Gravity versus radiation models: On the importance of scale and heterogeneity in commuting flows

A. Paolo Masucci,[1] Joan Serras,[1] Anders Johansson,[1,2] and Michael Batty[1]
[1]*Centre for Advanced Spatial Analysis, University College London, W1T 4TJ, London, United Kingdom*
[2]*Systems Centre, Department of Civil Engineering, University of Bristol, BS8 1UB, United Kingdom*
(Received 26 October 2012; revised manuscript received 18 June 2013; published 22 August 2013)

We test the recently introduced *radiation model* against the *gravity model* for the system composed of England and Wales, both for commuting patterns and for public transportation flows. The analysis is performed both at macroscopic scales, i.e., at the national scale, and at microscopic scales, i.e., at the city level. It is shown that the thermodynamic limit assumption for the original radiation model significantly underestimates the commuting flows for large cities. We then generalize the radiation model, introducing the correct normalization factor for finite systems. We show that even if the gravity model has a better overall performance the parameter-free radiation model gives competitive results, especially for large scales.

DOI: 10.1103/PhysRevE.88.022812 PACS number(s): 89.65.−s, 89.75.−k, 87.23.Ge

I. INTRODUCTION

In the last year, progress has been reported on theories providing a framework for human commuting patterns [1,2]. Both papers suggest that the main ingredient in a "universal" law predicting human mobility patterns is topological, i.e., it does not directly depend on metrical distance. This discovery aims to rewrite the assumptions that have been made during the last century on mobility patterns and, in particular, the traditional *gravity model* first suggested for use in human interaction systems by Carey (1859) [3] and popularized by Zipf in 1946 [4] and the *intervening opportunities model* introduced by Stouffer in 1940 [5]. It is worth noting that lately, purely topological relations have also been found to be leading components for the explanation of animal collective behavior [6].

In particular, in Ref. [1] a simple theory called the *radiation model*, based on diffusion dynamics, has been developed and the model appears to match experimental data well. The model gives exact analytical results and it has the additional desirable feature of being parameter-free, i.e., it has the characteristics of a universal theory.

In this contribution we use three different datasets in order to assess the universality, accuracy, and robustness of the proposed radiation model applied to human mobility and public transport infrastructure. The datasets we use are available as (i) a complete multimodal network for transportation in the United Kingdom, comprising the road network for bus and coach, the rail networks for tube and rail, and the airline networks for plane. The weights on these networks consist of the volumes of the transport (vehicles, trains, planes) from transport timetables; (ii) commuting patterns for England and Wales at ward level resolution from the 2001 Population Census; and (iii) population density data for the UK at ward level resolution, also from the Census.

Our first concern about the radiation model is the presumption of universality. In our interpretation, "universality" means that the model can be applied at all spatial scales, all time periods, and to different places. Regarding the system scale, we show that among cities, the radiation model is broadly accurate for commuting, while it is not accurate at all in forecasting both the transportation patterns between cities, or for the commuting flows within London. Regarding the applicability of the model to different countries, we notice that the radiation model is normalized to an infinite population system. We derive the correct normalization for finite systems and we show that it deviates from the one derived in Ref. [1] at the thermodynamic limit. This deviation is not really appreciable for large population systems at the scale of counties in the US, but it becomes relevant for smaller systems composed of much smaller but equivalent entities such as wards in the regions including England and Wales.

A. The gravity model

The gravity model is based on empirical evidence that the commuting between two places i and j, with origin population m_i and destination population n_j, is proportional to the product of these populations and inversely proportional to a power law of the distance between them. Many studies have been carried on such a model, where it is often subject to additional constraints on the generation and attractions of flows, and on the total travel distance (or cost) observed. These variants can be derived consistently using information minimizing or entropy maximizing procedures [7].

In our research we employ two models. One is a four-parameter one, which is the one also used in Ref. [1] and was first stated in this form by Alonso in 1976 [8]:

$$G_{ij} = A\frac{m_i^{\alpha} n_j^{\beta}}{r^{\gamma}}, \tag{1}$$

where A is a normalization factor and α, β, and γ are the parameters of the model, which can be determined by multiple regression analysis.

The second is a simpler perhaps more elegant model, that just carries the parameter as the exponent of the denominator and it is the one that is more frequently used in transportation modeling:

$$G_{ij} = A\frac{m_i n_j}{r^{\gamma}}. \tag{2}$$

B. The radiation model

The radiation model tracks its origin from a simple particle diffusion model, where particles are emitted at a given location and have a certain probability p of being absorbed by surrounding locations. It comes out that the probability for a particle to be absorbed is independent of p, but it depends only on the origin population m_i, the destination population n_j, and on the population in a circle whose center is the origin and radius the distance between the origin and the destination, minus the population at the origin and the population at the destination, s_{ij}. Then the number of commuters, which we call T_{ij}^{∞}, from location i to location j is estimated to be a fraction of the commuters from population i, T_i; that is,

$$T_{ij}^{\infty} = T_i \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}. \quad (3)$$

