Systematic comparison of trip distribution laws and models

Maxime Lenormand *, Aleix Bassolas, José J. Ramasco

Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus UIB, 07122 Palma de Mallorca, Spain

ARTICLE INFO

ABSTRACT

Trip distribution laws are basic for the travel demand characterization needed in transport and urban planning. Several approaches have been considered in the last years. One of them is the so-called gravity law, in which the number of trips is assumed to be related to the population at origin and destination and to decrease with the distance. The mathematical expression of this law resembles Newton's law of gravity, which explains its name. Another popular approach is inspired by the theory of intervening opportunities which argues that the distance has no effect on the destination choice, playing only the role of a surrogate for the number of intervening opportunities between them. In this paper, we perform a thorough comparison between these two approaches in their ability at estimating commuting flows by testing them against empirical trip data at different scales and coming from different countries. Different versions of the gravity and the intervening opportunities laws, including the recently proposed radiation law, are used to estimate the probability that an individual has to commute from one unit to another, called trip distribution law. Based on these probability distribution laws, the commuting networks are simulated with different trip distribution models. We show that the gravity law performs better than the intervening opportunities laws to estimate the commuting flows, to preserve the structure of the network and to fit the commuting distance distribution although it fails at predicting commuting flows at large distances. Finally, we show that the different approaches can be used in the absence of detailed data for calibration since their only parameter depends only on the scale of the geographic unit.

© 2016 Published by Elsevier Ltd.

1. Introduction

Everyday, billions of individuals around the world travel. These movements form a socio-economic complex network, backbone for the transport of people, goods, money, information or even diseases at different spatial scales. The study of such spatial networks is consequently the subject of an intensive scientific activity (Barthelemy, 2011). Some examples include the estimation of population flows (Murat, 2010; Gargiulo et al., 2012; Simini et al., 2012; Lenormand et al., 2012; Thomas and Tutert, 2013; Lenormand et al., 2014; Yang et al., 2014; Sagarra et al., 2015), transport planning and modeling (Rouwendal and Nijkamp, 2004; Ortúzar and Willumsen, 2011), spatial network analysis (De Montis et al., 2007, 2010), study of urban traffic (De Montis et al., 2007) and modeling of the spreading of infectious diseases (Viboud et al., 2006; Balcan et al., 2009; Tizzoni et al., 2014).

Trip distribution modeling is thus crucial for the prediction of population movements, but also for an explanatory purpose, in order to better understand the mechanisms of human mobility. There are two major approaches for the estimation of trip distribution at an aggregate level. The traditional gravity approach, in analogy with the Newton's law of gravitation, is based on the assumption that the amount of trips between two locations is related to their populations and decays with a function of the distance (Carey, 1858; Zipf, 1946; Wilson, 1970; Erlander and Stewart, 1990). In contrast to the gravity law, the Stouffer's law of intervening opportunities (Stouffer, 1940) hinges on the assumption that the number of opportunities plays a more important role in the location choices than the distance, particularly in the case of migration choices. The original law proposed by Stouffer has been reformulated by Schneider (1959) and extensively studied since then (Heanus and Pyers, 1966; Ruiter, 1967; Wilson, 1970; Haynes et al., 1973; Fik and Mulligan, 1990; Akwawua and Poller, 2001). The two approaches have been widely compared during the second half of the twentieth century (David, 1961; Pyers, 1966; Lawson and Dearinger, 1967; Zhao et al., 2001) showing that generally both approaches performed comparably. However, the simplicity of the mathematical form of the gravity approach appears to have weighted in its favor (Ortúzar and Willumsen, 2011). Indeed, the gravity approach has been extensively used in the past few decades to model, for instance, flows of population (Viboud et al., 2006; Griffith, 2009; Balcan et al., 2009; Murat, 2010; Gargiulo et al., 2012; Lenormand et al., 2012; Thomas and Tutert, 2013; Masucci et al., 2013; Liang et al., 2013; Lenormand et al., 2014; Tizzoni et al., 2014; Liu et al., 2014), spatial accessibility to health services (Luo and Wang, 2003), volume of international trade (Anderson, 1979; Bergstrand, 1985), traffic in transport networks (Jung et al., 2008; Kaluza et al., 2010) and phone communications (Krings et al., 2009).

---

* Corresponding author.
  E-mail address: maxime@ifisc.uib-csic.es (M. Lenormand).

However, the concept of intervening opportunities has recently regained in popularity thanks to the recently proposed radiation approach (Simini et al., 2012, 2013; Ren et al., 2014; Yang et al., 2014). This approach is inspired by a simple diffusion model where the amount of trips between two locations depends on their populations and the number of opportunities between them. The gravity law and the radiation law have been compared several times during the last years giving the superiority to either of the approaches depending on the study (Simini et al., 2012; Lenormand et al., 2012; Masucci et al., 2013; Liang et al., 2013; Yang et al., 2014). Two main issues can be identified in these comparisons. First, the inputs used to simulate the flows are not always identical. For example, in the comparison proposed in Masucci et al. (2013), the gravity law tested takes as input the population, whereas the radiation law is based on the number of jobs. Second, in all these studies, the models used to generate the trips from the radiation and the gravity laws are not constrained in the same way. The radiation models are always production constrained, this means that the number of trips, or at least an estimation of the number of trips generated by census unit, is preserved. The models used to generate the trips with the gravity laws can be either, unconstrained (Simini et al., 2012; Masucci et al., 2013), only the total number of trips is preserved or doubly constrained (Lenormand et al., 2012; Yang et al., 2014), both the trips produced and attracted by a census unit are preserved. Therefore, to fairly compare different approaches the same input data must be used and, most importantly, we need to differentiate the law, gravity or intervening opportunities, and the modeling framework used to generate the trips from this law. Indeed, both the gravity laws and the intervening opportunities laws can be expressed as a probability to move from one place to another, called trip distribution law, and based on these probability distributions, the total number of trips can then be simulated using different trip distribution models including different level of constraints.

In this work, we test and compare, in a systematic and rigorous way, gravity and intervening opportunities laws against commuting census data coming from six different countries using four different constrained models to generate the networks: unconstrained model, single constrained models (production or attraction) and the well-known doubly constrained model. For the gravity law, since the form of the distance decay functions may vary from one study to another (Fotheringham, 1981; Viboud et al., 2006; de Vries et al., 2009; Balcan et al., 2009; Barthelemy, 2011; Lenormand et al., 2014; Chen, 2015) both the power and the exponential forms are tested to model the impact of the distance. The intervening opportunities law is given by the Schneider's version of the Stouffer's original law as it is usually the case. We also considered two versions of the radiation law, the original free-parameter model (Simini et al., 2012) and the extended version proposed in Yang et al. (2014). The simulated networks are compared with the observed ones on different aspects showing that, globally, the gravity law with an exponential distance decay function outperforms the other laws in the estimation of commuting flows, the conservation of the commuting network structure and the fit of the commuting distance distribution even if it fails at predicting commuting flows at large distances. Finally, we show that the different laws can be used in absence of detailed data for calibration since their only parameter depends only on the scale of the geographic census unit.

2. Data

In this study, the trip distribution laws and models are tested against census commuting data of six countries: England and Wales, France, Italy, Mexico, Spain and the United States of America (hereafter called E&W, FRA, ITA, MEX, SPA and USA, respectively) and two cities: London and Paris (hereafter called LON and PAR, respectively).

- The England & Wales dataset comes from the 2001 Census in England and Wales made available by the Office for National Statistics (data

Table 1
Presentation of the datasets.

Case study	Number of units	Number of links	Number of Commuters
England & Wales	8846 wards	1,269,396	18,374,407
France	3645 cantons	462,838	12,193,058
Italy	7319 municipalities	419,556	8,973,671
Mexico	2456 municipalities	60,049	603,688
Spain	7950 municipalities	261,084	5,102,359
United State	3108 counties	161,522	34,097,929
London	4664 output areas	750,943	4,373,442
Paris	3185 municipalities	277,252	3,789,487

available online at https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&dataset=124).
- The French dataset was measured for the 1999 French Census by the French Statistical Institute (data available upon request at http://www.cmh.ens.fr/greco/adisp_eng.php).
- The Italian's commuting network was extracted from the 2001 Italian Census by the National Institute for Statistics (data available upon request at http://www.istat.it/it/archivio/139381).
- Data on commuting trips between Mexican's municipalities in 2011 are based on a microdata sample coming from the Mexican National Institute for Statistics (data available online at http://www3.inegi.org.mx/sistemas/microdatos/default2010.aspx).
- The Spanish dataset comes from the 2001 Spanish Census made available by the Spanish National Statistics Institute (data available upon request at http://www.ine.es/en/censo2001/index_en.html).
- Data on commuting trips between United States counties in 2000 comes from the United State Census Bureau (data available online at https://www.census.gov/population/www/cen2000/commuting/index.html).

