

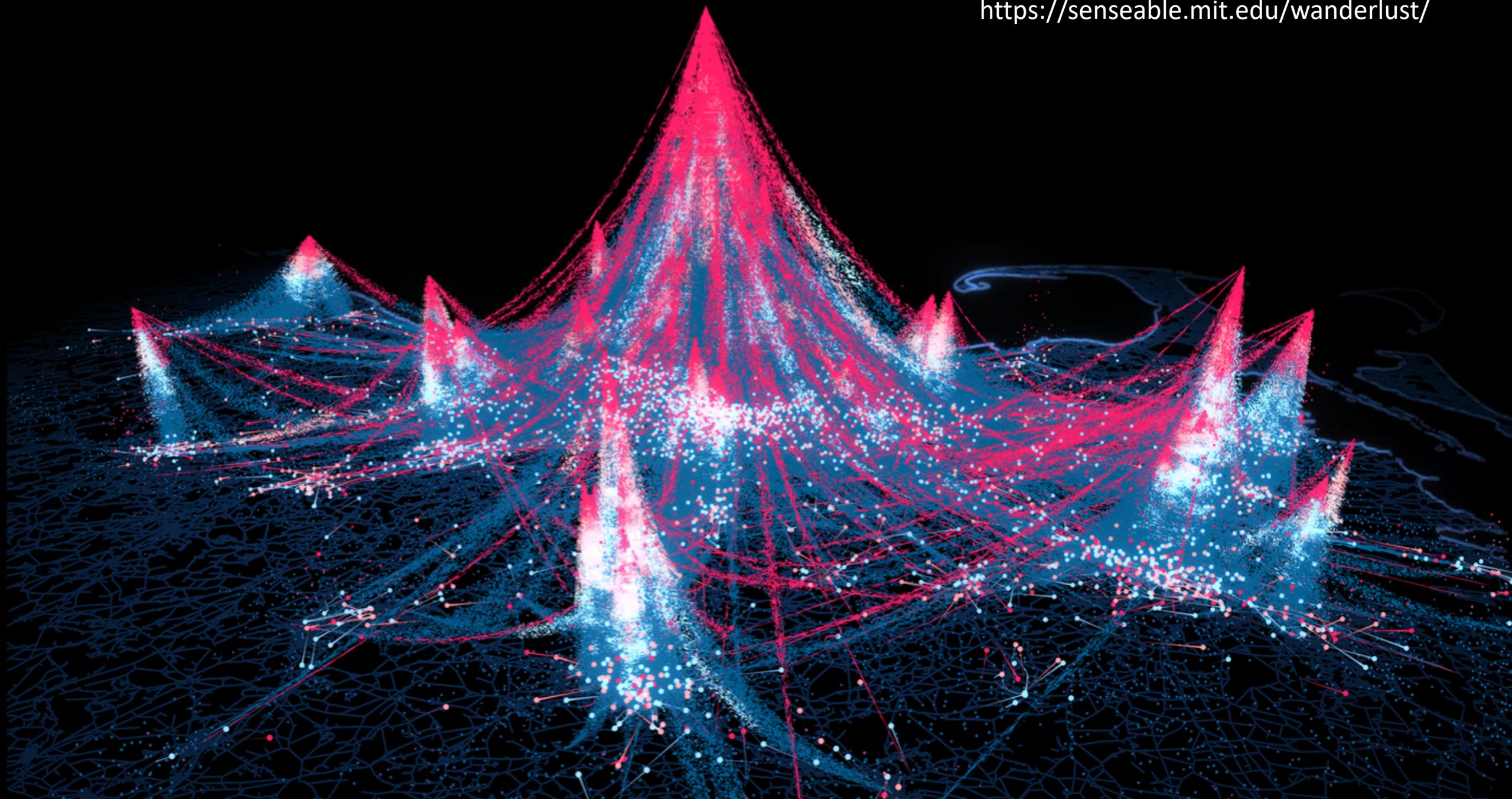
# The Universal Visitation Law of Human Mobility

Lei Dong

2021-06

Schläpfer, M., Dong, L.\*, O'Keeffe, K., Santi, P., Szell, M., Salat, H., Anklesaria, S., Vazifeh, M., Ratti, C., and West, G.B.  
The universal visitation law of human mobility. *Nature* **593**, 522–527 (2021).

<https://senseable.mit.edu/wanderlust/>



$$w_{ij} = C \frac{N_i^\alpha N_j^\gamma}{f(d_{ij})}$$

# Background

- Human Mobility (aggregated level)
  - Gravity law (Ravenstein, 1885; Zipf, 1946; +++)
  - Intervening opportunity model (Stouffer, ASR, 1940'; +)
  - Radiation model (Simini et al, Nature, 2012; ++)
  - Population weighted opportunity model (Yan et al, Interface, 2014; +)
  - Field model (Mazzoli et al, Nat Comm, 2019; +)
- Human Mobility (individual level)
  - Exploration and preferential return (EPR) model (Song et al, Nat Phys, 2010; +++)
- Human Mobility + Spatial Structure
  - Container model (Alessandretti et al, Nature, 2020)
  - More (Louail et al, Nat Comm, 2015; Barbosa et al, Phys Rep, 2018; ++)

# Background

第42卷 第4期  
2013年7月

电子科技大学学报  
Journal of University of Electronic Science and Technology of China

Vol.42 No.4  
Jul. 2013

## 人类行为时空特性的统计力学

周 涛<sup>1,2</sup>, 韩筱璞<sup>2,3</sup>, 闫小勇<sup>4,5</sup>, 杨紫陌<sup>1</sup>, 赵志丹<sup>1</sup>, 汪秉宏<sup>2,6</sup>

(1. 电子科技大学互联网科学中心 成都 611731; 2. 中国科学技术大学近代物理系 合肥 230026;  
3. 杭州师范大学阿里巴巴商学院 杭州 310036; 4. 石家庄铁道大学交通运输学院 石家庄 050043;  
5. 北京师范大学系统科学学院 北京 海淀区 100875; 6. 上海理工大学复杂系统科学研究中心 上海 杨浦区 200093)

**【摘要】**人类行为的定量化分析,特别是时空统计规律的挖掘和建模,是当前统计物理与复杂性科学的研究热点。对人类行为的深入理解,有助于解释若干复杂的社会经济现象,并在舆情监控、疾病防治、交通规划、呼叫服务、信息推荐等方面产生应用价值。该文综述人类行为时间和空间特性方面的研究进展,内容包括人类行为时间特性的实证分析和建模,人类行为空间特性的实证分析和建模,以及人类行为统计分析的应用研究。该文还将评述当前研究存在的亮点和不足,指出若干亟待解决的重大理论和实际问题。

**关 键 词** 人类动力学; 人类行为; 非泊松统计特性; 标度律

中图分类号 N94

文献标志码 A doi:10.3969/j.issn.1001-0548.2013.04.001

# Background



Physics Reports

Volume 734, 6 March 2018, Pages 1-74



## Human mobility: Models and applications

Hugo Barbosa <sup>a</sup>✉, Marc Barthelemy <sup>b, c</sup>✉, Gourab Ghoshal <sup>a</sup>✉, Charlotte R. James <sup>d</sup>✉, Maxime Lenormand <sup>e</sup>✉,  
Thomas Louail <sup>g</sup>✉, Ronaldo Menezes <sup>h</sup>✉, José J. Ramasco <sup>f</sup>✉, Filippo Simini <sup>d</sup>✉, Marcello Tomasini <sup>h</sup>✉

Show more ▾

+ Add to Mendeley   Share   Cite

<https://doi.org/10.1016/j.physrep.2018.01.001>

[Get rights and content](#)

### Abstract

Recent years have witnessed an explosion of extensive geolocated datasets related to human movement, enabling scientists to quantitatively study individual and collective mobility patterns, and to generate models that can capture and reproduce the spatiotemporal structures and regularities in human trajectories. The study of human mobility is especially important for applications such as estimating migratory flows, traffic forecasting, urban planning, and epidemic modeling. In this survey, we review the approaches developed to reproduce various mobility patterns, with the main focus on recent developments. This review can be used both as an introduction to the fundamental modeling principles of human mobility, and as a collection of technical methods applicable to specific mobility-related problems. The review organizes the subject by differentiating between individual and population mobility and also between short-range and long-range mobility. Throughout the text the description of the theory is intertwined with real-world applications.

