

URBAN FINDINGS

Modeling Urban Morphology by Unifying Diffusion-Limited Aggregation and Stochastic Gravitation

Diego Rybski¹ ^a, Yunfei Li² , Stefan Born³ , Jürgen P. Kropp² 

¹ Potsdam Institute for Climate Impact Research (PIK), Member of Leibniz Association, P.O. Box 60 12 03, D-14412 Potsdam, Germany; Department of Environmental Science Policy and Management, University of California Berkeley, 130 Mulford Hall #3114, Berkeley, CA 94720, USA; Complexity Science Hub Vienna, Josefstädterstrasse 39, A-1090 Vienna, Austria, ² Potsdam Institute for Climate Impact Research (PIK), Member of Leibniz Association, P.O. Box 60 12 03, D-14412 Potsdam, Germany; University of Potsdam, Institute of Earth and Environmental Science, Potsdam, Germany, ³ Technische Universität Berlin, Chair of Bioprocess Engineering and Institute of Mathematics, Strasse des 17. Juni 135, 10623 Berlin, Germany

Keywords: urban morphology, diffusion-limited aggregation, stochastic gravitation model, zipf's law for cities

<https://doi.org/10.32866/001c.22296>

Findings

More than 30 years ago, Diffusion-Limited Aggregation (DLA) has been studied as mechanism to generate structures sharing similarities with real-world cities. Recently, a stochastic gravitation model (SGM) has been proposed for the same purpose but representing a completely different mechanism. Both approaches have advantages and disadvantages, while e.g. the dendrites emerging via DLA are visually very different from real-world cities, the SGM does not preserve undeveloped locations close to the city center. Here we propose a unification of both mechanisms, i.e. a particle moves randomly and turns into an urban site with a probability that depends on the proximity to already developed sites. We study the cluster size distributions of the structures generated by both models and find that SGM generates more balanced distributions. We also propose a way to assess to which extent the largest cluster is a primate city and find that in both models, beyond certain parameter value, the size of the largest cluster becomes inconsistent with being drawn from the same distribution of remaining clusters.

Various models have been proposed to generate structures that share similarities with real-world cities. The correlated percolation model combines an exponential radial gradient with correlated random numbers (Makse, Havlin, and Stanley 1995). However, the largest city is too big and consequently omitted when Zipf's law for cities, i.e. a power-law size distribution with specific exponent, is validated. The model by Schweitzer & Steinbrink (1998) is based on two processes, the emergence of new clusters and cluster growth. In this model a similar problem is encountered – due to coagulation the largest cluster dominates the growth process. The authors avoid this by excluding the largest cluster from growth.

We consider the Stochastic Gravitation Model (SGM) (Rybski, García Cantú Ros, and Kropp 2013; Li, Rybski, and Kropp 2021; Rybski and Li 2021). In this approach, the probability that a site i is converted from non-urban ($w_i = 0$) to urban ($w_i = 1$) is given by

$$q_i \sim \frac{\sum w_j d_{i,j}^{-\gamma}}{\sum d_{i,j}^{-\gamma}} \quad (1)$$

where γ is the main parameter and $d_{i,j}$ is the Euclidean distance between sites i and j . The structures generated by the SGM also exhibit a largest cluster that is too large compared to the smaller ones.

The exponent γ characterizes what is also known as “friction of distance” or “friction of space” (Cliff, Martin, and Ord 1974). Couclelis (1996) describes a “[...] ‘glue’ holding cities together [...] to overcome the friction of distance for the purpose of efficient communication [...]”. In economic terms, distance translates into costs which are to be minimized (Friedmann 1956). Small γ allows new urban areas to be seeded further away and they do not have to be adjacent to old ones.

Here we propose a new model representing a combination of the SGM with Diffusion Limited Aggregation (DLA) (Witten Jr. and Sander 1981) which has previously been studied in the urban context (Fotheringham, Batty, and Longley 1989; Batty, Longley, and Fotheringham 1989)¹. The new model, which we call *Diffusion Limited Gravitation* (DLG), generates structures which visually seem less scattered than those from the SGM. Accordingly, we investigate to which extent the largest cluster outgrows the size distribution.

DLA starts with a seed “particle” at the center of the grid. A new particle performs a random walk starting far from the center. The particle stops once it reaches a cell adjacent to an occupied one. Then the procedure is repeated with a new particle.

We propose to combine the SGM with DLA. The SGM takes place on a square grid. As initial condition we use one occupied cell at the center of the system. The modeling is done according to the following iterative steps.

