EPJ .org

○ **EPJ Data Science**
a SpringerOpen Journal

REGULAR ARTICLE

Open Access

CrossMark

Commuting patterns: the flow and jump model and supporting data

Levente Varga[1], Géza Tóth[2] and Zoltán Néda[1*] ⓘ

*Correspondence:
zneda@phys.ubbcluj.ro
[1]Department of Physics,
Babeş-Bolyai University,
Cluj-Napoca, Romania
Full list of author information is
available at the end of the article

Abstract

A simple model, named the flow and jump model (FJM) is used for describing commuter fluxes at different distances. The model is based on a master equation which allows a local net probability flow and non-local jumps. FJM is in principle a one-parameter model, however it is found that by fixing this parameter we get a parameter free model, similar with the radiation model. We find that FJM offers an improved description for commuting data from USA, Italy and Hungary. For a special choice of the model parameter FJM leads to the radiation model.

Keywords: Human mobility models; Commuters data; Population density

1 Introduction

Commuter mobility patterns are in the focus of many recent studies. The problem by its nature belongs to the research field of human geography, sociology and economics. Nowadays however, researchers from many other fields became interested in the topic. The interest in such studies can be explained by the fact that many large electronic datasets became available for researchers, allowing to test both the assumptions and main results of the models. As a special case of human mobility, statisticians and data scientist are interested in universal patterns that govern the commuter fluxes at different spacial scales. Physicists and mathematicians are interested in simple models capable of explaining the observed patterns. A detailed review for the state of the art of the field of human mobility is given in the recent review article of Barbosa et al. [1].

Models for community fluxes, motivated phenomenologically by some simple socio-economic or probabilistic arguments, were proposed already in the early 1940 by Stouffer (*the intervening opportunities model*) [2] and by Block & Marschak in 1960 (*the random utility model*) [3]. Analogies with some classical physics phenomena were exploited by the very popular *gravity* and *generalized potential models* [4].

From modeling perspective a great leap in the understanding of human mobility patterns represented the *radiation model* introduced by Simini et al. [5]. In contrast with earlier models that were phenomenologically argued, the radiation model started from a basic socio-economic optimization assumption and derived a simple and compact formula for commuter fluxes. Relative to the earlier used models the compact result derived in [5] has also the advantage that it is parameter free. When compared however with real commuting and population data, the model contains an undetermined proportionality constant that

makes the connection between the population and available job number. In such sense one can argue that this model is a one free parameter model. Other models of similar complexity, built also on realistic assumptions are the population-weighted opportunities model (PWO) [6] and a novel version of it where also memory effects are considered [7]. Recently a new, parameter-free model was introduced by Liu and Yan [8]. Their basic assumption is that individuals select destination locations that present higher opportunity benefits than the ones at the origin and the intervening opportunities between the origin and destination.

The radiation model was generalized for continuous population distribution [9] and was also made more realistic by allowing a realistic job selection for the individuals. This new model, *the radiation model with selection* offered an improved fit for the commuting data in USA. Further improvement for the simple radiation model was considered by taking into account also the travel cost involved in commuting (see for example [10]). In the line of this model the *travel cost optimized radiation model* introduced by us recently [11] offered an improved description for the commuting fluxes in Hungary. The main drawback of all these later generalization to the radiation model is that the parameter-free beauty of the radiation model is lost.

Here we offer a further generalization for the original radiation model, and prove it's advantages relative to the earlier models using large-scale population density and commuter flux data from USA, Italy and Hungary. The nice aspect of this generalization is that our model is again a parameter-free model, since the only fitting parameter is fixed to a universal value.

2 Modelling framework

The *gravity model (GM)* is probably the most known approach to describe empirically the commuter fluxes between cities [12]. It is based on a phenomenological analogy with gravity, assuming that the interaction between two regions or cities depends in an inverse proportionality with the distance raised at a positive power and in direct proportionality with the size of the two regions/cities. Contrary with what is usually believed, GM is not only a simple analogy there are also theoretical arguments in favour of it. The oldest one is probably the one using the maximal entropy hypothesis [13, 14]. Other successful attempts are based on the principle of utility maximization in economics. Both deterministic [15, 16] and random utility theories [17] were considered.

In the most general form, the number of commuters $f_i(j)$ between cities i and j is written as:

$$f_i(j) = F(W_i) \frac{(W_j)^\alpha}{(r_{i,j})^\beta}. \tag{1}$$

We denoted here by W_i the population of the settlement i and by $r_{i,j}$ the distance between settlements i and j. $F(x)$ is an arbitrary monotonically increasing kernel function, and α and β are fitting exponents. From the $f_i(j)$ data one can also compute the $P_>^i(W)_{\mathrm{GM}}$ probability, that a worker living in location i commutes to a location that is outside of a disk containing a population W and centred at its home:

$$P_>^i(W)_{\mathrm{GM}} = 1 - \frac{\sum_{j \neq i}^{(w_i[j] < W)} f_i(j)}{\sum_j f_i(j)} = 1 - \frac{\sum_{j \neq i}^{(w_i[j] < W)} \frac{(W_j)^\alpha}{(r_{i,j})^\beta}}{\sum_j \frac{(W_j)^\alpha}{(r_{i,j})^\beta}}, \tag{2}$$

which is independent of the $F(x)$ kernel function. We denoted by $w_i[j]$ the total population inside a disk centred at location i and reaching to location j. Now, the $P_>(W)_{GM}$ probability that commuters travel to work at a distance where they pass a disk with population W is:

$$P_>(W)_{GM} = \langle P_>^i(W)_{GM} \rangle_i. \tag{3}$$

The GM model in such sense is a two parameter model and one has to determine the best α and β exponent values.