Complexity science

# Law of human travel uncovered

Laura Alessandretti & Sune Lehmann

An analysis of mobile-phone tracking data has revealed a universal pattern that describes the interplay between the distances travelled by humans on trips and the frequency with which those trips are made. See p.522

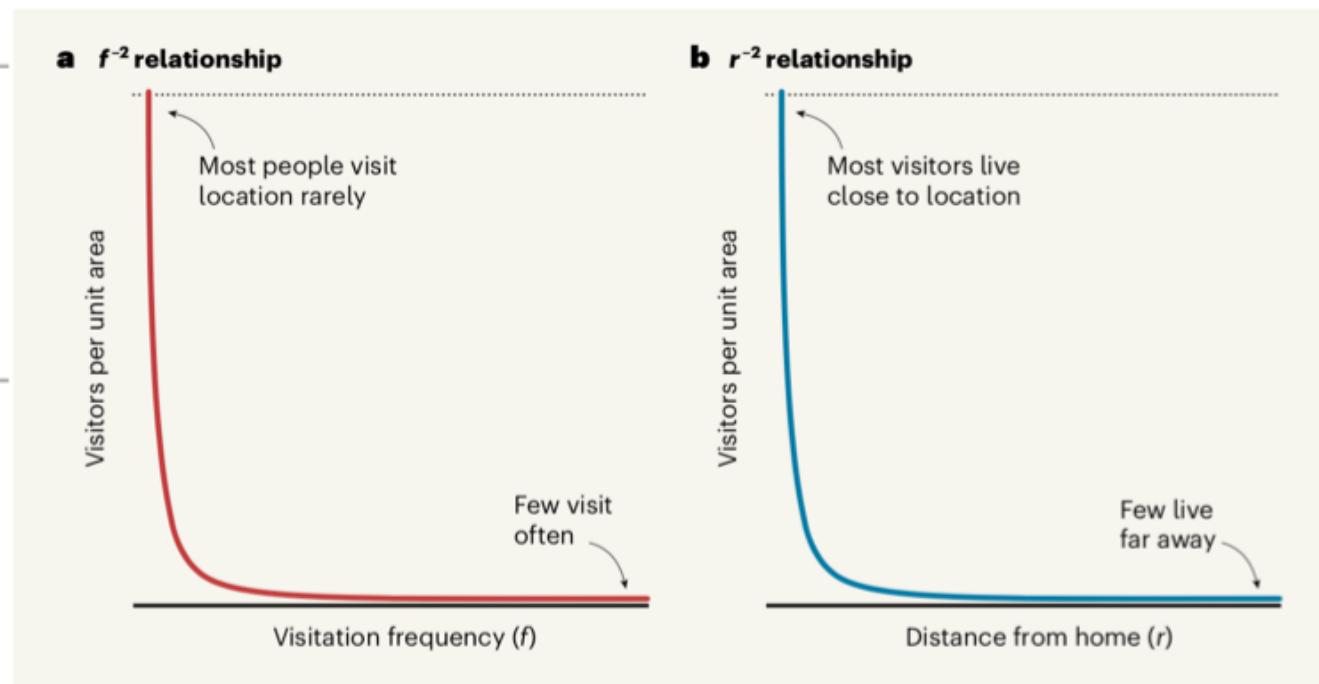
你转推了



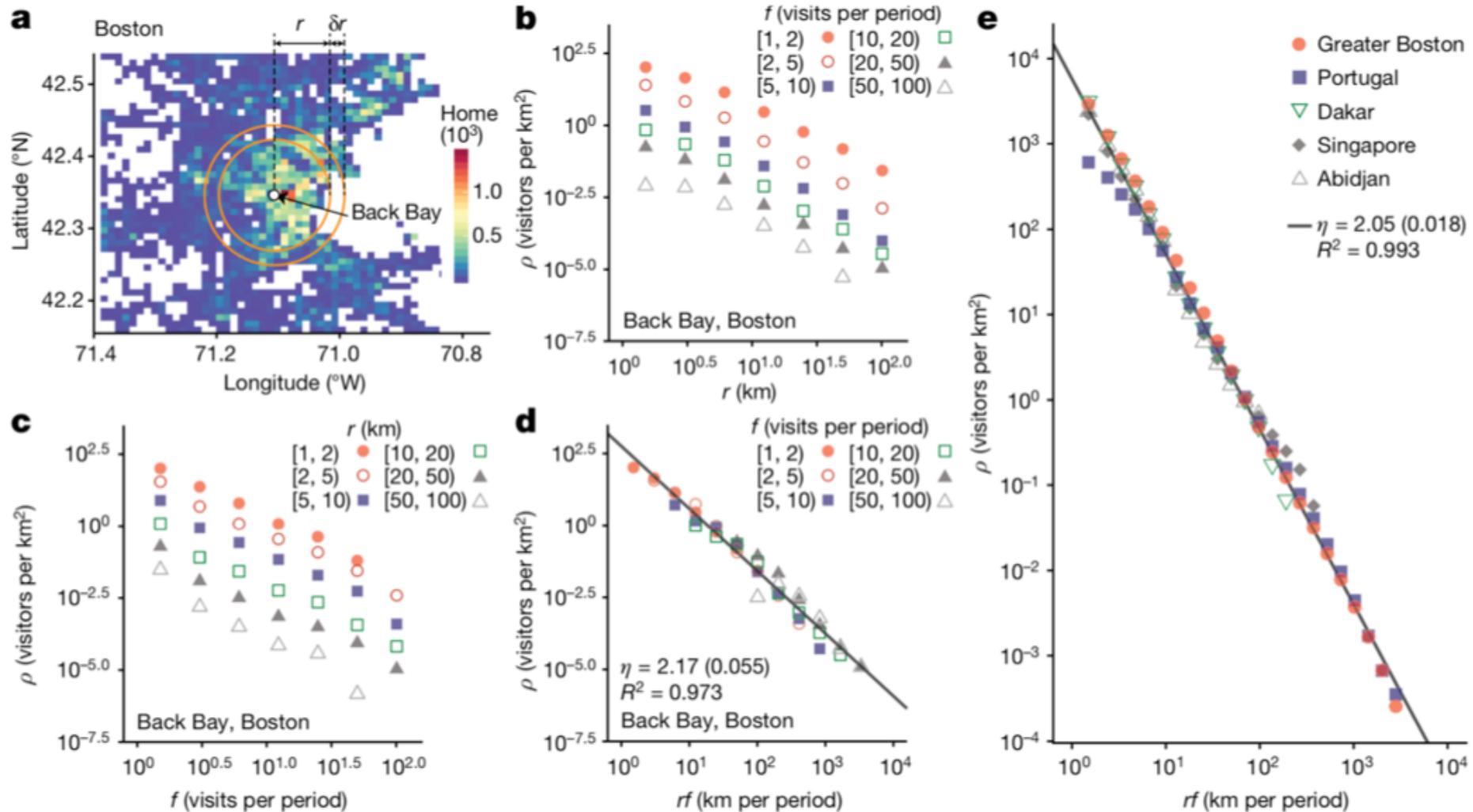
Laura Alessandretti @lau\_retti · 5月27日

...

A new universal law of Human Mobility uncovered in a new great paper out in @Nature today. A neat and robust finding across datasets worldwide. When I read the article I thought: "How could I not look for this and see it before?" :) Congratulations @mszll and all authors!



## Main Findings (1): rf-scaling



**Fig. 1 The universal distance–frequency distribution of population flows.** **a**, For each location, we count the number of visitors who are living at a distance  $[r, r + \delta r]$  away and are visiting with frequency  $f$ . **b**, For a fixed frequency  $f$ , the visitor flow to a specific location,  $\rho_i(r, f)$ , decreases with increasing distance  $r$ . **c**, When keeping the distance  $r$  fixed, the flow decreases similarly with increasing frequency  $f$ . **d**, Rescaled values collapse onto a single curve, making the flows dependent only on the single variable  $rf$ . The entire distance–frequency distribution is very well described by a power law of the form  $\rho_i(r, f) = \mu_i/(rf)^\eta$ , with exponent  $\eta \approx 2$ . **e**, Rescaled flows across all studied regions, demonstrating that the same scaling relation holds for radically different urban regions worldwide.

