributions under real traffic conditions: the effective OD travel speeds in Rio, the Bay Area and Boston decay considerably compared with those in free traffic conditions, whereas the speeds in Porto and Lisbon change less. We explore further these two different responses given the demand profiles of each city.

To that end, we analyse the experienced travel times per distances travelled in Fig. 3c. We observe a strong yet very simple relationship that pronounces the differences between the subject cities: Rio de Janeiro is the slowest city and is followed next by the Bay Area, and Porto is the fastest. All cities exhibit a linear relationship, with the exception of long-distance trips in Porto and Lisbon where a different regime appears for longer distances. To explain this observation, we model travel times by city-specific parameters describing the demand, the capacity and observed free traffic speeds. In doing so, we define

demand-to-supply ratio of a city as

$$\Gamma = \frac{\sum_{e \in E} l_e x_e}{\sum_{x_e > 0, e \in E} l_e C_e} \quad (5)$$

This dimensionless measure is a simple ratio of the total distance travelled by all vehicles to the upper bound of the total vehicle kilometres the road network can support per hour, thus capturing the load on the road infrastructure by bringing together trip distances, trip magnitudes, road capacities and the distances they span as shown in Table 1. Using this measure along with  $v_f$ , the average free travel speed of each city, we are able to better explain the linear relationship between travel time and distance by

$$t(d_r) = d_r \frac{(1 + \Gamma)^\alpha}{v_f} + \beta, \quad (6)$$



**Table 1 | A comparison of general properties of the subject cities.**

	City				
	Rio	SF Bay	Boston	Lisbon	Porto
Population (millions)	12.6	7.15	4.5	2.8	1.7
Area (1,000 km <sup>2</sup> )	4.6	18.1	4.6	2.9	2.0
Demand (vehicle km h <sup>-1</sup> )	3.1	9.1	5.4	2.9	1.1
Supply (vehicle km h <sup>-1</sup> )	17.6	43.0	39.7	25.5	11.7
Demand-to-supply ( $\Gamma$ )	0.18	0.21	0.14	0.11	0.09
Expansion factor	890	100	32	96	164
Vehicle usage (vehicle per person)	0.25	0.67	0.67	0.56	0.62

**Table 2 | Comparison of cost findings in the subject cities for the morning peak hour.**

(min)	City				
	Rio	SF Bay	Boston	Lisbon	Porto
FTT	20.6	21.1	19.3	22.4	15.3
<b>Loss</b>	<b>14.1</b>	<b>12.5</b>	<b>8.2</b>	<b>8.0</b>	<b>4.0</b>
UE	34.7	33.6	27.5	30.4	19.3
<b>Benefit</b>	<b>2.6</b>	<b>2.6</b>	<b>1.3</b>	<b>2.1</b>	<b>1.1</b>
SO	32.1	31.0	26.2	28.3	18.2
%S	18	21	16	27	28

FTT, free travel time; SO, social optimum; UE, user equilibrium; % S, percentage of total congestion attributed to selfish routing, defined as  $S = 100 \cdot \text{Benefit} / \text{Loss}$ . Bold rows indicate the loss of travel times from free travel times to socially optimal flows, and from socially optimal flows to user equilibrium flows for commuters, respectively.

where  $\alpha$ -values vary between 1.3 and 2.5, essentially describing the sensitivity of the city to the stress imposed by travel demand on its road infrastructure.

To untangle the particular ordering of cities in terms of speed and understand why some cities are more congested than others, we investigate a typical relationship in Fig. 3d, to test the common conception that cities with higher population densities tend to exhibit more heterogeneity in their demand profiles, and therefore tend to be more congested. For this purpose, we measure the ratio of the time lost in traffic to the travel time under free flow conditions, known as the traffic index, along with those measured for many other urban areas by TomTom, a leading GPS company. We consider the percentage of congestion, defined as the percentage of additional travel time due to traffic compared with free flow conditions, for different population densities in these various cities. We observe that Boston, Lisbon and Porto fall on the fit model, whereas the Bay Area and Rio demonstrates a significantly higher level of congestion. The outlier appearance of the Bay Area is a consequence of the arbitrary definitions of urban areas and its influence in population density as pointed out in ref. 4. To account for this, we plot the subdivisions of San Francisco and San Jose, which support the relationship, as they lie closer to the fit. Interestingly, the dimensionless demand-to-supply ratio  $\Gamma$  lacks this problem and presents a better linear trend with congestion for the five analysed urban areas as depicted in Fig. 3e, despite the broad behaviour of the traffic response. The two most congested cities have the highest ratios, the Bay Area closely followed by Rio de Janeiro, whereas Porto and Lisbon, the two least congested cities, have lower ratios.

