

# The scales of human mobility

<https://doi.org/10.1038/s41586-020-2909-1>

Laura Alessandretti<sup>1,2,3</sup>, Ulf Aslak<sup>1,2,3</sup> & Sune Lehmann<sup>1,2,3</sup>

Received: 3 February 2020

Accepted: 25 September 2020

Published online: 18 November 2020

 Check for updates

There is a contradiction at the heart of our current understanding of individual and collective mobility patterns. On the one hand, a highly influential body of literature on human mobility driven by analyses of massive empirical datasets finds that human movements show no evidence of characteristic spatial scales. There, human mobility is described as scale free<sup>1–3</sup>. On the other hand, geographically, the concept of scale—referring to meaningful levels of description from individual buildings to neighbourhoods, cities, regions and countries—is central for the description of various aspects of human behaviour, such as socioeconomic interactions, or political and cultural dynamics<sup>4,5</sup>. Here we resolve this apparent paradox by showing that day-to-day human mobility does indeed contain meaningful scales, corresponding to spatial ‘containers’ that restrict mobility behaviour. The scale-free results arise from aggregating displacements across containers. We present a simple model—which given a person’s trajectory—infers their neighbourhood, city and so on, as well as the sizes of these geographical containers. We find that the containers—characterizing the trajectories of more than 700,000 individuals—do indeed have typical sizes. We show that our model is also able to generate highly realistic trajectories and provides a way to understand the differences in mobility behaviour across countries, gender groups and urban–rural areas.

It is nearly impossible to underestimate the importance of establishing a quantitative foundation for our understanding of how individuals move from place to place in their everyday lives. Hundreds of millions of individuals spend billions of collective hours commuting every day<sup>6</sup>. Goods and food are transported through a global network using shared infrastructure<sup>7</sup>. Understanding mobility patterns helps us mitigate epidemic spreading<sup>8</sup>, assist in crisis management<sup>9</sup>, prepare for dramatic shifts in modes of transportation<sup>10</sup> and in many other cases<sup>6</sup>. For this reason, understanding the origin of scale-free distributions of displacements in empirical mobility traces is crucial, as this issue currently separates the large-scale data-driven human mobility research<sup>11</sup> from the community of human geography<sup>4,5</sup> and transportation research<sup>12</sup>.

Our mental representation of physical space has a hierarchical structure<sup>13</sup>. We describe space by referring to places<sup>4</sup>, meaningful spatial entities with associated typical size, or scale, from rooms and buildings—via neighbourhoods, cities and states—to nations and continents that are organized in a nested structure<sup>4,14–17</sup>. Geographical borders confine residential mobility<sup>18</sup> and collective mobility fluxes<sup>19</sup>. Commuting is characterized by a typical travel-time budget, and, as a consequence, there exist characteristic spatial scales that have evolved in connection with the progress of transportation<sup>20</sup>. Further, it has been conjectured that there are fundamental differences between forms of moving at different scales, from moving within a building to travelling across the globe<sup>6,12,16</sup>.

However, recent empirical research in the field of human mobility<sup>11</sup> has found no evidence for characteristic spatial scales in how people travel<sup>1–3,21</sup>. On the contrary, studies have shown that the distribution of displacement lengths  $\Delta r$  travelled by an individual has a power-law tail  $P(\Delta r) \approx \Delta r^{-\beta}$  over several orders of magnitude, where typically  $1 \leq \beta \leq 2$

(ref. <sup>22</sup>). Power-law distributions are also called scale free, because they are the only mathematical distribution to have no associated typical scale<sup>23</sup> (Supplementary Note 1).

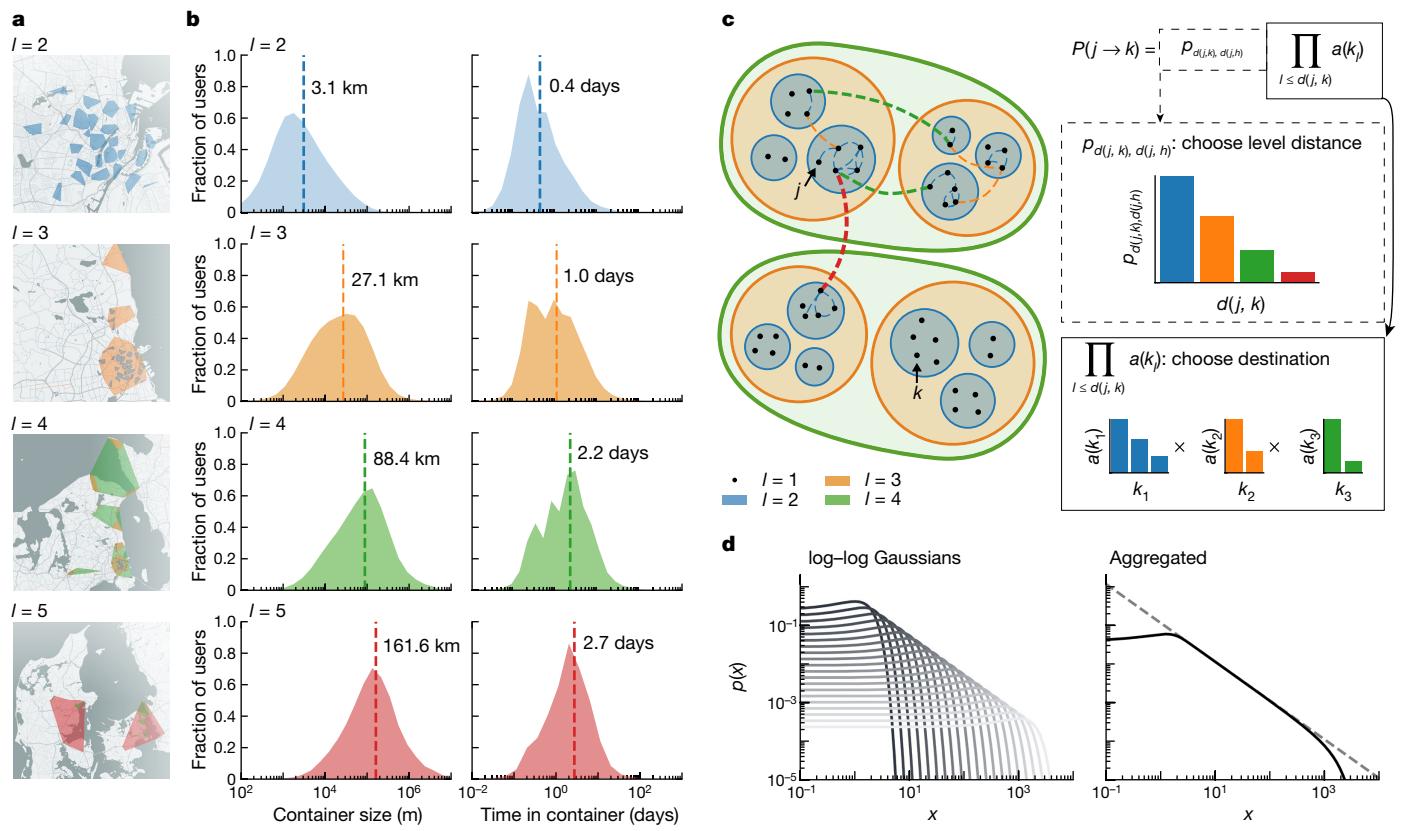
## Nested scales generate power laws

So the question becomes: How is it possible that our intuitive conception of space is clearly hierarchical and characterized by typical scales, when a broad range of empirical datasets, ranging from displacements of dollar bills<sup>1</sup> or cell-tower data<sup>2</sup> to public transportation systems<sup>24</sup>, and GPS data<sup>25,26</sup> all suggest that human mobility is scale free?

To explain this apparent contradiction, we propose that each typical scale of human mobility corresponds to a container of a certain mobility behaviour. These containers (rooms, buildings, neighbourhoods, cities, countries and so on) have typical sizes (Fig. 1a), and roughly correspond to the notion of places in geography<sup>4</sup>. The observed power law arises when we aggregate mobility behaviour within containers and mobility that transports a person between containers. Specifically, it is well known that mixtures of normal (or lognormal) distributions with different variances can generate power laws<sup>27</sup> (Fig. 1d). More specifically, we assume that for each individual, physical space is organized as a nested structure of containers. This structure relates, in part, to the organization of the transportation system<sup>20</sup> and to the concrete structure of our built environment<sup>16</sup> (Fig. 1a).

We propose that these nested containers affect how individuals move, and therefore can be inferred from the raw mobility data. Specifically, the amount of time spent within a container can depend on its hierarchical level. The connection between hierarchical level and mobility is supported by the literature, which shows that, for

<sup>1</sup>DTU Compute, Technical University of Denmark, Kongens Lyngby, Denmark. <sup>2</sup>Copenhagen Center for Social Data Science, University of Copenhagen, Copenhagen, Denmark. <sup>3</sup>These authors contributed equally: Laura Alessandretti, Ulf Aslak. <sup>✉</sup>e-mail: sljo@dtu.dk



**Fig. 1 | The scales of human mobility.** **a**, Example of containers for an individual living in Copenhagen, characterized by the size of containers in neighbourhoods (blue), cities (orange), urban agglomerations (green) and regions (red). Map data copyrighted OpenStreetMap contributors and available from <https://www.openstreetmap.org>. **b**, Distribution of container sizes (left) and median time spent in the same container (right) across individuals. Dashed lines correspond to medians. Results, shown here for containers at different hierarchical levels, are obtained by fitting the container model to the D1 dataset, consisting of approximately 700,000 anonymized

GPS traces of individuals distributed across the globe (see Extended Data Fig. 2 for dataset D2). **c**, Schematic representation of the container model. Individuals move between locations (black dots) inside a nested set of containers. The probability of transitioning between two locations  $j$  and  $k$  is the product of two factors, corresponding to choosing level distance and destination (see main text). **d**, Gaussian distributions with different variances (left) and their mixture (right) on a log–log scale. The dashed line (right) is a power law  $P(x) \approx x^{-\beta}$  with variable of interest  $x$  and exponent  $\beta = 1$  to guide the eye.

example, transitions between regions are more frequent than transitions between countries<sup>28</sup>.