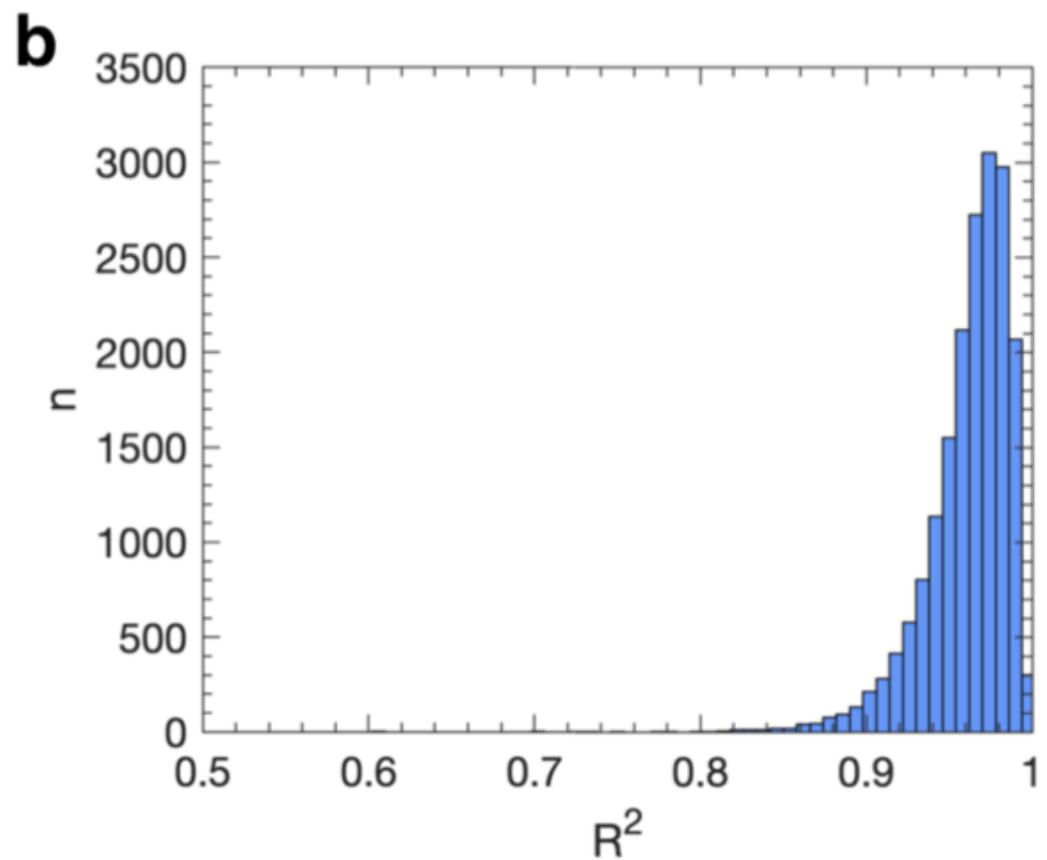
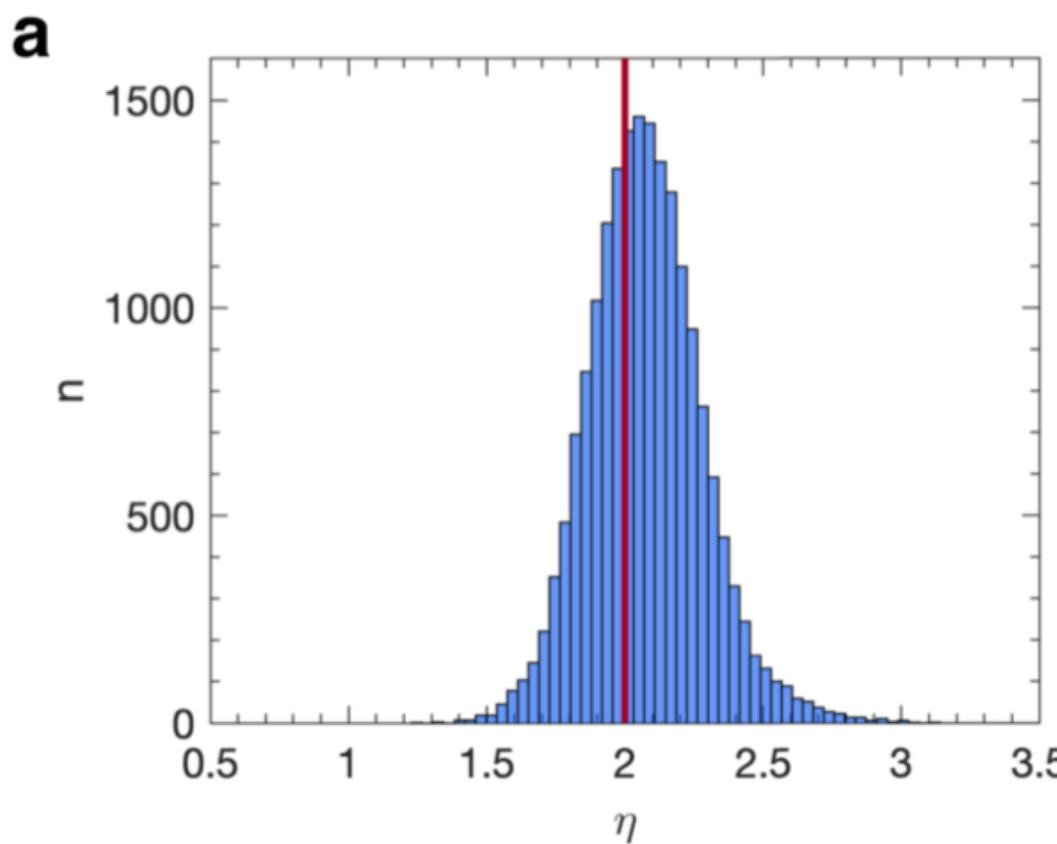
## Main Findings (1): rf-scaling

- Partitioned each geographical region into a high-resolution square grid (500m, 1km);
- Estimated the home location of each mobile phone user (the cell in which the user spent most of the time at night);
- Determined for each location  $i$  the set of unique users who visited the corresponding cell and grouped them according to **the distance  $r$  of their home location** and according to **their visitation frequency  $f$**  (number of days over a period  $T$  during which they visited for a minimum duration  $\tau$ ) ;
- To factor out the effects of area size, we normalized the resulting visitor counts,  $N_i(r, f)$ , by the area of their origin, giving  $\rho_i(r, f) = N_i(r, f) / A(r)$ , with  $A(r) \approx 2\pi r \delta r$  ;
- We will refer to the quantity  $\rho_i(r, f)$  as the ‘spectral’ flow as we essentially decompose an aggregate population flow from a given distance into its underlying frequency spectrum.
- We now show that  $\rho_i(r, f)$  does not depend on  $r$  and  $f$  separately but on the single rescaled variable  $rf$ .

$$\rho_i(r, f) = \frac{\mu_i}{(rf)^\eta},$$

The proportionality constant  $\mu_i$  determines the magnitude of the flows and thus reflects the location-specific ‘attractiveness’. The discovered scaling relation is truly remarkable as the regularity is mostly unaffected by location-specific conditions, including strong variations in surrounding population densities or in the level of economic or infrastructural development.

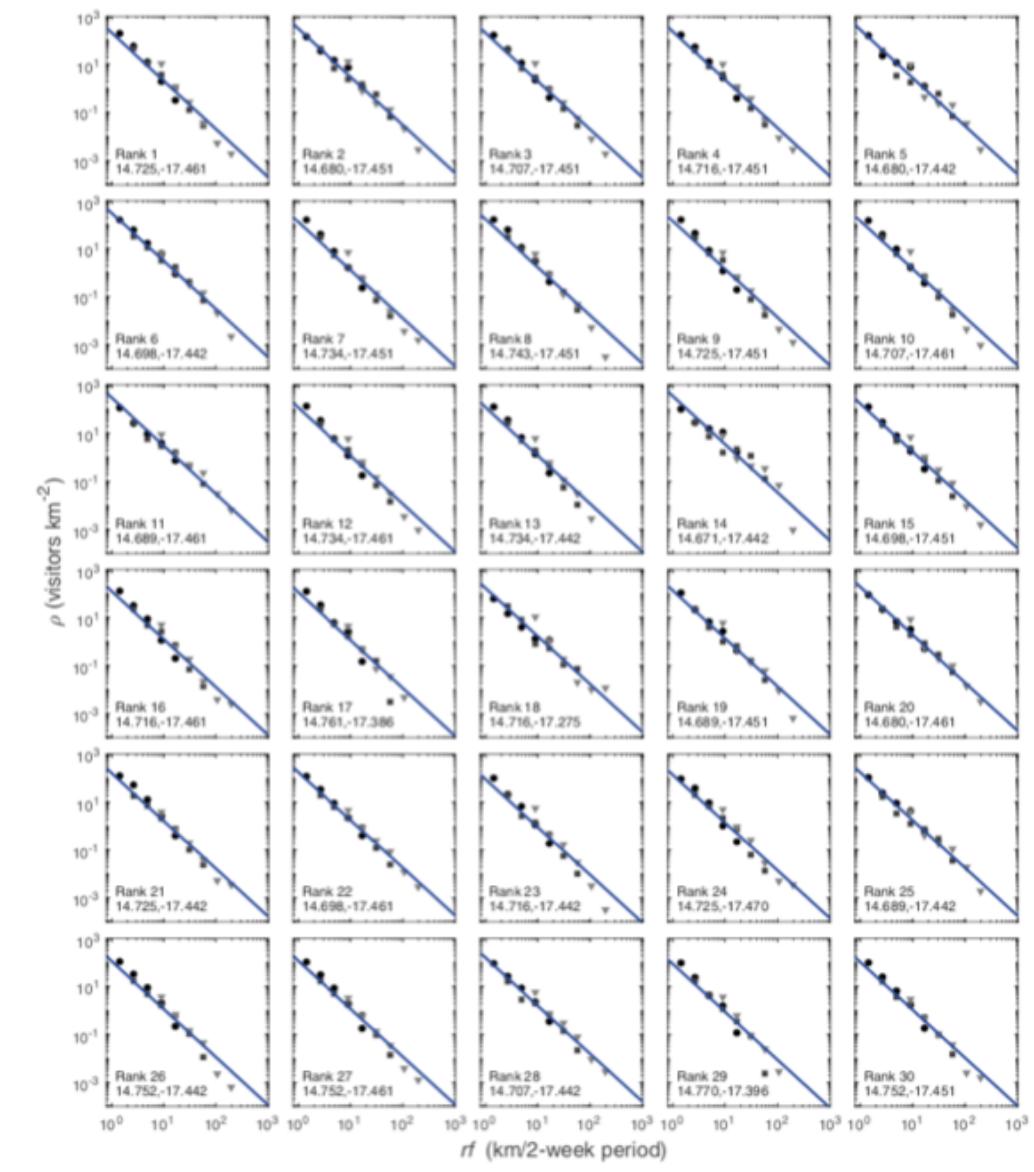
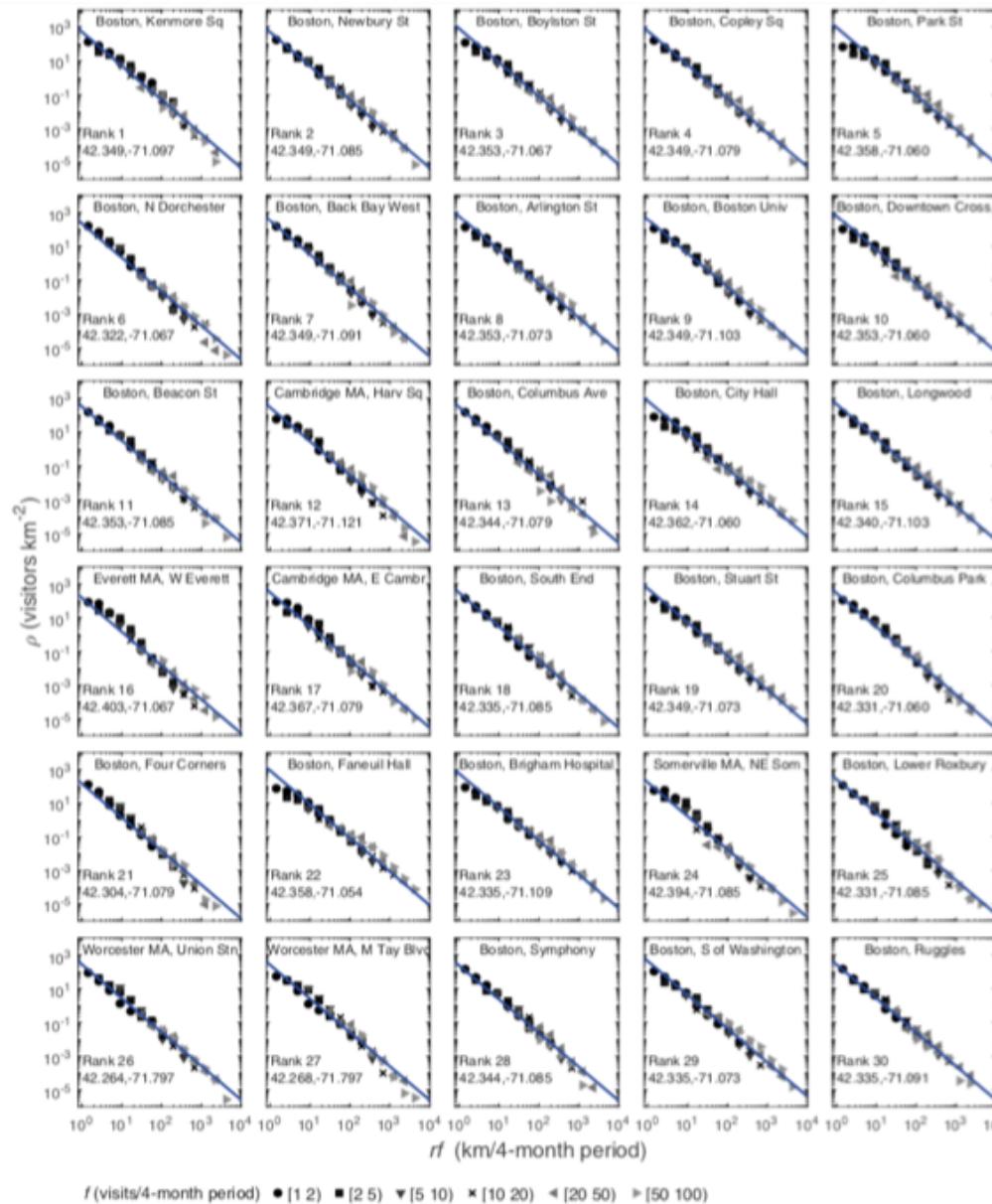
## Main Findings (1): rf-scaling



