INTERFACE

rsif.royalsocietypublishing.org

Research

Cite this article: Sun L, Jin JG, Axhausen KW, Lee D-H, Cebrian M. 2015 Quantifying long-term evolution of intra-urban spatial interactions. *J. R. Soc. Interface* **12**: 20141089. http://dx.doi.org/10.1098/rsif.2014.1089

Received: 1 October 2014
Accepted: 10 November 2014

Subject Areas:
mathematical physics, environmental science

Keywords:
complex networks, smart card, human mobility, urban evolution, transportation

Author for correspondence:
Lijun Sun
e-mail: lijun.sun@ivt.baug.ethz.ch

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsif.2014.1089 or via http://rsif.royalsocietypublishing.org.

THE ROYAL SOCIETY PUBLISHING

Quantifying long-term evolution of intra-urban spatial interactions

Lijun Sun[1,2], Jian Gang Jin[3], Kay W. Axhausen[1,4], Der-Horng Lee[2] and Manuel Cebrian[5]

[1]Future Cities Laboratory, Singapore-ETH Centre for Global Environmental Sustainability (SEC), 138602, Singapore
[2]Department of Civil and Environmental Engineering, National University of Singapore, 117576, Singapore
[3]School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China
[4]Institute for Transport Planning and Systems (IVT), ETH Zürich, Zürich 8093, Switzerland
[5]National Information and Communications Technology Australia, University of Melbourne, Melbourne, Victoria 3010, Australia

Understanding the long-term impact that changes in a city's transportation infrastructure have on its spatial interactions remains a challenge. The difficulty arises from the fact that the real impact may not be revealed in static or aggregated mobility measures, as these are remarkably robust to perturbations. More generally, the lack of longitudinal, cross-sectional data demonstrating the evolution of spatial interactions at a meaningful urban scale also hinders us from evaluating the sensitivity of movement indicators, limiting our capacity to understand the evolution of urban mobility in depth. Using very large mobility records distributed over 3 years, we quantify the impact of the completion of a metro line extension: the Circle Line (CCL) in Singapore. We find that the commonly used movement indicators are almost identical before and after the project was completed. However, in comparing the temporal community structure across years, we do observe significant differences in the spatial reorganization of the affected geographical areas. The completion of CCL enables travellers to re-identify their desired destinations collectively with lower transport cost, making the community structure more consistent. These changes in locality are dynamic and characterized over short timescales, offering us a different approach to identify and analyse the long-term impact of new infrastructures on cities and their evolution dynamics.

1. Introduction

Cities, as the core of modern society, are playing increasingly important roles through global urbanization, providing people with housing, transportation, communication and functional institutions for various social activities. Enabled by the transportation infrastructure layer, diverse social interactions among various entities shape a city's interaction layers, creating social economic outputs, which further spur the growth of the cities themselves [1–4]. Created by individuals' trips for work, school, shopping and other social activities, intra-urban movement is a crucial part of these spatial interactions. Intra-urban movements exhibit strong spatial and temporal patterns, which play an important role in urban planning and traffic forecasting [5–7]. Taking travel demand as an example, previous studies have focused on developing models to estimate current and predict future interaction intensity based on social and infrastructural input for planning purposes [8–13]. Although extensive efforts have been made using travel diary data collected from household surveys and interviews [6,7], studies on individual/collective movement still suffers from their small sampling shares, high cost, infrequent periodicity and limited accuracy. As a result, despite the observations telling us that near things are more related than distant things' geographically [14], studies trying to integrate the

infrastructure and interaction layers have remained limited due to the lack of detailed longitudinal data measuring change at high spatial and temporal resolutions.

The emerging individual-based large-scale datasets have allowed us to trace our daily behaviour pattern in detail, shifting our understanding on individual mobility from random to predictable [15–17]. The smart card data also show pronounced advantages in depicting the structure of collective encounter networks given the large share of public transit (capturing more than 63% of total mobility in Singapore) [18]. Besides revealing mobility regularity, these datasets also help us further understand human mobility-induced spatial interactions, which are crucial to various urban diffusion processes such as epidemic spreading, knowledge-spillover and social contagions [19–21]. Thus, recent studies demonstrate an increasing use of network theory to model diverse types of interactions (from transportation to digital communication) on multiple scales (state, country and global) [22], documenting the importance of such interactions among spatial units in shaping network typologies [23–27]. This paper analyses intra-urban movements using three one-week transit use data, including both bus and metro systems, from 3 years (11–15 April 2011; 19–23 March 2012; and 8–12 April 2013) in Singapore. We place special emphasis on the study of a key transportation infrastructure—the extended Circle Line (CCL; Mass Rapid Transit service), which provides us an ideal natural observation to investigate how such large infrastructures influence human mobility, and how the resulting spatial interactions shape the structural evolution of a city. The first stage of the CCL (the Eastern half; figure 4*d*, coloured orange) has been in operation since 2010. After the competition of the second Western half, the CCL has been fully operational (35.7 km in total; figure 4*e*) since October 2011. This metro line cost about 10 billion Singapore dollars (roughly 8 billion US dollars), carrying about half a million passengers daily today (0.37 million per day in stage 1 and 0.53 million per day after opening stage 2). It seems reasonable that such a large infrastructure project should affect local/global mobility patterns and city structure, by first influencing individual travel patterns, and then by prompting second-order effects such as new businesses and real estate. Indeed, such effects can be measured by looking at the usage patterns of the extension and the geographical space where it occurs [28]. However, these approaches suffer notable limitations when analysing city-wide movements. First, we miss how this local change affects global/city-scale patterns: we cannot tell whether the new service alters mobility patterns by merely looking at local travel patterns. Second, there might be changes of mobility patterns enabled by the new infrastructure we are not aware of yet. Taking advantage of emerging network analysis, this paper tries to present some steps to overcome this knowledge gap.