### A simple model identifies containers

We now describe the associated container model of mobility, a model that estimates a person's containers from their empirical mobility patterns (Fig. 1c). For each individual, we model physical space as a hierarchy of  $L$  levels, ordered from the smallest to largest (for example, individual locations to countries). At any level  $l$ , space is partitioned into topologically compact containers, with a characteristic size. For  $l < L$ , a container is fully included within a single parent container (for example, each neighbourhood is part of a single city). Hence, each geographical location  $k$  can be identified as a sequence of containers,  $k = (k_1, \dots, k_b, \dots, k_r)$ , where container  $k_l$  is included in  $k_{l+1}$ .

Next, consistent with most models of human mobility<sup>11,29</sup>, each container  $k_l$  is characterized by its probability to be selected within its parent container, its attractiveness  $a(k_l)$ . We define the level distance  $d(j, k)$  between locations  $j$  and  $k$  as the highest index at which the two sequences of containers describing  $j$  and  $k$  differ<sup>30</sup>. We model traces individually; each trace results in a unique hierarchical structure.

Based on the assumption that the amount of time spent in a container depends on its place in the hierarchy, we design a model of trajectories, where the probability of transitioning from location  $j$  to location  $k$

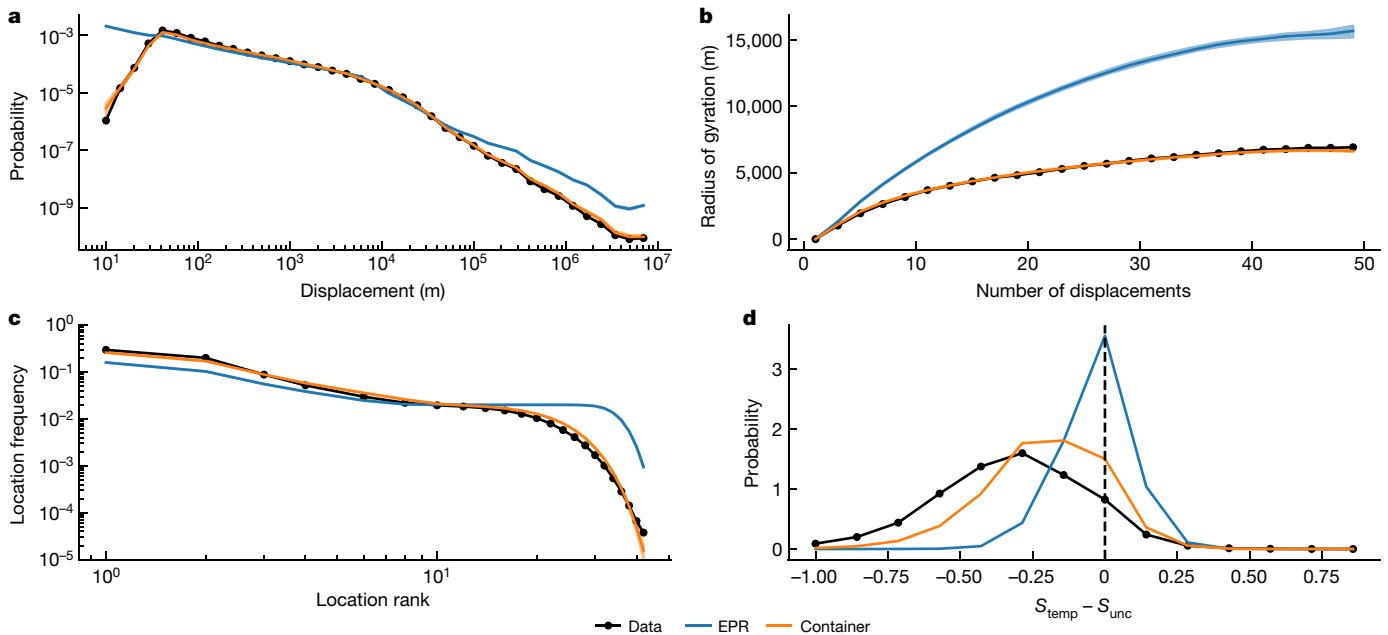
depends on the level distance between them. For an agent located in  $j$ , we model the probability of moving to  $k$  as the product of two factors:

$$P(j \rightarrow k) = p_{d(j,k), d(j,h)} \prod_{l \leq d(j,k)} a(k_l), \quad (1)$$

(see also 'Model description' in Methods). The first factor,  $p_{d(j,k), d(j,h)}$ , represents the probability of travelling at level distance  $d(j, k)$ , given that the current location  $j$  is at level distance  $d(j, h)$  from the individual home location,  $h$ . This probability follows a multinomial distribution, which must depend on level distance from home to account for the fact that higher-level transitions are more likely when individuals are not in the home container; for example, one is typically more likely to transition at the country scale, when not in the home country. The second factor  $\prod_{l \leq d(j,k)} a(k_l)$  is the probability of choosing a specific location  $k$  at that level distance, where  $a(k_l)$  is the attractiveness of a container at level  $l$  including location  $k$ .

### Scales of human mobility

We fit this container model to the individual GPS traces from two different datasets: dataset D1, which consists of traces of approximately 700,000 individuals distributed across the globe, and dataset D2, which consists of traces of approximately 1,000 students from the Technical University of Denmark (see 'Data description' in Methods).



**Fig. 2 | The container model generates realistic mobility traces.** **a**, The distribution of displacements for the entire population, computed for real traces (black line, dots), and traces generated by the container (orange filled area) and the EPR<sup>3</sup> (blue filled area) models. **b**, The median individual radius of gyration versus the number of displacements for data (black line, dots), and traces generated by the container (orange filled area) and the EPR (blue filled area) models. **c**, The average visitation frequency versus the rank of individuals' locations for real traces (black line, dots), and container (orange filled area) and the EPR (blue filled area) model traces. **d**, The distribution of the difference

between the temporal entropy  $S_{\text{temp}}$  and the uncorrelated entropy  $S_{\text{unc}}$  across individuals for real traces (black line, dots), and synthetic traces generated by the container (orange filled area) and the EPR (blue filled area) models. In **a**, **c** and **d**, the filled areas for synthetic traces include two standard deviations around the mean computed across 1,000 simulations for each user. In **b**, filled areas include the interquartile range. For each individual, we fitted the EPR and container models considering a training period of one year. The data used here for validation corresponds to the 50 individual displacements following the training period. Results are shown for a random sample of 9,000 individuals.

We fit the model using maximum likelihood estimation (see 'Likelihood optimization' in Methods). For each individual, the fitting procedure outputs the most likely hierarchical spatial structure, along with attractiveness of containers and probabilities of travelling at a given level distance. We find that empirical individual mobility traces are characterized, on average, by four hierarchical levels. In contrast, synthetic traces generated by the current state-of-the-art models, for example, the exploration and preferential return (EPR) model<sup>3</sup> and its variations<sup>31</sup>, are best described by a single hierarchical level grouping individual stop locations (Extended Data Fig. 5). In both datasets of GPS traces, our model finds characteristic sizes of containers. The sizes of containers—defined as the maximum distance between two locations in a container at a given level—are not broad, but well described by a lognormal distribution across the population. Our results are robust across datasets (Extended Data Table 1). We argue that the characteristic sizes of containers are precisely the 'scales' of human mobility.

These typical sizes of containers can be characterized by the median value  $e^{\mu_l}$ , of the lognormal distributions with log-mean  $\mu_l$  and log-standard deviation  $\sigma_l$  (ref. <sup>32</sup>), for each hierarchical level  $l$ . We find  $e^{\mu_2}=3.089 \pm 0.006$  km,  $e^{\mu_3}=27.064 \pm 0.006$  km,  $e^{\mu_4}=88.442 \pm 0.022$  km and  $e^{\mu_5}=161.634 \pm 0.049$  km (Fig. 1b, Extended Data Table 3). The coefficient of variation  $C_l = \sqrt{e^{\sigma_l^2}} - 1$  (ref. <sup>33</sup>), characterizing the relative dispersion of the lognormal distribution, is included in the range [2.721, 3.042] for  $l$  in the range [2, 5].

The median time spent within the same container at a given level is also well described by a lognormal distribution (Fig. 1c, Extended Data Table 2), implying that there are characteristic temporal scales associated with spatial scales.