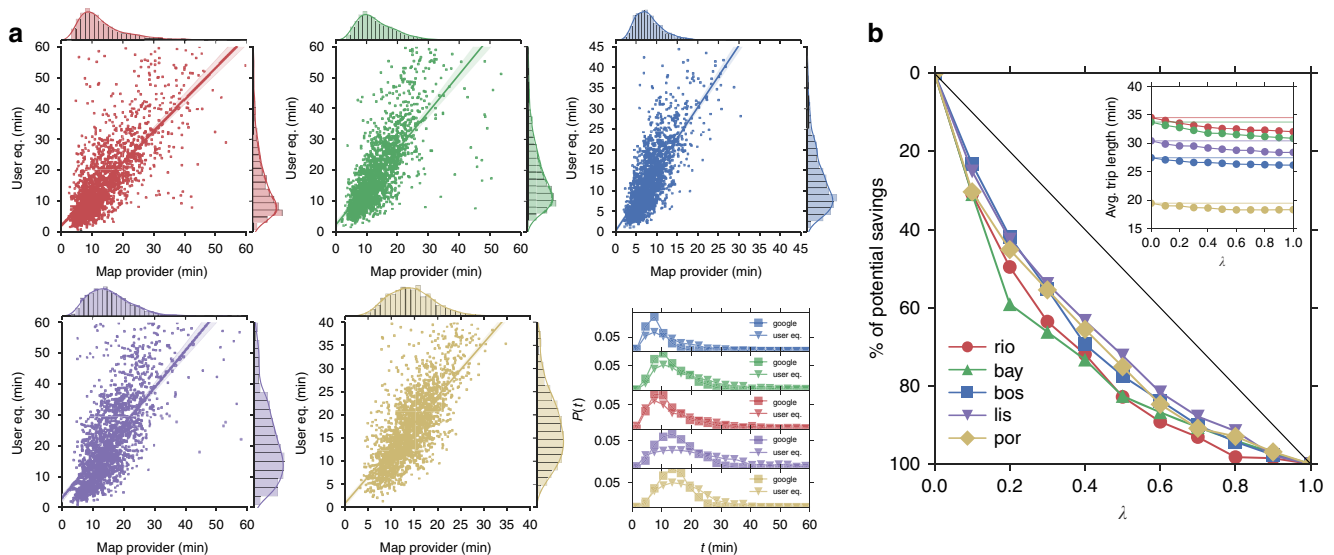
To finalize our analysis, in Fig. 3f we measure how population densities are spatially distributed from the most densely populated region in each of the subject cities based on the

chosen administrative level. The results show different spatial distributions in the population density of the five cities. First, it verifies the expected effect of higher population densities in increasing congestion. It also highlights the importance of the spatial distribution around the highest density point. Lisbon and Porto present densities of population below 500 people per km<sup>2</sup> for distances of  $r > 20$  km, whereas the other three cities stabilize in values  $> 1,000$  people per km<sup>2</sup>. These differences can explain the two types of responses in the effective travel speeds presented in Fig. 3b, where Lisbon and Porto belong to a city type of lower density. Taking these results together, we observe that congestion increases with  $\Gamma$  and appears to be influenced by the spatial distribution of population density and its gradient.