2. Results

To sketch the geographical structure of intra-urban movement from collective transit use, we divided the island city–state Singapore into zones of 500×500 m, indexing all transit journeys with their origin–destination (OD) pairs (see the electronic supplementary material for defining spatial zones and transit journey). By aggregating all transit journeys across weekdays using OD indexes in 2012, we obtained a directed and weighted spatial interaction network, which demonstrates symmetrical in/out-degree distributions (figure 1*a*,*b*). This network—with spatial zones as its vertices and the time-resolved commuting flows as its edges—is the basis of our analyses. To explore its statistical properties, we first measure interaction intensity w_{ij} as total passenger flow travelling from zone i to j, finding that when $w_{ij} \geq 10$, the tail of distribution $P(w_{ij})$ is well characterized by a power law $P(w_{ij}) \sim w_{ij}^{-\beta}$ with an exponent $\beta = 1.744 \pm 0.002$ (figure 1*c*) using statistical tools provided in [29]. This indicates that intra-urban movement displays a strong heterogeneity: most ODs have small flows, but a few ODs involve massive demands. Note that a similar scaling is reported for London Underground [5], suggesting that such scaling of interaction intensity might be a fundamental property of urban spatial interactions. To quantify the importance of individual zones in shaping the aggregate interaction network, we measured in/out-strength of vertex i as its total inflows and outflows ($s_i^{\text{in}} = \sum_{i=1}^{N} w_{ji}$ and $s_i^{\text{out}} = \sum_{j=1}^{N} w_{ij}$) on weekdays, respectively. As figure 1*d* shows, the spatial distribution of zone strength also exhibits a strong heterogeneity across the city. Although it is difficult to identify a deterministic function to characterize zone strength distributions $P(s^{\text{in}})$ and $P(s^{\text{out}})$, we still find the same heavy tail characterizing both of them and observe an intrinsic balance across all zones (figure 1*e*,*f*). Different from a previous study about the dependency of vertex strength s on vertex degree k in the worldwide airport network [23], we found an exponential increase $s(k) \sim \exp(\lambda k)$ instead of a power law $s(k) \sim k^{\beta}$ (with $\beta > 1$), suggesting that the network demonstrates a much faster increase of s than k (see the electronic supplementary material, figure S1). In fact, considering the strong heterogeneity of w_{ij}, degree k, which merely counts the presence of edges and exhibits a saturation process with time, is not appropriate to capture the time-varying architecture and the backbone of this interaction network (electronic supplementary material, figures S2 and S3). To check whether these properties hold over time, we applied these analyses for year 2011 and 2013 as well. Despite the completion of the CCL during the study period, we found indistinguishable aggregate properties for the 3 years.

A key property in understanding the dynamics of a spatial network is its community structure, defined as vertex partitions which have more connections within themselves than between each other. The importance of community lies in revealing the intermediate scales of network organization and identifying hidden structure in network theory. However, in practice the detection of communities is a difficult task to apply. The state-of-the-art approach to find these partitions is to maximize modularity Q [30]

$$Q = \frac{1}{W} \sum_{ij} (w_{ij} - w'_{ij}) \delta(c_i, c_j), \tag{2.1}$$

where $W = \sum_i s_i^{\text{in}} = \sum_j s_j^{\text{out}}$ is the total network traffic, w'_{ij} is the expected interaction intensity estimated from a null model and δ is an indicator function: $\delta(c_i, c_j) = 1$ if zone i and j belong to the same partition and $\delta(c_i, c_j) = 0$ otherwise. The approach has been successfully applied on various spatial interaction networks at different resolutions, such as identifying effective administrative borders and finding hidden structures of countries and cities [24–27]. In practice, an appropriate null model is crucial to get a meaningful expectation w'_{ij} to reveal corresponding network structural attributes. Without accounting for spatial attributes, we adopted the default null model for