Having shown that we can infer information on geographical scales directly from the raw data, we now demonstrate the usefulness of this novel description of mobility patterns. We approach this task in two

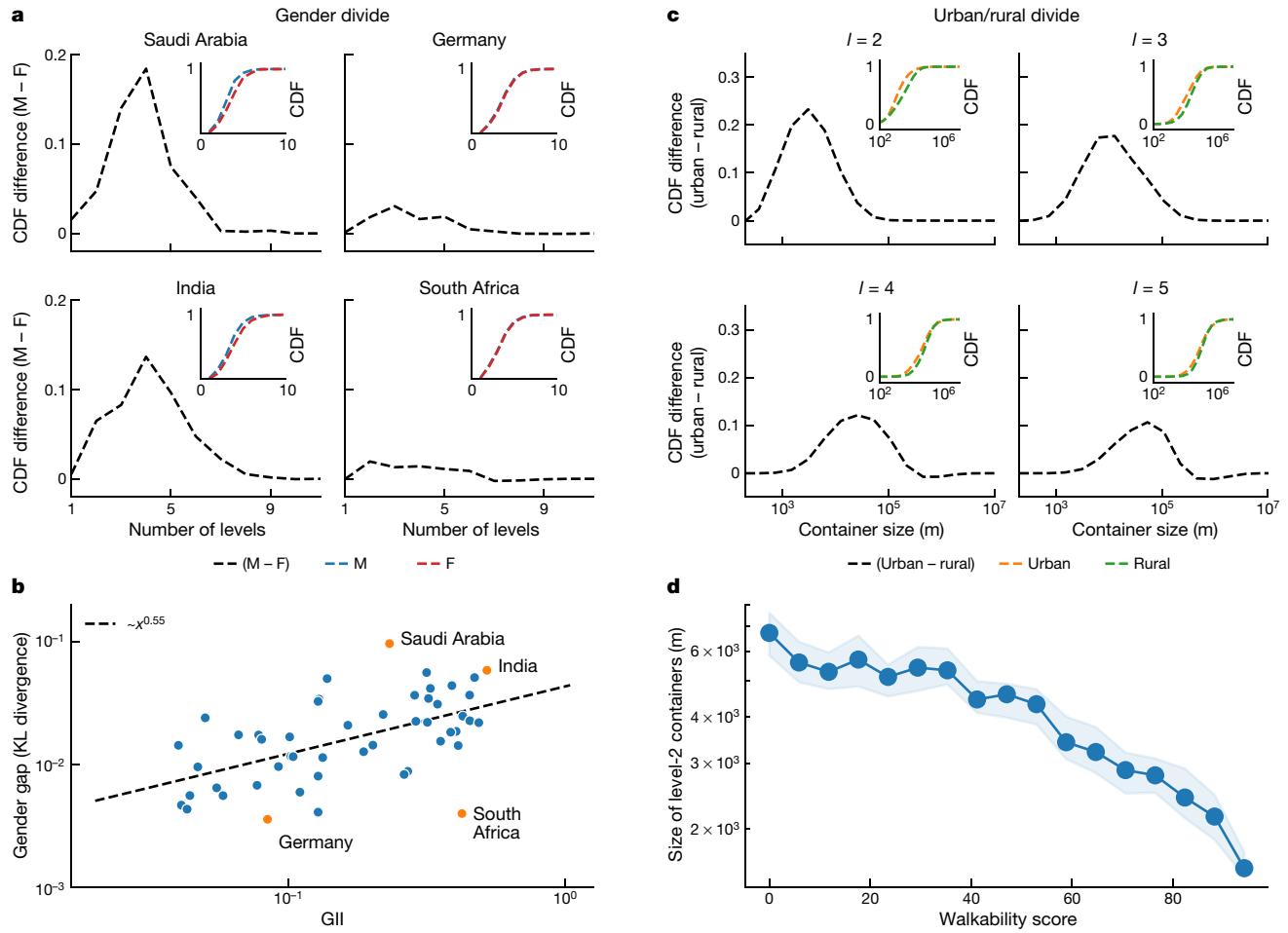
steps. First, we argue that the hierarchical description generated by the container model generalizes to unseen data without overfitting, while providing a more expressive and nuanced description of mobility relative to state-of-the-art models according to unbiased performance estimates. Second, drawing on demographic and environmental data, we show that the container model produces results that converge with existing literature on gender differences, urban/rural divides and walkability scores.

### Validating through generation of traces

First, we explore the ability of the container model to capture key features of empirical mobility patterns and compare it with state-of-the-art models. The container model allows us to generate synthetic traces. The realistic nature of these trajectories can be verified by comparing the statistical properties of synthetic and real sequences of locations (Fig. 2). For each individual, we fit the container model parameters using a portion of the entire trace with length one year (see 'Likelihood optimization' in Methods), and we then generate 1,000 synthetic sequences of 50 displacements (see 'Generation of traces' in Methods). Now, we can compare these synthetic traces with actual traces of the same length, collected in the one-year window subsequent to training. Thus, there is no overlap between the data we used to fit the model and the data we used to validate the model. Comparing synthetic traces to unseen data provides an unbiased performance estimate, which allows us to compare model performance across multiple models and confirm that the container model does not overfit (Supplementary Note 4).

We focus on four key properties of mobility in the generated data: distribution of displacements, evolution of radius of gyration, time allocation among locations and entropy.

Considering the distribution of displacement lengths between consecutive locations, a widely studied property of mobility traces<sup>22</sup>,



**Fig. 3 | Socio-demographic differences and heterogeneity in scales.** **a**, The cumulative distribution function (CDF) of number of levels for males ( $M$ , blue dashed line, inset) and females ( $F$ , red dashed line, inset), and the difference between the two ( $M - F$ , black dashed line). Results are shown for the four countries with the largest (left, Saudi Arabia and India) and the smallest (right, Germany and South Africa) gender gap, measured as the Kullback–Leibler (KL) divergence. **b**, The gender gap in number of levels, computed as the KL divergence between the number of levels for males and females, versus the GII<sup>39</sup>. Each dot represents a different country and the orange dots are the

countries shown in **a**. The black dashed line is a power law fit  $P(x) \approx x^\beta$  with  $\beta = 0.55$ . **c**, The cumulative distribution of container sizes for individuals living in urban (orange dashed line, inset) and rural (green dashed line, inset) areas, and the difference between the two (black dashed line). Results are shown for hierarchical levels from 2 to 5. **d**, The size of containers at level 2 (with level 1 corresponding to individual locations) versus the walkability score<sup>38</sup> around an individual's home location (blue dots). The shaded area corresponds to the 50% interquartile range computed by bootstrapping 500 samples of individuals for each value of the walkability score.

the likelihood ratio test<sup>34</sup> shows that the container model provides a significantly better description of the data than the EPR model<sup>3</sup> and its variations (Fig. 2a, Extended Data Fig. 4, Extended Data Table 4; with  $P \ll 0.01$ ).

Next, the radius of gyration<sup>2</sup> (see ‘Metrics’ in Methods) quantifies the spatial extent of an individual’s mobility. Here we find that while the evolution of individuals’ radius of gyration  $r_g(t)$  over time  $t$  is well described by a logarithmic growth in all cases—real<sup>2</sup>, EPR<sup>3</sup> and the container model (Fig. 2b)—only the fit  $r_g(t) = a + b \log(t)$  with parameters  $a$  and  $b$  for the container model traces is consistent with the real data within errors (Supplementary Note 4).

We characterize the way in which individuals allocate time among locations (Fig. 2c), and find that the distribution of location frequencies is better described by the container model, compared with the EPR model, under the likelihood ratio test<sup>34</sup> (with  $P \ll 0.01$ ).

The final property of synthetic traces is the individual difference between the uncorrelated entropy  $S_{\text{unc}}$ , which characterizes the heterogeneity of visitation patterns, and the temporal entropy,  $S_{\text{temp}}$ , which depends not only on the frequency of visitation but also on the order in which locations were visited<sup>35</sup> (see ‘Metrics’ in Methods). The likelihood ratio test<sup>34</sup> shows that the distribution of  $S_{\text{unc}} - S_{\text{temp}}$  is better described

by the container model, compared with the EPR model (with  $P \ll 0.01$ ). The result that the container model provides a better description of mobility compared with the state-of-the-art models holds also when considering a comprehensive<sup>11</sup> set of six state-of-the-art individual-level models (Supplementary Note 4).

## Validating through demographics and built environment

Now, we aggregate users based on demographics and contextual features and explore the characteristics of containers for each subgroup of users, to underscore how the container model reveals patterns that have strong support in the existing literature. We focus on three factors that describe heterogeneity in mobility behaviour: gender<sup>36</sup>, level of urbanization<sup>37</sup> and walkability score<sup>38</sup> in the area surrounding one’s home location. First, we find that gender differences can partly explain the observed heterogeneity, in line with previous findings<sup>36</sup>, although not in all of the countries under study (Fig. 3a, Supplementary Table 1). A novel finding concerns the fact that in 21 out of 53 countries, females are characterized by a significantly larger number of hierarchical levels than males ( $P \leq 0.05$ ), while the opposite is not the case for any country

# Article

(Supplementary Table 2). As a key observation inviting further research, we find that the difference between genders across countries, measured as the Kullback–Leibler divergence between the distributions of number of levels, is strongly correlated with the Gender Inequality Index (GII)<sup>39</sup>. The GII measures the percentage of potential human development loss due to gender inequality (Spearman correlation  $\rho = 0.69$ ,  $P < 10^{-6}$ ; Fig. 3b). Turning next to the size of containers, we find that in 48 out of 53 countries, the containers characterizing the mobility of females are smaller compared with males (Supplementary Note 2,  $P \leq 0.05$ ), in line with previous research<sup>36</sup>.

Second, we find that the urban/rural divide partly explains differences in mobility patterns, in line with the fact that rural areas are characterized by limited accessibility<sup>16,40</sup>. Individuals living in rural areas (see ‘Other data’ in Methods for definitions) have significantly larger containers compared with urban individuals ( $P \leq 0.05$ ), and this difference is more pronounced at the lowest hierarchical levels (Fig. 3c, Supplementary Table 3).

Finally, we find that the walkability score around an individual’s home location correlates negatively with the size of containers at the lowest hierarchical level (Spearman correlation  $\rho = -0.96$ ,  $P < 10^{-9}$ ; Fig. 3d), in line with the finding that improved walkability increases accessibility to goods, services and activities<sup>41</sup>. The correlation between walkability and container size is significant up to the third level of description (Supplementary Note 2,  $P < 0.01$ ).