Each case study is divided into n census units of different spatial scale: from the Output Area in London with an average surface of 1.68 km^2 to the counties in the United States with an average surface of 2596.8 km^2. See Table 1 for a detailed description of the datasets.

Figs. 1 and 2 display the centroids of the census units for the eight case studies. For each unit, the statistical offices provide the following information:

- T_{ij}, the number of trips between the census units i and j (i.e. number of individuals living in i and working in j);
- d_{ij}, the great-circle distance between the unit i and the unit j computed with the Haversine formula;
- m_i, the number of inhabitants in unit i.

In this work we consider only inter-unit flows (i.e. $T_{ii} = 0$), mainly because it is not possible to estimate intra-units flows with the radiation laws.[1] We note $N = \sum_{i,j=1}^{n} T_{ij}$ the total number of commuters, $O_i = \sum_{j=1}^{n} T_{ij}$ the number of out-commuters (i.e. number of individuals living in i and working in another census unit) and $D_j = \sum_{i=1}^{n} T_{ij}$ the number of in-commuters (i.e. number of individuals working in j and living in another census unit).

3. Comparison of trip distribution laws and models

The purpose of the trip distribution models is to split the total number of trips N in order to generate a trip table $\tilde{T} = (\tilde{T}_{ij})_{1 \leq i,j \leq n}$ of the estimated number of trips form each census area to every other. Note that

---

[1] Note that it is possible to estimate intra-unit flows with the gravity laws by approximating intra-unit distances with, for example, half the square root of the unit's area or half the average distance to the nearest neighbors.
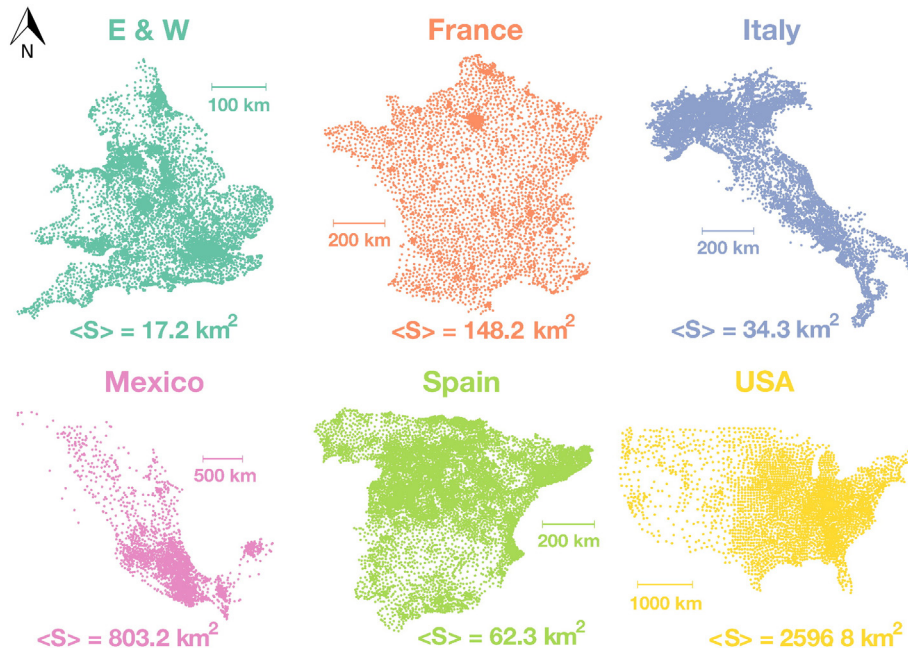
Fig. 1. Position of the units' centroids for the six countries. $\langle S \rangle$ represents the average surface of the census units (i.e. municipalities, counties or wards).

by trip we are referring to commuting travels from home to work, there is a return trip not considered in $\tilde{T}$ and N is also equivalent to the number of unique commuters. The trip distribution depends on, on one hand, the characteristics of the census units and the way they are spatially distributed, and, on the other hand, the level of constraints required by the model. Therefore, to fairly compare different trip distribution modeling approaches we have to consider separately the law used to calculate the probability to observe a trip between two census units, called trip distribution law, and the trip distribution model used to generate the trip allocation from this law.

3.1. Gravity and intervening opportunities laws

The purpose of this study is to test the capacity of both the gravity and the intervening opportunities approaches to estimate the probability p_{ij} that out of all the possible travels in the system we have one between the census unit i and j. This probability is asymmetric in i and j as the flows themselves, and, by convention, the self-loops are excluded of the analysis $p_{ii} = 0$. This probability is normalized to all possible couples of origins and destinations, $\sum_{i,j=1}^{n} p_{ij} = 1$. Note that p_{ij} does not refer

Fig. 2. Position of the units' centroids around London (left) and Paris (right). The black contours represent the boundaries of the Greater London Authority (left) and the french *département* Ile de France (right). $\langle S \rangle$ represents the average unit surface.

to the conditional probability of a trip starting in i finishes in j $\mathbb{P}(1|i,j)$. There exists a relation between both of them:

$$p_{ij} = \mathbb{P}(i)\, \mathbb{P}(1|i,j) \qquad (1)$$

where $\mathbb{P}(i)$ stands for the probability of a trip starting in i. $\mathbb{P}(1|i,j)$ will appear later for the intervening opportunities laws as a function of the populations of origin m_i, destination m_j and the number of opportunities between them s_{ij}, $\mathbb{P}(1|m_i,m_j,s_{ij})$, but the basis of our analysis will be p_{ij}.

3.1.1. Gravity laws

In the simplest form of the gravity approach, the probability of commuting between two units i and j is proportional to the product of the origin population m_i and destination population m_j, and inversely proportional to the travel cost between the two units:

$$p_{ij} \propto m_i m_j f(d_{ij}), \quad i \neq j \qquad (2)$$

The travel cost between i and j is usually modeled with an exponential distance decay function,

$$f(d_{ij}) = e^{-\beta\, d_{ij}} \qquad (3)$$

or a power distance decay function,

$$f(d_{ij}) = d_{ij}^{-\beta} \qquad (4)$$

As mentioned in Barthelemy (2011), the form of the distance decay function can change according to the dataset, therefore, both the exponential and the power forms are considered in this study. In both cases, the importance of the distance in commuting choices is adjusted with a parameter β with observed data.

3.1.2. Intervening opportunities laws

In the intervening opportunity approach, the probability of commuting between two units i and j is proportional to the origin population m_i and to the conditional probability that a commuter living in unit i with population m_i is attracted to unit j with population m_j, given that there are s_{ij} job opportunities in between. The conditional probability
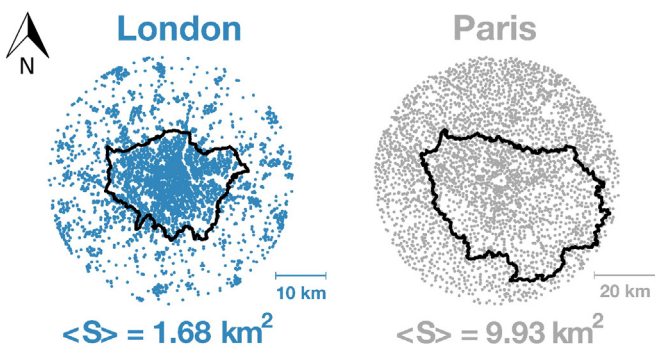
$\mathbb{P}(1|m_i, m_j, s_{ij})$ needs to be normalized to ensure that all the trips end in the region of interest.

$$p_{ij} \propto m_i \frac{\mathbb{P}(1|m_i, m_j, s_{ij})}{\sum_{k=1}^{n} \mathbb{P}(1|m_i, m_k, s_{ik})}, \quad i \neq j \tag{5}$$

In the Schneider's version of the intervening opportunities approach the conditional probability is given by

$$\mathbb{P}(1|m_i, m_j, s_{ij}) = e^{-\gamma s_{ij}} - e^{-\gamma(s_{ij} + m_j)} \tag{6}$$

where s_{ij} is the number of opportunities (approximated by the population in this case) in a circle of radius d_{ij} centered in i (excluding the source and destination). The parameter γ can be seen as a constant probability of accepting an opportunity destination. Note that in this version the number of opportunities m_i at the origin is not taken into account.