The *original radiation model (RM)* [5] is based on the simple assumption that jobseekers are optimizing their income by accepting the closest job offer that offers a better salary than the one available at their current address. Assuming a $p_\leq(z)$ distribution function for the incomes in the studied society the probability $P_>(z|n)$ that a person with income z refuses the closest n jobs is:

$$P_>(z|n) = \left[p_\leq(z) \right]^n. \tag{4}$$

By using the probability density function for incomes, the probability of not accepting the closest n jobs, $P_>(n)$, can be calculated as:

$$
\begin{aligned}
P_>(n) &= \int_0^\infty P_>(z|n) p(z)\, dz = \int_0^\infty P_>(z|n) \frac{\partial p_\leq(z)}{\partial z}\, dz \\
&= \int_0^1 \left[p_\leq(z) \right]^n dp_\leq(z) = \frac{1}{n+1}.
\end{aligned} \tag{5}
$$

Accepting now the hypothesis that the number of job openings in a territory is proportional with the W population ($n = \mu W$), the radiation model predicts the probability that a person commutes to a location that is outside of a disk centered on its current location and containing a population W:

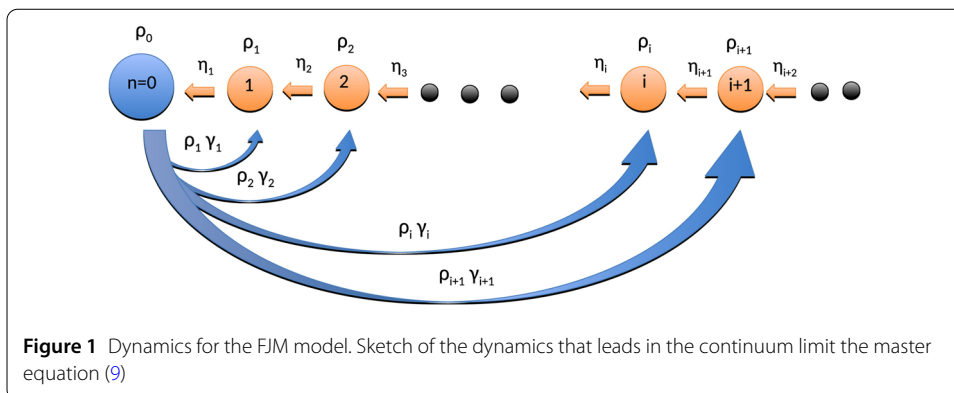$$P_>(W)_{RM} = \frac{1}{\mu W + 1}. \tag{6}$$

It is interesting to note that the hypothesis $n \propto W$ can be proven on real-life data using job advertisement and population data. This is also done in the Results section.

Assuming that jobseekers are willing to accept jobs (or they are aware of the jobs) only with a probability q, the above presented simple argument can be generalized [9] (*radiation model with selection (RMwS)*), leading to a result with two fitting parameters (q, μ):

$$P_>(W)_{RMwS} = \frac{1 - (1-q)^{\mu W + 1}}{(\mu W + 1)q}. \tag{7}$$

The *travel cost optimized radiation model (TCORM)* takes into account the fact that travel costs are distance dependent so in addition with the transited jobs the travel distance, r has to be considered when applying the arguments used in the radiation model. Assuming an exponential distribution kernel for the income distribution, and repeating the arguments from the original radiation model [11] one arrives again to a result with two fitting parameters:

$$P_>(W)_{TCORM} = \frac{1 + \lambda \sqrt{W}}{\mu W + 1}. \tag{8}$$

Figure 1 Dynamics for the FJM model. Sketch of the dynamics that leads in the continuum limit the master equation (9)

The λ fitting parameter incorporates both the value of μ, the value of a proportionality constant between the travelled distance and cost of travel and a third constant governing the shape of an assumed exponential-type income distribution [11].

Here we introduce yet another model, offering another one-parameter alternative for the simple RM model. Our alternative, dynamical approach is based on simple master equation for the $\rho(n,t) = -dP_>(n,t)/dn$ probability density and reproduces as a specific case the results of the RM model. We name the model as *Flow and Jump Model (FJM)*. Following the assumptions of the recently introduced *growth and reset type models* (for a review please consult [18]) we assume now an inverse process: a backward probability flow supplemented by a jump process from the origin to any state with a given n value. The discrete version of the process is depicted in Fig. 1. The continuous master equation has the form:

$$\frac{d\rho(n,t)}{dt} = \frac{\partial(\eta(n)\rho(n,t))}{\partial n} + \left[\gamma(n)\rho(n,t)\right]\rho(0,t). \tag{9}$$

The above master equation describes a process where there is a local net probability density flow from each state towards the $n = 0$ state and a jump probability from the origin ($n = 0$) to an n state. For the state dependent $\eta(n)$ and $\gamma(n)$ rates we consider now simple kernels which makes sense for the commuting process. Definitely the transitions $0 \rightarrow n$ governed by the $\gamma(n)\rho(n,t)$ rates describes the probability that workers choose a commuting job. $\gamma(n)$ should decrease with distance (or correspondingly with n) and the proportionality with $\rho(n,t)$ suggests that where are already many commuters there should also be many good jobs, so it is attractive to commuters.