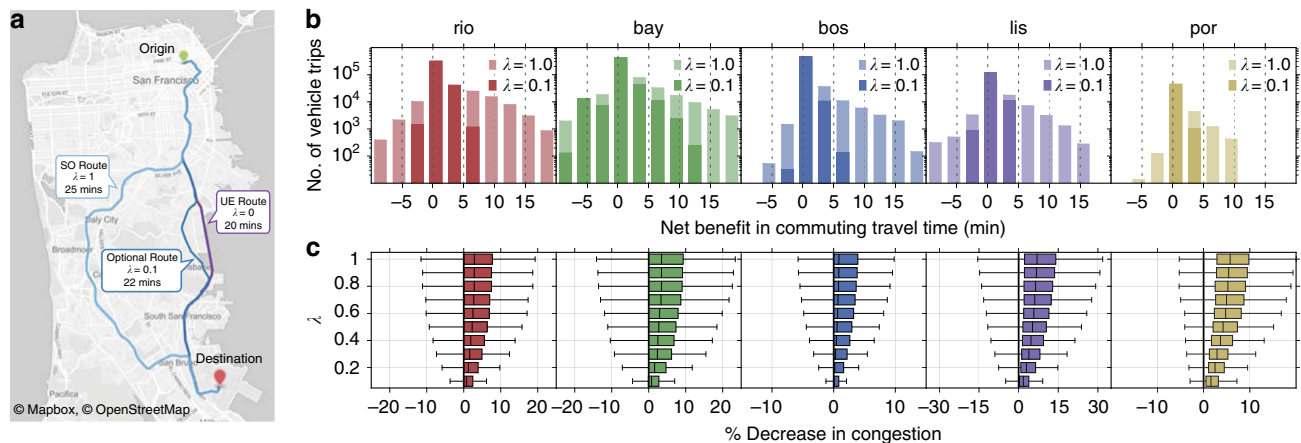
**Selfish routing.** In this section, we compare the travel times for commuters in free flow, socially optimal and user equilibrium flow configurations. Our findings in the five subject cities are outlined in Table 2. Although the estimated free travel time averages are similar, congestion plays a significant role: Lisbon commuters lose 2.1 min on average by selfish routing preferences. Rio de Janeiro exhibits an average loss of 2.6 min on average incurred by selfish routing. The results show that on average 15–30% of total minutes lost in congestion is caused solely by selfish routing.

Although a more nuanced methodology incorporating stochastic traffic assignment and probabilistic OD matrices would probably improve validation, our formulation and central findings would remain robust, as they are based on aggregate and endogenous, albeit simplified, behaviour of our system. Furthermore, a principled and singular validation source does not exist for our cities; we instead use an online map provider as a validation benchmark. Although the validation data are also the product of internal models and estimations, it is of value as they are obtained from an independent data source to ours. In Fig. 4a, we compare the distributions of obtained travel times with those obtained from the map provider in the morning peak hour between 7:30 and 8:30 h for 2,000 OD pairs with the highest commuting flows (see Supplementary Table 3 for statistics related to the regressions). There is an overall overestimation of travel times, which strengthens the notion that route choice in reality might not be a perfect user equilibrium or a social optimum, but somewhere in between. Neither the provider’s nor our findings are expected to have accurate travel time variability, as these comparisons are estimates of typical travel times for the given OD pairs and they act as a first step towards the validation of our estimated travel times based on the assigned traffic flows obtained from the phone data.

**Weight of social good.** In assessing the effects of socially aware routing behaviour for the subject cities, we calculate the average commuting time for various levels of  $\lambda$ . The inset of Fig. 4b depicts the decrease in average commuting travel times for increasing  $\lambda$  in all five cities, ranging from an average of 1–3 min. More importantly, the shape of the curves indicate that even modest social consideration weights can realize a significant portion of the potential savings. Figure 4b collapses these curves to represent realized potential savings as a percentage to exhibit a striking similarity between the five cities in terms of response to socially aware routing. To assess the economies of such routing behaviour, we measure the Gini index of the obtained curves; by definition, higher values of  $G$  indicate higher savings for smaller levels of social good weight. Our findings show that  $G$  ranges from 30–40%:  $G_{\text{rio}} = 41\%$ ,  $G_{\text{bay}} = 42\%$ ,  $G_{\text{bos}} = 33\%$ ,  $G_{\text{lis}} = 30\%$  and  $G_{\text{por}} = 34\%$ . These findings indicate congested cities benefit



**Figure 4 | Travel time comparisons and potential savings.** (a) Comparison of travel times and their distributions between user equilibrium versus routes obtained from the online map provider. OD samples consist of 2,000 OD pairs with the highest commuting flow magnitudes for each city. (b) The percentage of potential savings in average commuting times for the five cities for varying levels of social good weight of routing. (inset: the travel time savings represented in actual minutes).