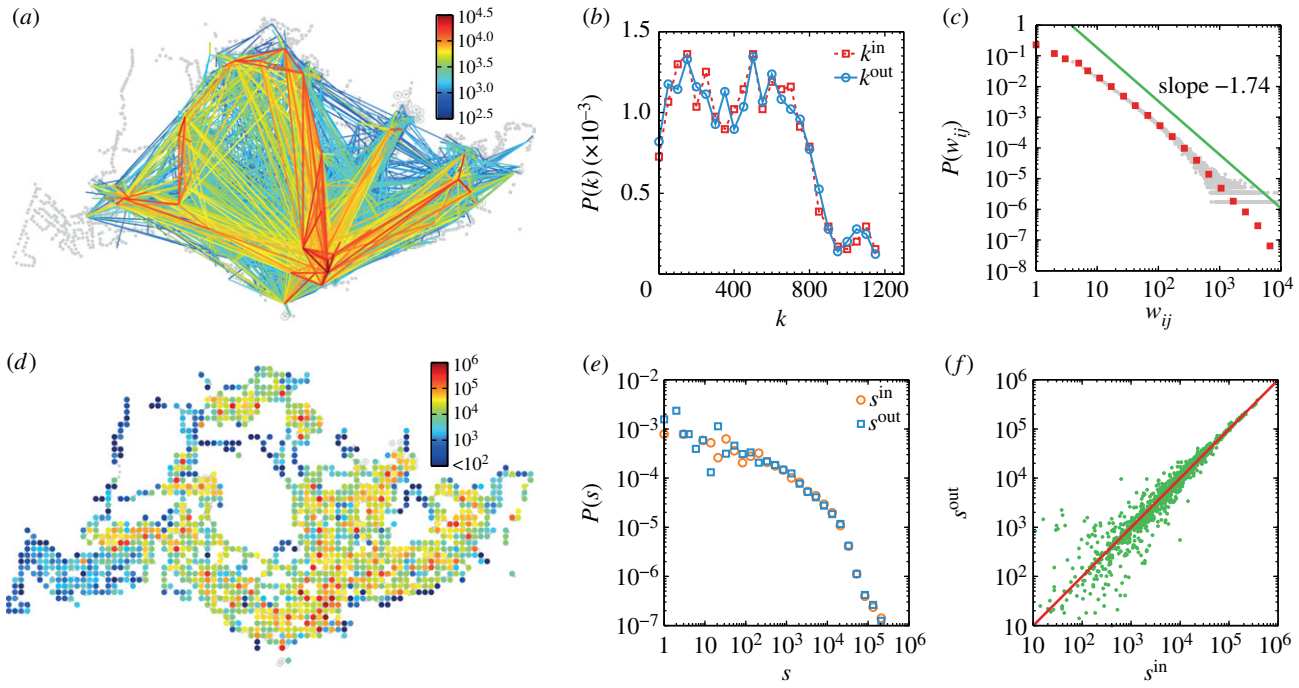

Figure 1. Structure of intra-urban public transport movements in Singapore. (*a*) Aggregated spatial interaction network across weekdays (from 19 to 23 March 2012). For simplicity, we only show the top 1% of edges with highest total interaction intensity $w_{ij} + w_{ji}$ in an undirected manner. The grey markers show the spatial locations of transit infrastructures, including both bus stops (dots) and metro stations (circles). (*b*) In/out-degree distribution $P(k^{in})$ and $P(k^{out})$ measured on the aggregated network. (*c*) Probability density function $P(w_{ij})$ of interaction intensity across all OD pairs. Grey dots show the original histogram and red squares correspond to a log-binned histogram. As a guide, the green line shows a power law with an exponent $\beta = 1.74$. (*d*) The spatial distribution of total strength $s_i = s_i^{in} + s_i^{out}$ of each cell. (*e*) Probability density function $P(s^{in})$ of in-strength and $P(s^{out})$ of out-strength. Both of them exhibit heavy-tailed properties. (*f*) Symmetrical plot of s^{in} and s^{out}. Most dots are scattered around the red line, suggesting the homogeneous inflow and outflow spatially.

defining modularity, i.e. $w'_{ij} = s_i^{out} s_j^{in}/W$ (see the electronic supplementary material, figure S2). Note that one may replace the null model for special considerations; for example, w'_{ij} is determined using a gravity model to exclude the dependency on distance in [26].

To find partitions that maximize modularity Q, we applied the well-established Louvain method on the aggregated network across weekdays [31]. Although the detection process employs only interaction intensity (without using any geographical information), we still observe a clear spatial consistency from 2011 to 2013 (see the electronic supplementary material, figure S4), suggesting that collective movements are remarkably constant across years. Yet, we may miss the temporal evolution of these community structures by merely analysing the aggregated network. To explore this evolution over the day, we grouped all transit journeys according to their transaction times (see the electronic supplementary material for details). Based on these temporally grouped journeys, we created a series of sub-networks and applied the community detection processes on each of them. In figure 2*a*, we summed intra- and inter-community flows and measured their temporal variations using 2012 data. Although community structure is well defined spatially and intra interactions are much stronger than inter interactions ($\langle w_{in}\rangle > \langle w_{out}\rangle$), the total number of journeys crossing communities is still higher than that of intra community trips ($\sum w_{in} < \sum w_{out}$), suggesting that people are not confined to a spatial community but show wider destination choices. Likewise, we repeated this analysis for the other years for comparison, finding that temporal variations of modularity Q are essentially indistinguishable across years (figure 2*b* and electronic supplementary material, figures S6–S8). This indicates that temporal intra-urban movements (or collective mobility) might be comparable as well. Given the clear geographical consistence of these communities, modularity Q is actually a measure of spatial mobility patterns embedded in temporal activities. To explore the variability of collective mobility over time, we measured the average trip distance $\langle d\rangle$ for each sub-network. Not surprisingly, we also found high degree of similarity of $\langle d\rangle$ across years (figure 2*c*), indicating that temporal travel behaviours of populations are essentially comparable over 3 years as well. Next, we compared Q and $\langle d\rangle$ jointly to explore how collective movement shapes the time-varying community structure. If communities are well identified as spatial clusters, the longer people travel, the higher the chance that one jumps out of a community. However, the relation is unclear without the assumption. For example, taking only structured metro trips will give rise to high modularity and long travel distance, whereas short and random local trips will lead to low modularity and short distance. We find that they are linked by a universal and substantial negative correlation $Q \sim \alpha\langle d\rangle$ (figure 2*d*; with $\rho_{2011} = -0.9718$ [$p < 10^{-4}$], $\rho_{2012} = -0.9623$ [$p < 10^{-4}$] and $\rho_{2013} = -0.9763$ [$p < 10^{-4}$]); and once again, we observed similar structural patterns across years, independently of the completion of the CCL. Therefore, using time-varying travel displacement as an indicator, we confirmed that temporal variation of collective movement plays a crucial role in expressing the dynamic community structures.

To explore how distance affects systematic/occasional travel behaviour, we quantified the variation of two distance-related diversity measures. Without considering individual identity, we first quantify the degree of heterogeneity of