## Discussion

The paradigm of power-law descriptions does not stand entirely unchallenged within the quantitative analysis literature. For example, it has been argued that exponential or lognormal functions may be more suitable to describe the distributions of displacements within cities<sup>22</sup>, hinting that human mobility may not be completely free of scales. Until now, however, the nature of the probability distribution of displacements has been unclear<sup>22,26</sup>. For example, it has been suggested that scaling laws could be the signature of Lévy flights—a type of random walk with scale-free step-size attributed to animal foraging<sup>42</sup>—but Lévy flights do not reproduce all statistical properties of human trajectories<sup>3</sup>. It has also been proposed that the structure of the transportation system<sup>26,43,44</sup>, where each mode of transportation corresponds to a typical distance travelled, could explain the observed scaling laws. Intracity displacements considering all transportation modes, however, are not scale-free distributed, as this theory would suggest<sup>21,24</sup>. Owing to the lack of agreement on the functional form of distribution of displacements, many state-of-the-art agent-based models of individual mobility focus on temporal aspects<sup>11</sup>, including the interplay between exploration and exploitation<sup>3,45</sup>, recency and memory effects<sup>25,31,46</sup>, and weekly and circadian rhythms<sup>47</sup>. With few exceptions<sup>43</sup>, these models do not account for effects due to the spatial distribution of locations.

Here we have proposed a model in which human mobility is organized according to a hierarchical structure of spatial containers, corresponding to the notion of places in geography (see equation (1)). Under this model, the observed power-law data arise by merging mobility within containers with mobility that transports a person between containers. The container model focuses on a specific aspect of mobility, and neglects other important features, including temporal visitation patterns, exploration<sup>3</sup> and the structural connectedness of geographical spaces (for example, through transportation networks)<sup>48–50</sup>. These could be incorporated in future versions of the model. Fitting the model to trajectories collected in two distinct datasets, consisting of approximately 700,000 GPS traces of individuals distributed across the world, we found that—across individuals—the containers have typical sizes, representing the ‘scales’ of human mobility. We showed that our model allows for better understanding of mobility behaviour and improves on the state of the art in modelling.

## Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2909-1>.

1. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462–465 (2006).
2. González, M. C., Hidalgo, C. A. & Barabási, A.-L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).
3. Song, C., Koren, T., Wang, P. & Barabási, A.-L. Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818–823 (2010).
4. Paasi, A. Place and region: looking through the prism of scale. *Prog. Hum. Geogr.* **28**, 536–546 (2004).
5. Marston, S. A. The social construction of scale. *Prog. Hum. Geogr.* **24**, 219–242 (2000).
6. Cresswell, T. *On the Move: Mobility in the Modern Western World* (Taylor & Francis, 2006).
7. Kaluza, P., Kölzsch, A., Gastner, M. T. & Blasius, B. The complex network of global cargo ship movements. *J. R. Soc. Interface* **7**, 1093–1103 (2010).
8. Kraemer, M. U. et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
9. Song, X., Zhang, Q., Sekimoto, Y. & Shibusaki, R. Prediction of human emergency behavior and their mobility following large-scale disaster. In *Proc. 20th ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* 5–14 (ACM, 2014).
10. Becker, F. & Axhausen, K. W. Literature review on surveys investigating the acceptance of automated vehicles. *Transportation* **44**, 1293–1306 (2017).
11. Barbosa, H. et al. Human mobility: models and applications. *Phys. Rep.* **734**, 1–74 (2018).
12. Larsen, J. & Urry, J. *Mobilities, Networks, Geographies* (Routledge, 2016).
13. Hirtle, S. C. & Jonides, J. Evidence of hierarchies in cognitive maps. *Mem. Cognit.* **13**, 208–217 (1985).
14. Von Thünen, J. H. *Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie* Vol. 13 (G Fischer, 1910).
15. Christaller, W. *Die zentralen Orte in Süddeutschland: eine ökonomisch-geographische Untersuchung über die Gesetzmässigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen* (Wissenschaftliche Buchgesellschaft, 1980).
16. Berry, B. J. L. *Geography of Market Centers and Retail Distribution* (Prentice Hall, 1967).
17. Alonso, W. et al. *Location and Land Use. Toward a General Theory of Land Rent* (Harvard Univ. Press, 1964).
18. Cadwallader, M. *Migration and Residential Mobility: Macro and Micro Approaches* (Univ. Wisconsin Press, 1992).
19. Thiernam, C., Theis, F., Grady, D., Brune, R. & Brockmann, D. The structure of borders in a small world. *PLoS ONE* **5**, e15422 (2010).
20. Marchetti, C. Anthropological invariants in travel behavior. *Technol. Forecast. Soc. Change* **47**, 75–88 (1994).
21. 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); correction **7**, <https://doi.org/10.1371/annotation/ca85bf7a-7922-47d5-8fbf-bcdf25af8c72> (2020).
22. Alessandretti, L., Sapiezynski, P., Lehmann, S. & Baronchelli, A. Multi-scale spatio-temporal analysis of human mobility. *PLoS ONE* **12**, e0171686 (2017).
23. Newman, M. E. Power laws, Pareto distributions and Zipf’s law. *Contemp. Phys.* **46**, 323–351 (2005).
24. Liang, X., Zhao, J., Dong, L. & Xu, K. Unraveling the origin of exponential law in intra-urban human mobility. *Sci. Rep.* **3**, 2983 (2013).
25. Alessandretti, L., Sapiezynski, P., Sekara, V., Lehmann, S. & Baronchelli, A. Evidence for a conserved quantity in human mobility. *Nat. Hum. Behav.* **2**, 485–491 (2018).
26. Gallotti, R., Bazzani, A., Rambaldi, S. & Barthelemy, M. A stochastic model of randomly accelerated walkers for human mobility. *Nat. Commun.* **7**, 12600 (2016).
27. Gheorghiu, S. & Coppens, M.-O. Heterogeneity explains features of “anomalous” thermodynamics and statistics. *Proc. Natl Acad. Sci. USA* **101**, 15852–15856 (2004).
28. Amini, A., Kung, K., Kang, C., Sobolevsky, S. & Ratti, C. The impact of social segregation on human mobility in developing and industrialized regions. *EPJ Data Sci.* **3**, 6 (2014).
29. Fotheringham, A. S. A new set of spatial-interaction models: the theory of competing destinations. *Environ. Plan. A* **15**, 15–36 (1983).
30. Saracı, S., Doğan, N. & Doğan, İ. Comparison of hierarchical cluster analysis methods by cophenetic correlation. *J. Inequal. Appl.* **2013**, 203 (2013).
31. Barbosa, H., de Lima-Neto, F. B., Esvukoff, A. & Menezes, R. The effect of recency to human mobility. *EPJ Data Sci.* **4**, 21 (2015).
32. Gaddum, J. H. Lognormal distributions. *Nature* **156**, 463–466 (1945).
33. Romeo, M., Da Costa, V. & Bardou, F. Broad distribution effects in sums of lognormal random variables. *Eur. Phys. J. B* **32**, 513–525 (2003).
34. Clauset, A., Shalizi, C. R. & Newman, M. E. Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703 (2009).
35. Song, C., Qu, Z., Blumm, N. & Barabási, A.-L. Limits of predictability in human mobility. *Science* **327**, 1018–1021 (2010).
36. Gauvin, L. et al. Gender gaps in urban mobility. *Humanit. Soc. Sci. Commun.* **7**, 11 (2020).
37. Breheny, M. The compact city and transport energy consumption. *Trans. Inst. Br. Geogr.* **20**, 81–101 (1995).
38. Carr, L. J., Dunsiger, S. I. & Marcus, B. H. Walk Score™ as a global estimate of neighborhood walkability. *Am. J. Prev. Med.* **39**, 460–463 (2010).
39. Gaye, A. et al. *Measuring Key Disparities in Human Development: The Gender Inequality Index* Human Development Research Paper 46 (UNDP, 2010).

40. Velaga, N. R., Beecroft, M., Nelson, J. D., Corsar, D. & Edwards, P. Transport poverty meets the digital divide: accessibility and connectivity in rural communities. *J. Transp. Geogr.* **21**, 102–112 (2012).
41. Litman, T. A. Economic value of walkability. *Transp. Res. Rec.* **1828**, 3–11 (2003).
42. Baronchelli, A. & Radicchi, F. Lévy flights in human behavior and cognition. *Chaos Solitons Fractals* **56**, 101–105 (2013).
43. Han, X.-P., Hao, Q., Wang, B.-H. & Zhou, T. Origin of the scaling law in human mobility: hierarchy of traffic systems. *Phys. Rev. E* **83**, 036117 (2011).
44. Zhao, K., Musolesi, M., Hui, P., Rao, W. & Tarkoma, S. Explaining the power-law distribution of human mobility through transportation modality decomposition. *Sci. Rep.* **5**, 9136 (2015).
45. Pappalardo, L. et al. Returners and explorers dichotomy in human mobility. *Nat. Commun.* **6**, 8166 (2015).
46. Szell, M., Sinatra, R., Petri, G., Thurner, S. & Latora, V. Understanding mobility in a social petri dish. *Sci. Rep.* **2**, 457 (2012).
47. Jiang, S. et al. The TimeGeo modeling framework for urban mobility without travel surveys. *Proc. Natl Acad. Sci. USA* **113**, E5370–E5378 (2016); correction **113**, E7137 (2016).
48. Pumain, D. in *Hierarchy in Natural and Social Sciences* (ed. Pumain, D.) 169–222 (Springer, 2006).
49. Batty, M. in *Hierarchy in Natural and Social Sciences* (ed. Pumain, D.) 143–168 (Springer, 2006).
50. Arcalate, E. et al. Cities and regions in Britain through hierarchical percolation. *R. Soc. Open Sci.* **3**, 150691 (2016).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

# Article

## Methods

### Data description and pre-processing

**Mobility data.** Our analyses are based on two mobile-phone datasets collecting high-resolution human trajectories. The study procedure follows the guidelines provided by the Danish Data Protection Agency.