The most interesting aspect of Eq. (3) is that it is independent of the distance and that it is parameter-free. Nevertheless Eq. (3) has been derived in the thermodynamic limit, which is for an infinite system. It is easy to show that for a finite system the normalization brings us to a slightly different form of the radiation model, that is,

$$T_{ij} = \frac{T_{ij}^{\infty}}{1 - \frac{m_i}{M}} = \frac{T_i}{1 - \frac{m_i}{M}} \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})}, \quad (4)$$

where $M = \sum_i m_i$ is the total sample population and we have $T_{ij} \to T_{ij}^{\infty}$ for $M \to \infty$.

In a finite system T_{ij}^{∞} underestimates the commuting flows by a factor $\frac{1}{1-\frac{m_i}{M}}$. For a very large system with uniform population Eq. (3) is a very good approximation, but actually the city size distribution is not uniform for it usually follows a very heterogeneous skewed distribution, such as Zipf's law [9].

To understand the deviations of Eq. (3) from Eq. (4), we measure the factor $F = \frac{1}{1-\frac{m_i}{M}} - 1$ for the dataset used in Ref. [1] and for a smaller system: the region composed of England and Wales. In the former case the US system is very large. The analysis is performed at the county level and that reduces the population heterogeneity of the system. We find that the largest deviation is in the flows from Anderson county and this is of around $F \approx 6\%$. This is not a particularly large deviation, but the same measure for England and Wales, for example, brings a deviation for the commuting flows from London $F \approx 17\%$, which is a considerably larger deviation.

As we have shown that Eq. (3) is not universal, but scale dependent, a better choice for our investigation of UK commuting patterns is Eq. (4).

In Ref. [1] T_i is considered to be proportional to m_i, which is a good estimate, while in our analysis we derive its value directly from the commuting network w_{ij}, i.e., $T_i = \sum_j w_{ij}$, which is in network theory terminology the out strength of location i [10].

Moreover in Ref. [1] the model is based on job opportunities that are considered to be proportional to population. In fact, Eq. (4) can be rewritten in network theory terminology. Hence given that w_{ij} are the elements of the weighted directional adjacency matrix representing the commuting between locations i and j, we define the out strength of vertex i as $s_i^{\text{out}} = \sum_i w_{ij}$, and the in strength as $s_i^{\text{in}} = \sum_j w_{ij}$. Then we have

$$w_{ij} = \frac{s_i^{\text{out}}}{1 - \frac{s_i^{\text{in}}}{\sum_{ij} w_{ij}}} \frac{s_i^{\text{in}} s_j^{\text{in}}}{\left(s_i^{\text{in}} + S_{ij}\right)\left(s_i^{\text{in}} + s_j^{\text{in}} + S_{ij}\right)}. \quad (5)$$

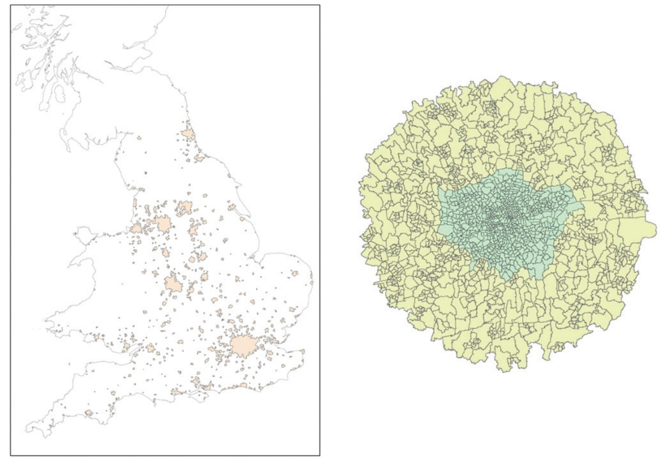

FIG. 1. (Color online) The geographical areas considered in the present analysis. Left panel: cities of England and Wales. Right panel: wards of the GLA (Greater London Authority) and the surrounding Outer Metropolitan Region.

Here $S_{ij} = \sum_{k \in K_{ij}} s_k^{\text{in}}$ and $K_{ij} = \{\forall k : d(i,k) > 0, d(i,k) < d(i,j)\}$ where $d(i,j)$ is the distance between i and j. Equation (5) is an interesting relation between the commuting flows of the network which can be verified in itself. In fact the in strength of a given vertex represents the job opportunities in that location, since it quantifies exactly the number of people going to work in that location.

II. DATA ANALYSIS

In this section we test the models defined in Eqs. (1), (2), and (4) against empirical data. In the Sec II A we analyze the commuting between the cities of England and Wales (see left panel of Fig. 1), thereby simulating the models at macroscales, while in Sec. II B we analyze the commuting between the wards of London (see right panel of Fig. 1), simulating the models at microscales.