**Extended Data Fig. 2 | Empirical power-law exponents of the distance-frequency distribution.** **a**, Histogram of the exponents for all locations in the Greater Boston area. The values were determined using ordinary least squares

minimization to a linear relation of the logarithmically transformed variables. The red line shows  $\eta = 2$ , consistent with our theoretical argument. **b**, Corresponding histogram of the  $R^2$  values.

# Main Findings (1): rf-scaling



**Extended Data Fig. 3 | Universality of the scaling relation  $\rho \sim (rf)^{-2}$  across Greater Boston.** The panels depict the data for individual locations (500 m × 500 m grid cells), ranked according to the total number of visitors from neighbouring cells. Shown are locations of rank 1–30 (from top left to bottom right). The geographic coordinates of each location (latitude and longitude of the centre point of the grid cell) are indicated. The straight lines denote the inverse square of rf (slope = -2), consistent with our theoretical argument.

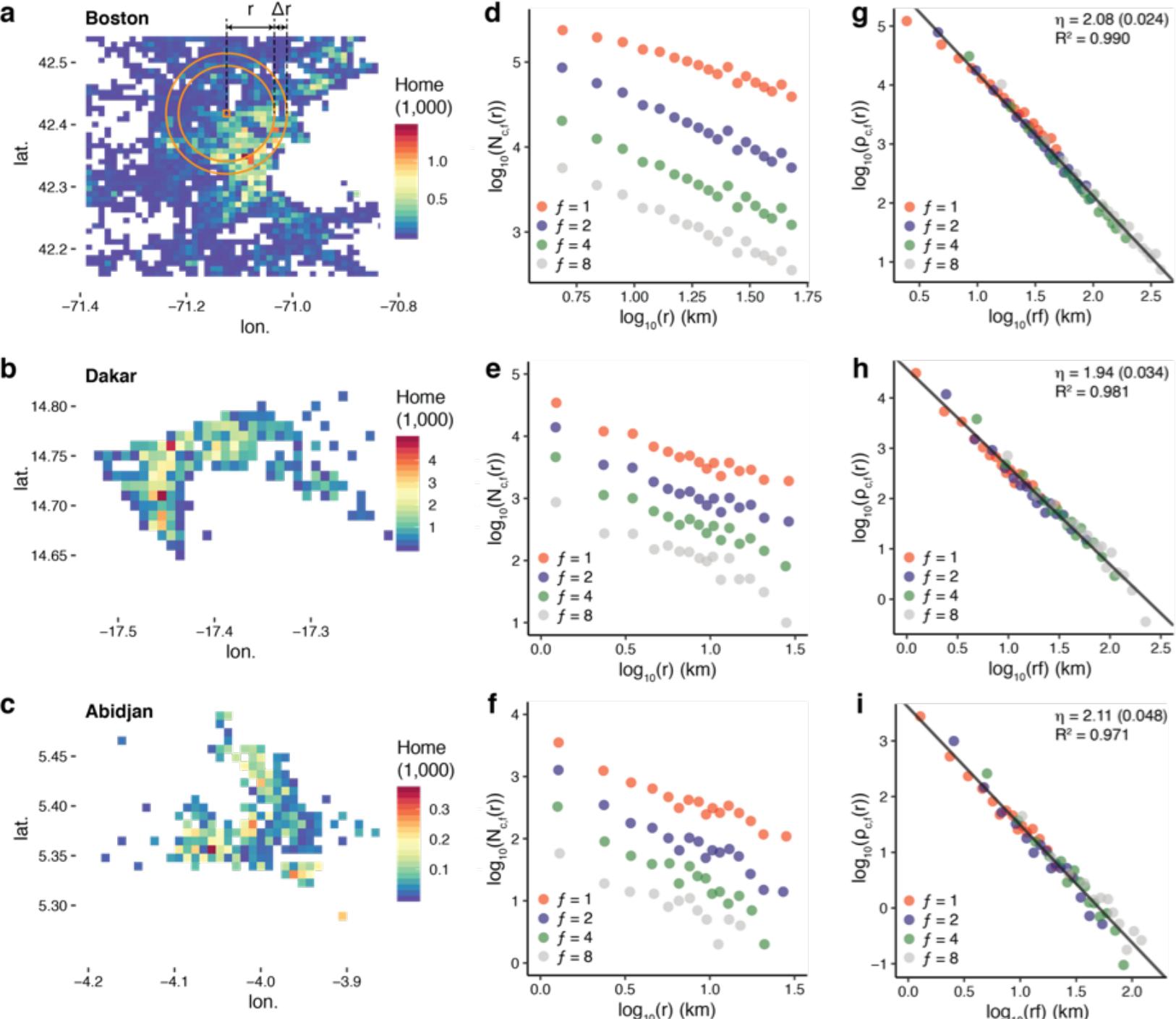
The geographic coordinates of each location (latitude and longitude of the centre point of the grid cell) are indicated. The straight lines denote the inverse square of rf (slope = -2), consistent with our theoretical argument.

**Extended Data Fig. 5 | Universality of the scaling relation  $\rho \sim (rf)^{-2}$  across Dakar.** The panels depict the data for individual locations (1 km × 1 km grid cells), ranked according to the total number of visitors from neighbouring cells. Shown are locations of rank 1–30 (from top left to bottom right).

The geographic coordinates of each location (latitude and longitude of the centre point of the grid cell) are indicated. The straight lines denote the inverse square of rf (slope = -2), consistent with our theoretical argument.