For more details about such dynamical equation, their stability and stationarity please consult [18]. As shown in [18], the stationary solution ($d\rho_s(n,t)/dt = 0$) of (9) is:

$$\rho_s(n) = \frac{\eta(0)\rho_s(0)}{\eta(n)} e^{-\int_0^n \frac{\gamma(x)\rho_s(0)}{\eta(x)} dx}. \tag{10}$$

The $\rho_s(0)$ value is obtained from the normalization condition:

$$\int_0^\infty \rho_s(n) \, dn = 1. \tag{11}$$

For $\eta(n)$ and $\gamma(n)$ rates we consider now the simplest kernels which makes sense for the commuting process. For $\gamma(n)\rho_s(0)$ the simplest choice that avoids also the divergence in

$n = 0$ is an inverse proportionality:

$$\gamma(n)\rho_s(0) = \frac{C}{n + 1}. \tag{12}$$

C is a constant which fixes also the time unit in the dynamical equation (9). The backward flow characterizes the tendency of the commuters to search for appropriate jobs that are closer to their living places, accepting with a bigger probability jobs that will approach them to their home. This net flow is described by the $\eta(n)$ terms. The simplest choice that leads to a final equilibrium distribution is:

$$\eta(n) = \eta = \text{Const.} \tag{13}$$

For the above $\gamma(n)$ and $\eta(n)$ kernels (equations (12) and (13), respectively), and assuming $a = C/\eta > 1$ the solution (10) writes as

$$\rho_s(n) = (a - 1)(n + 1)^{-a}, \tag{14}$$

which is a scaling Tsallis–Pareto (or Lomax) type distribution [19]. This probability density leads to the $P_>(n, t)$ probability:

$$P_>(n, t) = \int_n^\infty \rho_s(x)\, dx = (1 + n)^{(1-a)}. \tag{15}$$

With the assumption $n(r) = \mu W(r)$ we get a slightly modified expectation for $P_>(W)$
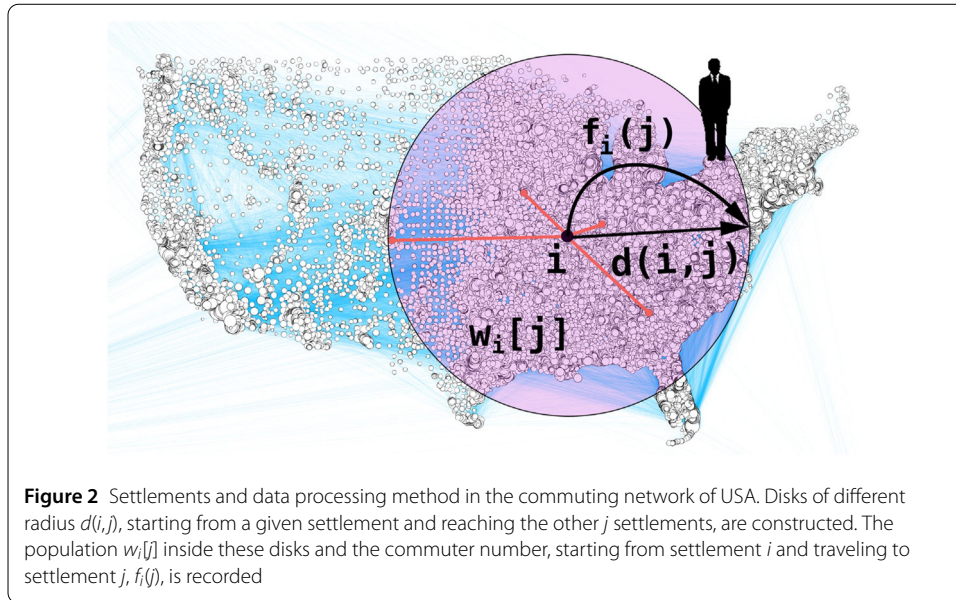
$$P_>(W)_{\text{FJM}} = \frac{1}{(\mu W + 1)^{(a-1)}}. \tag{16}$$

In the followings we demonstrate on real commuting data that the FJM model with the universal choice $a = 7/4$ offers a much improved fit for the real commuting data. For the specific case $a = 2$ one gets back the original radiation model. In principle the model is a two-parameter one, however if we admit the universality of a it becomes similarly with RM a one-parameter model.