More recently, Simini et al. (2012) reformulated the Stouffer's intervening opportunities law in terms of radiation and absorption processes. This model is inspired by a diffusion model where each individual living in an unit i has a certain probability of being "absorbed" by another unit j according to the spatial distribution of opportunities. The original radiation model is free of parameters and, therefore, it does not require calibration. The conditional probability $\mathbb{P}(1|m_i, m_j, s_{ij})$ is expressed as:

$$\mathbb{P}(1|m_i, m_j, s_{ij}) = \frac{m_i m_j}{(m_i + s_{ij})(m_i + m_j + s_{ij})} \tag{7}$$

This conditional probability needs to be normalized because the probability for an individual living in a census unit i of being absorbed by another census unit is not equal to 1 in case of finite system but equal to $1 - \frac{m_i}{M}$ where M is the total population (Masucci et al., 2013). Some recent works have shown that the model fails to describe human mobility compared to more classic approaches particularly on a small scale (Lenormand et al., 2012; Masucci et al., 2013; Liang et al., 2013). To circumvent these limitations, an extended radiation model has been proposed by Yang et al. (2014). In this extended version, the probability $\mathbb{P}(1|m_i, m_j, s_{ij})$ is derived under the survival analysis framework introducing a parameter a to control the effect of the number of job opportunities between the source and the destination on the job selection,

$$\mathbb{P}(1|m_i, m_j, s_{ij}) = \frac{\left[(m_i + m_j + s_{ij})^\alpha - (m_i + s_{ij})^\alpha\right](m_i^\alpha + 1)}{\left[(m_i + s_{ij})^\alpha + 1\right]\left[(m_i + m_j + s_{ij})^\alpha + 1\right]} \tag{8}$$

3.2. Constrained models

After the description of the probabilistic laws, the next step is to materialize the people commuting. The purpose is to generate the commuting network $\tilde{T} = (\tilde{T}_{ij})_{1 \leq i,j \leq n}$ by drawing at random N trips from the trip distribution law $(p_{ij})_{1 \leq i,j \leq n}$ respecting different level of constraints according to the model. We are going to consider four different types of models:

1. *Unconstrained model.* The only constraint of this model is to ensure that the total number of trips $\tilde{N}$ generated by the model is equal to the total number of trips N observed in the data. In this model, the N trips are randomly sample from the multinomial distribution

$$\mathcal{M}\left(N, \left(p_{ij}\right)_{1 \leq i,j \leq n}\right) \tag{9}$$

2. *Production constrained model.* This model ensures that the number of trips "produced" by a census unit is preserved. For each unit i, O_i trips are produced from the multinomial distribution

$$\mathcal{M}\left(O_i, \left(\frac{P_{ij}}{\sum_{k=1}^{n} P_{ik}}\right)_{1 \leq j \leq n}\right) \tag{10}$$

3. *Attraction constrained model.* This model ensures that the number of trips "attracted" by a unit is preserved. For each census unit j, D_j trips are attracted from the multinomial distribution

$$\mathcal{M}\left(D_i, \left(\frac{P_{ij}}{\sum_{k=1}^{n} P_{kj}}\right)_{1 \leq i \leq n}\right) \tag{11}$$

4. *Doubly constrained model.* This model, also called production–attraction constrained model ensures that both the trips attracted and generated by a census unit are preserved using two balancing factors K_i and K_j calibrated with the *Iterative Proportional Fitting* procedure (Deming and Stephan, 1940). The relation between K_i, K_j, p_{ij} and the trip flows is given by

$$\begin{cases} \tilde{T}_{ij} = K_i K_j p_{ij} \\ \sum_{j=1}^{n} \tilde{T}_{ij} = O_i, \quad \sum_{i=1}^{n} \tilde{T}_{ij} = D_j \end{cases} \tag{12}$$

Unlike the unconstrained and single constrained models, the doubly constrained model is a deterministic model. Therefore, the simulated network $\tilde{T}$ is a fully connected network in which the flows are real numbers instead of integers. This can be problematic since we want to study the capacity of both the gravity and the radiation approaches to preserve the topological structure of the original network. To bypass this limitation N trips are randomly sample from the multinomial distribution,

$$\mathcal{M}\left(N, \left(\frac{\tilde{T}_{ij}}{\sum_{k,l=1}^{n} \tilde{T}_{kl}}\right)_{1 \leq i,j \leq n}\right) \tag{13}$$

3.3. Goodness-of-fit measures

3.3.1. Common part of commuters
We calibrate the parameters β, γ and a using the common part of commuters (CPC) introduced in (Gargiulo et al., 2012; Lenormand et al., 2012):

$$CPC(T, \tilde{T}) = \frac{2 \sum i, j = 1n \min(T_{ij}, \tilde{T}_{ij})}{\sum_{i,j=1}^{n} T_{ij} + \sum_{i,j=1}^{n} \tilde{T}_{ij}}. \tag{14}$$

This indicator is based on the Sørensen index (Sørensen, 1948). It varies from 0, when no agreement is found, to 1, when the two networks are identical. In our case, the total number of commuters N is preserved, therefore the Eq. (14) can be simplified to

$$CPC(T, \tilde{T}) = \frac{\sum_{i,j=1}^{n} \min(T_{ij}, \tilde{T}_{ij})}{N} = 1 - \frac{1}{2} \frac{\sum_{i,j=1}^{n} |T_{ij} - \tilde{T}_{ij}|}{N} \tag{15}$$

which represents the percentage of good prediction as defined in Lenormand and Deffuant (2013).

In order to assess the robustness of the results regarding the choice of goodness-of-fit measures, we also test the results obtained with the normalized root mean square error,

$$NRMSE(T, \tilde{T}) = \frac{\sum_{i,j=1}^{n} (T_{ij} - \tilde{T}_{ij})^2}{\sum_{i,j=1}^{n} T_{ij}} \tag{16}$$

and the information gain statistic,

$$I\left(T, \tilde{T}\right) = \sum_{i,j=1}^{n} \frac{T_{ij}}{N} \ln\left(\frac{T_{ij}}{\tilde{T}_{ij}}\right) \tag{17}$$

3.3.2. Common part of links

The ability of the models to recover the topological structure of the original network can be assessed with the common part of links (CPL) defined as

$$CPL\left(T, \tilde{T}\right) = \frac{2 \sum_{i,j=1}^{n} 1_{T_{ij}>0} \cdot 1_{\tilde{T}_{ij}>0}}{\sum_{i,j=1}^{n} 1_{T_{ij}>0} + \sum_{i,j=1}^{n} 1_{\tilde{T}_{ij}>0}} \tag{18}$$

where 1_x is equal to one if the condition X is fulfilled and zero otherwise. The common part of links measures the proportion of links in common between the simulated and the observed networks (i.e. links such as $T_{ij}>0$ and $\tilde{T}_{ij}>0$). It is null if there is no link in common and one if both networks are topologically equivalent.

3.3.3. Common part of commuters according to the distance

In order to measure the similarity between the observed commuting distance distribution and the ones simulated with the models, we introduce the common part of commuters according to the distance (CPC$_d$). Let us consider N_k the number of individuals having a commuting distance in the bin between $2k - 2$ and $2k$ kms. The CPC$_d$ is equal to the CPC based on N_k instead of T_{ij}

$$CPC_d\left(T, \tilde{T}\right) = \frac{\sum k = 1\infty \, min\left(N_k, \tilde{N}_k\right)}{N}. \tag{19}$$

4. Results

In this section, we compare the five laws: gravity with an exponential or a power distance decay function, the Schneider's intervening opportunities law and the original and the extended radiation laws. We test these laws against empirical data coming from eight different case studies using four constrained models to estimate the flows. For each constrained model, the parameters β, γ and a are calibrated so as to maximize the CPC. Since the models are stochastic, we consider an average CPC value measured over 100 replications of the trip distribution. Similarly, all the goodness-of-fit measures are obtained by calculating the average measured over 100 network replications. It is important to note that the networks generated with the constrained models are very stable, the stochasticity of the models does not affect the statistical properties of the network. Therefore, the goodness-of-fit measures does not vary much with the different realizations of the multinomial sampling. For example, within the 100 network instances for all models and case studies, the CPC varies, at most, by 0.09% around the average.