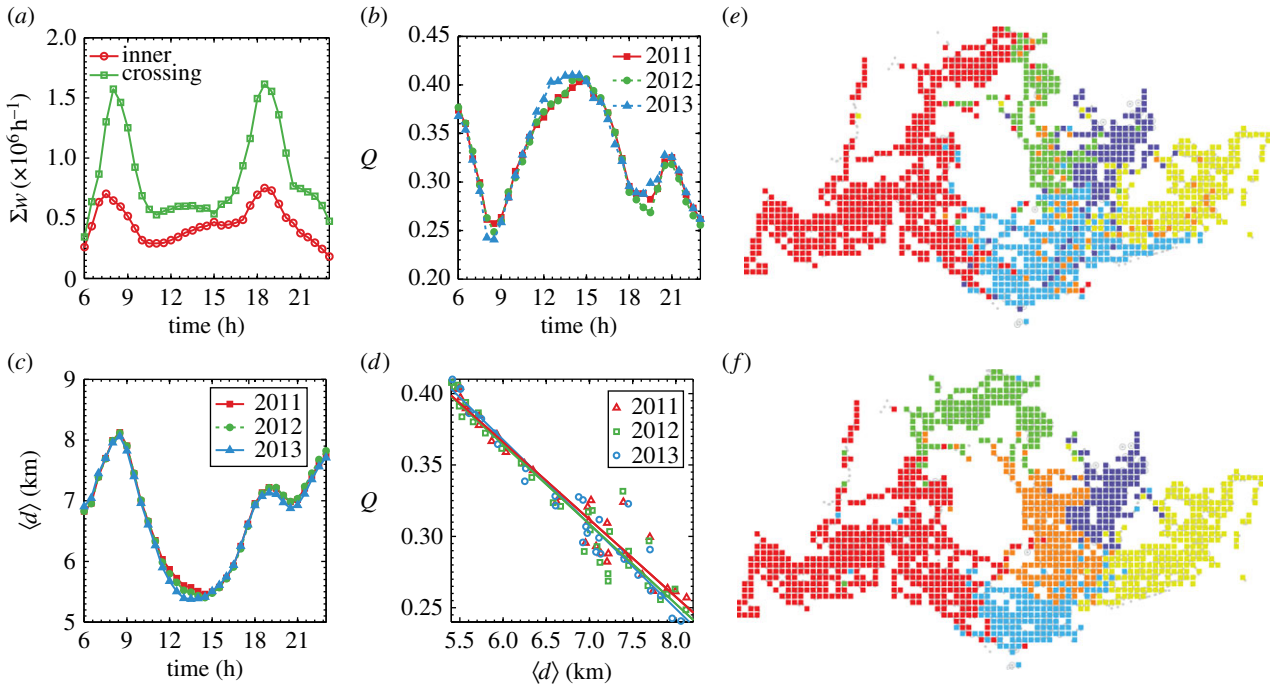

Figure 2. Structure and temporal variation of intra-urban spatial interactions. (*a*) Temporal variation of total movements within communities (red-circled line) and across communities (green-squared line) in the year 2012. The summation of these two parts is the total transit flows. (*b*) Temporal change of modularity Q in 3 years. Modularity Q is characterized by a wave with pronounced troughs (at 08.30 and 19.30) and peaks (at 14.30 and 21.00). (*c*) Temporal variation of average trip distance $\langle d \rangle$ in 3 years. (*d*) The scatter plot of temporal modularity Q and average trip distance $\langle d \rangle$. The solid lines represents linear regressions of $Q \sim \alpha \langle d \rangle$, with $\alpha_{11} = -0.69 \pm 0.03$, $\alpha_{12} = -0.70 \pm 0.04$ and $\alpha_{13} = -0.72 \pm 0.03$ (the subscripts represent years), respectively. (*e*,*f*) Community structures identified for 08.30 and 14.30 in 2012.

bi-directional flows by measuring a diversity index of OD (i, j) as $\theta_{ij} = \max\{w_{ij}, w_{ji}\}/\min\{w_{ij}, w_{ji}\}$ (for OD pairs with $w_{ij}, w_{ji} > 0$). If the collective attribute θ_{ij} is independent of distance, we expect to find θ_{ij} being characterized by a determined distribution free from d_{ij}. However, on the contrary, we do find a significant and consistent reduction of $\langle \theta_{ij} \rangle$ with d_{ij}, indicating that bi-directional transit flows at a collective level become more and more balanced with distance. Still, this is not sufficient to show homogeneity at the individual level. Taking advantage of the anonymized card ID, we further investigated the similarity of users travelling between i and j for each day using the Jaccard index $\lambda_{ij} = |W_{ij} \cap W_{ji}|/|W_{ij} \cup W_{ji}|$, where W_{ij} represents the set of individuals travelling from i to j; thus, λ_{ij} is close to one if all individuals travel symmetrically each day, and zero if no one returns to previously visited locations (i.e. $W_{ij} \cap W_{ji} = \emptyset$). After measuring λ_{ij} for all OD pairs across weekdays, we show the dependence of $\langle \lambda_{ij} \rangle$ on d_{ij} in figure 3*b*, finding another consistent (increasing) trend across 3 years. Therefore, despite previous observations characterizing exploration/preferential return behaviour [32], we found that shorter travel distances are associated with higher exploration; and correspondingly, previously visited locations are more preferred for longer distance journeys. Moreover, these properties are also stable across years, independent of the new metro line.

Large-scale human mobility patterns have been described by three indicators: trip distance distribution $P(d)$, temporal variation of radius of gyration $r_g(t)$ and number of visited locations $S(t)$. However, the duration of stay at one location, as another important attribute in understanding why people move rather than why people stay, is seldom considered in the literature. To investigate recurrence and periodicity of travel behaviour and the patterns of stay in terms of both exploration/preferential return behaviours, we classified transit usage based on the pattern [33]: round journeys (with two trips $i \to j$ and $j \to i$) and trip chains with two trips ($i \to j$ and $j \to k$, where $i \neq k$), and measured the duration of stay at zone j for each. In figure 3*c*,*d*, we show the change of average duration of stay for both round journeys ($\langle t_R \rangle$) and exploratory trip chains ($\langle t_E \rangle$) as a function of distance d_{ij}, finding that both $\langle t_R \rangle$ and $\langle t_E \rangle$ display an consistent increase with d in the beginning and reach saturation after $d = 10$ km. In comparing them, we find that t_R is significantly longer than t_E ($p < 10^{-4}$, Wilcoxon rank-sum test), suggesting that people tend to stay longer at the destination of round trips. To further distinguish t_R from t_E, we group journeys with similar travel distance d_{ij} and determine the distribution $P(t_R|d)$ of stay duration (by measuring the interval between journey $i \to j$ and journey $j \to i$) and $P(t_E|d)$ (by measuring the interval between two journeys $i \to j$ and $j \to k$) for each group. As the insets of figure 3*c*,*d* show, both $P(t_R|d)$ and $P(t_E|d)$ can be approximately characterized by a mixture distribution of a short stopover (secondary activity around 1 h, such as shopping, eating and leisure activities) and a long primary activity (around 10 h, such as work and school). We note that the key difference between t_R and t_E is their composition: the proportion of primary activity in return journeys is significantly higher than that of a trip chain. Thus, figure 3*c*,*d* also implies a strong correlation between trip purpose and travel distance, further confirming the role of distance in shaping spatial interaction structure. Taken together, we show that travel distance not only determines the balance of intra-urban movement and a traveller's exploration/preferential return behaviour, but also with the type of journey. Not surprisingly, we once again observe a clear consistency of t_R and t_E across the 3 years.