**Figure 5 | Benefit and congestion decrease distributions for different weights of social good.** (a) A depiction of three route alternatives with the corresponding travel times for a trip from Union Square to San Francisco Airport for  $\lambda = 0$ ,  $\lambda = 0.2$  and  $\lambda = 1$ , respectively. (b) Counts of vehicle trips and observed travel time benefits for  $\lambda = 1$  and  $\lambda = 0.1$ . Negative benefits refer to increase in travel times for vehicles sacrificing for the social good. The spread of the distributions increase for higher  $\lambda$ . (c) The response of distributions of percentage decrease in time lost to congestion to increasing values of  $\lambda$ . The skewness towards positive values of congestion decrease indicate movement towards more optimal configurations. Maps under © OpenStreetMap contributors BY-SA.

more from incorporating social good considerations into routing behaviour.

**Travel time benefit distributions.** In the previous section we characterized the percentage of potential savings that can be obtained for increasing levels of social consideration. However, these benefits are achieved at the expense of time of drivers who adjust their commute for the benefit of others. The unwillingness to give up time is the defining factor in drivers’ failure to reach an optimal state on their own. This highlights the importance of fairness of the distribution of who has to sacrifice versus who benefits in terms of both the success potential of the implementation of policies or a reward/punishment reinforcement schema. Figure 5a demonstrates one such schema, where drivers

are shown a route that corresponds to a choice, which might result in a travel time sacrifice.

Our findings, in accordance with the results of the previous sections, indicate a net bias towards benefits, meaning the number of drivers who benefit outnumber those who sacrifice. Figure 5b summarizes the benefit distributions for the five cities for  $\lambda = 0.1$  and  $\lambda = 1$ . The former exhibits a less spread distribution than the latter but the skewness remains inherent to the distributions. Although the average benefits described in the previous sections appear small, it should be noted that 10-min benefits can be observed for tens of thousands of vehicles. Figure 5c describes in more detail how the positive skewness evolves for increasing social consideration. For higher  $\lambda$ , the % decrease in congestion distributions are shifted towards positive values, indicating a net benefit. This result demonstrates the

potential of incentive schemes, which could compensate the few drivers who sacrifice under consideration of social good.

## Discussion

The economic and social costs of congestion are crippling. In addition to the overall loss of time, congestion underlies many major economic and urban issues such as increased gas consumption, infrastructure deterioration and CO<sub>2</sub> emissions. In this work, we use massive amounts of data to estimate peak hour travel demand and understand travel times. We then explore the power of information-based routing on congestion alleviation.

Our findings suggest very interesting similarities in the behaviour of the five subject cities to explain congestion and potential benefits of social routing. Commuting distances follow a lognormal distribution and free travel speeds are normally distributed. A city's unique congestion fingerprint is strongly related to measurable characteristics. The population density and its spatial distribution together with the  $\Gamma$  parameter of demand-to-supply ratio are the two driving factors of the observed congestion in a diverse range of cities. Further, given the current state of traffic, we then estimate how centralized routing schemes using the power of information would reach possible benefits in travel times. Such information is important, as it allows the assessment of the upper bounds of routing policies; if effective in implementation, it would influence the traffic on a city scale. In practice, this would imply that we could have similar routing applications that we use today with the incorporation of demand profiles, to provide routes that are not necessarily the shortest but also the best for decreasing overall congestion.