4.1. Estimation of commuting flows

Fig. 3 displays the common part of commuters obtained with the different laws and models for the eight case studies. Globally, the gravity laws give better results than the intervening opportunities laws. For the gravity laws, the results improve with the exponential rather than with the power distance decay function. For the intervening opportunities laws, the extended radiation law outperforms the original one and achieves slightly better results than the Schneider law. In the top left panel, we observe the results for the unconstrained model. In this case, the extended radiation law and the Schneider law give better results than the gravity ones for most case studies. However, these better
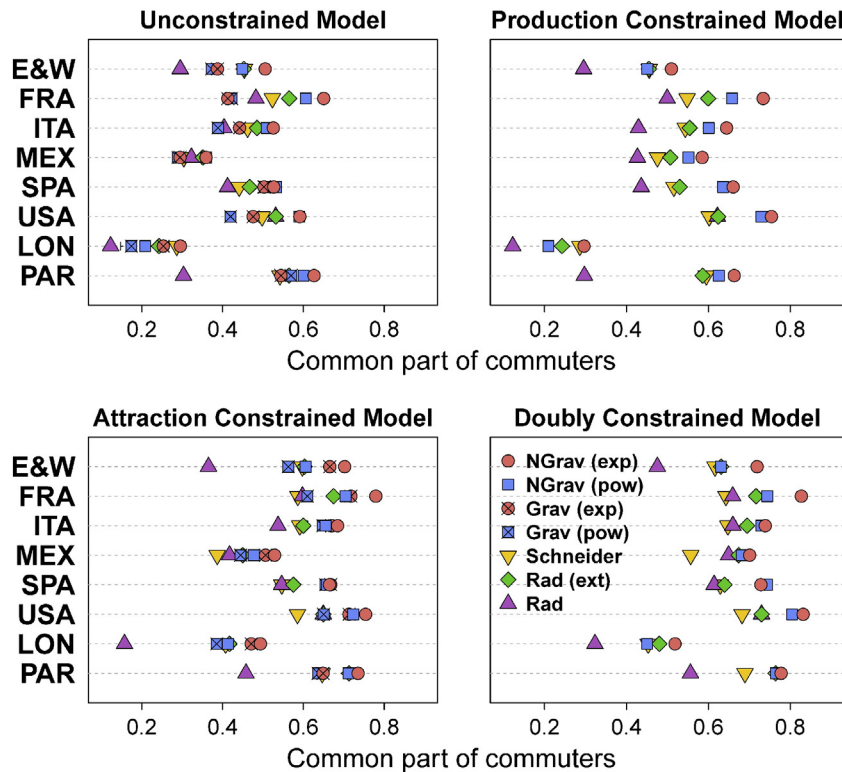
Fig. 3. Common part of commuters according to the unconstrained models, the gravity and intervening opportunities laws for the eight case studies. The circles represent the normalized gravity law with the exponential distance decay function (the circles with a cross inside represent the original version); the squares represent the normalized gravity law with the power distance decay function (the squares with a cross inside represent the original version); the point down triangles represent the Schneider's intervening opportunities law; the green diamonds represent the extended radiation law; the purple triangles represent the original radiation law. Error bars represent the minimum and the maximum values observed in the 100 realizations but in most cases they are too close to the average to be seen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performances are due to the normalization factor used in Eq. (5). Indeed, this normalization implies that the probability of having a trip originating in a census unit i is proportional to the population of i, which is not necessarily the case for the gravity laws. If we use the same type of normalization for the gravity trip distribution law p_{ij} (Eq. (20)), we observe that the "normalized" gravity laws give better results than the intervening opportunities laws. In the following, we will refer to the normalized version when mentioning the gravity law.

$$p_{ij} \propto m_i \frac{m_j f(d_{ij})}{\sum_{k=1}^{n} m_k f(d_{ik})}, \quad i \neq j \qquad (20)$$

To compare the constrained model performances, we plot in Fig. 4a the CPC obtained with the four models according to the laws averaged over the eight case studies. As expected, more constrained the model is, higher the CPC becomes. Unconstrained models are able to reproduce on average around 45% of the observed commuting network against 65% for the doubly constrained model. It is interesting to note that, the attraction constrained model gives better results than the production constrained model. This can be explained by the fact that the job demand is easier to estimate than the job offer, which can be related to extra economic questions. This is in agreement with the results obtained with a uniform distribution ($p_{ij} \propto 1$) plotted in Fig. 3d.

Although the results obtained with the normalized root mean square error and the information gain statistic are very similar to the ones obtained with the CPC, it is worth noting that globally the extended radiation law gives smaller normalized root mean square error values than the normalized gravity laws with the unconstrained

model (see Table 2 for more details about the laws exhibiting the best performances).

4.2. Structure of the commuting network

We consider next the capacity of the gravity and the intervening opportunities laws to recover the structure of the empirical commuting networks. Fig. 4b shows the average common part of links obtained with the different laws and models. We observe that the gravity law with an exponential distance decay function outperforms the other laws when the unconstrained and the single constrained models are used to generate the flows. However, when the doubly constrained models is considered, very similar results are obtained except for the Schneider law and the original version of the radiation law. In any case, the common part of links never exceed 0.55, this can be explained by the fact that, globally, the different laws fail at reproducing the number of links. Indeed, as it can be seen in Fig. 5, which displays the ratio between the number of links generated with the models and the observed ones, the radiation law and the exponential gravity law tend to underestimate the number of links whereas the extended radiation law and the power gravity law overestimate it. The flow networks generated with the Schneider law have globally a number of links closer to the observed values than the networks generated with the other laws.

4.3. Commuting distance distribution

Another important feature to study is the commuting distance distribution. Fig. 4c shows the average common part of commuters according to the distance obtained with the different models and laws. The results obtained with the exponential gravity law are slightly better than the
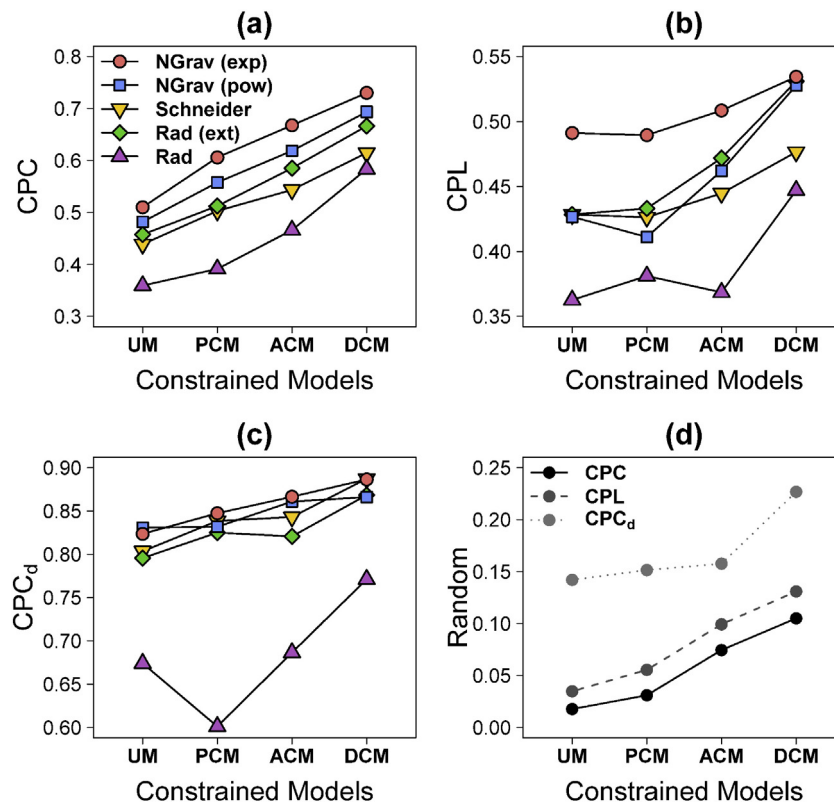
Fig. 4. Performance of the unconstrained model (UM), the production constrained model (PCM), the attraction constrained model (ACM) and the doubly constrained model (DCM) according to the gravity and the intervening opportunities laws (a)–(c) and a uniform distribution (d). (a) Average CPC. (b) Average CPL. (c) Average CPC$_d$. The red circles represent the normalized gravity law with the exponential distance decay function; the blue squares represent the normalized gravity law with the power distance decay function; the point down triangles represent the Schneider's intervening opportunities law; the green diamonds represent the extended radiation law; the purple triangles represent the original radiation law. The grey point down triangles represent the uniform distribution, form dark to light grey, the CPC, the CPL and the CPC$_d$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Law exhibiting the best performances according to the inputs, case studies, constrained models and goodness-of-fit measures.