Previous analyses has shown that intra-urban movement is consistent over the 3 years, exhibiting patterns independent of

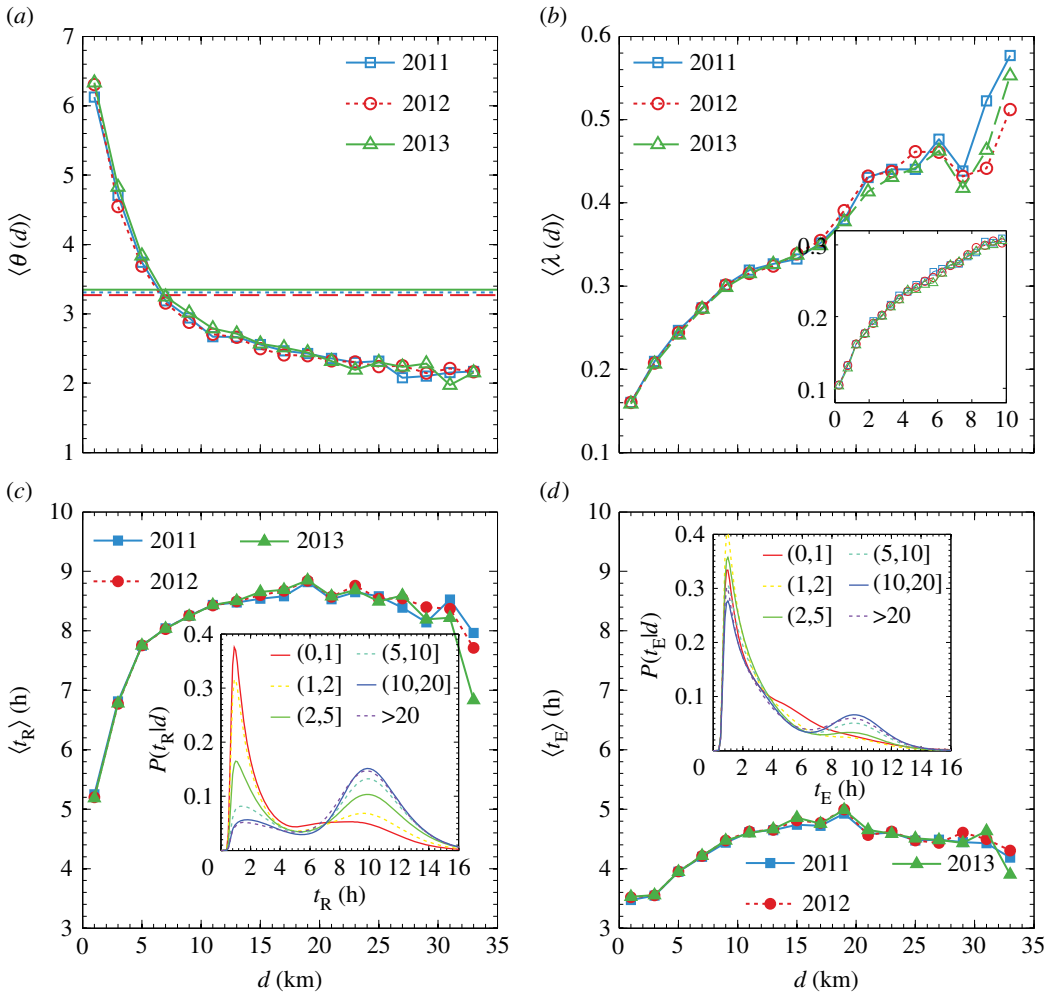

Figure 3. The impact of distance on spatial interaction patterns. (*a*) The average diversity index θ_{ij} with given displacement d_{ij}. The dashed line demonstrations a null model assuming that θ is independent of d. (*b*) The average correlation $\langle\lambda_{ij}\rangle$ with d_{ij}. The inset illustrates the change of $\langle\lambda\rangle$ when $0 < d \leq 10$ km. (*c*) Stay duration $t_R(d)$ in between round trips (i, j) and (j, i), averaged over d_{ij} for both years. The inset shows the distribution $P(t_R|d)$ for groups characterized by different d using kernel smoothing. (*d*) The same plot as in panel (*c*), however, for the stay duration $t_E(d)$ in between trip chain. The inset shows the corresponding conditional distribution $P(t_E|d)$.