We find that routing solutions that mimic socially optimal configurations, that is,  $\lambda = 1$ , have a limit of decreasing time lost in congestion by up to 30%. This is in contrast with the effectiveness of direct and costly interventions where 1% target decrease in demand can achieve 18% decrease in travel times<sup>18</sup>. Although in both scenarios the collective benefits for the whole city can be significant (15–30% decrease), the observed time benefits the average individual receives are marginal, ranging from 1 to 3 min. Furthermore, these times are below the travel time variability based on events, weather conditions or traffic lights. Our findings indicate that in the best-case scenario, time savings would be imperceptible for the majority of the drivers. From this, it is clear that such routing solutions cannot fix the traffic problem for individual drivers but rather would contribute to the city as a whole. The advantage is that in the context of the implied routing application, the number of vehicles sacrificing their travel time is significantly smaller than the number of those that benefit. Lower levels of weight towards social good will also moderate the magnitude of benefits and losses, consequently making the policies fairer and easier to implement.

Open work in this subject contains, but is not limited to, a more generalized bottom-up approach to comparison of cities that includes various modes of transportation to demonstrate their similarities, differences and their consequences. As the volume, the variety and the resolution of data increase along with the expected disruptions from connected self-driving cars and similar technologies, this front of research will become more relevant to facilitate the study and planning for the future of urban mobility. With more updated demand models extracted from communication technologies, understanding the network effects on congestion will become easier to pinpoint and address. In addition, planning tasks on urban mobility previously difficult to tackle may now be addressed at lower costs and with much larger samples of the population. For example, a thorough analysis of how travel time and congestion is distributed among

the population and its split by income and other sociodemographic characteristics remains an open front.

## Methods

**Mobile phone data.** Mobile phone data sets, also referred to as CDRs, used in this study consist of at least 3 weeks of records of all mobile phone users of a particular carrier across each subject city. Each individual CDR consists of a hashed user identification string, a timestamp and the location of the activity. The spatial granularity of the data varies between cell tower level, where calls are mapped to tower locations and distributed uniformly within the Voronoi cell that it forms, and triangulated geographical coordinate pairs, where each call has a unique pair of coordinates accurate to within a few hundred metres. Market shares associated with the carriers that provide the data also vary (see Supplementary Figs 1 and 2, and Supplementary Note 1).

**Census and travel survey data.** At the census tract (or equivalent) scale, we obtain the population, vehicle usage rate and median income of residents in that area. For US cities, the American Community Survey provides this data on the level of census tracts (each containing roughly 5,000 people). Census data are obtained for Brazil through IBGE (Instituto Brasileiro de Geografia e Estatística) and for Portugal through the Instituto de Nacional de Estatística. All cities analysed in this work have varying spatial resolutions of the census information. Wherever possible, we obtain the most recent travel demand model or survey from the subject city and compare the results with those output by our methods. We use the 2011 Massachusetts Household Travel Survey for Boston, 2,000 Bay Area Transportation Survey for the Bay Area and a recent transportation model output provided by the local government for Rio de Janeiro. For Lisbon, the most recent estimates from the MIT-Portugal UrbanSim LUT model that uses the 1994 Lisbon transportation survey as input are used. We found no recent travel survey or model for Porto (see Supplementary Note 2).

**Extraction of validated OD information.** Traditional modelling approaches to OD information use data obtained from travel surveys, possibly combined with land-use and point-of-interest information, to generate estimates of trip production and attraction for locations. Although new data sources such as CDRs do not provide the same detailed demographic and contextual information about individuals or trips, they do provide many high-resolution data points over a far longer observation period. Mobile phones offer good, but imperfect measurements of geographic position due to the uncertainty of the location estimates and the non-uniform sampling frequency (see Supplementary Fig. 5 and Supplementary Note 3 for procedures to generate OD matrices and more descriptive information). For further questions and inquiries about the OD data, please contact the corresponding author.

**Road networks.** For many cities in the United States, detailed road network data are made available by local or state transportation authorities. These data sets generally are well maintained; however, many properties are often incomplete or missing entirely. For this purpose, we infer required road characteristics to build realistic and routable networks using OpenStreetMap, an open-source crowd sourced mapping tool (see Supplementary Note 4).

**Traffic flow and travel time.** Relating travel performance to traffic conditions has been a long-standing problem in transportation. Many different characterizations exist, ranging from conical volume-delay functions to more complex approaches (see Supplementary Fig. 4 and Supplementary Note 5).