Inputs	Case study	Model	CPC	CPL	CPC$_d$	NRMSE	I
Population	E&W	UM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	IO
Population	FRA	UM	NGrav (exp)	NGrav (exp)	NGrav (exp)	Rad (ext)	NGrav (exp)
Population	ITA	UM	NGrav (exp)	NGrav (exp)	IO	Rad (ext)	NGrav (exp)
Population	MEX	UM	NGrav (pow)	NGrav (exp)	Rad	Rad (ext)	NGrav (exp)
Population	SPA	UM	NGrav (pow)	NGrav (exp)	NGrav (pow)	Rad (ext)	NGrav (exp)
Population	USA	UM	NGrav (exp)	NGrav (exp)	NGrav (pow)	Rad (ext)	NGrav (exp)
Population	LON	UM	NGrav (exp)	IO	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	PAR	UM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (pow)	NGrav (exp)
Population	E&W	PCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	IO
Population	FRA	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	ITA	PCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
Population	MEX	PCM	NGrav (exp)	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)
Population	SPA	PCM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	USA	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	LON	PCM	NGrav (exp)	IO	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	PAR	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	E&W	ACM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
Population	FRA	ACM	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	ITA	ACM	NGrav (exp)	NGrav (exp)	IO	NGrav (pow)	NGrav (exp)
Population	MEX	ACM	NGrav (exp)	NGrav (exp)	NGrav (exp)	Rad (ext)	NGrav (exp)
Population	SPA	ACM	NGrav (exp)	NGrav (pow)	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	USA	ACM	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	LON	ACM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	PAR	ACM	NGrav (exp)	NGrav (exp)	NGrav (pow)	IO	NGrav (exp)
Population	E&W	DCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
Population	FRA	DCM	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	ITA	DCM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
Population	MEX	DCM	NGrav (exp)	NGrav (pow)	Rad (ext)	NGrav (exp)	NGrav (exp)
Population	SPA	DCM	NGrav (pow)	NGrav (pow)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	USA	DCM	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	LON	DCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
Population	PAR	DCM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	E&W	UM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	FRA	UM	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	ITA	UM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	MEX	UM	NGrav (exp)	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)
In/out flows	SPA	UM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	USA	UM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	LON	UM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	PAR	UM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	E&W	PCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	FRA	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	ITA	PCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	MEX	PCM	NGrav (exp)	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)
In/out flows	SPA	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	USA	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	LON	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	PAR	PCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	E&W	ACM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	FRA	ACM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	ITA	ACM	NGrav (exp)	NGrav (exp)	IO	NGrav (pow)	NGrav (exp)
In/out flows	MEX	ACM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	SPA	ACM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	USA	ACM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	LON	ACM	NGrav (exp)	Rad (ext)	IO	NGrav (exp)	NGrav (exp)
In/out flows	PAR	ACM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)
In/out flows	E&W	DCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	FRA	DCM	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	ITA	DCM	NGrav (exp)	NGrav (exp)	IO	NGrav (exp)	NGrav (exp)
In/out flows	MEX	DCM	NGrav (exp)	NGrav (pow)	Rad (ext)	NGrav (exp)	NGrav (exp)
In/out flows	SPA	DCM	NGrav (pow)	NGrav (pow)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	USA	DCM	NGrav (exp)	Rad (ext)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	LON	DCM	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)	NGrav (exp)
In/out flows	PAR	DCM	NGrav (exp)	NGrav (exp)	NGrav (pow)	NGrav (exp)	NGrav (exp)

ones obtained with the other laws. However, the results are globally good, and except the original radiation law, the gravity and intervening opportunities laws are able to reproduce more than 80% of the commuting distances.

To go further, we plot in Fig. 6 the observed and the simulated commuting distance distributions obtained with the production constrained model in France and United States. We can clearly see that the exponential gravity law is better for estimating commuting distances which are below a certain threshold equal to 50 km in France and 150 km in United States. After this threshold, it fails at estimating the commuting flows as it is the case for the Schneider's intervening opportunities law. On the contrary, the radiation laws and the gravity law with a power distance decay function are able to estimate commuting flows at large distances. However, we have to keep in mind that the proportion of commuters traveling such long distances are less than 6% in France and 5% in United States. Besides, one can legitimately wonder
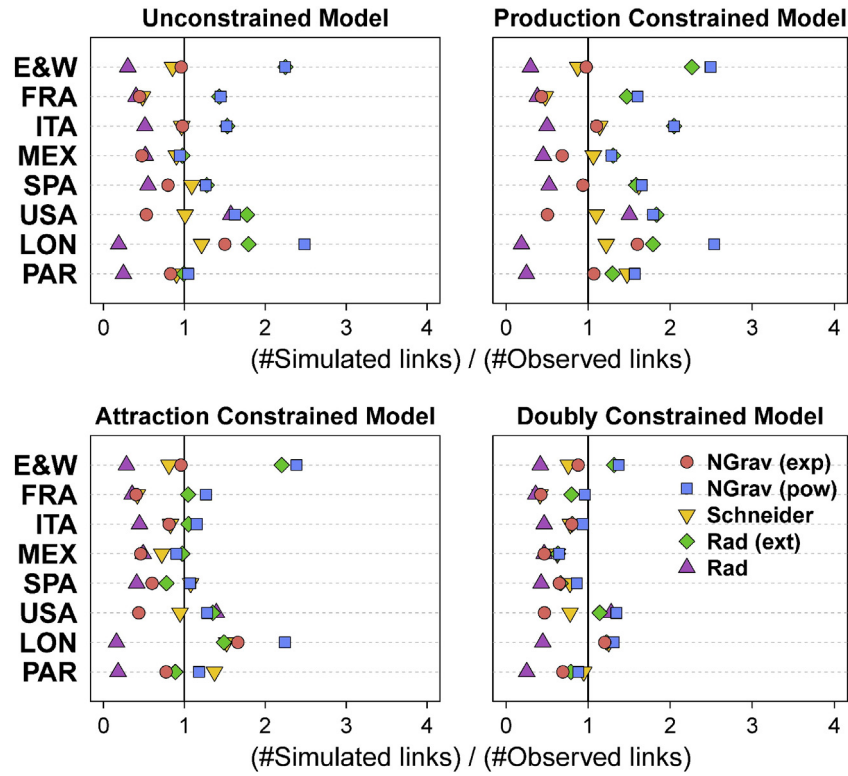
Fig. 5. Ratio between the simulated and the observed number of links according to the unconstrained models, the gravity and intervening opportunities laws for the eight case studies. The red circles represent the normalized gravity law with the exponential distance decay function; the blue squares represent the normalized gravity law with the power distance decay function; the point down triangles represent the Schneider's intervening opportunities law; the green diamonds represent the extended radiation law; the purple triangles represent the original radiation law. Error bars represent the minimum and the maximum but in most cases they are too close to the average to be seen. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

whether these long travels are repeated twice per day or if they may be an artifact of the way in which the census information is collected.