the key infrastructure project, suggesting that human mobility may display universal patterns invariant to the change of transportation infrastructures. However, as mentioned, with time-varying interactions during a day, the resulting structural communities should be changing simultaneously over time (such as the continuous changing of boarders and emergence of new communities). This effect offers us more insight into how spatial interactions shape the city. To quantify it, we define neighbourhood variability $\gamma_i(t_1, t_2)$ of zone i between the sub-networks at time t_1 and t_2 as

$$\gamma_i(t_1, t_2) \equiv 1 - \frac{|C_i(t_1) \cap C_i(t_2)|}{|C_i(t_1) \cup C_i(t_2)|}, \tag{2.2}$$

where $C_i(t)$ represents the community which zone i belongs to at time j. Hence, $\gamma_i(t_1, t_2)$ is close to one if the intersection contains only i and zero if C_i does not change from time t_1 to t_2. Using continuous observations during 1 day, we quantified the overall spatial evolution of zone i by calculating mutability index ϕ_i as average neighbourhood variability from t_0 till $t_{\max}$ [34]

$$\phi_i \equiv \frac{\sum_{t=t_0}^{t_{\max}-1} \gamma_i(t, t+1)}{t_{\max} - t_0}, \tag{2.3}$$

where t_0 is the time step when we start observing the temporal evolution and $t_{\max}$ is the last time step. Thus, ϕ_i quantifies the overall evolution of community structure for each individual zone, characterizing the robustness/fragility of the spatial community to which zone i belongs. Essentially, ϕ_i quantifies community transition rate of zone i when other zones are also changing simultaneously. A high ϕ_i typically indicates that zone i is attached to diverse communities over time and vice versa. In other words, the mutability index ϕ_i used here could also be interpreted as a measure to quantify the diversity of temporal community attachment. By setting $t_0 = 6$ and $t_{\max} = 23$ (from 06.00 to 23.00 in 1 h intervals, see the electronic supplementary material for details), we determined the value of $\phi_{i,y}$ for each individual zone i in year y. As shown in figure 4*a–c*, we find that mutability displays clear and comparable spatial patterns: one can easily distinguish the borders between regions with different mutability. Particularly, the central/southern area displays a higher mutability and the western/eastern parts are generally stable across 3 years.

To further compare mutability across years, we calculated $\Delta\phi$ (as $\Delta\phi^1 = \phi_{i,12} - \phi_{i,11}$ and $\Delta\phi^2 = \phi_{i,13} - \phi_{i,12}$) and map the results in figure 4*d*,*e*, respectively. Notably, although the temporal change of Q, $\langle d\rangle$ and other collective mobility indicators are essentially indistinguishable, we do observe a significant difference when using $\Delta\phi^1$ as an indicator, while not much difference is observed when measuring $\Delta\phi^2$. We think that the completion of the CCL appears as a main factor for this difference

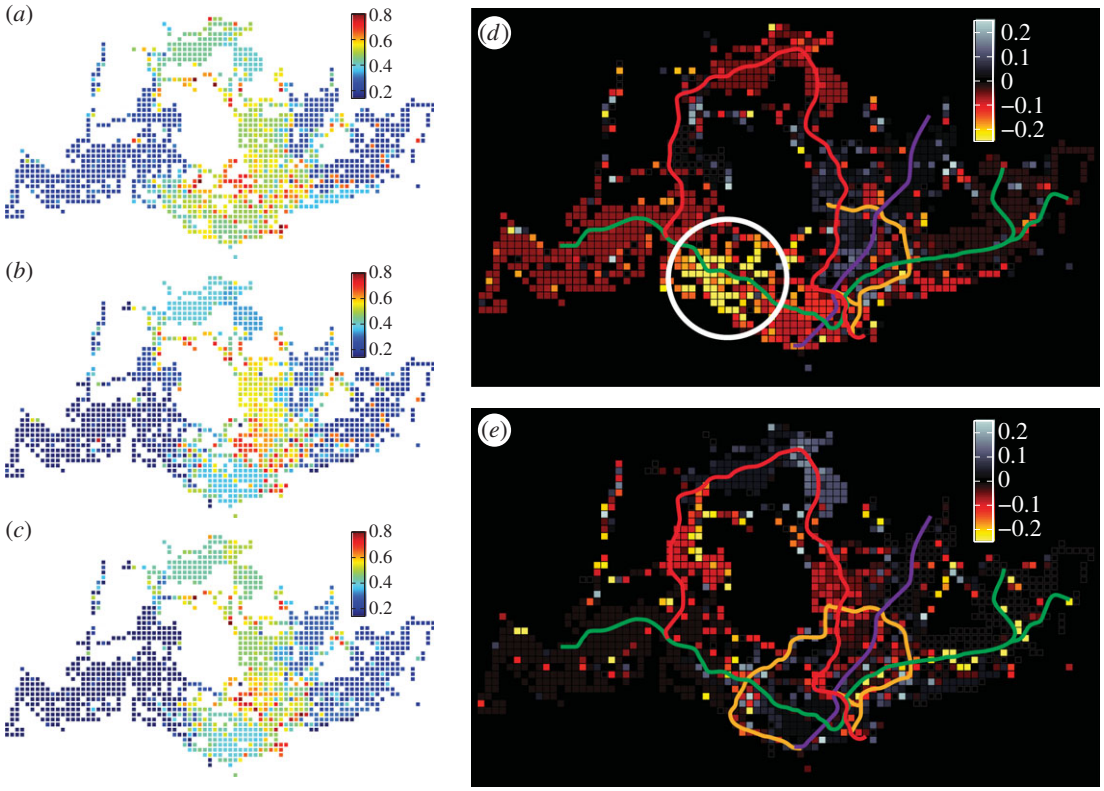

Figure 4. Spatial distribution of mutability ϕ_i from 06.00 to 21.00: (*a*–*c*) $\phi_{i,11}$, $\phi_{i,12}$ and $\phi_{i,13}$ in the year 2011, 2012 and 2013, respectively. (*d*) The difference of mutability from 2011 to 2012 $\Delta\phi^1 = \phi_{i,12} - \phi_{i,11}$. (*e*) The difference of mutability from 2012 to 2013 $\Delta\phi^2 = \phi_{i,13} - \phi_{i,12}$. Given that the full completion of CCL was in October 2011, panel (*a*) shows the mutability of stage 1 while both (*b*,*c*) show the mutability of stage 2. Thus, panel (*d*) actually shows the change of mutability from stage 1 to stage 2, while both $\phi_{i,12}$ and $\phi_{i,13}$ are in stage 2 for panel (*e*).