**Traffic assignment.** Traffic assignment is a mature domain that aims to bring together travel demand with road infrastructure, to better understand traffic, and has been studied extensively by urban and transportation planners. In this work, we follow an efficient, static, origin-based assignment algorithm that focuses on the equilibration of a directed acyclic graph structure emanating from every origin node (see Supplementary Fig. 6 and Supplementary Note 6).

## References

- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Shleifer, A. Growth in cities. *Working Paper 3787* (National Bureau of Economic Research, 1991).
- Batty, M. The size, scale, and shape of cities. *Science* **319**, 769–771 (2008).
- Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306 (2007).
- Arcaute, E. *et al.* Constructing cities, deconstructing scaling laws. *J. R. Soc. Interface* **12**, 20140745 (2015).
- Bettencourt, L. M. A. The origins of scaling in cities. *Science* **340**, 1438–1441 (2013).
- Hernando, A., Hernando, R. & Plastino, A. Space-time correlations in urban sprawl. *J. R. Soc. Interface* **11**, 20130930 (2014).
- Jacobs, J. *The Death and Life of Great American Cities* (Vintage, 1961).



8. González, M. C., Hidalgo, C. A. & Barabási, A. -L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).
9. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462–465 (2006).
10. Song, C., Koren, T., Wang, P. & Barabási, A.-L. Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818–823 (2010).
11. Song, C., Qu, Z., Blumm, N. & Barabási, A.-L. Limits of predictability in human mobility. *Science* **327**, 1018–1021 (2010).
12. de Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M. & Blondel, V. D. Unique in the crowd: The privacy bounds of human mobility. *Sci. Rep.* **3**, 1376 (2013).
13. Stouffer, S. A. Intervening opportunities: a theory relating mobility and distance. *Am. Sociol. Rev.* **5**, 845–867 (1940).
14. Ren, Y., Ercey-Ravasz, M., Wang, P., González, M. C. & Toroczkai, Z. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. *Nat. Commun.* **5**, 5347 (2014).
15. Simini, F., González, M. C., Maritan, A. & Barabási, A.-L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).
16. Yan, X.-Y., Zhao, C., Fan, Y., Di, Z. & Wang, W.-X. Universal predictability of mobility patterns in cities. *J. R. Soc. Interface* **11**, 20140834 (2014).
17. Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z. & González, M. C. Unravelling daily human mobility motifs. *J. R. Soc. Interface* **10**, 20130246 (2013).
18. Wang, P., Hunter, T., Bayen, A. M., Schechtner, K. & González, M. C. Understanding road usage patterns in urban areas. *Sci. Rep.* **2**, 1001 (2012).
19. Ortúzar, J. D. & Willumsen, L. G. *Modelling Transport* (John Wiley & Sons, 1994).
20. Balmer, M. *et al. Agent-Based Simulation of Travel Demand: Structure and Computational Performance of MATSim-T* (ETH, Eidgenössische Technische Hochschule Zürich, IVT Institut für Verkehrsplanung und Transportsysteme, 2008).
21. Toole, J. L. *et al.* The path most traveled: Travel demand estimation using big data resources. *Transport. Res. C Emerg. Technol.* **58**, 162–177 (2015).
22. Alexander, L., Jiang, S., Murga, M. & Gonzalez, M. C. Origin-destination trips by purpose and time of day inferred from mobile phone data. *Transport. Res. C Emerg. Technol.* **58**, Part B, 240–250 (2015).
23. Çolak, S., Alexander, L. P., Alvim, B. G., Mehndiretta, S. R. & González, M. C. Analyzing cell phone location data for urban travel: current methods, limitations, and opportunities. *Transport. Res. Rec. J. Transport. Res. Board* **2526**, 126–135 (2015).
24. Louf, R. & Barthelemy, M. How congestion shapes cities: from mobility patterns to scaling. *Sci. Rep.* **4**, 5561 (2014).