A. Macroscopic analysis: England and Wales

In this section, we test the gravity model defined in Eqs. (1) and (2) and the radiation model defined in Eq. (4) against the empirical data for the cities of England and Wales. In this study, city clusters have been defined via a two step process using the population for the 8850 Census Area Statistics (CAS) of wards in England and Wales from the 2001 Population Census [11]. In the first step, those wards with population density above 14 persons per hectare are selected from the rest; in the second step, adjacent selected wards have been grouped to form a total of 535 city clusters [12]. We show these cities in the left panel of Fig. 1.

On the top of this sociogeographical dataset, we analyze data for the commuting between these cities using the 2001 Census Journey to Work data which specifies, for all surveyed commuters, the origin ward they travel from,their home location, and their destination ward—their work location. From this data, we have calculated the number of commuters between all pairs of cities in England and Wales.

For this study, we have also used data for the number of trains and buses moving between these cities. This information has been derived from timetable data held by the National Public Transport Data Repository. This data includes all public transport services running in England, Wales, and Scotland between the 5th and the 11th of October 2009. The data is composed by two data sets: the NaPTAN (National Public Transport Access Nodes) dataset and the TransXChange files. The former includes all public transport nodes categorized by travel mode and geolocated in space. The latter one has a series of transport modal files for each county within England, Wales, and Scotland (143 counties in total), with information on all services running within the county. The travel modes included are air, train, bus, coach, metro, and ferry. Each service includes routing information as a series of NaPTAN referenced stops each with its corresponding departure and waiting time. In this paper, we have deduced the number of trains and buses operating between all pairs of cities on a typical working day (24 h) by first assigning a ward area to each bus and train stop via spatial point-in-polygon queries, and then extending this assignment to city areas.

It is worth noting that in Ref. [1] the analysis has been made over US counties that are artificial units, while in this analysis we consider cities as natural entities for commuting. The different choice is not merely speculative, since counties have different physics and statistical properties than cities. It is well known that the city size distribution follows Zipf's law [9]. That means that city size distribution has a fat tail characterized by the scale of very large cities. The representation of the system in terms of counties introduces an artificial cutoff in the tail of the distribution, cutting down the tail, as we show in Fig. 2. It is sufficient to think of the fact that New York City is made up of five different counties (boroughs), so that in a county level analysis its population is split between those five counties.

In the top panels of Fig. 3, we show the analysis for the flows of commuters in England and Wales. In these top panels, we show the comparison between the real data (x axis) and the data elaborated by the model (y axis). On top of that we show the average value (circles), with standard deviation bars and the line $y = x$, that shows where the model meets the real data. The gravity model parameters are estimated via multiple regression analysis. Further details are given in the next section, but for now we can say that, based on R^2 estimates, the gravity model

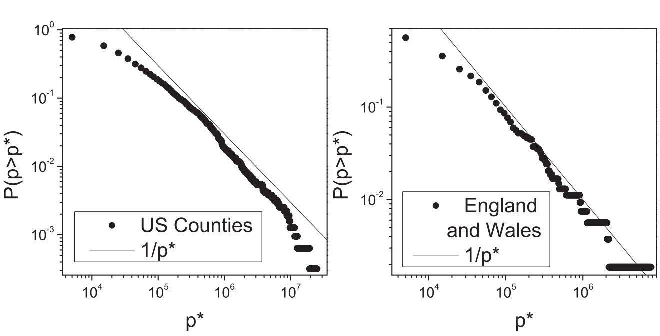

FIG. 2. In the left panel, the cumulative frequency distribution for the population size $P(p > p*)$ for the US counties analyzed in Ref. [1]. In the right panel, the cumulative frequency distribution for the population size $P(p > p*)$ for the cities of England and Wales.

of Eq. (1) performs better, followed by the radiation model, while the gravity model of Eq. (2) has $R^2 = 0$ (see Table I for R^2 values).

In the following panels of Fig. 3, we can see the correlations of the commuting flows with three sensitive quantities: the distance r in the left panel, the destination population n_j in the central panel, and the population in the circle centered on the origin population, with radius r, s_{ij}. Commuter flows are strongly correlated with all these quantities. All the plots are in log-log scale, so that apparently the correlations are in the form of power laws.

Both the radiation and the gravity model perform well in reproducing these correlations, while the gravity model II fails to reproduce the correlations with s_{ij}. Also we can see that on average both the radiation and the gravity model catch the real data, so that the main difference between the two resides in the fluctuations that will be discussed in the next section. Nevertheless, comparing the model results, we have to keep in mind that the gravity model of Eq. (1) has three independent parameters plus a normalization factor. In light of this, we have to consider the parameter-free radiation model as performing quite competently, even if the R^2 is not that high.

In the middle panels of Fig. 3, we show the analysis for the flows of trains in England and Wales. In this case the model that performs better is the gravity model but its performance is not satisfactory anyway (see Table I).