3 Data source and format

For *USA* we processed a complete commuter and population database. We analyzed estimated population census data between 2006 and 2010 [20] using $Q = 73,803$ settlements (nodes) (white circles in Fig. 2) and 4,156,426 commuter routes (edges) (blue lines between white circles in Fig. 2). We use the same dataset as the one used in [21], where the authors attempted a region-like geographic division of USA based on commuting patterns. For studying the geographical population distribution we used a database from years between 2006 and 2010 giving the estimated population of continental USA divided in 11,078,286 cells of 1 km^2 area [22]. We detail now the three different data subsets that were constructed by us and are the input for our calculations:

a. *Settlement data* where the *settlement code* and their *latitudes* and *longitudes* are given. In the case of USA, the total number of settlements is $Q = 73,803$. These geographical locations are the source and targets for commuting. The data is in the format given below:

Figure 2 Settlements and data processing method in the commuting network of USA. Disks of different radius $d(i,j)$, starting from a given settlement and reaching the other j settlements, are constructed. The population $w_i[j]$ inside these disks and the commuter number, starting from settlement i and traveling to settlement j, $f_i(j)$, is recorded

set. code	lat	lon
1	32.4771763256	−86.4901731173
2	32.474292121	−86.4733798888
3	32.4754563613	−86.460168641

b. *Commuting data*, containing the source and targets for 4,156,426 directed travels to work. The data has the following structure: the first and second column contains the *source* and *target settlements* code and the third column gives the *number of commuters*. Below we illustrate the format of this data:

source set.	target set.	num. of com.
9719	9719	20,950
9703	9719	785
29,719	29,719	540
69,719	69,719	490
69,719	69,720	480
9711	9719	465

c. *Population distribution data*. The original dataset contains 11,078,286 square like cells of 1 km^2 area with its *population*, the *latitude* and *longitude* for the middle point. In order to speed up our calculations we have spatially renormalized this data and obtained a less accurate resolution with 4 km^2 size cells. This is done by collapsing the data of four neighbouring cells and averaging their latitudinal and longitudinal coordinates. As result we ended up with 1,230,920 cells containing a total population $W = 308{,}745{,}231$. The data we have worked with has the following structure:

pop. num.	lat	lon
18.0	51.8642065666	−176.664361722
30.0	51.8621521667	−176.6534376
9.0	51.8700427111	−176.644826767
0.0	51.8704367889	−176.633367733
7.0	51.8785901	−176.629460933
219.0	51.8383112778	−176.512803256

From the above three datasets one can compute the $P_>(W)$ dependency. For USA we have used yet another dataset to prove the linear proportionality between the number of job openings and total population for a geographical region. For this we obtained the number of listed jobs for each state of the continental USA using the site [23]. In the day we have processed the data (12.02.2018) we found a total of 2,596,391 jobs. The population of the states was obtained using the estimate between 2006 and 2011, available on the Internet [22].

Apart of the large-scale data available for USA we have used two smaller-size datasets for Hungary and Italy. These two additional datasets contain the same three data subsets: settlement data, commuting data and population distribution data. The population distribution data was used in its original form with cells of sizes 1 km^2.

For *Hungary* we used the same commuting data as in [11]. Commuting data is between Q = 3176 settlements, it contains 81,664 commuter routes [24] and the spatial distribution of population is for the W = 9,972,000 total inhabitants [25] as measured in the 2011 population census.

The data for *Italy* contains Q = 8093 settlements, 556,120 commuter routes and it is from the Italian population census from 2011 [26]. The total population W = 55,605,065 is mapped in cells of 1 km^2 area [27].

4 Data processing

During the data processing, we select one by one the settlements i as source for commuting and construct the disks with radius $d(i,j)$, reaching to the target settlement j. This is illustrated schematically in Fig. 2. We count the total population $w_i[j]$ inside this disk and record the number of commuters $f_i(j)$ starting from settlement i and traveling to settlement j.

Having the data $d(i,j)$, $f_i(j)$, and $w_i[j]$ for all the settlement pairs (i,j) we compute the experimental $P_>(W)$ probabilities.

The number of commuters that have their residence in settlement i are denoted by N_i.

$$N_i = \sum_{j=1}^{Q} f_i(j). \tag{17}$$

We ordered the settlements according to their distance relative to i. Let $h_i^{[k]}$ be the index of the settlement that is the kth one in this row (for example, $h_i^{[1]}$ is the index of the settlement that is the closest to settlement i and $h_i^{[2]}$ is the index of the settlement that is the second closest to i). We denote by $s(i,w)$ the smallest number of settlements for which the population inside a disk centered in i becomes larger (or equal) than w.
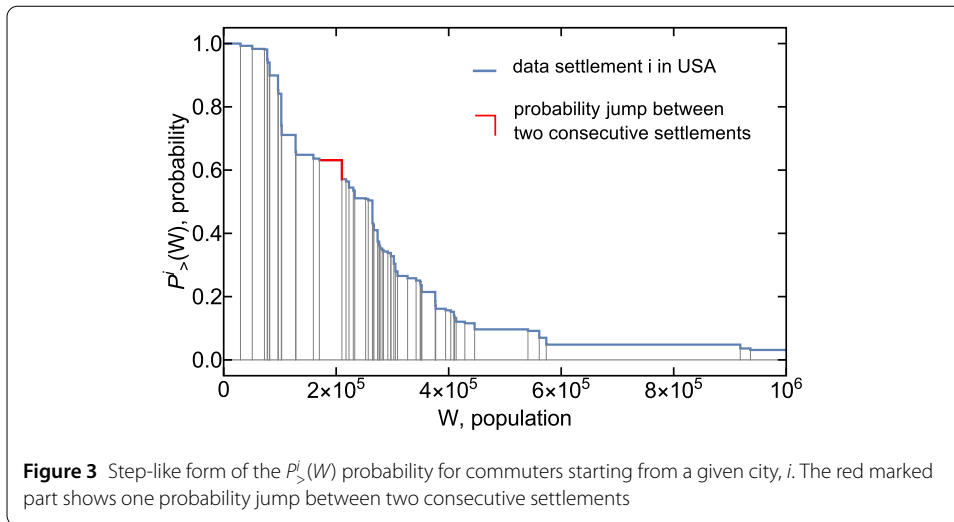
Mathematically:

$$s(i,w) \in \{1,2,3,\ldots,Q\} \tag{18}$$

and satisfies for any $w \le W$:

$$w_i\big[h_i^{[s(i,w)-1]}\big] < w \le w_i\big[h_i^{[s(i,w)]}\big]. \tag{19}$$

(We record here that Q denotes the total number of settlements and W is the total population in the studied territory.) If no such number exists, then we will consider $s(i,w) = Q$.