(figure 4*d*). As mentioned, only the right half of the CCL was in operation in 2011 (stage 1), while the full metro line has come into service since October 2011 (stage 2). Given the definition of ϕ_i, the implication of $\Delta\phi_i > 0$ are twofold. On one hand, for those zones that have not changed their membership during a day, such as most zones in the eastern/western community, the decrease of ϕ_i suggests that the community to which zone i belongs becomes more consistent over time. On the other hand, for those zones that changed their membership during a given time, a decreasing ϕ_i implies that zone i changes less frequently and strengthened its dependency to the attached community. In fact, the completion of stage 2 (the western half of the circle) enables travellers to find their desired destination choices collectively in a structured manner with lower cost, instead of making diverse choices individually. In this sense, the completion of stage 2 of CCL may make zones in southern area (the white circle in figure 4*d*) more accessible to either the western community or the central community (as shown in the electronic supplementary material, figures S6–S8). We next perform statistical tests on $\Delta\phi_i$ for zones within the white circle (radius 4 km, 124 zones nearby the extended CCL) and zones outside the circle (1170 zones). We find that $\Delta\phi_i^1$ in the nearby area are significantly lower than others ($p < 10^{-4}$; left side Wilcoxon rank-sum test), while there is no clear evidence to show $\Delta\phi_i^2$ in the nearby area are different from the others ($p = 0.213$; Wilcoxon rank-sum test).

3. Discussion

Understanding spatial interactions is crucial to urban planning, traffic forecasting and various mobility-related urban diffusion processes such as epidemic spreading and social contagions. More importantly, coupled with the transportation infrastructure layers of a city, social economic outputs are shaped by these interactions [35,36]. Although the study of mobility has a long history, previous works were almost all based on modelling these interactions for planning purposes owing to a lack of longitudinal data detailed enough across both spatial and temporal scales. Yet, the evolution of urban structure with these temporal interactions is merely revealed: it remains a challenge to distinguish the natural variability in the city's mobility from large deviations, using either coarse-grained or short-timescale mobility data.

Taking advantage of population-scale smart card datasets spanning 3 years, we study the structure of Singapore's intra-urban interaction network and present how it is influenced by a key transportation infrastructure project (the CCL in this case). Despite that Singapore has been a dynamic, fast-changing city, we show that human mobility displays invariant aggregate patterns across the years, even when seeing a large infrastructure project. As a city evolves over the years, how can we distinguish large deviations in mobility from statistical fluctuations in a city's mobility patterns? Our study suggests that commonly used tools and statistics do not offer sufficient sensitivity to identify key changes in the city's mobility structure.

We present evidence for this by first examining the temporal community structure that emerges from collective travel behaviour, and study its variation across years. Using modularity as an indicator, we find that the community structure varies consistently with the spatial–temporal characteristics of collective mobility, indicating that distance acts as a powerful constraint to keep universal mobility patterns in place over 3 year period, and therefore, does not

allow us to discriminate the impact of the completion of the CCL. Moreover, we found that both modularity Q and average journey distance $\langle d \rangle$ demonstrate clear and consistent temporal homogeneity, exhibiting remarkable robustness to the competition of the extended CCL as well. Taking stay-durations as another indicator quantifying human mobility, we showed that travel distance not only determines individual's exploration and preferential return behaviour, but also associates with one's purpose of travelling. However, none of these indicators help us identify the global impact of CCL.

Notably, despite other structural and behavioural dynamic indicators being almost consistent and indistinguishable over the long-term, we do observe a significant difference of mutability. Mutability, which is defined as the average ratio of community members changed across time, emerges as a highly sensitive tool to understand the position of individual zones in the overall evolution and real-time evolving borders of community structure. In fact, it is sensitive to the differences in mobility caused by major transportation infrastructure change, showing the evolving borders in community structure, and with this the way people interact to shape, sustain or reform a city.

Our findings and analysis framework offers analytical tools to better sense the evolution of mobility patterns in cities, providing insights for urban planning, modelling and understanding the evolution of cities through the coupling of infrastructure and interaction networks. Given that ICT is being fast embedded in our daily-life, spatial–temporal digital traces helping us to follow cities will become available in various forms, overcoming the limits of field surveys and interviews. In the near future, much urban data which are generated in real-time will become available for urban planning, improving the quality of life. Despite the privacy concerns, taking full advantage of such data in planning would surely help us further understand urban dynamics and make our cities smarter. Taken together, our study offers a quantitative and general strategy to understand the dynamic evolution in multiple temporal scales and serve as a basis to further track and model such evolution [37].

Acknowledgements. We thank M. González for discussions and comments on the manuscript. We thank Singapore's Land Transport Authority for providing the smart card data.

Funding statement. This study was supported by National Research Foundation of Singapore, which is the funding authority of the Future Cities Laboratory, Singapore-ETH Centre. M.C. is funded by the Australian Government as represented by the Department of Broadband, Communications and Digital Economy and the Australian Research Council through the ICT Centre of Excellence programme.