In the bottom panels of Fig. 3, we show the analysis for the flows of buses in England and Wales. In this case, interestingly enough the radiation model outperforms the gravity model, even if the overall result is of a generally quite poor performance, while the gravity model II again has $R^2 = 0$. From the correlation analysis we can see that both the models can grab the average behavior of the bus transportation system, while they fail to reproduce the distance and the s_{ij} correlations for very large scales. Nevertheless the poor results are again given by the difficulty in reproducing the large fluctuations.

1. Fluctuation analysis

In the previous section we saw that the analysed commuting models produce reasonable R^2 values just for the commuting case in UK and Wales for the gravity model and the radiation model. Qualitatively, we see from Fig. 3 that on average the models catch the real data behavior and the sensitive parameter correlations in a striking manner. Nevertheless the R^2 values are not that good, especially in the case of the radiation model. To understand this, we perform a fluctuations analysis, based on the Sørensen-Dice coefficient [13,14]:

$$E^{\text{Sørensen}} \equiv \frac{2\sum_{i,j}\min\left(T_{ij}^{\text{model}}, T_{ij}^{\text{empirical}}\right)}{\sum_{i,j} T_{ij}^{\text{empirical}} + \sum_{i,j} T_{ij}^{\text{model}}}. \tag{6}$$

$E^{\text{Sørensen}}$ is a similarity index that ranges from 0 to 1, where it is 0 when there is no match between empirical and modeled data, and 1 when there is a complete match.

In Fig. 4, we show the error analysis as a function of two sensitive parameters, the distance and the destination population. In the top panels of Fig. 4, we show the Sørensen-Dice coefficient $E^{\text{Sørensen}}$ in different locations in the phase space made up by distance, populations at destination, and