Figure 3 Step-like form of the $P_>^i(W)$ probability for commuters starting from a given city, *i*. The red marked part shows one probability jump between two consecutive settlements

The probability that commuters from *i* are a transiting a disk with population *W* inside it can be written as:

$$P_>^i(W) = \frac{1}{N_i} \left[N_i - \sum_{k=1}^{s(i,W)} f_i\big(h_i^{[k]}\big) \right]. \tag{20}$$

Due to the discrete nature of the settlements this has a step-like structure for a given commuting source *i*, as it is illustrated on Fig. 3 for a given town in USA.

After obtaining these probabilities for all settlements, we constructed the desired $P_>(W)$ probability by averaging over all *i* settlements:

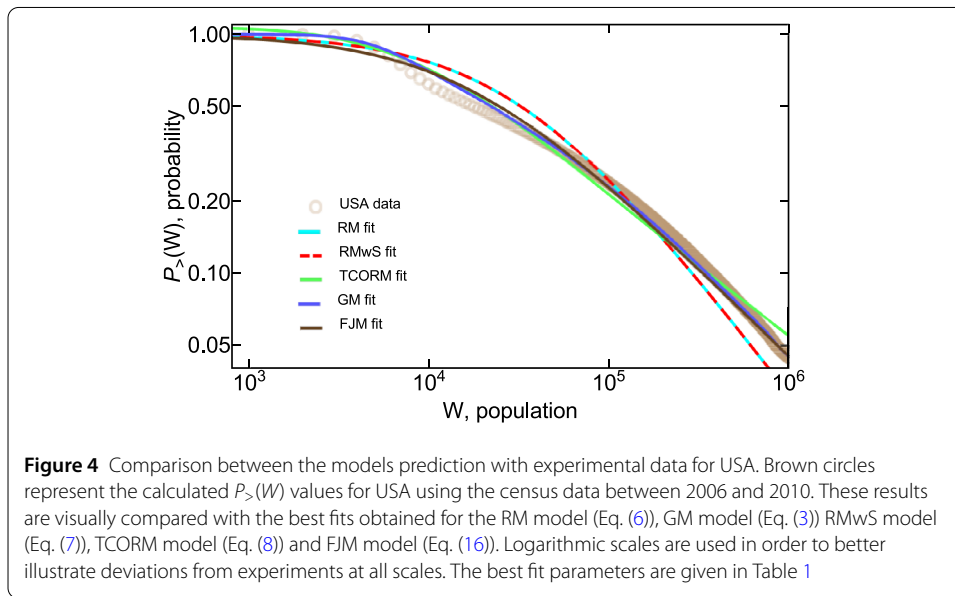$$P_>(W) = \frac{1}{Q} \sum_{i=1}^{Q} P_>^i(W). \tag{21}$$

As expected, averaging will result in a smoother curve, showing the experimental trend for the probability that is in our focus.

5 Results and discussions

We show first the experimentally obtained $P_>(W)$ probability for the large USA dataset in comparison with the best fit results obtained from the GM model (3), the original RM model (6), the RMwS model (7), the TCORM model (8) and our novel FJM model (16). In the FJM model we have fixed the $a = 7/4$ parameter for all studied datasets, so in principle the only free parameter of this model is μ. Boundary effects become important for large W values (the disks centred on the settlements become largely incomplete due to the fact that they extend over the borders of continental USA). To minimize these effects we considered the data only up to $W_{\max} = 1{,}000{,}000$. Also, to eliminate very short commuting routes (where commuting is questionable) we have imposed a lower threshold of $W_{\min} = 1000$. Fitting was realized on the $[W_{\min}, W_{\max}]$ interval using the nonlinear fitting features of the Wolfram Mathematica® software. For the GM model, equation (3) does not lead to a compact functional form, so fitting was realised by considering a progressive mesh method for various α and β values in the interval $\alpha \in [-1.0, 2.5]$ and $\beta \in [-1.0, 2.5]$. The best fit

Table 1 Fitting parameters for the USA commuting data, considering the models given by equations (6), (7), (8), (3) and (16). The R^2 values and fitted curves suggests that the FJM and GM models performs better than the simple RM, RMwS, or TCORM model. For the FJM model $a = 7/4$ was fixed

	RM	RMwS		TCORM		GM		FJM
	μ	μ	q	μ	λ	α	β	μ
USA data	0.0000308	0.0000308	1.0	0.000119	0.0056	1.2	1.2	0.000062
R^2	0.971	0.971		0.992		0.993		0.993

Figure 4 Comparison between the models prediction with experimental data for USA. Brown circles represent the calculated $P_>(W)$ values for USA using the census data between 2006 and 2010. These results are visually compared with the best fits obtained for the RM model (Eq. (6)), GM model (Eq. (3)) RMwS model (Eq. (7)), TCORM model (Eq. (8)) and FJM model (Eq. (16)). Logarithmic scales are used in order to better illustrate deviations from experiments at all scales. The best fit parameters are given in Table 1

parameters and the goodness of the fits (R^2 correlation coefficient) are summarized in Table 1.