4.4. Robustness against changes in the inputs

In Eqs. (2) and (5), the population is used as input instead of the out-flows O_i and the inflows D_j, which are usually preferred since they are a more faithful reflection of the job demand and offer. The job demand and offer are considered to be related to the population but the proportion is rarely direct (it needs to be adjusted with an exponent) and according to the case study, the fit can be bad. In order to assess the robustness of the results to changes in the input data, we consider the results obtained with the gravity law (Eq. (21)) and the general intervening opportunities law (Eq. (22)) based on the in and out flows. In the case of the intervening opportunities laws, s_{ij} is the number of in-commuters in a circle of radius d_{ij} centered in i (excluding the source
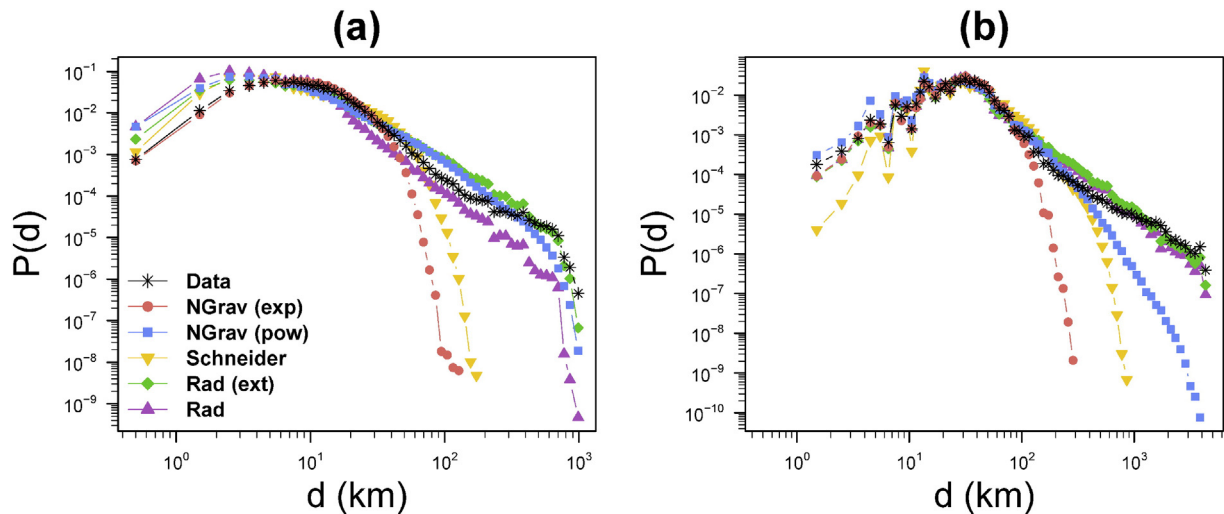
Fig. 6. Probability density function of the commuting distance distribution observed in the data and simulated with the production constrained model. (a) France and (b) United States. The red circles represent the normalized gravity law with the exponential distance decay function; the blue squares represent the normalized gravity law with the power distance decay function; the point down triangles represent the Schneider's intervening opportunities law; the green diamonds represent the extended radiation law; the purple triangles represent the original radiation law. The black stars represent the census data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
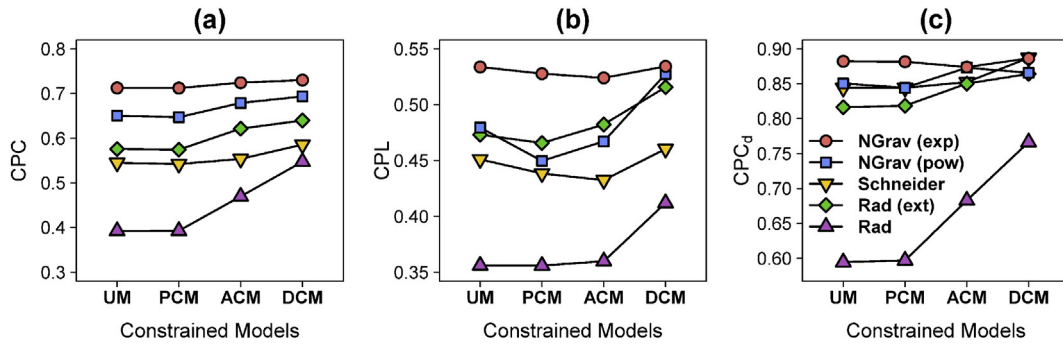
Fig. 7. Performance of the constrained models according to the gravity and the intervening opportunities laws. (a) Average CPC. (b) Average CPL. (c) Average CPC$_d$. The red circles represent the normalized gravity law with the exponential distance decay function; the blue squares represent the normalized gravity law with the power distance decay function; the point down triangles represent the Schneider's intervening opportunities law; the green diamonds represent the extended radiation law; the purple triangles represent the original radiation law. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and destination) and the role of the populations in the gravity law is taken by O_i and D_j. To be more specific, the gravity law becomes:

$$p_{ij} \propto O_i \frac{D_j f(d_{ij})}{\sum_{k=1}^{n} D_k f(d_{ik})}, \quad i \neq j \tag{21}$$

while the intervening opportunities law can be written as

$$p_{ij} \propto O_i \frac{\mathbb{P}(1|D_i, D_j, s_{ij})}{\sum k = 1 n \mathbb{P}(1|D_i, D_k, s_{ik})}, \quad i \neq j. \tag{22}$$

Fig. 7 displays the average CPC, CPL and CPC$_d$ obtained with the four models according to the laws averaged over the eight case studies. As it can be seen on these plots the results observed in Fig. 4 are quite stable to changes in the input data.

4.5. Parameter calibration in the absence of detailed data

An important issue with the estimation of commuting flows is the calibration of the parameters. Indeed, how to calibrate the parameters β, γ and a in the absence of detailed data? This problem has already been tackled in previous studies (Balcan et al., 2009; Lenormand et al., 2012; Yang et al., 2014). In Lenormand et al. (2012), the authors have shown that, in the case of the exponential form of the gravity law, the value of β can be directly inferred from the average census unit surface with the relationship $\beta = 0.3 <S>^{-0.18}$. Similarly, Yang et al. (2014) proposed to estimate the value of a in the extended radiation law with the average spatial scale $l = \sqrt{<S>}$ using the functional relationship $\alpha = 0.0085\, l^{1.33}$.

In Fig. 8, we plot the calibrated value of β, γ and a obtained with the laws based on the population as a function of the average census unit surface $<S>$ for the four constrained models. Fig. 8a shows the
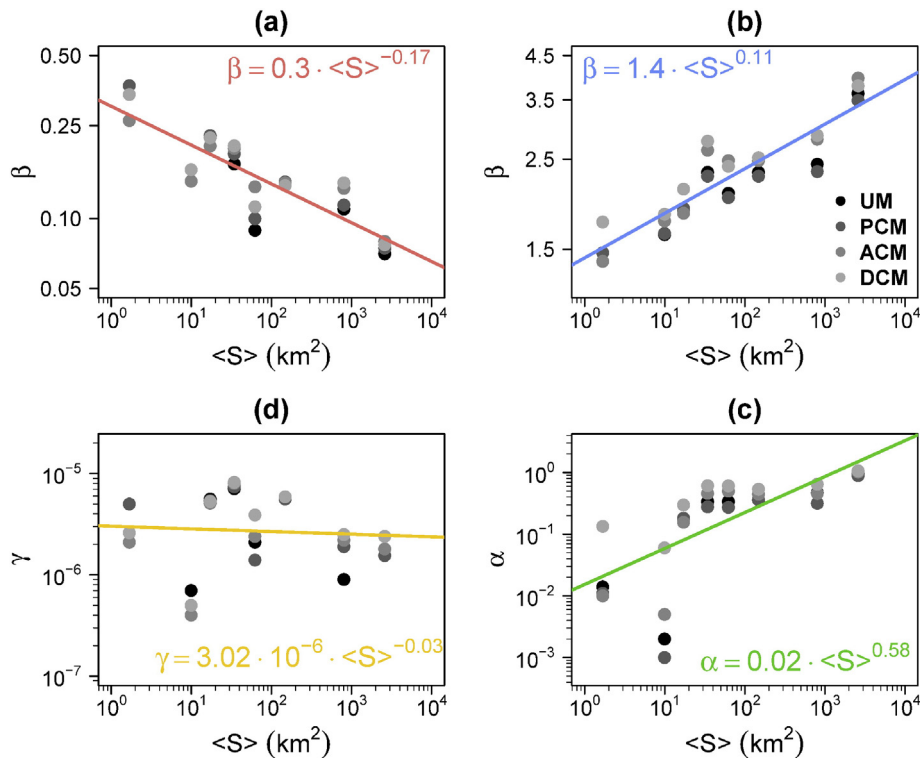

Fig. 8. Parameter value as a function of the average unit surface. (a) Normalized gravity laws with an exponential distance decay function. (b) Normalized gravity laws with a power distance decay function. (c) Schneider's intervening opportunities law. (d) Extended radiation law.

relationship obtained with the gravity law with an exponential distance decay function. We observe that the coefficients of the relationship are the same than the ones obtained in Lenormand et al. (2012). This is not surprising since three datasets out of the six used here coincide. In this case, the value of β decreases with larger spatial scales. This can be explained by the fact that β in the exponential form of the gravity law is inversely proportional to the average commuting distance and such distance increases with the average unit surface since the shorter distance trips are excluded (Fig. 9b). Fig. 8b displays the same relationship for the power form of the gravity law, in this case the value of β increases with the scale to fit the tail of the commuting distance distribution. In fact, we observe in the data that, globally, the steepness of the curve (measured with the Pearson's Kurtosis) increases with the scale (Fig. 9c). Fig. 8c shows the results obtained with the parameter γ of the Schneider intervening opportunities law. The value of γ seems to decrease slightly with the scale but the existence of a relationship between the two variables is not significant. Finally, we plot in Fig. 8d the relationship between the parameter a of the extended radiation law and the average unit surface, the exponent obtained is similar to the one reported in Yang et al. (2014). In the extended version of the radiation law, the parameter a controls the effect of the number of job opportunities between home and work on the job selection. In particular,

for a given number of job opportunities, higher the value of a, higher the probability of not accepting a job among these opportunities. This implies that a is directly proportional to the average commuting distance and, by extension, to the average unit surface (Fig. 9b). As mentioned in Yang et al. (2014), the value of a is also influenced by the heterogeneity of the distribution of opportunities. As it can be seen in Fig. 8d, the three case studies presenting the largest deviation from the regression line are also the most heterogeneous ones (Paris, Spain and Italy which have the second, fourth and fifth smallest average unit surface, respectively).