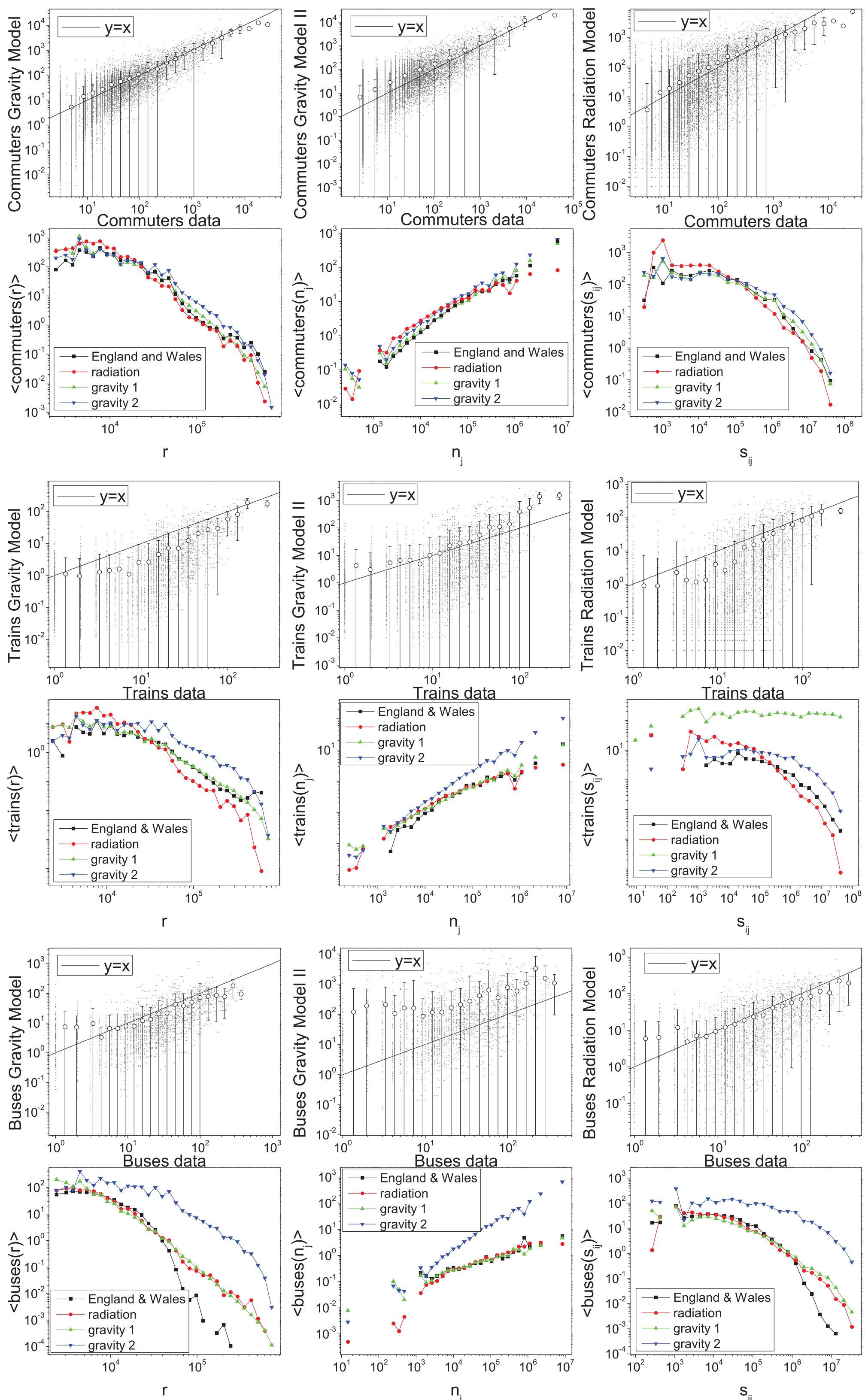

FIG. 3. (Color online) Analysis results using England and Wales city clusters. Top panel: Census 2001 commuter flows (parameters for the gravity model: $\alpha = 0.81$, $\beta = 1.03$, $\gamma = 2.42$; parameters for the gravity model II: $\gamma = 1.54$). Middle panel: trains flows (parameters for the gravity model: $\alpha = 0.76$, $\beta = 0.76$, $\gamma = 1.67$; parameters for the gravity model II: $\gamma = 1.44$). Bottom panel: bus flows (parameters for the gravity model: $\alpha = 0.61$, $\beta = 0.61$, $\gamma = 2.63$; parameters for the gravity model II: $\gamma = 1.91$).

TABLE I. R^2 calculated for the different models for England and Wales.

R^2	Gravity I	Gravity II	Radiation
Commuters	0.67	0.00	0.36
Trains	0.39	0.00	0.00
Buses	0.11	0.00	0.32

empirical flows. It is possible to see how the gravity model performs quite well for short and moderate distances, while it overestimates the flows for distances larger than 100 km. On the other hand, we see how the radiation model underestimates the flows over the entire phase space. In the bottom panel of Fig. 4, we show the comparison of the two models within the same phase space for the Sørensen-Dice coefficient $E^{\text{Sørensen}}$, where the phase space is black when the gravity model performs better than the radiation model and it is gray otherwise. From this panel it is possible to see how the gravity model performs better for short and moderate distances, where the majority of flows are concentrated, while the radiation model can better predict the commuting flows for very large distances with small and moderate population at destination.

B. Small scale analysis: London

In this section we perform an analysis on the extended Greater London Authority area at ward level. In order to do so we consider all the wards in the GLA (Greater London Authority), plus the wards in the outer metropolitan area of

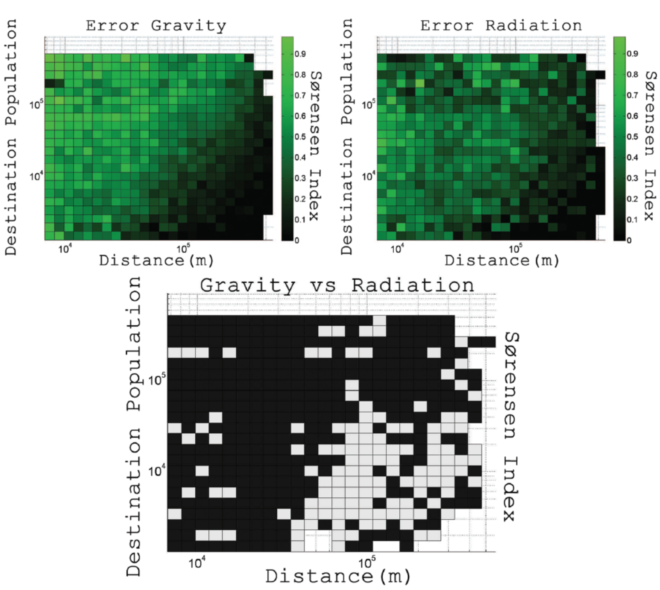

FIG. 4. (Color online) Error of flow estimates compared to empirical commuter flows for cities of England and Wales as a function of distance and population at destination. In the top panels, we show the Sørensen-Dice coefficient of Eq. (6) for the gravity (left) and the radiation models (right). The bottom panel shows areas where the gravity model performs better (black) and areas where the radiation model performs better (light gray) using the Sørensen-Dice coefficient, $E^{\text{Sørensen}}$. Wherever the two models perform equally well, there is a gap in the plot (shown as white cells without borders).

TABLE II. R^2 calculated for the different models for London.

R^2	Gravity I	Gravity II	Radiation
Commuters	0.07	0.06	0.00
Buses	0.22	0.21	0.00

the GLA for a radius of approximately 60 km that accounts for commuting inside of the GLA (see Fig. 1, left panel).

Our dataset, as already discussed in the previous section, is derived from the 2001 Census of Population with the Journey to Work data also from this census; It gives the ward population and ward to ward commuting flows. Moreover we test the models also with the vehicular transportation, considering the number of buses traveling from ward to ward which have already been calculated (see previous section).