# Main Findings (1): rf-scaling

Some early results



## Main Findings (2): Conserved effective travel distance per visitor

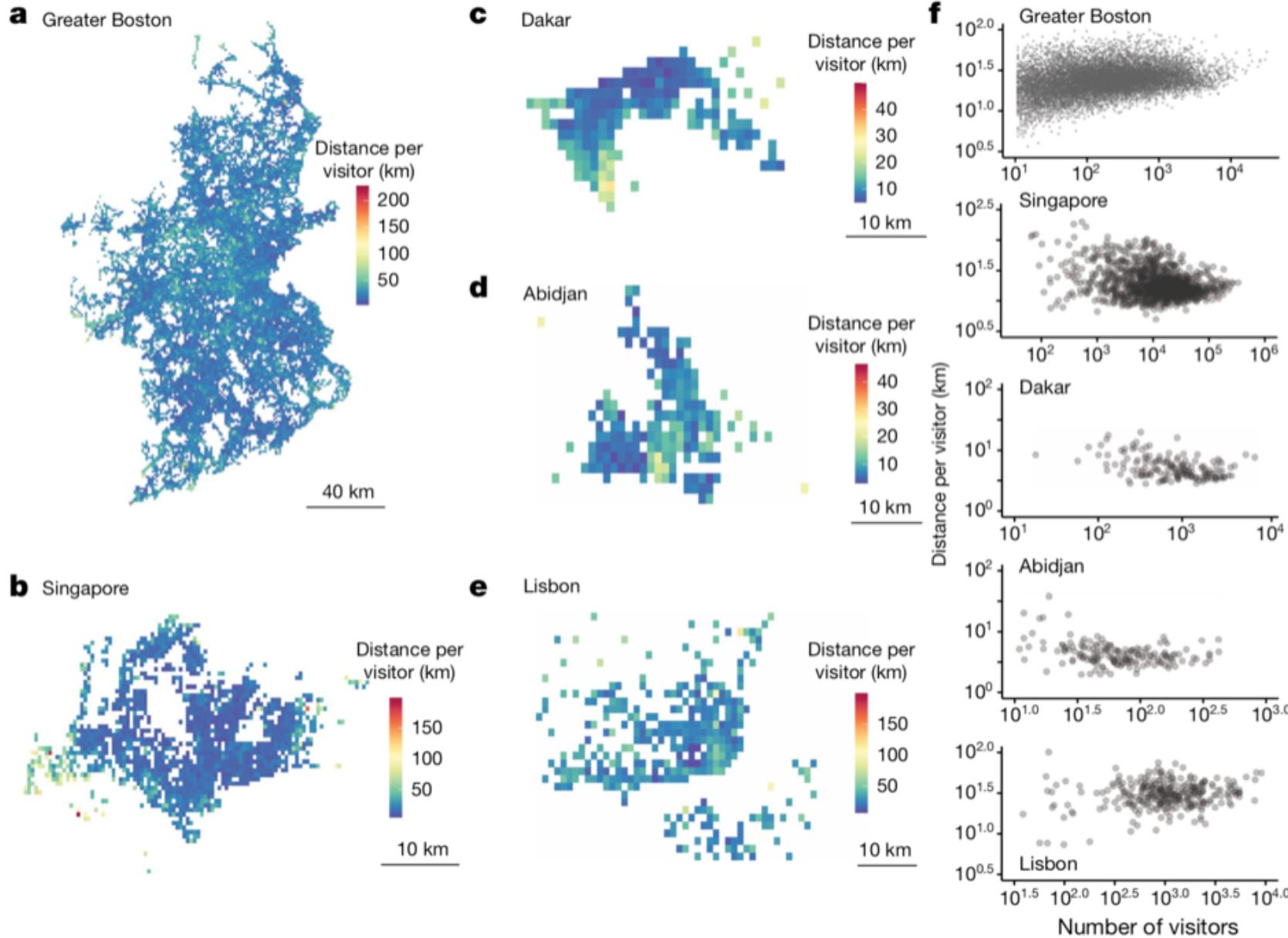
For a given location  $i$ , the effective travel distance per visitor during the observation period is

$\langle d \rangle_i = d_i^{\text{tot}} / N_i^{\text{tot}}$ , where  $d_i^{\text{tot}}$  is the effective distance travelled by all visitors during the time period  $T$ .

$$\langle d \rangle_i = \frac{\sum_f \iint r f T \rho_i(r, f) r \ dr \ d\varphi}{\sum_f \iint \rho_i(r, f) r \ dr \ d\varphi} = \frac{\sum_f \int T(rf)^{-1} r \ dr}{\sum_f \int (rf)^{-2} r \ dr}$$

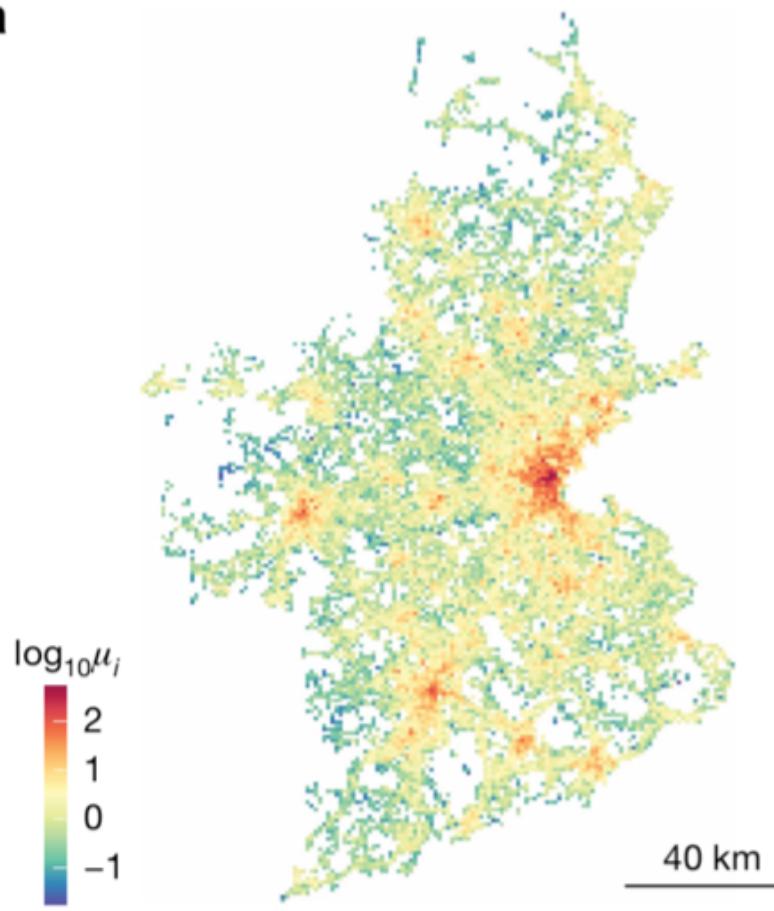
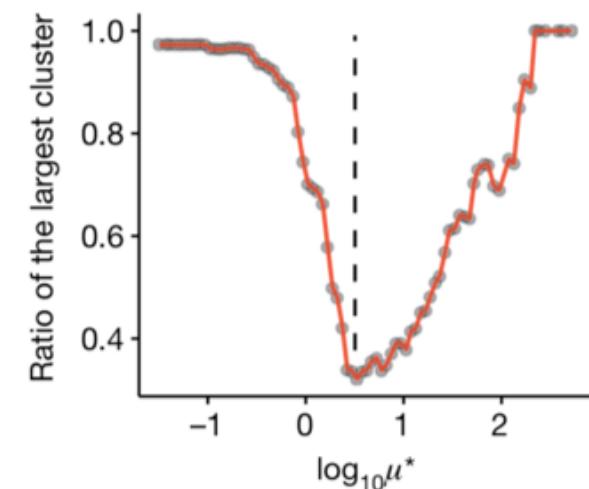
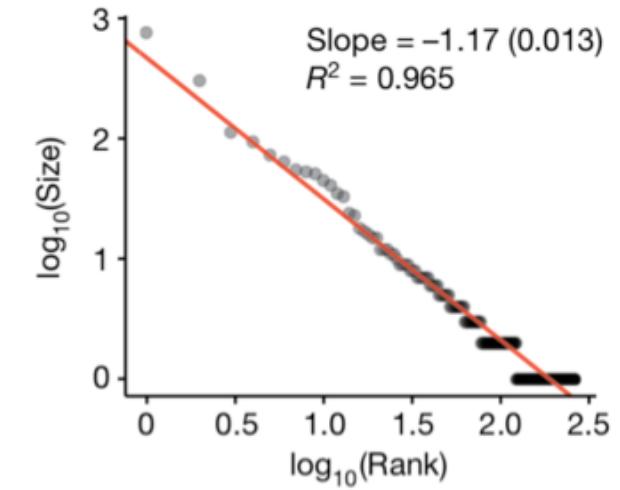
$\langle d \rangle_i$  does not depend on  $\mu_i$