As in Lenormand et al. (2012), it is possible to assess the quality of the parameter estimation by measuring its impact on the CPC. The idea is to measure for each law, model and case study, the difference between the CPC obtained with the calibrated value of the parameter and the CPC obtained with the estimated one. The parameter value is estimated with the regression model obtained with the laws based on the population and the difference between the original CPC and the "estimated" one is measured with the absolute percentage error (i.e. absolute error as a percent of the original CPC value). In order to assess the robustness of the estimation in changes in the input we have also measured the CPC percentage error obtained with an estimation of the parameters for the laws based on the in/out flows. Note that in
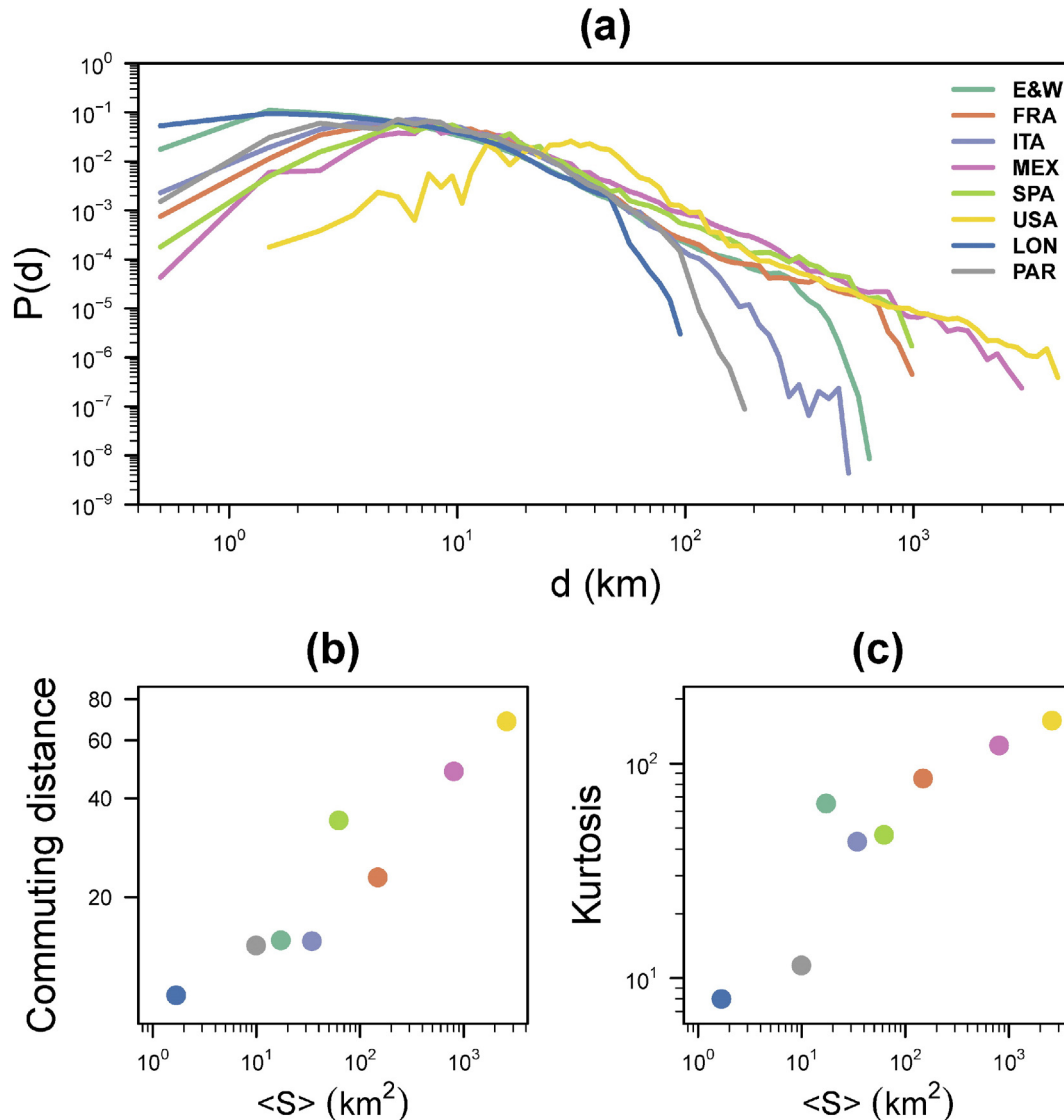
Fig. 9. Observed commuting distance distributions. (a) Probability density function of the commuting distance distribution according to the case study. (b) Average commuting distance as a function of the average unit surface. (c) Pearson's measure of Kurtosis as a function of the average unit surface.
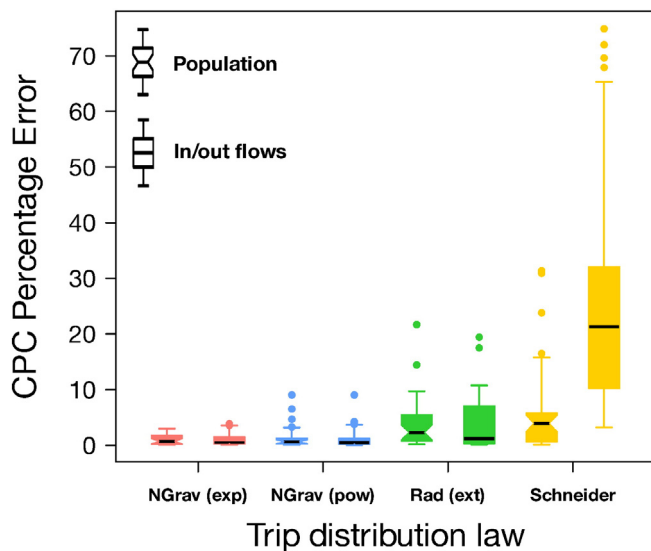
Fig. 10. CPC absolute percentage error. Boxplots of the absolute percentage error between the CPC obtained with a calibrated value of the parameters and the CPC obtained with values estimated with the regression models obtained with the laws based on the population. The notched and classic boxplots represent the percentage error obtained with the laws based on the population and the number of in/out flows, respectively. The boxplot is composed of the first decile, the lower hinge, the median, the upper hinge and the last decile.

this case the parameters' estimation come also from regression models obtained with the laws based on the population. The results are presented in Fig. 10. The CPC percentage errors obtained with the gravity laws are globally small and robust to the change of inputs. They vary at most by 4% of the original CPC values for the exponential form and 10% for the power form. Similar results are obtained for the extended radiation law where the majority of the errors are below 10% and vary at most by 22% of the original CPC values. This means that for these laws the parameter value can be directly inferred from the scale, and thus, commuting networks at different scales can be generated without requiring detailed data for calibration. The situation is different for the Schneider's intervening opportunities law very sensible to change in inputs. For the law based on the population, the errors obtained for the CPC are reasonable, the majority of them are below 10%. However when we try to estimate the value of γ for the law based on in/out flows with a regression model obtained with the law based on the population the CPC percentage error increases dramatically, meaning that the value of γ is highly dependent on the variable uses as a surrogate measure of the number of "real" opportunities.

5. Discussion

In summary, we have compared different versions of the gravity and the intervening opportunities laws. These two approaches have already been compared in the past but using different inputs, number of parameters and/or type of constraints. For this reason, the aim of this work has been to bring some light into the discussion by systematically comparing the intervening opportunities and the gravity laws taking care of dissociating the probabilistic laws and the constrained models used to generate the trip networks. We have shown that, globally, the gravity approach outperforms the intervening opportunities approach to estimate the commuting flows but also to preserve the commuting network structure and to fit of the commuting distance distribution. More particularly the gravity law with the exponential distance decay function give better results than the other laws even if it fails at estimating commuting flows at large distances. The reason for this is that most of the travels are short-range, which are better captured by the gravity law with exponential decay in the distance. The large distance commuting trips

are few and probably associated with weekly rather than daily commuting. To handle these different types of mobility, it may be necessary to investigate further the nature of the trips and to consider even mixed models for different displacement lengths. The superiority of the gravity law is very robust to the choice of goodness-of-fit measure and to the change of input. Regarding a more practical issue which is the calibration of the parameters without detailed data, we have shown that the parameter values can be estimated with the average unit surface. We also demonstrated that, except for the Schneider's intervening opportunities law, this estimation is robust to changes in input data. This allows for a direct estimation of the commuting flows even in the absence of detailed data for calibration.

Although more research is needed to investigate the link between mobility, distances and intervening opportunities for other types of movements such as migrations, tourism or freight distribution, the distance seems to play a more important role than the number of intervening opportunities in work location choices. More specifically, the superiority of the gravity approach seems to be due to its flexibility, and, what was considered as a weakness by Simini et al. (2012), the lack of theoretical guidance to choose the distance-decay function, emerges as a strength. Indeed, people do not choose their place of work as they choose their new place of residence, therefore, having the possibility of adjusting the effect of the distance in the decision process is clearly an advantage which does not apply to the intervening opportunities approach in its present form.