In Table II we show the results of the R^2 test for the different models. We can observe straightaway that the models all perform rather badly, implying that the structure of a metropolis is more complex than the one forecasted by both the radiation and by gravity models.

In the top panels of Fig. 5 we show the analysis for the commuting patterns, i.e., the models against the real data. We perform a multiple regression analysis to find the best fit with the data for Eq. (1), whose results are shown in the figure caption.

In the second from top, left panel of Fig. 5, we show the average number of commuters in London as a function of the distance. The plot shows that real data decay faster than a power law with the distance, and this behavior is captured by none of the gravity models, which tend to follow a power law behavior. On the other hand, the radiation model forecasts a good amount of commuting for short distances and a rapid decay, but this does not reproduce the data well either.

In the adjacent panel we show the correlations between the commuting flows and the destination population. For London, this is counterintuitive, since the correlation analysis shows a few large peaks for wards with very small population. This phenomena resides in the fact that the wards where most of the jobs are concentrated in London are not residential wards. This evidence would let us think that the approximation ward population/ward employment is not valid for London and that we should take this bias into account in our analysis.

In the right panel of the same figure we show the correlations of the number of commuters and s_{ij}. There are hints of a strong dependency of the commuting flows from this quantity, even if this dependency is weaker than the one reproduced by the radiation model.

In the bottom panels of Fig. 5, we show the results for the analysis on the bus flows in the GLA. In Table II the R^2 values are displayed and we can see that the models do not perform very well, but still better than for the commuters case. The correlations for the number of buses with the distance display an exponential tail, which has not been picked up by any of the models. As for the commuter case, we see the strongest correlations with distance and s_{ij}, while the correlations with destination populations are ill defined.

One could argue that the poor results obtained applying the commuting models to the London intracommuting flows could