The statistics is in favour of the FJM and GM model. The best fits drawn on Fig. 4 suggests visually the same conclusion. The fact that FJM over performs the approximation given by the RM model for $a = 7/4$ is not surprising, since it has one more parameter: a. We will show however that one can fix this parameter and get also an excellent fit on other datasets as well (Italy and Hungary).The clear improvement in fitting the data relative to the RMwS and TCORM models is however a great leap forward since these models offer a two-parameter fit. It is important to notice the fact that *GM* offers also a good fit. This is again a two-parameter fit, but we will show in the followings on other datasets, that one cannot fix any of these parameters and remain with a fit quality that is comparable with FJM.

For the sake of completeness we also show for USA that our hypothesis, according to which the number of job openings in a geographical region is linearly proportional with the population. On Fig. 5 we plot the total number of job openings for different states as a function of the population of the state. The straight-line trend confirms our hypothesis.

The FJM model for $a = 7/4$ works well also for the commuting data processed for Hungary and Italy. The goodness of the fits and the best fit parameters are shown in Table 2. The visual picture for $P_>(W)$ and best fits offered by the models are plotted in Fig. 6.

We show the same results also for the GM model. The obtained best fit parameters are given in Table 3 and the results are plotted on Fig. 7. We learn form here that GM offers a good description also for Hungary but it fails for the Italy data, and suggests that one
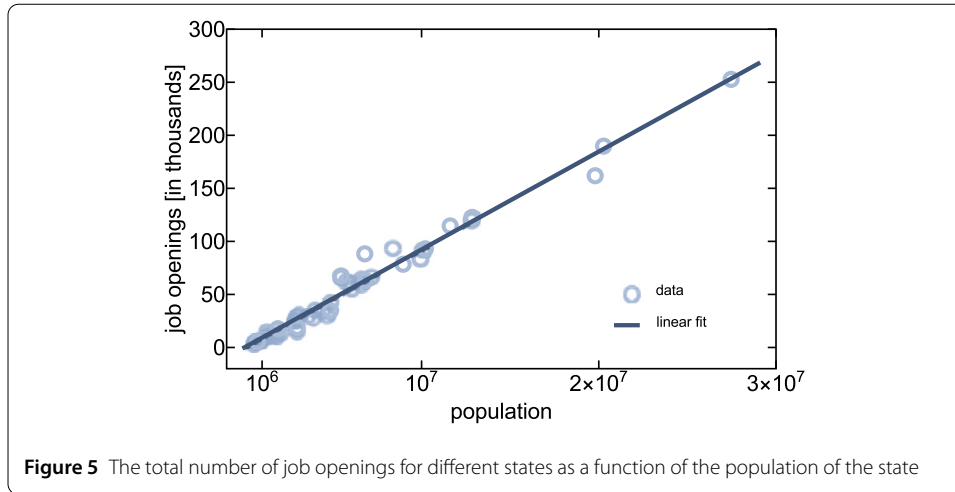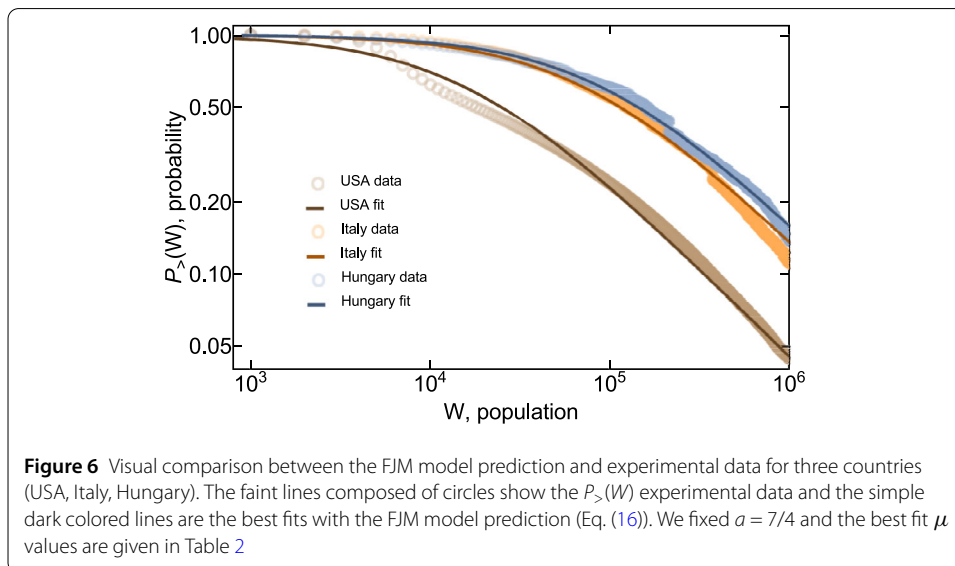
Figure 5 The total number of job openings for different states as a function of the population of the state