The D1 dataset contains anonymized GPS location data for approximately 5,000,000 individuals collected by a global smartphone and electronics company between 2017 and 2019 (Extended Data Fig. 1). The data consist of anonymized users who self-reported their age, gender, height, weight and country of residence. Data were extracted through a smartphone app. All data analysis was carried out in accordance with the European Union's General Data Protection Regulation 2016/679 (GDPR) and the regulations set out by the Danish Data Protection Agency. We selected approximately 700,000 individuals with at least one year of data and whose position is known, every day, at least 50% of the time. Individuals are located across the world and are aged between 18 and 80 years old, with an average age of 36 years. About one-third of individuals are female. Gender and age were provided by the users at the time of registration. Data are not collected at a fixed sampling rate. Instead, the location estimate is updated when there is a change in the motion state of the device (if the accelerometer registers a change). Location estimation error is below 100 m for 93% of data points. Informed consent was obtained for all study participants.

The D2 data were collected as part of an experiment that took place between September 2013 and September 2015<sup>51</sup>. The experiment involved 851 Technical University of Denmark students (about 22% female and about 78% male), typically aged between 19 and 21 years old. Participants' position over time was estimated from a combination of GPS and WiFi information, resulting in samples every 1–2 min. The location estimation error was below 50 m in 95% of the cases. Data collection was approved by the Danish Data Protection Agency. All participants provided informed consent.

The data to produce Fig. 1a are the location trajectory of one of the authors. We pre-processed all trajectories to obtain stop locations using the Infostop algorithm<sup>52</sup>. We used the following algorithm parameters:  $r_1 = 30$  m,  $r_2 = 30$  m,  $\text{min\_staying\_time} = 10$  min,  $\text{max\_time\_between} = 24$  h. Results are robust with respect to variation of these parameters (Supplementary Note 1).

**Other data.** We collected data on the walkability score in the area surrounding individuals' home locations using the WalkScore<sup>38</sup> application programming interface (<https://www.walkscore.com/professional/walk-score-apis.php>). We collected data for 11,511 individuals living in New Zealand, Australia, Canada and the United States, for which WalkScore data were available.

Data on the urbanization level in the area surrounding individuals' home locations is based on the GHS Settlement Model grid<sup>53</sup> that delineates and classifies settlement typologies via a logic of population size, population and built-up area densities. This classification categorizes areas in urban areas, towns and rural areas. In our analysis, we merged towns and cities into a single category. Data can be downloaded from: <https://ghsl.jrc.ec.europa.eu/data.php>.

The GII dataset can be downloaded from: <http://hdr.undp.org/en/content/gender-inequality-index-gii>. We used data for 2017.

### The container model

**Model description.** The container model models the trace of an agent transitioning between locations in space. The model is specified by three sets of parameters that can be either simulated to generate synthetic traces or estimated for an empirical trace through maximum likelihood estimation. The model contains the following.

(1) A hierarchical structure  $H$  with  $L$  levels, where each level consists of containers encapsulating locations. Accordingly, each location  $k$

can be described as a sequence of containers encapsulated within each other,  $\mathbf{k} = (k_1, \dots, k_l, \dots, k_L)$ , where levels are ordered from the most fine grained  $l=1$  to the most coarse grained  $l=L$ . In analogy, a restaurant can be described as a sequence corresponding to the building, the neighbourhood, the city and so on where it is located. At each level in the hierarchy, containers have comparable size. In the simplest form, this structure is a nested grid (Supplementary Note 3).

(2) The collection of these containers' attractivenesses,  $a$ . The attractiveness  $a(k_i)$  is the probability of visiting container  $k_i$  among all containers encapsulated within  $k_{i+1}$ . Accordingly,  $\sum_{k_j \in k_{i+1}} a(k_j) = 1$ .

(3) The  $L \times L$  matrix,  $p$ , characterizing the probability of travelling at a certain level distance. Each row in  $p$  is a probability vector that describes the probabilities  $p(d, d_h)$  of travelling at level distance  $d$  when the level distance from home is  $d_h$ . Here, home is defined as the location with largest attractiveness at all levels. By level distance we mean the so-called cophenetic distance<sup>30</sup>: the highest level in the hierarchy one travels to reach the destination. It is necessary to maintain separate probability distributions for each level distance from home. This is, for example, because travelling at the highest level distance (for example, intercontinentally) is unlikely when one is near home, but comparatively likely when on a different continent.

Under the container model, each transition is the result of a two-stage decision process. First, the individual selects at which level distance to travel. Then, she selects a specific destination based on container attractiveness. Specifically, an individual located in  $j$ , chooses destination location  $k$  with probability:

$$P_{H,a,p}(j \rightarrow k) = p_{d(j,k), d(j,h)} \frac{a(k_{d(j,k)})}{1 - a(j_{d(j,k)})} \prod_{l=1}^{d(j,k)-1} a(k_l)$$

The first factor,  $p_{d(j,k), d(j,h)}$ , is the probability of travelling at level distance  $d(j, k)$ . The second factor  $\frac{a(k_{d(j,k)})}{1 - a(j_{d(j,k)})}$  is the probability of choosing container  $k_{d(j,k)}$ . Such a container is found at level  $d(j, k)$  in the hierarchy and has attractiveness  $a(k_{d(j,k)})$ . The renormalization  $1 - a(j_{d(j,k)})$  accounts for the fact that container  $j_{d(j,k)}$  cannot be selected (this detail is not present in the main text for readability). The third factor  $\prod_{l=1}^{d(j,k)-1} a(k_l)$  is the probability of picking all other containers  $k_l$  that encapsulate location  $k$ , for any level in the hierarchy lower than  $d(j, k)$ . Note that the way we model destination choice in a hierarchical fashion connects to the class of choice models called nested logit models<sup>54</sup>. The nested structure of the physical space in the container model relates, in part, to the organization of the transportation system<sup>20,55,56</sup> and to the concrete structure of our built environment<sup>16,49</sup>. The importance of these contexts are also gradually being recognized in the human mobility literature, where early studies focused on large datasets, but did not consider the effect of contextual information, for example, transportation type or other mobility characteristics, which can introduce heterogeneity<sup>21,36,44,57–62</sup>.

**Generating traces.** We model transitions as a two-step decision process. Thus, we can simulate synthetic trajectories given a hierarchical description  $H$ , container attractivenesses  $a$  and the probability matrix  $p$ , (either designed or obtained by fitting the container model to an empirical trace). We simulate the mobility of an agent by the following algorithm. To guide the reader we offer an example at each step, describing an agent travelling across a hierarchy where levels correspond to countries, cities, neighbourhoods, buildings and locations.

(1) Initialize the agent in a random location,  $j$ , at level distance  $d(j, h)$  from the home location. Example: the agent is initialized in location  $j$  located in a different country than her home country.

(2) Select a level distance  $l^*$  that the agent should travel at, by drawing from the multinomial distribution,  $p_{d(j,h)}$ . Example: the agent chooses to travel at the city distance.

(3) Select a destination,  $k$ :

(a) At level  $l^*$ , select a container  $k_{l^*}$ , by drawing from the attractiveness distribution over the containers encapsulated in  $j_{l+1}j_{l^*}$ , cannot be selected in this process, so  $k_{l^*} \neq j_{l^*}$ . Example: the agent chooses the destination city among other cities in the same country where she is currently located.

(b) At level  $(l^*-1)$ , select a container  $k_{l^*-1}$  encapsulated within  $k_{l^*}$ , by drawing from the attractiveness distribution over containers in  $k_{l^*}$ . Continue this process until level 1 is reached. Example: the agent picks a neighbourhood, then a building and then a location encapsulated within the destination city chosen in the previous step.

(4) Repeat steps (2) and (3) for any desired number of displacements.

**Likelihood optimization.** We can fit the container model to an empirical trace and obtain the model parameters  $H$ ,  $a$  and  $p$ , using maximum likelihood estimation. We write the likelihood that a sequence of individual locations  $T = [k(0), \dots, k(i), \dots, k(n_T)]$ , where  $i$  is the sequence index and  $n_T$  is the length of the sequence, was generated by an instance of the container model specified by  $H$ ,  $a$  and  $p$ , as:

$$\mathcal{L}(H, a, p | T) = \prod_{i=0}^{n_T-1} P_{H,a,p}(k(i-1) \rightarrow k(i)),$$

where  $P_{H,a,p}(k(i-1) \rightarrow k(i))$  is the probability of a transition to occur.

Unlike spatial clustering methodologies, this method allows us to identify ‘containers’, structures that are not only compact in size but also contain mobility behaviour. This optimization task, however, is computationally expensive; therefore, we approach the problem according to the following heuristic.

First, we note that, when  $n_T$  is large and  $H$  is chosen,  $p$  and  $a$  are trivial to estimate. In fact, for  $n_T \rightarrow \infty$ , element  $p_{d,d_h}$  of matrix  $p$  equals the fraction of transitions covering a level distance  $d$  among all transitions starting at level distance  $d_h$  from home. Similarly, for  $n_T \rightarrow \infty$ , the attractiveness of a container equals the fraction of times such container is selected among containers in the same parent container.

Thus, for long enough traces, we can estimate the maximum likelihood by exploring different choices of  $H$  only, where  $H$  is effectively a spatial hierarchical partition of individual locations. To ensure that clusters are compact, we choose  $H$  among the solutions of the complete linkage hierarchical clustering algorithm<sup>63</sup>.

First, we run the complete-linkage algorithm for the set of locations in sequence  $T$ . The algorithm initializes each location as a separate cluster. It then iteratively joins the two clusters whose union has the smallest diameter, defined as the maximum distance between two locations in a cluster, and stores the clustering solution. It runs for  $N$  iterations, where  $N$  is the number of locations (and possible clustering solutions). In the final iteration, all locations form one cluster. The result of the complete-linkage algorithm can be visualized as a dendrogram and queried for clusters at any cut distance (Extended Data Fig. 3a).