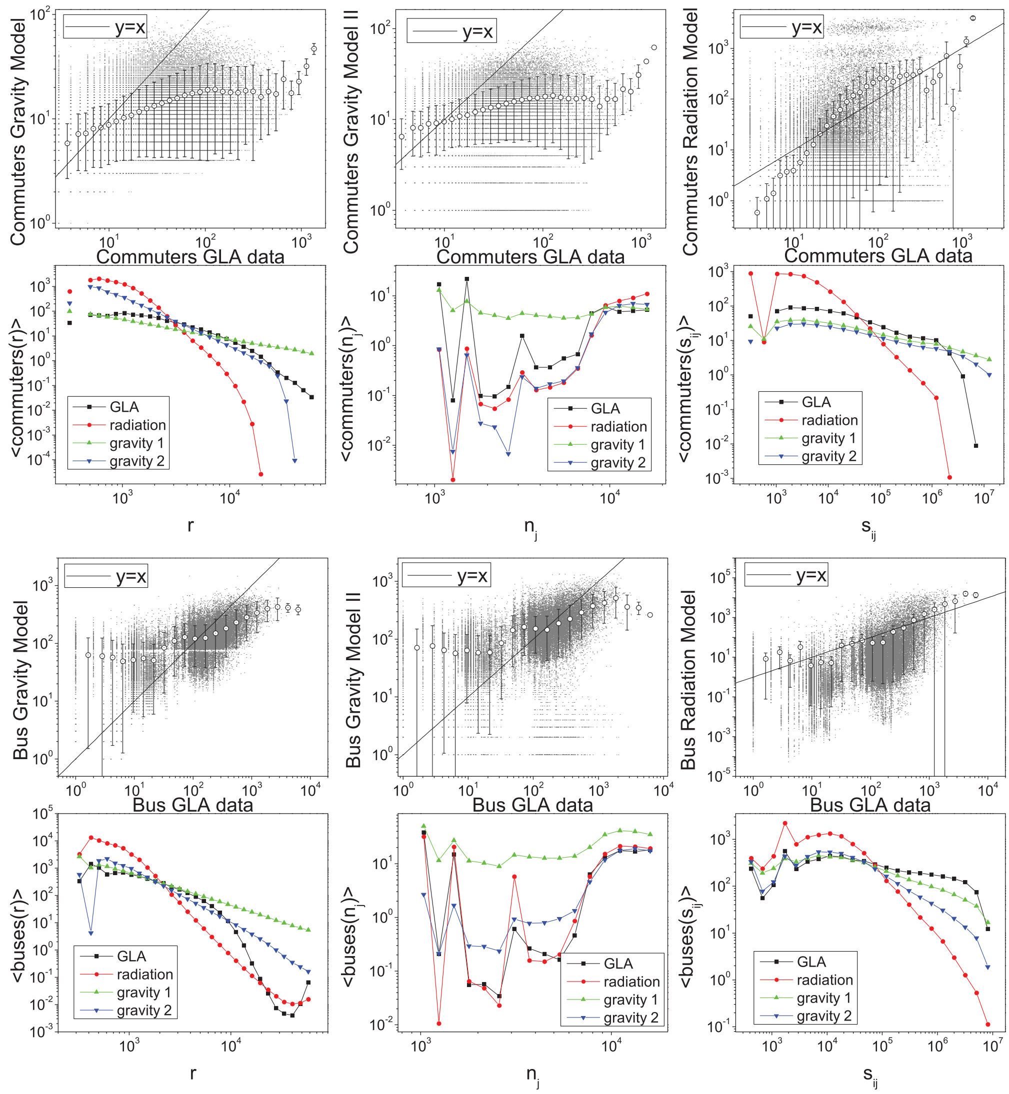

FIG. 5. (Color online) Top panels: commuter flows analysis for the GLA (parameters for the gravity model: $\alpha = 0.45$, $\beta = -0.21$, $\gamma = 0.82$; parameters for the gravity model II: $\gamma = 1.14$). Bottom panels: bus flows analysis for GLA (parameters for the gravity model: $\alpha = 0.06$, $\beta = 0.09$, $\gamma = 1.20$; parameters for the gravity model II: $\gamma = 2.39$).

reside in the approximation validity of the employment data with the population size, in the case of the London wards.

To address this question, in the top left panel of Fig. 6, we show the frequency distribution $P(p)$ of the population size p for London's wards. This is well fitted by a Gaussian distribution centered around 11 300 people. This is not a surprise since the ward boundaries have been designed to have approximatively the same population size.

In the top right panel of Fig. 6, we show instead the cumulative frequency distribution $P(s^{\text{in}} > s^*)$ for the number of people s^{in} working in a given ward. This is a skew distribution with a broad tail, well fitted by a power law $P(s^{\text{in}} > s^*) \propto s^{\gamma+1}$, with exponent $\gamma = -2.24$. This reflects the fact that the approximation population size/employment data, that has been shown to be valid in the case of counties in the US, is not valid in the case of London's wards. In particular, we see that employment s^{in} follow a distribution that suggests a complex and hierarchical organization for these resources within the city.

In fact, from the bottom left panel of Fig. 6, where we measure the average number of employees $\langle s^{\text{in}}(p)\rangle$ as a function of the ward population size, we can see that apart from some nontrivial deviations for small population size, there are no significant correlations. These deviations are related to the fact that the most significant employment locations in London often have a very small population.

We can now check whether Eq. (5) could be a more appropriate choice in order to describe commuting flows inside of a city, instead of Eq. (4). In the bottom right panel of Fig. 6, we show the results of Eq. (5) applied to the commuting between GLA wards, versus the real commuting flows, in the same style and notation of Fig. 5. We notice that the plot is very similar to the one obtained using Eq. (4) and the $R^2 = 0.00$ tells us that using Eq. (5) instead of Eq. (4) does not improve the goodness of the fit. This implies that failure of the radiation model in forecasting urban commuting flows does not reside in approximating population size/employment, but in the complexity of the system [15].

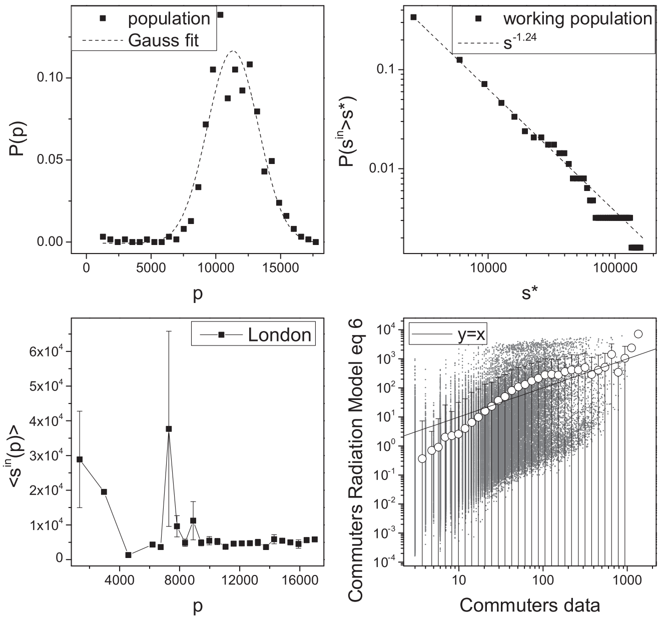

FIG. 6. Top-left panel: frequency distribution of wards population size $P(p)$ in the GLA. Top-right panel: cumulative frequency distribution for the working population $P(s^{\text{in}} > s^*)$ in a given ward. Bottom left panel: average number of employees $\langle s^{\text{in}}(p)\rangle$ as a function of the ward population size. Bottom right panel: commuters flow analysis for the GLA; in this case the flows are modeled via Eq. (5).