References

1. Bettencourt LM, Lobo J, Helbing D, Kuhnert C, West GB. 2007 Growth, innovation, scaling, and the pace of life in cities. *Proc. Natl Acad. Sci. USA* **104**, 7301–7306. (doi:10.1073/pnas.0610172104)
2. Batty M. 2008 The size, scale, and shape of cities. *Science* **319**, 769–771. (doi:10.1126/science.1151419)
3. Bettencourt LM. 2013 The origins of scaling in cities. *Science* **340**, 1438–1441. (doi:10.1126/science.1235823)
4. Pan W, Ghoshal G, Krumme C, Cebrian M, Pentland A. 2013 Urban characteristics attributable to density-driven tie formation. *Nat. Commun.* **4**, 1961. (doi:10.1038/ncomms2961)
5. Roth C, Kang SM, Batty M, Barthélemy M. 2011 Structure of urban movements: polycentric activity and entangled hierarchical flows. *PLoS ONE* **6**, e15923. (doi:10.1371/journal.pone.0015923)
6. Schlich R, Axhausen KW. 2003 Habitual travel behaviour: evidence from a six-week travel diary. *Transportation* **30**, 13–36. (doi:10.1023/A:1021230507071)
7. Axhausen KW, Zimmermann A, Schönfelder S, Rindsfüser G, Haupt T. 2002 Observing the rhythms of daily life: a six-week travel diary. *Transportation* **29**, 95–124. (doi:10.1023/A:1014247822322)
8. Zipf GK. 1946 The P1P2/D hypothesis: on the intercity movement of persons. *Am. Sociol. Rev.* **11**, 677–686. (doi:10.2307/2087063)
9. Erlander S, Stewart NF. 1990 *The gravity model in transportation analysis: theory and extensions*. Utrecht, The Netherlands: VSP.
10. Stouffer SA. 1940 Intervening opportunities: a theory relating mobility and distance. *Am. Sociol. Rev.* **5**, 845–867. (doi:10.2307/2084520)
11. Wilson AG. 1969 The use of entropy maximising models, in the theory of trip distribution, mode split and route split. *J. Transp. Econ. Policy* **111**, 108–126. (doi:10.2307/20052128)
12. Simini F, González MC, Maritan A, Barabási AL. 2012 A universal model for mobility and migration patterns. *Nature* **484**, 96–100. (doi:10.1038/nature10856)
13. Yan XY, Zhao C, Fan Y, Di Z, Wang WX. 2014 Universal predictability of mobility patterns in cities. *J. R. Soc. Interface* **11**, 20140834. (doi:10.1098/rsif.2014.0834)
14. Tobler WR. 1970 A computer movie simulating urban growth in the detroit region. *Econ. Geogr.* **46**, 234–240. (doi:10.2307/143141)
15. González MC, Hidalgo CA, Barabási AL. 2008 Understanding individual human mobility patterns. *Nature* **453**, 779–782. (doi:10.1038/nature06958)
16. Song C, Qu Z, Blumm N, Barabási AL. 2010 Limits of predictability in human mobility. *Science* **327**, 1018–1021. (doi:10.1126/science.1177170)
17. de Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. 2013 Unique in the crowd: the privacy bounds of human mobility. *Sci. Rep.* **3**, 1376. (doi:10.1038/srep01376)
18. Sun L, Axhausen KW, Lee DH, Huang X. 2013 Understanding metropolitan patterns of daily encounters. *Proc. Natl Acad. Sci. USA* **110**, 13 774–13 779. (doi:10.1073/pnas.1306440110)
19. Christakis NA, Fowler JH. 2013 Social contagion theory: examining dynamic social networks and human behavior. *Stat. Med.* **32**, 556–577. (doi:10.1002/sim.5408)
20. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. 2009 Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl Acad. Sci. USA* **106**, 21 484–21 489. (doi:10.1073/pnas.0906910106)
21. Sun L, Axhausen KW, Lee DH, Cebrian M. 2014 Efficient detection of contagious outbreaks in massive metropolitan encounter networks. *Sci. Rep.* **4**, 5099. (doi:10.1038/srep05099)
22. Barthélemy M. 2011 Spatial networks. *Phys. Rep.* **499**, 1–101. (doi:10.1016/j.physrep.2010.11.002)
23. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. 2004 The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101**, 3747–3752. (doi:10.1073/pnas.0400087101)
24. Guimerà R, Mossa S, Turtschi A, Amaral LA. 2005 The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proc. Natl Acad. Sci. USA* **102**, 7794–7799. (doi:10.1073/pnas.0407994102)
25. Ratti C, Sobolevsky S, Calabrese F, Andris C, Reades J, Martino M, Claxton R, Strogatz SH. 2010 Redrawing the map of Great Britain from a network of human interactions. *PLoS ONE* **5**, e14248. (doi:10.1371/journal.pone.0014248)
26. Expert P, Evans TS, Blondel VD, Lambiotte R. 2011 Uncovering space-independent communities in spatial networks. *Proc. Natl Acad. Sci. USA* **108**, 7663–7668. (doi:10.1073/pnas.1018962108)
27. De Montis A, Caschili S, Chessa A. 2013 Commuter networks and community detection: a method for planning sub regional areas. *Eur. Phys. J. Spec. Top.* **215**, 75–91. (doi:10.1140/epjst/e2013-01716-4)
28. Sung H, Choi K, Lee S, Cheon S. 2014 Exploring the impacts of land use by service coverage and station-level

accessibility on rail transit ridership. *J. Transp. Geogr.* **36**, 134–140. (doi:10.1016/j.jtrangeo.2014.03.013)
29. Alstott J, Bullmore E, Plenz D. 2014 powerlaw: a python package for analysis of heavy-tailed distributions. *PLoS ONE* **9**, e85777. (doi:10.1371/journal.pone.0085777)
30. Newman M. 2004 Analysis of weighted networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **70**, 056131. (doi:10.1103/PhysRevE.70.056131)
31. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. 2008 Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008. (doi:10.1088/1742-5468/2008/10/P10008)
32. Song C, Koren T, Wang P, Barabási AL. 2010 Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818–823. (doi:10.1038/nphys1760)
33. Schneider CM, Belik V, Couronné T, Smoreda Z, González MC. 2013 Unravelling daily human mobility motifs. *J. R. Soc. Interface* **10**, 20130246. (doi:10.1098/rsif.2013.0246)
34. Palla G, Barabási AL, Vicsek T. 2007 Quantifying social group evolution. *Nature* **446**, 664–667. (doi:10.1038/nature05670)
35. Fujita M, Krugman P, Venables AJ. 1999 *The spatial economy: cities, regions and international trade*. Cambridge, MA: MIT Press.
36. Eagle N, Macy M, Claxton R. 2010 Network diversity and economic development. *Science* **328**, 1029–1031. (doi:10.1126/science.1186605)
37. Hopcroft J, Khan O, Kulis B, Selman B. 2004 Tracking evolving communities in large linked networks. *Proc. Natl Acad. Sci. USA* **110**, 5249–5253. (doi:10.1073/pnas.0307750100)