## Main Findings (2): Conserved effective travel distance per visitor



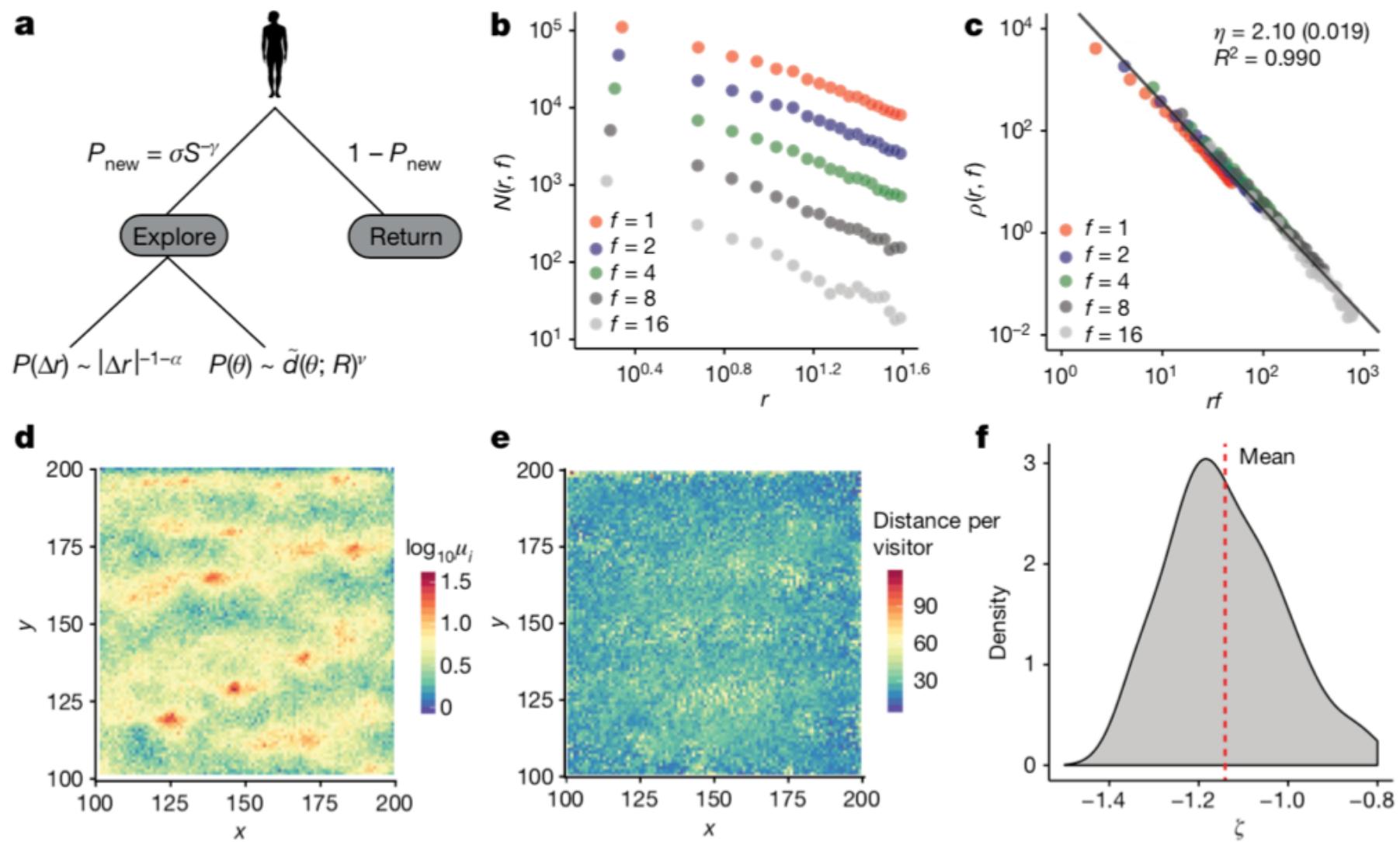
**Fig. 2 Constant effective travel distance per visitor.** a–f, The average effective distance,  $\langle d \rangle_i$ , covered by an individual over time to visit a given location  $i$  is largely invariant across space and independent of the attractiveness of the location (in terms of number of visitors) in Greater Boston (a), Singapore (b), Dakar (c), Abidjan (d) and Lisbon (e). The  $R^2$  values of the linear regression between the number of visitors and average distance per visitor, as shown in the scatter plots (f), are very small. Notice that there are some ‘anomalous’ locations that are associated with larger effective travel distances. In the majority of cases, these locations correspond to ports (for example, Singapore and Dakar) or tourism attractions (Lisbon) and thus have an intrinsic reason to attract visitors from particularly far away.

## Main Findings (3): Spatial clusters with an area distribution that follows Zipf's Law

**a****c****b****d**

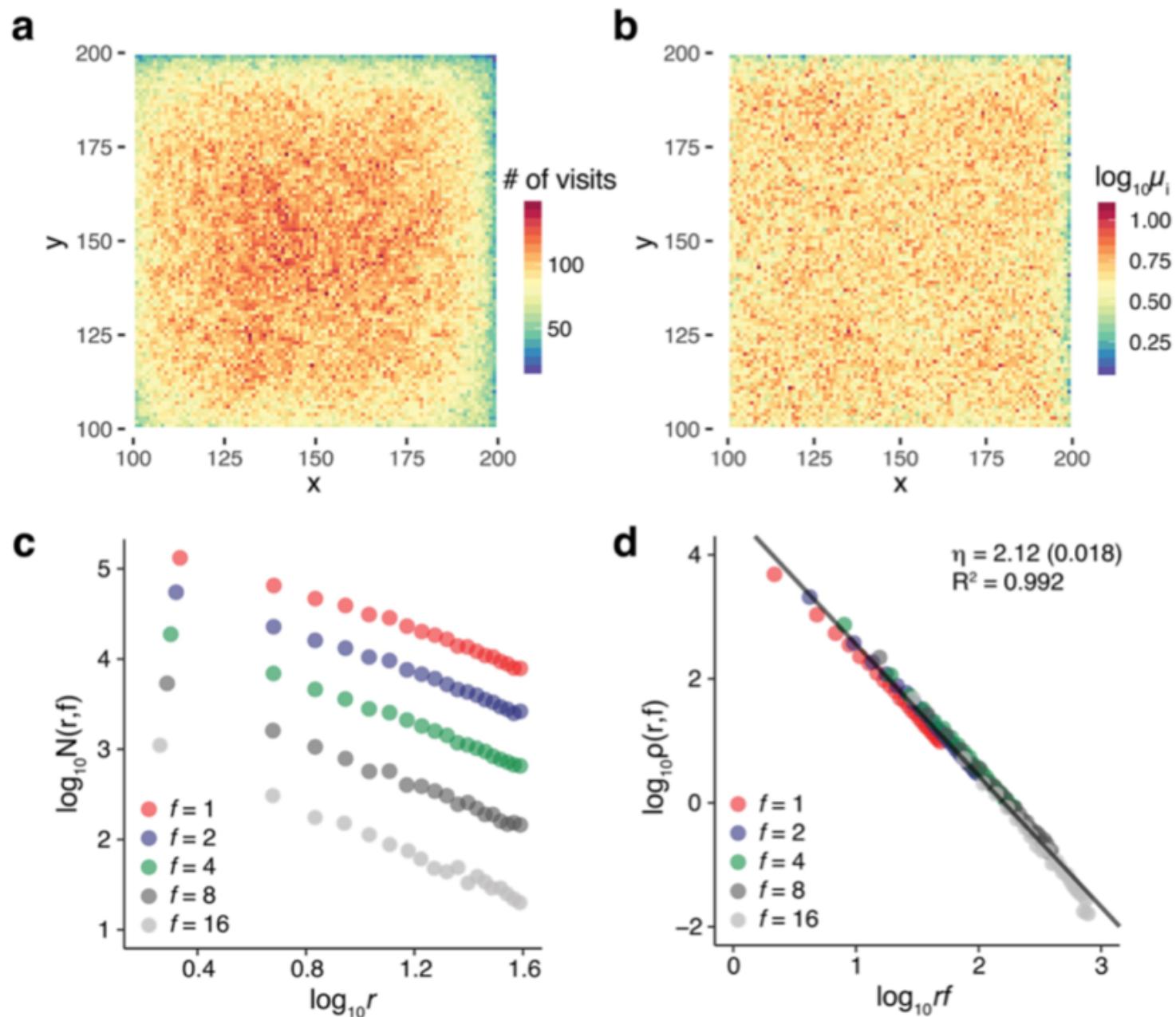
**Fig. 3 Spatial structure of the location-specific attractiveness.** **a**, Geographic distribution of the attractiveness values  $\mu_i$  across the Greater Boston area. **b**, Area ratio of the largest spatial cluster to all spatial clusters versus the minimum attractiveness threshold  $\mu^*$  as derived through the CCA. **c**, Detected clusters at the critical value of  $\mu^*$ , where the area ratio of the largest cluster is minimized. **d**, Rank-size distribution at the critical value of  $\mu^*$ . Consistent with Zipf's law, the data are well approximated by a power law with exponent  $\zeta = -1.17$ .

## Model (PEPR)



**Fig. 4 Microscopic model of spectral population flows.** **a**, Schematic of the PEPR model. At each time step, an individual (agent) decides to explore a previously unvisited location with probability  $P_{\text{new}}$ . The radial distance  $\Delta r$  and direction  $\vartheta$  of this displacement are drawn from random distributions that capture the characteristic jump-size distribution of human trajectories and the propensity to explore popular areas. With probability  $1 - P_{\text{new}}$  the agent returns to a previously visited location. **b–f**, Results for  $10^5$  agents on a regular square grid. The area distribution of the spatial clusters follows Zipf's law (derived through the application of the CCA to a total of 50 model realizations, leading to a power-law scaling with exponent  $\zeta = -1.14 \pm 0.13$ ), which is again in agreement with the data (**f**).