Table 2 Fitting parameters and goodness of the fits shown in Fig. 6, considering the functional form given by equation (16) and fixing $a = 7/4$

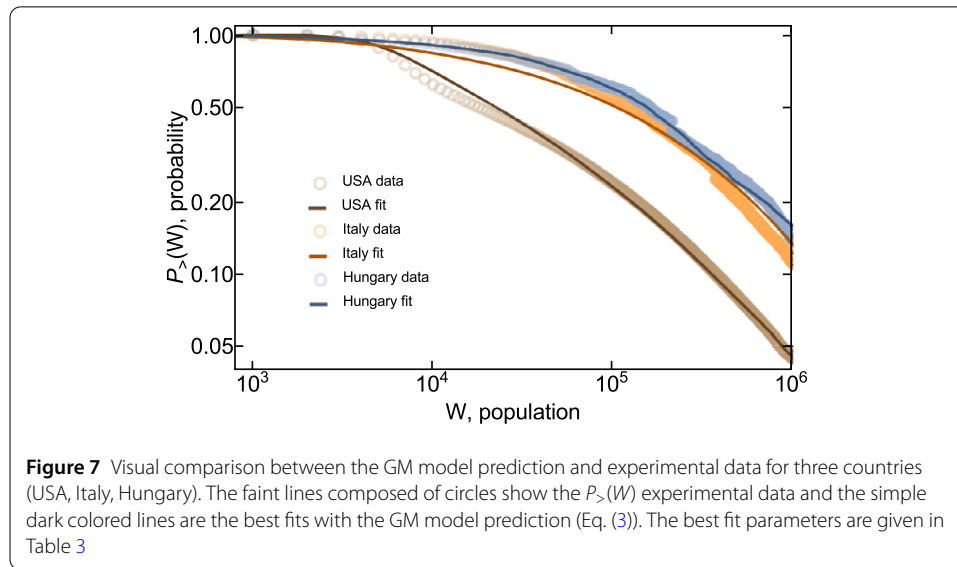|	USA	Italy	Hungary
μ	0.000062	0.000013	0.000011
R^2	0.993	0.997	0.998

Figure 6 Visual comparison between the FJM model prediction and experimental data for three countries (USA, Italy, Hungary). The faint lines composed of circles show the $P_>(W)$ experimental data and the simple dark colored lines are the best fits with the FJM model prediction (Eq. (16)). We fixed $a = 7/4$ and the best fit μ values are given in Table 2

cannot consider a universal value for α and β so that all datasets are reasonable well fitted. The negative value obtained for the α is more than strange, and suggests again, that the GM model is seemingly not appropriate for fitting the Italian commuting data.

6 Conclusions

In order to describe the statistics of commuter fluxes at different distances we introduced the FJM model based on a mean-field type dynamical approach. The model takes into account indirectly that commuting to larger distances is costly and less probable. Relative to the classical models it offers an improved fit for commuter fluxes in USA, Hungary

Table 3 Fitting parameters and goodness of the fits shown in Fig. 7, considering the functional form given by the GM model (Eq. (3))

	USA	Italy	Hungary
α	1.2	−0.4	1.4
β	1.2	0.3	1.6
R^2	0.993	0.96	0.996

Figure 7 Visual comparison between the GM model prediction and experimental data for three countries (USA, Italy, Hungary). The faint lines composed of circles show the $P_>(W)$ experimental data and the simple dark colored lines are the best fits with the GM model prediction (Eq. (3)). The best fit parameters are given in Table 3

and Italy. The probability that commuters are traveling for their jobs over a population W is compactly given by equation (16). The model is a two-parameter one, although we have shown that one parameter can be fixed, so that all studied datasets are reasonable well explained. In such sense the model becomes similarly with the RM model a one-parameter one, and improves the RM model in a considerable manner.

In order to comment on the results obtained for USA, Italy and Hungary we review from Table 2 the best fit parameter μ obtained with the FJM model. The parameter μ characterises both the availability of jobs per population and the attractiveness of these jobs to jobseekers. A higher value of μ suggests that there are many jobs relative to the population, jobseekers are aware of them and consider them for a potential commuting. A smaller μ value suggests that the number of available jobs per population is smaller and jobseekers are very selective for commuting. The obtained fitting parameter for μ are in good agreement with the given heuristic justifications and confirms the known social and economic profile of USA, Italy and Hungary. Commuting is more common in USA relative to Europe and there are more available commuting jobs per population. Related to the value of the a exponent in equation (16), one can also draw some interesting conclusions. The difference from the original radiation model (where we have $a = 2$) suggests an already known issue, i.e. commuters are selective, not all available jobs are acceptable for them and travel cost has to be taken into account in accepting a commuting job [7–11]. Due to this the C/η value is smaller than the one for a simple salary optimization mechanism where the commuters accept the closest job that improves their salary at home (assumption of RM). This can be done either by lowering the C constant or by increasing the value of η,

or changing both of them. The seemingly universal value of $a = 7/4$ remains however a puzzle motivating further studies.

In conclusion, we believe that the FJM model proposed in the present study lies on simple and reasonable assumptions and the studied experimental data supports it's predictions.

List of abbreviations
GM, Gravity Model; RM, Radiation Model; RMwS, Radiation Model with Selection; TCORM, Travel Cost Optimized Radiation Model; FJM, Flow and Jump Model.