The objective of this work has been to establish the basis for a fair and systematic comparison separating probabilistic laws and different degrees of constraint trip generation models. Our results emphasize the importance of identifying and separating the different processes involved in the estimation of flows between locations for the comparison of spatial interaction models. Indeed, the use of these models in contexts such as urban and infrastructure planning, where large investments are at stake, imposes the need for the selection of the aptest model before taking decisions based on its results. The software package to generate spatial networks using the approach described in the paper can be downloaded from https://github.com/maximelenormand/Trip-distribution-laws-and-models.

References

Akwawua, S., Poller, J.A., 2001. The development of an intervening opportunities model with spatial dominance effects. J. Geogr. Syst. 3, 69–86.

Anderson, J., 1979. A theoretical foundation for the gravity equation. Am. Econ. Rev. 69, 106–116.

Balcan, D., Colizza, V., Goncalves, B., Hud, H., Ramasco, J., Vespignani, A., 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. Proc. Natl. Acad. Sci. U. S. A. 106, 21,484–21,489.

Barthelemy, M., 2011. Spatial networks. Phys. Rep. 499, 1–101.

Bergstrand, J., 1985. The gravity equation in international trade: some microeconomic foundations and empirical evidence. Rev. Econ. Stat. 67, 474–481.

Carey, H.C., 1858. Principles of Social Science. Lippincott.

Chen, Y, 2015. The distance-decay function of geographical gravity model: power law or exponential law? Chaos, Solitons Fractals 77, 174–189.

David, K.W., 1961. Comparison of trip distribution by opportunity model and gravity model. Pittsburgh Area Transportation Study.

De Montis, A., Barthelemy, M., Chessa, A., Vespignani, A., 2007. The structure of interurban traffic: a weighted network analysis. Environ. Plan. B Plan. Des. 34, 905–924.

De Montis, A, Chessa, A, Campagna, M, Caschili, S, Deplano, G, 2010. Modeling commuting systems through a complex network analysis: a study of the italian islands of sardinia and sicily. J. Transp. Land Use 2, 39–55.

Deming, W.E., Stephan, F.F., 1940. On a least squares adjustment of a sample frequency table when the expected marginal totals are known. Ann. Math. Stat. 11, 427–444.

Erlander, S, Stewart, NF, 1990. The gravity model in transportation analysis: theory and extensions. Topics in transportation, VSP, Utrecht, The Netherlands.

Fik, T.J., Mulligan, G.F., 1990. Spatial flows and competing central places: toward a general theory of hierarchical interaction. Environ. Plan. A 22, 527–549.

Fotheringham, A., 1981. Spatial structure and distance-decay parameters. Ann. Assoc. Am. Geogr. 71, 425–436.

Gargiulo, F., Lenormand, M., Huet, S., Baqueiro Espinosa, O., 2012. Commuting network model: getting to the essentials. J. Artif. Soc. Soc. Simul. 15, 13.

Griffith, D.A., 2009. Modeling spatial autocorrelation in spatial interaction data: empirical evidence from 2002 Germany journey-to-work flows. J. Geogr. Syst. 11, 117–140.

Haynes, K.E., Poston, D.L.J., Schnirring, P., 1973. Intermetropolitan migration in high and low opportunity areas: indirect tests of the distance and intervening opportunities hypotheses. Econ. Geogr. 49, 68–73.

Heanus, K., Pyers, C., 1966. A comparative evaluation of trip distribution procedures. Public Roads 34, 43–51.

Jung, W.S., Wang, F., Stanley, H., 2008. Gravity model in the Korean highway. Europhys. Lett. 81, 48,005.

Kaluza, P., Kölzsch, A., Gastner, M.T., Blasius, B., 2010. The complex network of global cargo ship movements. J. R. Soc. Interface 7, 1097–1103.

Krings, G., Calabrese, F., Ratti, C., Blondel, V.D., 2009. Urban gravity: a model for inter-city telecommunication flows. J. Stat. Mech: Theory Exp. 2009, (L07003 +).

Lawson, M., Dearinger, J., 1967. A comparison of four work trip distribution models. Proc. Am. Soc. Civ. Eng. 93, 1–25.

Lenormand, M., Deffuant, G., 2013. Generating a synthetic population of individuals in households: sample-free vs sample-based methods. J. Artif. Soc. Soc. Simul. 14, 12.

Lenormand, M., Huet, S., Gargiulo, F., 2014. Generating French virtual commuting network at municipality level. J. Transp. Land Use 7, 43–55.

Lenormand, M, Huet, S, Gargiulo, F, Deffuant, G, 2012. A universal model of commuting networks. PLoS One 7, e45985.

Liang, X, Zhao, J, Dong, L, Xu, K, 2013. Unraveling the origin of exponential law in intra-urban human mobility. Sci. Report. 3, 2983.

Liu, Y, Sui, Z, Kang, C, Gao, Y, 2014. Uncovering patterns of inter-urban trip and spatial interaction from social media check-in data. PLoS One 9, e86026.

Luo, W., Wang, F., 2003. Measures of spatial accessibility to health care in a GIS environment: synthesis and a case study in the Chicago region. Environ. Plan. B Plan. Des. 30, 865–884.

Masucci, A., Serras, J., Johansson, A., Batty, M., 2013. Gravity versus radiation models: on the importance of scale and heterogeneity in commuting flows. Phys. Rev. E 88, 022,812.

Murat, C.H., 2010. Sample size needed for calibrating trip distribution and behavior of the gravity model. J. Transp. Geogr. 18, 183–190.

Ortúzar, J., Willumsen, L., 2011. Modeling Transport. John Wiley and Sons Ltd., New York.

Pyers, C., 1966. Evaluation of intervening opportunities trip distribution models. Highw. Res. Rec. 114, 71–88.

Ren, Y., Ercsey-Ravasz, M., Wang, P., González, M.C., Toroczkai, Z., 2014. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. Nat. Commun. 5, 5347.

Rouwendal, J., Nijkamp, P., 2004. Living in two worlds: a review of home-to-work decisions. Growth Chang. 35, 287–303.

Ruiter, E.R., 1967. Toward a better understanding of the intervening opportunities model. Transp. Res. 1, 47–56.

Sagarra, O., Szell, M., Santi, P., Diaz-Guilera, A., Ratti, C., 2015. Supersampling and network reconstruction of urban mobility. PLoS One 10, e0134508.

Schneider, M., 1959. Gravity models and trip distribution theory. Pap. Reg. Sci. Assoc. 5, 51–58.

Simini, F., González, M.C., Maritan, A., Barabasi, A.L., 2012. A universal model for mobility and migration patterns. Nature 484, 96–100.

Simini, F., Maritan, A., Néda, Z., 2013. Human mobility in a continuum approach. PLoS One 8, e60069.

Sørensen, T., 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. Biol. Skr. 5, 1–34.

Stouffer, S.A., 1940. Intervening opportunities: a theory relating mobility and distance. Am. Sociol. Rev. 5, 845–867.

Thomas, T., Tutert, S., 2013. An empirical model for trip distribution of commuters in the netherlands: transferability in time and space reconsidered. J. Transp. Geogr. 26, 158–165.

Tizzoni, M., Bajardi, P., Decuyper, A., Kon Kam King, G., Schneider, C.M., Blondel, V., Smoreda, Z., González, M., Colizza, V., 2014. On the use of human mobility proxies for modeling epidemics. PLoS Comput. Biol. 10, e1003716.

Viboud, C., Bjørnstad, O.N., Smith, D.L., Simonsen, L., Miller, M.A., Grenfell, B.T., 2006. Synchrony, waves, and spatial hierarchies in the spread of influenza. Science 312, 447–451.

de Vries, J, Nijkamp, P, Rietveld, P, 2009. Exponential or power distance-decay for commuting? an alternative specification. Environ. Plan. A 41, 461–480.

Wilson, A., 1970. Entropy in urban and regional modelling. Pion, London.

Yang, Y., Herrera, C., Eagle, N., González, M.C., 2014. Limits of predictability in commuting flows in the absence of data for calibration. Sci. Report. 4, 5662.

Zhao, F., Chow, L.F., Li, M.T., Gan, A., L., S.D., 2001. Refinement of FSUTMS trip distribution methodology. Technical ReportTechnical Memorandum 3. Florida International University.

Zipf, G.K., 1946. The P1 P2/D hypothesis: on the intercity movement of persons. Am. Sociol. Rev. 11, 677–686.