## Model (EPR)

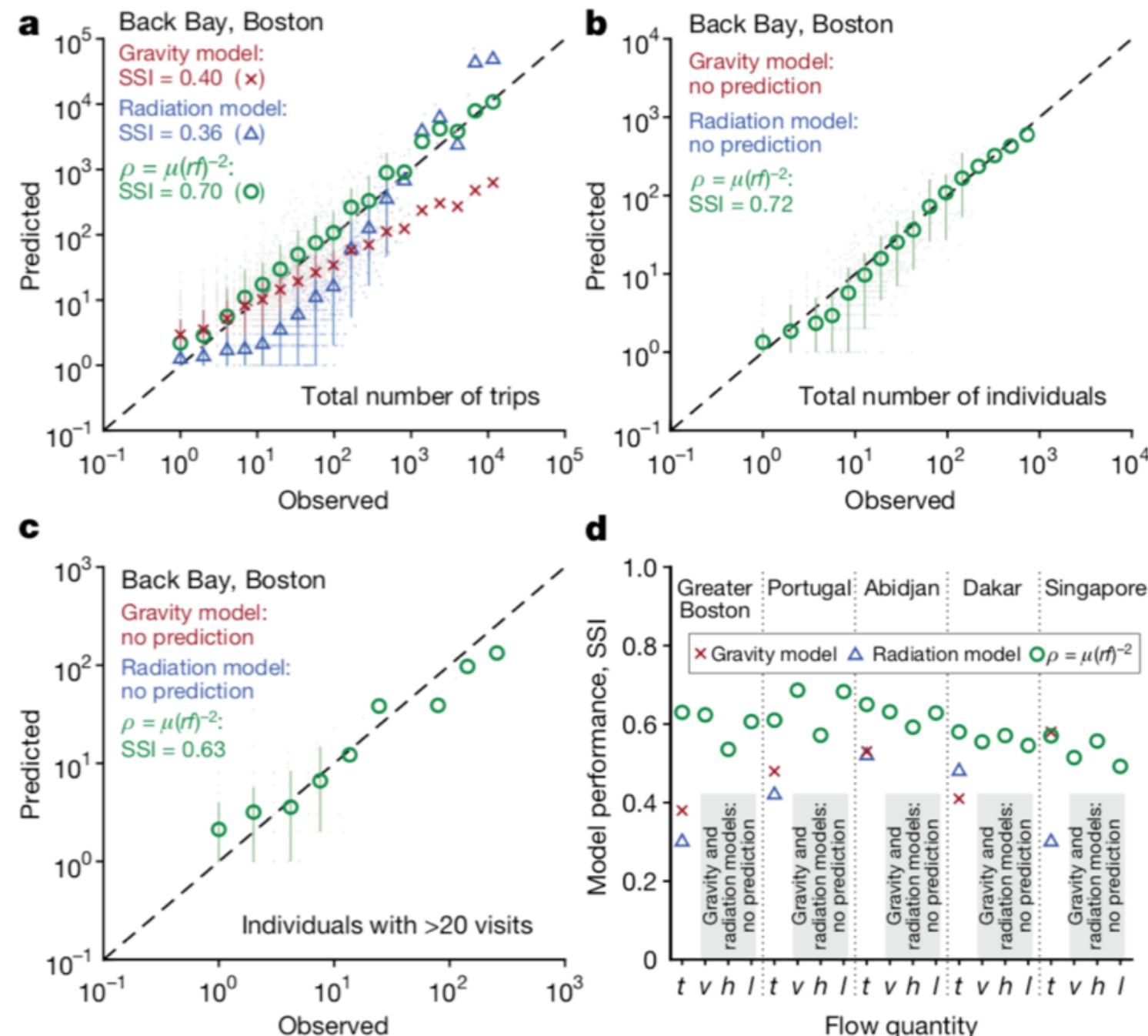


**Extended Data Fig. 7 | Simulation results of the EPR model.** **a, b,** Generated number of visits (**a**) and attractiveness values  $\mu_i$  (**b**). **c, d,** The EPR model generates the  $r^f$  scaling of the population flows with a scaling exponent that is in remarkable agreement with the data. The generated visitor counts,  $N(r, f)$ ,

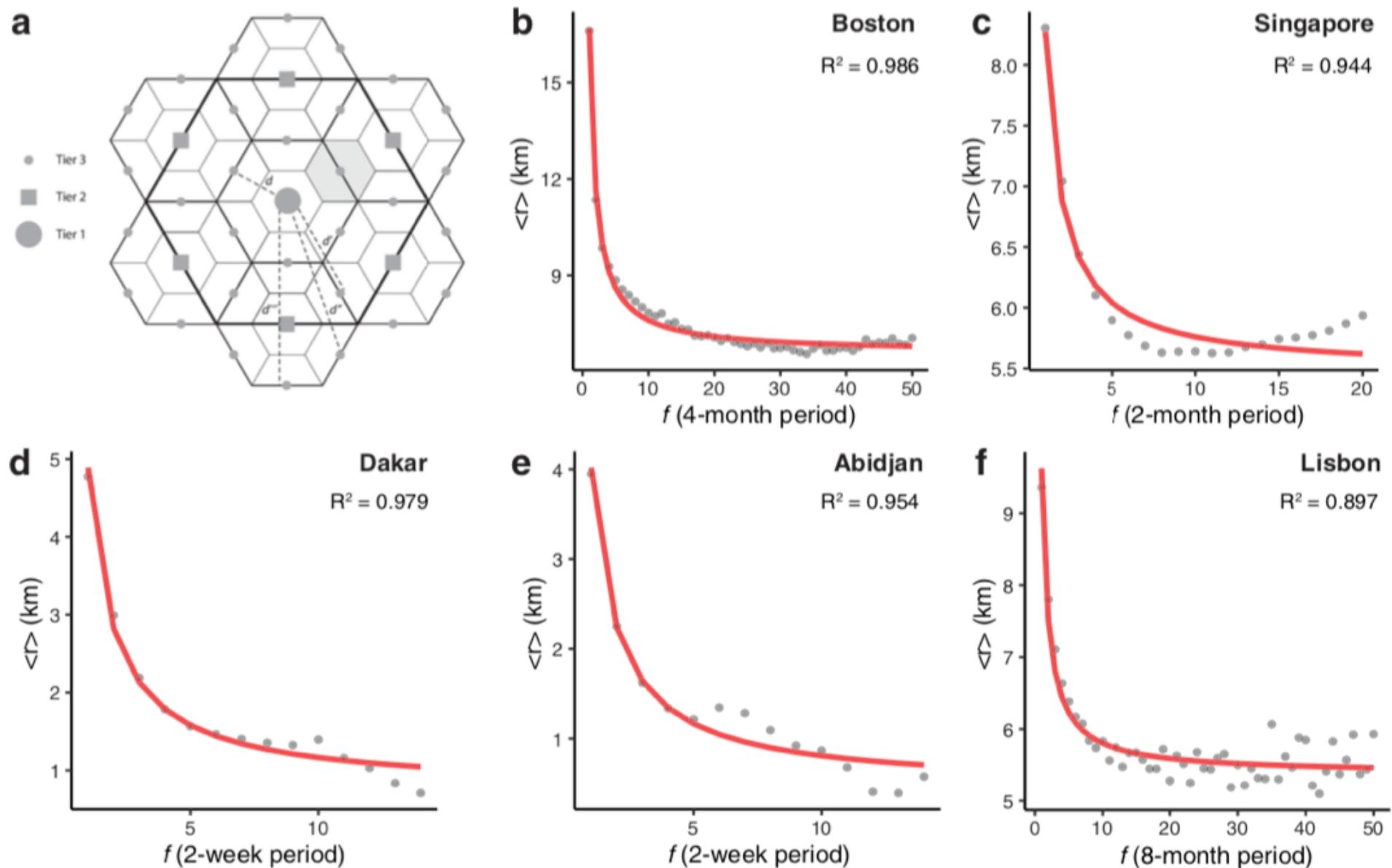
are shown in **c**, and the resulting  $r^f$ 's scaling of the spectral flows,  $\rho(r, f)$ , is shown in **d**. The generated attractiveness values  $\mu_i$  are rather homogeneous and uniform across space, which is in contrast to the empirical data (**b**). Model parameters are taken from Song et al.<sup>32</sup> (Methods).

## Application

**Fig. 5 Predicting the flows between individual locations.** **a**, Predictions for the observed trips to Back Bay West, Boston, derived from the gravity law and the radiation model compared with predictions based on the *rf*-scaling framework. Symbols are mean values for each bin and lines are the 0.1–0.9 quantiles, clearly showing that the *rf*-scaling framework systematically outperforms the existing models. The performance of each model is further quantified based on the SSI. **b**, Number of unique visitors. The fitting parameters of the gravity law from the number of trips (**a**) do not allow the prediction of the number of individuals. The radiation model does not provide a prediction of the number of visitors either, because it assigns only one destination location to each individual. It is therefore unable to explicitly consider the fact that an individual may visit several different locations. **c**, Number of high-frequency visitors. **d**, Systematic comparison over all considered locations in the studied world regions for number of trips (*t*), number of visitors (*v*), number of high-frequency visitors (*h*) and number of low-frequency visitors (*l*). The gravity model (calibrated for *t*) and the radiation model are unable to predict *v*, *h* or *l*. The *rf* scaling overcomes this limitation.



## Additional Results



**Extended Data Fig. 9 | CPT and radius of attraction.** **a**, Schematics of CPT, showing the spatial arrangement of three tiers of centres (see Supplementary Information for details). This hierarchical arrangement of central places results in the most efficient transport network. **b–f**, Average travel distance per visit ( $\langle r \rangle_f$ ) to perform activities with fixed visiting frequency  $f$  across all locations in

Greater Boston (**b**), Singapore (**c**), Dakar (**d**), Abidjan (**e**) and Lisbon (**f**). We find a clear inverse relation,  $\langle r \rangle_f \propto 1/f$ . The quantity  $\langle r \rangle_f$  can be interpreted as the characteristic distance associated with the level of specialization of the functions provided by the locations.