We then proceed to find the hierarchical partition  $H^*$  corresponding to the maximum likelihood  $\mathcal{L}^*$ . Exhaustive search would require computing the likelihood for all possible partitions  $H$ . When we let  $L$  range from 1 to  $N$ , we arrive at the total number of possible partitions by the following logic: for  $L = 1$ , the dendrogram is cut zero times so there is one partition, which has only individual locations and no containers; for  $L = 2$  the dendrogram is cut once, so there are  $N$  partitions because there are  $N$  ways to cut the dendrogram; for  $L = 3$ , there are  $\frac{N(N-1)}{2}$  ways to cut the dendrogram two times, and so on. The set of all possible partitions then has size  $\sum_{L=1}^N \binom{N}{L}$ . We define a heuristic to reduce the set of candidate partitions  $H$ , by optimizing the likelihood one level at a time (Extended Data Fig. 3b). The algorithm works as follows. First, we compute the likelihood  $\mathcal{L}_1$  of  $T$  in the case  $L = 1$ , corresponding to having no containers. Then, we test the  $N$  possible partitions corresponding to cutting the dendrogram one time (that is,  $L = 2$ ), by

computing the corresponding likelihoods. We find the cut  $C_2$  of the dendrogram resulting in the maximum likelihood  $\mathcal{L}_2$ . If  $\mathcal{L}_1$  is significantly larger than  $\mathcal{L}_2$  (tested by bootstrapping, with  $P \leq 0.01$ ), we assign  $\mathcal{L}^* = \mathcal{L}_1$ , conclude that  $H^*$  has only one level (individual locations) and stop the algorithm. Otherwise, we explore the set of partitions corresponding to two cuts of the dendrogram (that is,  $L = 3$ ), where one of them is  $C_2$ , and find the cut  $C_3$  that yields the maximum likelihood  $\mathcal{L}_3$ . We compare  $\mathcal{L}_2$  and  $\mathcal{L}_3$ , and stop the algorithm if  $\mathcal{L}_2$  is significantly larger than  $\mathcal{L}_3$  (tested by bootstrapping, with  $P \leq 0.01$ ). We proceed for increasing values of  $L$ , until  $L = N$  or there is no significant improvement in likelihood (tested by bootstrapping, with  $P \leq 0.01$ ). In the worst-case scenario, we explore  $N!$  partitions. We validate the one-level-at-a-time algorithm against synthetic data (Extended Data Fig. 3, Supplementary Note 3). We find that the algorithm recovers the correct number of hierarchical levels about 95% of the time. The similarity between the correct and recovered hierarchical structure, measured as their cophenetic correlation<sup>30</sup> has median value 1 (the cophenetic correlation is the correlation between the cophenetic distance computed for all pairs of locations according to two different hierarchical descriptions, and thus is 1 for identical descriptions). The median absolute error relative to the estimation of the matrix of probabilities  $p$  is 0.03.

## Model validation

**Metrics.** In Fig. 2, we compare synthetic and real traces by computing quantities characterizing individual trajectories.

The radius of gyration for an individual  $u$  is defined as:

$$r_g^u = \sqrt{\frac{1}{N} \sum_{n=0}^N (r_n^u - r_{CM,n}^u)^2},$$

where  $N$  is the total number of displacements (50 in our analysis),  $r_n^u$  is the position of  $u$  after  $n$  displacements,  $r_{CM,n}^u$  is its centre of mass after  $n$  displacements, defined as:

$$r_{CM,n}^u = \frac{1}{n} \sum_{j=0}^n r_j^u.$$

The uncorrelated entropy  $S_{unc}$  is defined as:

$$S_{unc} = - \sum_{i=0}^{N_l} P(i) \log_2(P(i)),$$

where  $P(i)$  is the probability of visiting location  $i$ , and  $N_l$  is the total number of locations. The temporal entropy  $S_{temp}$ , is defined as

$$S_{temp} = - \sum_{T_{i'} \in T_i} P(T_{i'}) \log_2(P(T_{i'})),$$

where  $P(T_{i'})$  is the probability of finding a particular time-ordered subsequence  $T_{i'}$  in the trajectory  $T_i$ . We estimate  $S_{temp}$  using the method described by Sekara et al.<sup>64</sup>

**EPR model.** We generate EPR synthetic traces as follows. First, we fit the model parameters<sup>9</sup> and determine, for each individual, the number of visited locations  $S$  as well as the number of visits  $f_i$  per location  $i$  using traces with one-year duration. Then, we generate traces using the model described in Song et al.<sup>9</sup>. At each new displacement, an individual explores a new place with probability  $\rho S^{-\gamma}$  and exploits a previously known location with the complementary probability, where  $\rho$  and  $\gamma$  are parameters of the model. In the first case, she chooses a place at distance  $\Delta r$ , extracted from a power-law distribution  $P(\Delta r) \approx \Delta r^{-\beta}$ . In the latter case, she chooses a previously visited location  $i$  with probability proportional to the number of visits  $f_i$ . See Supplementary Note 4 for further details and the implementation of other models.

# Article

## Data availability

Derived data that support the findings of this study are available in DTU Data with the identifier <https://doi.org/10.11583/DTU.12941993.v1>. Additional data related to this paper may be requested from the authors. Raw data for dataset D1 are not publicly available to preserve individuals' privacy under the European General Data Protection Regulation. Raw data for dataset D2 are not publicly available due to privacy considerations, but are available to researchers who meet the criteria for access to confidential data, sign a confidentiality agreement and agree to work under supervision in Copenhagen. Please direct your queries to the corresponding author. Source data are provided with this paper.

## Code availability

Code is available at [https://github.com/lalessan/scales\\_human\\_mobility/](https://github.com/lalessan/scales_human_mobility/).

51. Stopczynski, A. et al. Measuring large-scale social networks with high resolution. *PLoS ONE* **9**, e95978 (2014).
52. Aslak, U. & Alessandretti, L. Infostop: scalable stop-location detection in multi-user mobility data. Preprint at <https://arxiv.org/abs/2003.14370> (2020).
53. Pesaresi, M. et al. *Operating Procedure for the Production of the Global Human Settlement Layer from Landsat Data of the Epochs 1975, 1990, 2000, and 2014* (Publications Office of the European Union, 2016).
54. Train, K. E. *Discrete Choice Methods with Simulation* (Cambridge Univ. Press, 2009).
55. Zahavi, Y. & Ryan, J. The stability of travel components over time. *Traffic Eng. Control* **750**, 19–26 (1978).
56. Miller, H. J. Tobler's first law and spatial analysis. *Ann. Assoc. Am. Geogr.* **94**, 284–289 (2004).
57. Kraemer, M. U. G. et al. Mapping global variation in human mobility. *Nat. Hum. Behav.* **4**, 800–810 (2020).
58. Steele, J. E. et al. Mapping poverty using mobile phone and satellite data. *J. R. Soc. Interface* **14**, 20160690 (2017).
59. Lu, X. et al. Detecting climate adaptation with mobile network data in Bangladesh: anomalies in communication, mobility and consumption patterns during cyclone Mahasen. *Climatic Change* **138**, 505–519 (2016).
60. Lu, X., Bengtsson, L. & Holme, P. Predictability of population displacement after the 2010 Haiti earthquake. *Proc. Natl Acad. Sci. USA* **109**, 11576–11581 (2012).
61. Weiss, D. J. et al. A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336 (2018).
62. Althoff, T. et al. Large-scale physical activity data reveal worldwide activity inequality. *Nature* **547**, 336–339 (2017).
63. Everitt, B. S., Landau, S., Leese, M. & Stahl, D. Hierarchical clustering. *Cluster Anal.* **5**, 71–110 (2011).
64. Sekara, V., Stopczynski, A. & Lehmann, S. Fundamental structures of dynamic social networks. *Proc. Natl Acad. Sci. USA* **113**, 9977–9982 (2016).

**Acknowledgements** We thank J. Sellergren for being the best; F. Simini and L. K. Hansen for providing insightful comments; and M. C. Gonzalez for help with auxiliary datasets. The work was supported in part by the Villum Foundation and the Danish Council for Independent Research.

**Author contributions** L.A., U.A. and S.L. designed the study and the model. L.A. and U.A. performed the analyses and implemented the model. L.A., U.A. and S.L. analysed the results and wrote the paper.

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

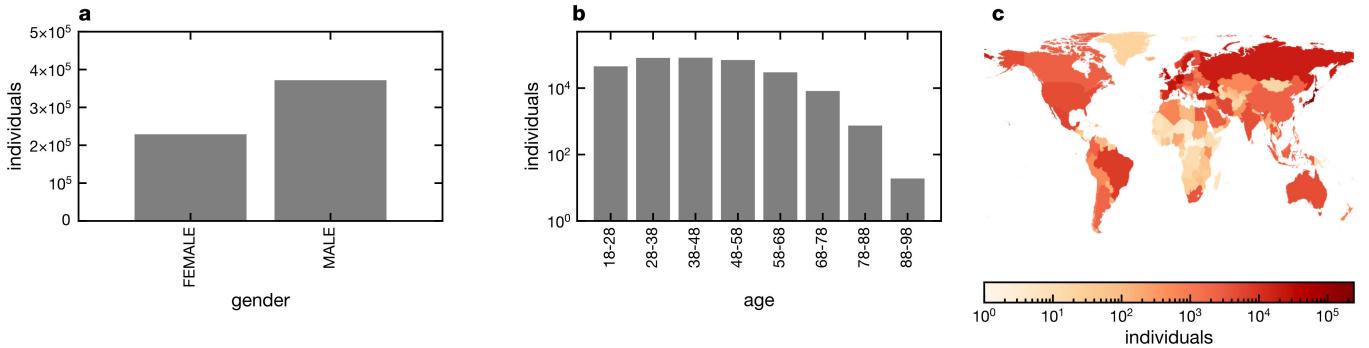
### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41586-020-2909-1>.