25. Noulas, A., Scellato, S., Lambiotte, R., Pontil, M. & Mascolo, C. A tale of many cities: universal patterns in human urban mobility. *PLoS ONE* **7**, e37027 (2012).
26. Louail, T. *et al.* Uncovering the spatial structure of mobility networks. *Nat. Commun.* **6**, 6007 (2015).
27. Lämmer, S., Gehlsen, B. & Helbing, D. Scaling laws in the spatial structure of urban road networks. *Phys. A Stat. Mech. Appl.* **363**, 89–95 (2006).
28. Rosvall, M., Trusina, A., Minnhagen, P. & Sneppen, K. Networks and cities: An information perspective. *Phys. Rev. Lett.* **94**, 028701 (2005).
29. Barthelemy, M. Spatial networks. *Phys. Rep.* **499**, 1–101 (2011).
30. Braess, D., Nagurney, A. & Wakolbinger, T. On a paradox of traffic planning. *Transport. Sci.* **39**, 446–450 (2005).
31. Van Huyck, J. B., Battalio, R. C. & Beil, R. O. Tacit coordination games, strategic uncertainty, and coordination failure. *Am. Econ. Rev.* **80**, 234–248 (1990).
32. Roughgarden, T. & Tardos, É. How bad is selfish routing? *JACM* **49**, 236–259 (2002).
33. Roughgarden, T. *Selfish Routing and the Price of Anarchy* (MIT Press, 2005).
34. Roughgarden, T. & Tardos, É. Bounding the inefficiency of equilibria in nonatomic congestion games. *Games Econ. Behav.* **47**, 389–403 (2004).
35. Vickrey, W. S. Congestion theory and transport investment. *Am. Econ. Rev.* **59**, 251–260 (1969).
36. Wardrop, J. G. in *Proceedings of the Institution of Civil Engineers* vol. **1**, 325–378, Part 2 (1952).
37. Boyce, D. E., Mahmassani, H. S. & Nagurney, A. A retrospective on Beckmann, McGuire and Winsten's Studies in the economics of transportation. *Pap. Reg. Sci.* **84**, 85–103 (2005).
38. Youn, H., Gastner, M. T. & Jeong, H. Price of anarchy in transportation networks: Efficiency and optimality control. *Phys. Rev. Lett.* **101**, 128701 (2008).
39. Sheffi, Y. *Urban Transportation Networks* (Prentice-Hall, Englewood Cliffs, NJ, 1985).
40. Correa, J. R., Schulz, A. S. & Stier-Moses, N. E. A geometric approach to the price of anarchy in nonatomic congestion games. *Games Econ. Behav.* **64**, 457–469 (2008).
41. Pigou, A. C. *The Economics of Welfare* (Palgrave Macmillan, 2013).
42. Smith, M. The marginal cost taxation of a transportation network. *Transport. Res. B Methodol.* **13**, 237–242 (1979).
43. Blondel, V. D., Decuyper, A. & Krings, G. A survey of results on mobile phone datasets analysis. *EPJ Data Sci.* **4**, 10 (2015).
44. Ahuja, R. K., Magnanti, T. L. & Orlin, J. B. *Network Flows: Theory, Algorithms, and Applications* 1st edn (Prentice Hall, 1993).
45. Monderer, D. & Shapley, L. S. Potential games. *Games Econ. Behav.* **14**, 124–143 (1996).
46. Beckmann, M., McGuire, C. & Weinstein, C. *Studies in the Economics of Transportation* (Yale Univ. Press, 1956).
47. Chen, P.-A. & Kempe, D. in *Proceedings of the 9th ACM Conference on Electronic Commerce* 140–149 (ACM, 2008).
48. Bertaud, A. The spatial organization of cities. *Deliberate Outcome or Unforeseen Consequence, Background Paper to World Development Report* (2003).

## Acknowledgements

We thank Saurabh Amin for stimulating discussions and helpful suggestions and Airspace for the data provided. The research was partly funded by the World Bank, Ford, the Department of Transportation's grant of the New England UTC Y25, the MIT Portugal Program, the MIT-Brazil seed Grants Program and the Center for Complex Engineering Systems at KACST-MIT, and A.L. was funded by the Vest Scholarship.

## Author contributions

S.Ç. and A.L. processed and analysed the data. S.Ç. and M.C.G. designed the study and wrote the manuscript. All authors read, commented and approved the final version of the manuscript.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Çolak, S. *et al.* Understanding congested travel in urban areas. *Nat. Commun.* **7**:10793 doi: 10.1038/ncomms10793 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>