## Additional Results

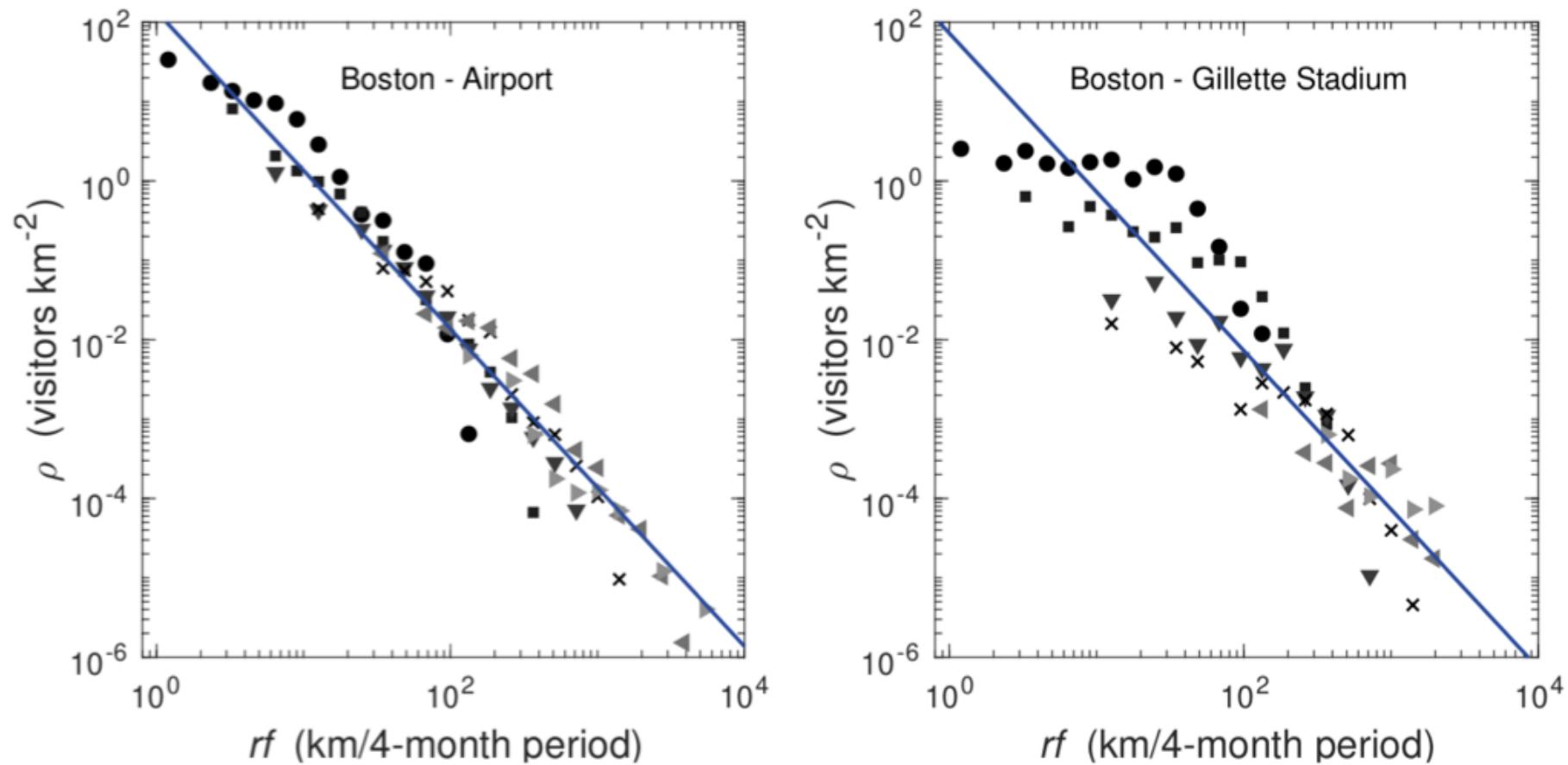


Figure S19: **Exemplary locations with ‘anomalous’ behaviour.** Symbols denote different visitation frequencies and are as in Fig. S4. Straight lines denote the inverse square law (slope = -2).

## Additional Results

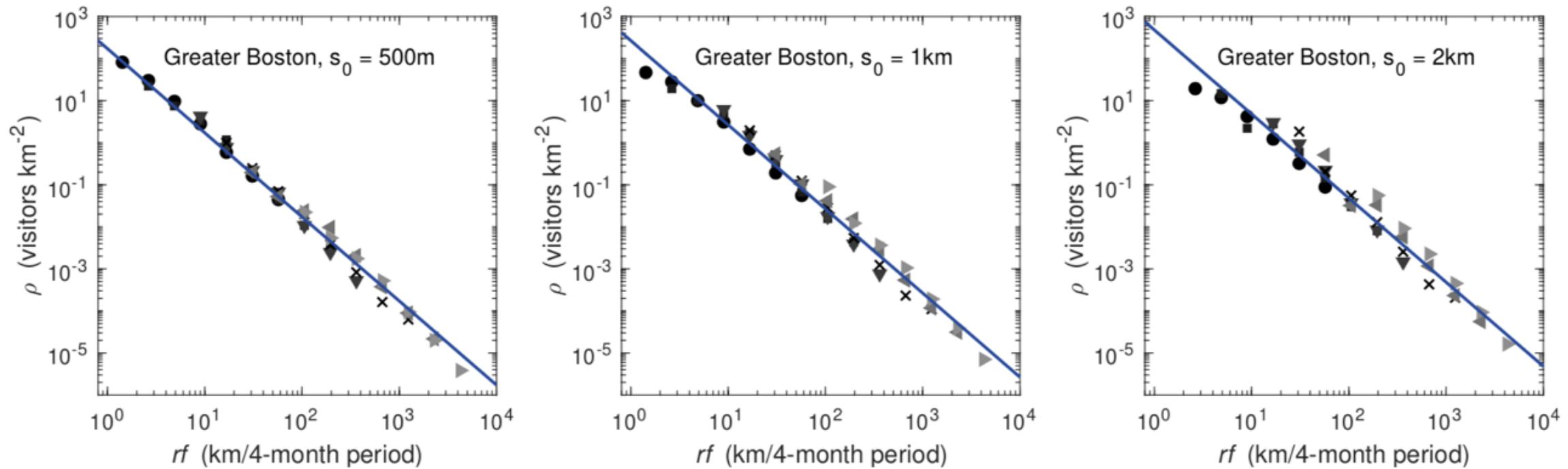


Figure S23: **Robustness of the scaling relation  $\rho \propto (rf)^{-2}$  against variations in the size of the grid cells,  $500\text{ m} \leq s_0 \leq 2\text{ km}$ .** Values are averages over the 250 most visited locations. Symbols denote different frequency bins and are as in Fig. S4. Straight lines are the inverse square law (slope = -2).

## Additional Results

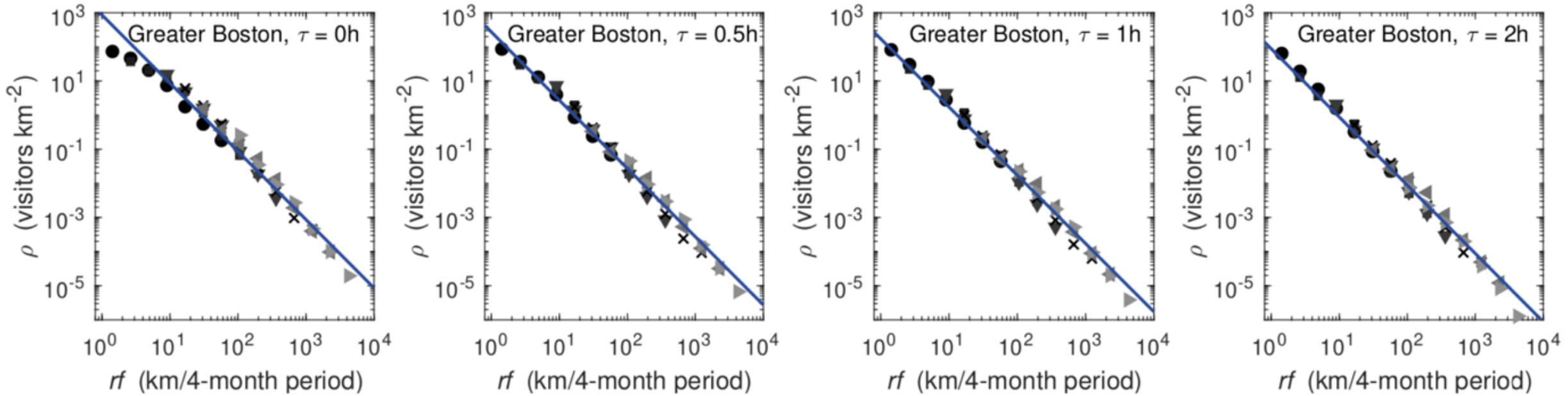


Figure S26: **Robustness of the scaling relation  $\rho \propto (rf)^{-2}$  against variations in the minimum stay time,  $\tau$ .** Symbols denote different frequency bins and are as in Fig. S4. Values are averages over the 250 most visited locations. Straight lines are the inverse square law (slope = -2).