**Correspondence and requests for materials** should be addressed to S.L.

**Peer review information** *Nature* thanks Elsa Arcaute, Denise Pumain and José Ramasco for their contribution to the peer review of this work. Peer review reports are available.

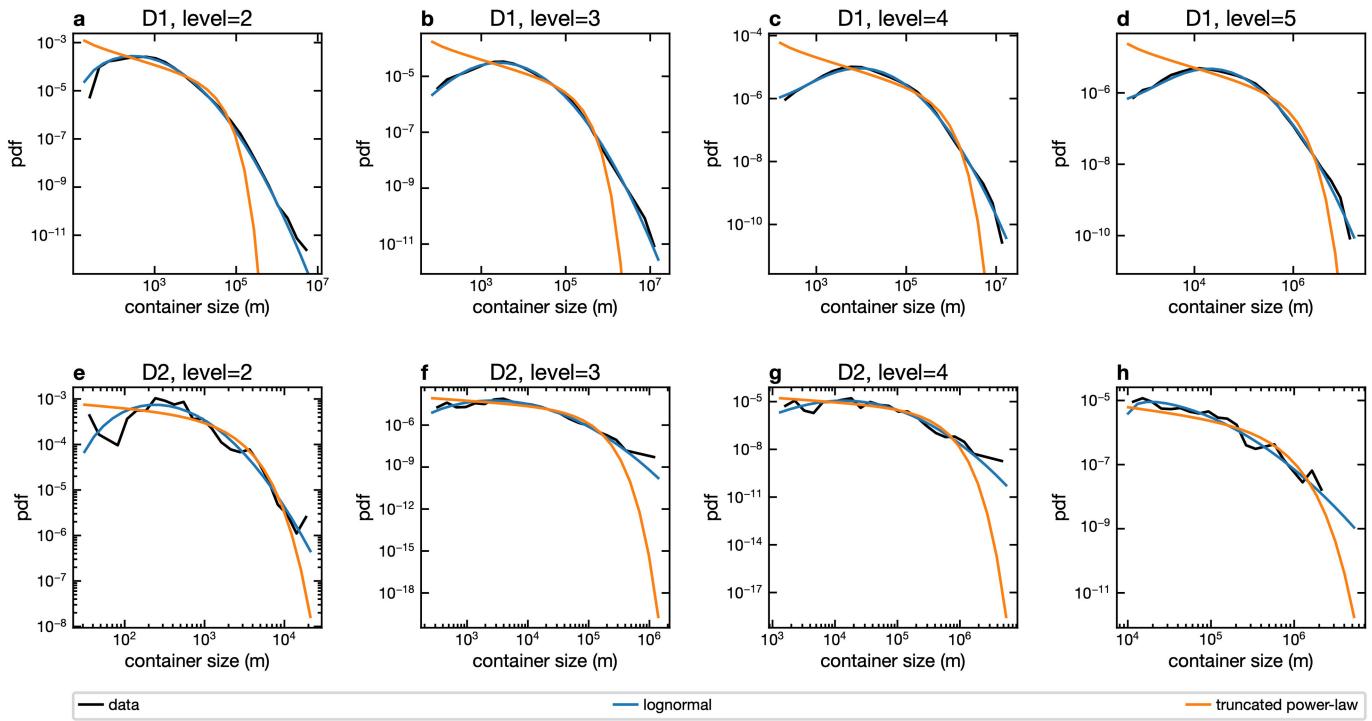
**Reprints and permissions information** is available at <http://www.nature.com/reprints>.



**Extended Data Fig. 1 | The D1 dataset.** **a**, Number of individuals for each gender. **b**, Number of individuals per age group. **c**, Number of individuals per country (see colour scale). We considered the 600,817 individuals in our dataset with at least one year of data, and whose time coverage (the fraction of

time an individual position is known) was higher than 50% at any given day. For these individuals, we considered one year of data with highest median time coverage. Map data from the GADM Database of Global Administrative Areas, version 3.6, available at <http://www.gadm.org>.