III. CONCLUSIONS

Human mobility is an outstanding problem in science. In more than one century of active work and observation, the gravity model has been considered the best option to model such a phenomenon [7]. The appearance of a new statistical model based on physical science [1] has reopened the debate on this topic. In particular, the apparent independence of the radiation model from metrical distance and its property of being parameter-free is a significant and desirable change from past practice. The model needs to be tested in many different circumstances so that its wider applicability can be assessed.

In this paper we address the reliability of the radiation model against the gravity model for large scale commuting and transportation networks in England and Wales and for the intraurban commuting and transportation network for the London region.

The first thing we notice is that both models fail to describe human mobility within London. In this sense we argue that commuting at the city scale still lacks a valid model and that further research is required to understand the mechanism behind urban mobility. In fact, the phenomena of sociogeographical segregation [16] and residential/business ward specialization [1] are key drivers in determining the structure of flows and the density of population in the city and these are not reflected by these statistical models [15].

For England and Wales, we first introduce the correct normalization for finite systems in the radiation model. Such a normalization affects the flows from London by a factor of 17%. Then we notice that the models are not very good in describing transportation data, such as bus and train flows, while they can be considered acceptable for modeling the commuting flows. The gravity model II of Eq. (2) fails to describe commuting models, and confirms that commuting correlations with population at origin and destination is not just linear. The gravity model is satisfactory in describing the commuting flows and surely much better than the radiation model, even if the latter has the advantage of being parameter-free, which turns out to be useful in cases where there is no data available to estimate any parameters.

Nevertheless from the fluctuation analysis it emerges that there is a consistent portion of the distance/destination population phase space where the radiation model gives better estimates of the gravity model in terms of the Sørensen-Dice coefficient. This means that for large distances and small and moderate destination population scales, the principles of the radiation model are reliable and that mobility patterns can be approached by a diffusion model where intervening opportunities on the commuting paths prevail on the distance of such paths. However, the modest overall radiation model performance in terms of R^2 indicates that more research on the subject has to be done in order to improve the model reliability.

Other ways to represent the commuting system are possible. For example, if we were to grid all the data thereby strictly defining population and employment as density measures, this would change the dynamics of the gravity and radiation models in that they have been originally specified to deal with counts of activity data such as population and employments rather than their densities. Moreover the tradition in this field is to work with data that is available in administrative units rather than approximate that data on a grid because these units reflect changes in the spatial system over time. We believe that the best way to conduct this study is to consider urban conglomerations as the natural entities involved in commuting flows. This choice relates to a well settled tradition in statistical physics that consider cities as well defined entities, such as in Zipf's and Gibrat's law.

ACKNOWLEDGMENTS

J.S. and A.P.M. were partially funded by the EPSRC SCALE project (EP/G057737/1) and M.B. by the ERC MECHANICITY project (249393 ERC-2009-AdG). Further, we would like to thank the anonymous reviewers for their constructive feedback that has improved the paper, especially the suggestion to use the Sørensen-Dice coefficient as an alternative error metric.

[1] F. Simini, M. C. Gonzlez, A. Maritan, and A. L. Barabsi, Nature (London) **484**, 96 (2012).

[2] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo, PLoS ONE **7**, e37027 (2011).

[3] H. C. Carey, *Principles of Social Science* (J. B. Lippincott & Co., Philadelphia, 1959).

[4] G. K. Zipf, Am. Sociol. Rev. **11**, 677 (1946).

[5] S. A. Stouffer, Am. Sociol. Rev. **5**, 845 (1940).
[6] M. Ballerini, N. Cabibbo, R. Candelier, A. Cavagna, E. Cisbani, I. Giardina, V. Lecomte, A. Orlandi, G. Parisi, A. Procaccini, M. Viale, and V. Zdravkovic, Proc. Natl. Acad. Sci. USA **105**, 1232 (2008).
[7] A. G. Wilson, *Entropy in Urban and Regional Modeling* (Pion Limited, London, 1970).
[8] W. Alonso, Working Paper No. 266, Institute of Urban and Regional Development, University of California, Berkeley, CA, 1976 (unpublished).
[9] G. K. Zipf, *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology* (Addison-Wesley, Reading, MA, 1949).
[10] S. N. Dorogovtsev and J. F. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW* (Oxford University Press, Oxford, 2003).
[11] See http://www.neighbourhood.statistics.gov.uk/ (last visited 25-06-12).
[12] E. Arcaute, E. Hatna, P. Ferguson, H. Youn, A. Johansson, and M. Batty, arXiv:1301.1674.
[13] T. A. Sørensen, Biol. Skr. - K. Dan. Vidensk. Selsk. **5**, 1 (1948).
[14] L. R. Dice, Ecology **26**, 297 (1945).
[15] P. Wang, T. Hunter, A. M. Bayen, K. Schechtner, and M. C. González, Sci. Rep. **2**, 1001 (2012).
[16] H. Theil and A. J. Finizza, J. Math. Sociol. **1**, 187 (1971).