Availability of data and materials
The used data is available on the Internet by following the links from the "Data source and format" and "Data processing" sections. The processed data in the format indicated in the "Data source and format" section is available in the Figshare repository, doi:10.6084/m9.figshare.6151130, URL: https://figshare.com/s/b86965bb06ce018f52bf. In this repository one will find: the figs_data.zip file containing the data used for plotting the Figures, the Hungary.zip, Italy.zip and USA.zip files containing the processed data for Hungary, Italy and USA, respectively. The fig2.gdf file is a GUESS Graph Data Format file, which is editable in a simple text editor.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
ZN designed the study, elaborated the FJM model and wrote up the first version of the manuscript. LV analyzed the data and draw the figures. GT collected the data, interpreted them and putted in the desired form. All authors worked on the final version of the manuscript. All authors read and approved the final manuscript.

Authors' information
Z.N. is professor of theoretical physics, working in the area of interdisciplinary applications of statistical physics. He uses both analytical and computational models to understand complex phenomena from physics, economics, biology and sociology. G.T. is a senior investigator at the Hungarian Central Statistical Office. He is specialist in geographical data collection, in handling and processing large geographical datasets. L.V. is a PhD student in computational physics with a strong background in computer science and informatics.

Author details
[1]Department of Physics, Babeş-Bolyai University, Cluj-Napoca, Romania. [2]Hungarian Central Statistical Office, Budapest, Hungary.

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

References
1. Barbosa-Filho H, Barthélemy M, Ghoshal G, James RC, Lenormand M, Louail T, Menezes R, Ramasco JJ, Simini F, Tomasini M (2018) Human mobility: Models and Applications. Physics Reports. https://doi.org/10.1016/j.physrep.2018.01.001
2. Stouffer SA (1940) Intervening opportunities: a theory relating mobility and distance. Am Sociol Rev 5:845–867
3. Block H, Marschak J (1960) Random orderings and stochastic theories of responses. In: Contributions to probability and statistics, vol 2 pp 97–132
4. Lukermann F, Porter PW (1960) Gravity and potential models in economic geography. Ann Assoc Am Geogr 50:493–504
5. Simini F, González MC, Maritan A, Barabási AL (2012) A universal model for mobility and migration patterns. Nature 484:96–100
6. Yan XY, Zhao C, Fan Y, Di Z, Wang WX (2014) Universal predictability of mobility patterns in cities. J R Soc Interface 11:20140834
7. Yan XY, Wang WX, Gao ZY, Lai YC (2017) Universal model of individual and population mobility on diverse spatial scales. Nat Commun 8:1639
8. Liu E, Yan XY New parameter-free mobility model. Preprint. arXiv:1808.06363
9. Simini F, Maritan A, Néda Z (2013) Human mobility in a continuum approach. PLoS ONE 8(3):e60069
10. Ren Y, Ercsey-Ravasz M, Wang P, Gonzales MC, Toroczkai Z (2014) Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. Nat Commun 5:5347

11. Varga L, Tóth G, Néda Z (2017) An improved radiation model and its applicability for understanding commuting patterns in Hungary. Reg Statist 6(2):27–38
12. Stefanouli M, Polyzos S (2017) Gravity vs radiation model: two approaches on commuting in Greece. Transp Res Proc 24:65–72
13. Wilson AG (1967) A statistical theory of spatial distribution models. Transp Res 1:253–269
14. Hua CI, Porell F (1979) A critical review of the development of the gravity model. Int Reg Sci Rev 4(2):97–126
15. Sheppard ES (1978) Theoretical underpinnings of the gravity hypothesis. Geogr Anal 10(4):386–402
16. Niedercorn JH, Bechdolt BV (1969) An economic derivation of the "gravity law" of spatial interaction. J Regional Sci 9(2):273–282
17. Domencich T, McFadden DL (2015) Urban travel demand: a behavioral analysis. North-Holland, Amsterdam
18. Biró TS, Néda Z (2018) Unidirectional random growth with resetting. Physica A 499:355–361
19. Thurner S, Kyriakopoulos F, Tallis C (2007) Unified model for network dynamics exhibiting nonextensive statistics. Phys Rev E 76:036111
20. CTPP 2006–2010 Census Tract Flows, Commuting data, American Community Survey. https://www.fhwa.dot.gov/planning/census_issues/ctpp/data_products/2006-2010_tract_flows/
21. Dash Nelson G, Rae A (2016) An economic geography of the United States: from commutes to megaregions. PLoS ONE 11(11):e0166083
22. 2006–2010 Population distribution, American Community Survey. https://www.census.gov/geo/maps-data/data/tiger-data.html
23. 2018 USA job openings accessed at 10.02.2018. https://www.indeed.com/
24. 2011 Census Tract Flow, Commuting data, Hungary. http://www.ksh.hu
25. 2011 Population distribution, Hungary. http://ec.europa.eu/eurostat/cache/GISCO/geodatafiles/GEOSTAT-grid-POP-1K-2011-V2-0-1.zip
26. 2011 Census Tract Flow, Commuting data, Italy. http://www.istat.it/storage/cartografia/matrici_pendolarismo/matrici_pendolarismo_2011.zip
27. 2011 Population distribution, Italy. http://ec.europa.eu/eurostat/cache/GISCO/geodatafiles/GEOSTAT-grid-POP-1K-2011-V2-0-1.zip