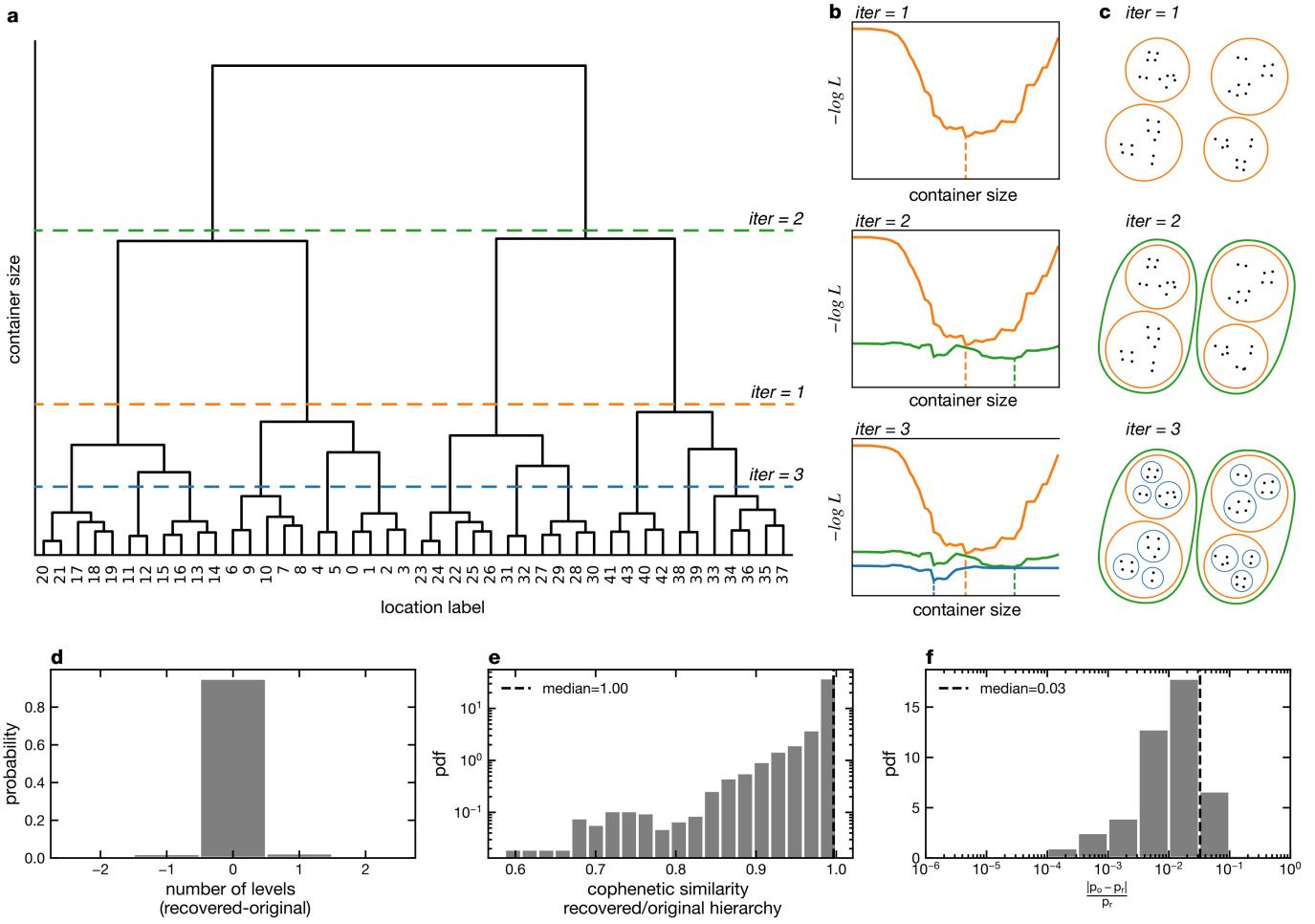
# Article



## Extended Data Fig. 2 | Distribution of container sizes at different levels.

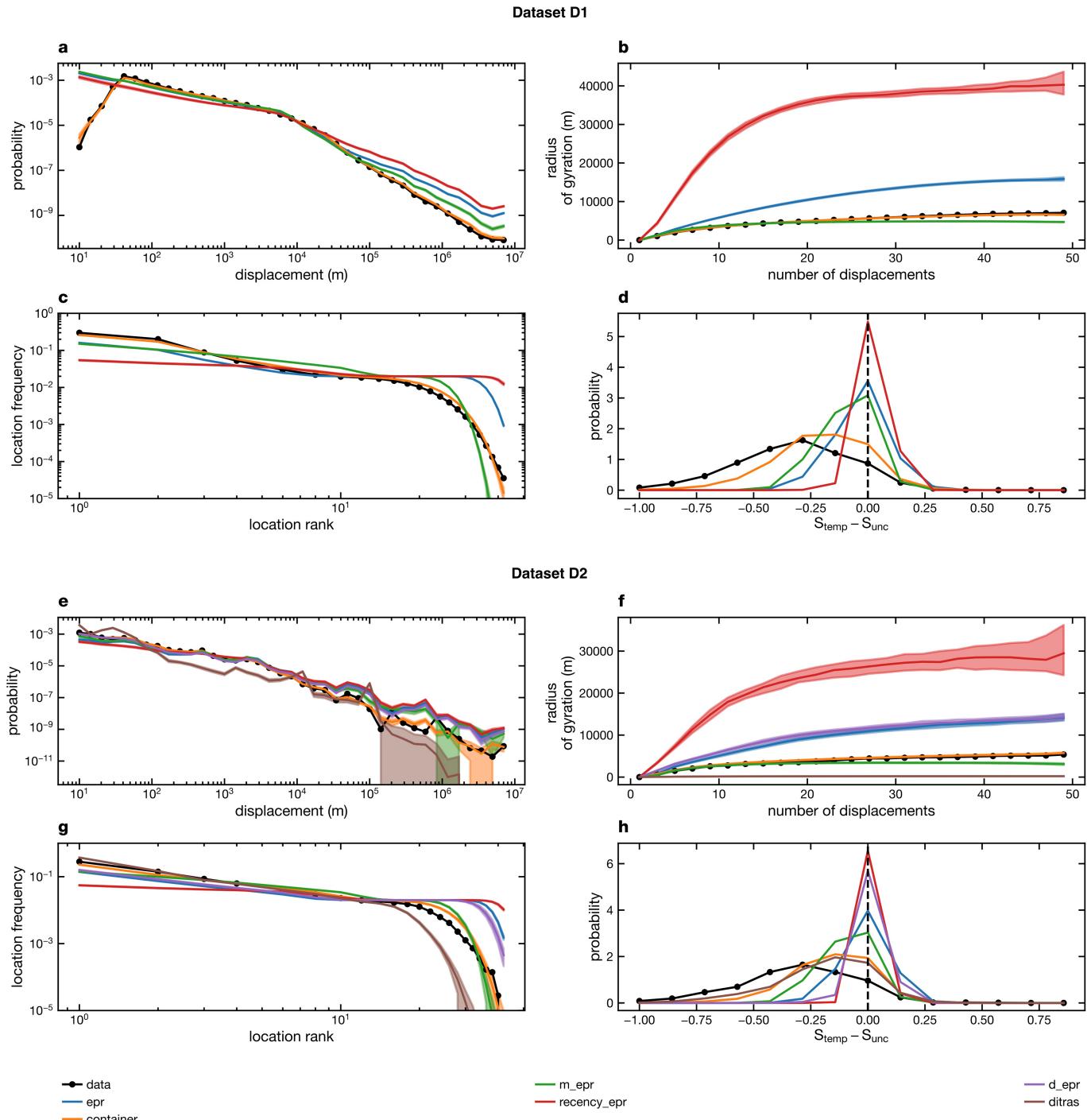
**a–h**, Distribution of individual container sizes at hierarchical levels 2 (**a**, **e**), 3 (**b**, **f**), 4 (**c**, **g**) and 5 (**d**, **h**) (black line) and the corresponding lognormal (blue line)

and truncated power-law (orange line) fits. Results are shown for the D1 (**a–d**) and D2 (**e–h**) datasets.



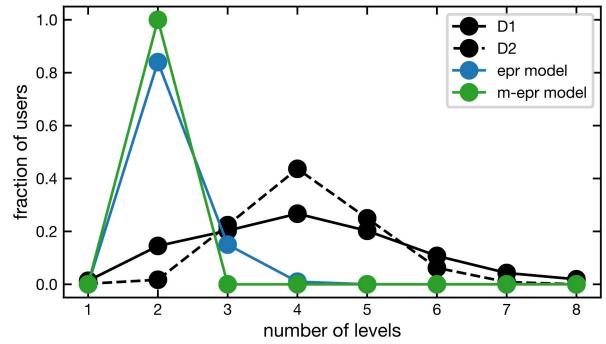
**Extended Data Fig. 3 | Schematic description and validation of the likelihood optimization algorithm.** We find the hierarchical partitioning corresponding to a sequence of locations as follows. **a**, Individual locations are iteratively merged to form clusters via the complete linkage algorithm. Here the output of the algorithm is visualized as a dendrogram. **b**, We add levels to the hierarchical partition by maximizing the likelihood of the container model one level at a time: at the first iteration (top), we find the container size ( $x$ -axis) corresponding to the dendrogram cut (dashed line) that minimizes the negative likelihood ( $y$ -axis), if any. We proceed by adding more dendrogram cuts (middle and bottom), and thus hierarchical levels, until the likelihood can no longer be improved. **c**, The dendrogram cuts correspond to a hierarchical partitioning of individual locations. We evaluate the ability of the algorithm to recover the original parameters using 5,000 synthetic traces of 3,000 locations.

**d**, Distribution of the difference between the number of recovered and original levels. The difference is 0 in 70% of the cases. **e**, Probability density associated with the cophenetic similarity between the original and recovered hierarchical structure. The dashed line corresponds to the median value. **f**, Probability density associated to the relative difference  $|p_o - p_r| / p_r$  between original ( $p_o$ ) and recovered ( $p_r$ ) entries of the matrix  $p$ .



**Extended Data Fig. 4 | The container model generates realistic synthetic traces.** **a, e.** The distribution of displacements for the entire population. **b, f.** The median individual radius of gyration versus the number of displacements. **c, g.** The average visitation frequency versus the rank of individuals' locations. **d, h.** The distribution of the difference between the real entropy  $S_{temp}$  and the uncorrelated entropy  $S_{unc}$  across individuals. Results are shown for real traces (black line, dots), and traces generated by various models

(see legend), for dataset D1 (a-d) and D2 (e-h). In a, c, d, e, g and h, the filled areas for the synthetic traces include two standard deviations around the mean computed across 1,000 simulations for each user. In b and f, the filled areas include the interquartile range. For each individual, we fitted the models considering a training period of one year. The data used here for validation corresponds to the 50 individual displacements following the training period.



**Extended Data Fig. 5 | Number of hierarchical levels recovered from traces.**

Distribution of the number of hierarchical levels found by the container model for trajectories in the D1 dataset (plain black line), the D2 dataset (dashed black line), and 1,000 synthetic traces generated by the EPR model<sup>9</sup> (blue line) and the memory EPR (m-EPR) model (green line).

# Article

**Extended Data Table 1 | The distribution of container sizes is not scale free**

Dataset	D1					D2										
	30 m				20 m				30 m				50 m			
Location size	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
Hierarchical level	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
truncated power-law	114347***	66452***	47085***	28175***	56***	85***	58**	26***	53***	51***	85***	26***	38***	48***	69***	32***
power-law	713820***	508332***	314747***	166193***	230***	573***	288***	46	256***	244***	248***	70*	185***	303***	282***	385***
log-logistic	5070***	7208***	2657***	461***	5*	-1	-1	3	7**	4*	-2	-1	0	0	0	1
log-gamma	2811817***	1529163***	613320***	496578***	451***	761***	540***	169***	428***	422***	652***	200***	321***	431***	630***	280***
log-laplace	26175***	32457***	15927***	6022***	34***	14*	9	5	32***	29***	15**	2	31***	10	19**	14***
log-weibull	553359***	377751***	277157***	154326***	213***	295***	198***	83***	191***	204***	298***	94***	118***	166***	249***	138***

The log-likelihood ratio  $R$  (ref. <sup>34</sup>) comparing the lognormal to other distributions (one per row) as a model for the distribution of container sizes. When  $R$  is positive, the lognormal distribution has higher likelihood compared with the alternative, and vice versa. The table reports also the  $P$  values associated with  $R$  (\* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ ). Results are shown at different hierarchical levels (rows), for dataset D2 and for dataset D1 under different choices of the parameter characterizing the typical size of individual locations<sup>52</sup>.

**Extended Data Table 2 | The distribution of time spent within container is not scale free**

Dataset	D1				D2											
	30 m				20 m				30 m				50 m			
Location size	2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
truncated power-law	176821***	104424***	50225***	22965***	129***	66**	174***	89***	52***	78**	122***	95***	65***	42***	233***	64**

The log-likelihood ratio  $R$  (ref.<sup>34</sup>) comparing the lognormal to the truncated power-law distribution as a model for the distribution of time spent within a container before transitioning to a different one. When  $R$  is positive, the lognormal distribution has higher likelihood compared with the alternative, and vice versa. The table reports also the  $P$  values associated with  $R$  (\* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ ). Results are shown at different hierarchical levels (rows), for dataset D2 and for dataset D1 under different choices of the parameter characterizing the typical size of individual locations<sup>52</sup>.

# Article

**Extended Data Table 3 | Characteristics of the lognormal distributions of container sizes**

$l$	$\mu_l$	$\sigma_l$	median, $e^{\mu_l}$	mode	coeff. variation
2	8.036±0.002	1.526±0.001	3089±6	301±1	3.042±0.007
3	10.206±0.002	1.524±0.002	27064±58	2653±14	3.033±0.008
4	11.390±0.002	1.519±0.002	88442±217	8791±51	3.010±0.009
5	11.993±0.003	1.459±0.002	161634±498	19233±136	2.721±0.010

The parameters  $\mu_l$  and  $\sigma_l$  characterizing the lognormal distributions of container sizes at level  $l$ . We report also the median  $e^{\mu_l}$ , the mode and the coefficient of variation  $\sqrt{e^{\sigma_l^2} - 1}$  defined as the fraction between the standard deviation and the mean<sup>33</sup>.

**Extended Data Table 4 | The container model describes unseen data better than other individual mobility models**

Dataset	D1			D2		
	$P(\Delta r)$	$P(f_i)$	$P(S_{unc} - S_{temp})$	$P(\Delta r)$	$P(f_i)$	$P(S_{unc} - S_{temp})$
EPR	81659***	93296***	44490***	2082***	5963***	7533***
m-EPR	367763***	87617***	47140***	28675***	2275***	3315***
d-EPR			557*	3610***	4885***	
DITRAS			122684***	11498***	14630***	
rec-EPR	253090***	497676***	108554***	7625***	25709***	27410***

The log-likelihood ratio  $R$  (ref. <sup>34</sup>) comparing the likelihood of the container model to other models (one per row). When  $R$  is positive, the container model has higher likelihood compared with the alternative, and vice versa. The table reports also the  $P$  values associated with  $R$  (\* $P \leq 0.05$ , \*\* $P \leq 0.01$ , \*\*\* $P \leq 0.001$ ). Results are shown for the D1 and D2 datasets. We report the results obtained considering different properties of the trajectories: the probability of displacement length  $P(\Delta r)$ , the probability of number of visits per location  $P(f_i)$ , and the probability of the difference between the uncorrelated and temporal entropy  $P(S_{unc} - S_{temp})$ .