1. Calculate q_i according to Eq. (1) and normalize so that $\max(q) = 1$.
2. A random walker starts from an arbitrary cell (can be close to the origin; it can cross urban cells).
3. The random walker stops with probability q_i from Eq. (1). I.e. it stops at site i if a random number $[0, 1]$ is smaller than q_i . The cell is converted to urban, $w_i = 1$.
4. Repeat from step 1.

Analogous to the original SGM, also DLG involves the γ parameter which determines how strongly the probability of urbanization drops with the distance. Accordingly, we run the models with different parameter values, i.e. 10 realizations for each. For DLG we use 13 γ -values within

¹ For completeness, we would like to mention the Dielectric Breakdown Model (Niemeyer, Pietronero, and Wiesmann 1984) that was also used for generating urban forms (Batty and Longley 1988; Batty 1991).

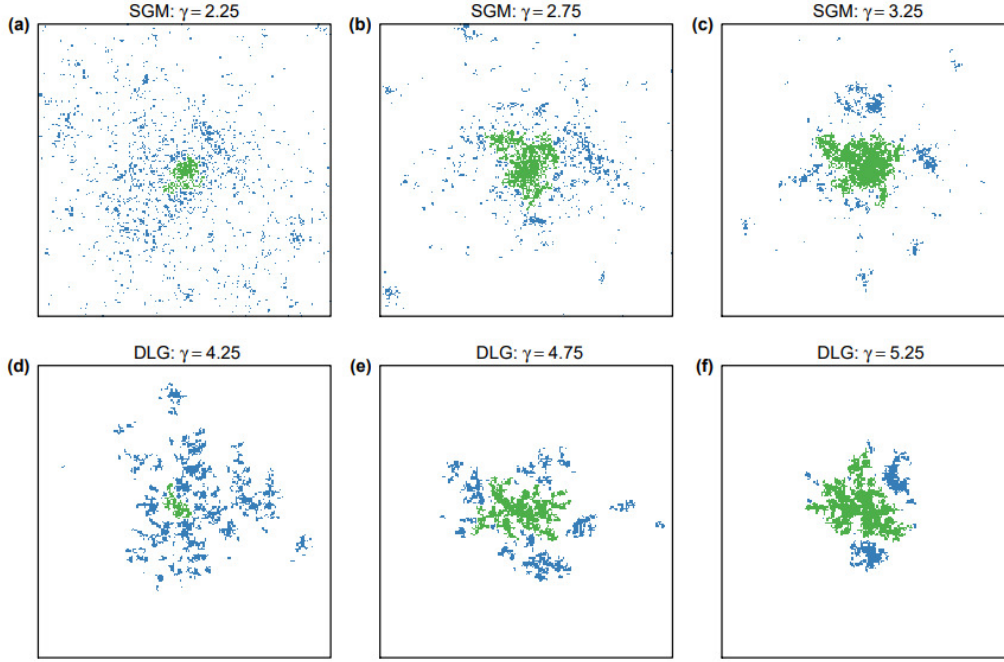


Figure 1. Examples of structures generated with (a)-(c) the SGM and (d)-(f) DLG for different values of the model parameters indicated on top of the panels

In all cases 2,000 urban pixels (5 %) are shown. Green color indicates the central clusters while blue color indicates other occupied pixels.

{4.00, 4.25, 4.50, ..., 7.00}. As in each iteration one pixel is added, the number of pixels corresponds to the number of iterations. For SGM the γ -values are {2.00, 2.25, 2.50, ..., 5.00}. Additionally, for the SGM a normalization factor C was used to ensure $\sum Cq_i = 10$, which implies that the number of pixels approximately increases by 10 during each iteration. We store the model output after every 10th (DLG) or every iteration (SGM). We use a system size of 200×200 pixels.

[Figure 1](#) shows illustrative examples of both models. It can be seen that small γ lead to more scattered and a larger number of small clusters. Vice versa, with increasing γ the largest cluster becomes dominant. Moreover, for the DLG some of the dendrite structure, similar to the original DLA, can be seen, e.g. [Figure 1\(e\)](#).

As detailed above, previously it was found that the largest cluster outgrows the other ones in various models. [Figure 1\(d\)](#) suggests a more balanced result, i.e. the largest cluster is relatively small compared to abundant smaller clusters. Accordingly, in the following we want to analyze if DLG generates Zipf-distributed clusters and to what extent the largest cluster outgrows.

Zipf's law for cities (Auerbach 1913; Zipf [1949] 2012; Berry and Okulicz-Kozaryn 2012; Rybski 2013; Nitsch 2005; Soo 2005; Cottineau 2017; Rozenfeld et al. 2011; Ribeiro et al. 2021) can be written as

$$p(A) \sim A^{-\zeta}, \quad (2)$$

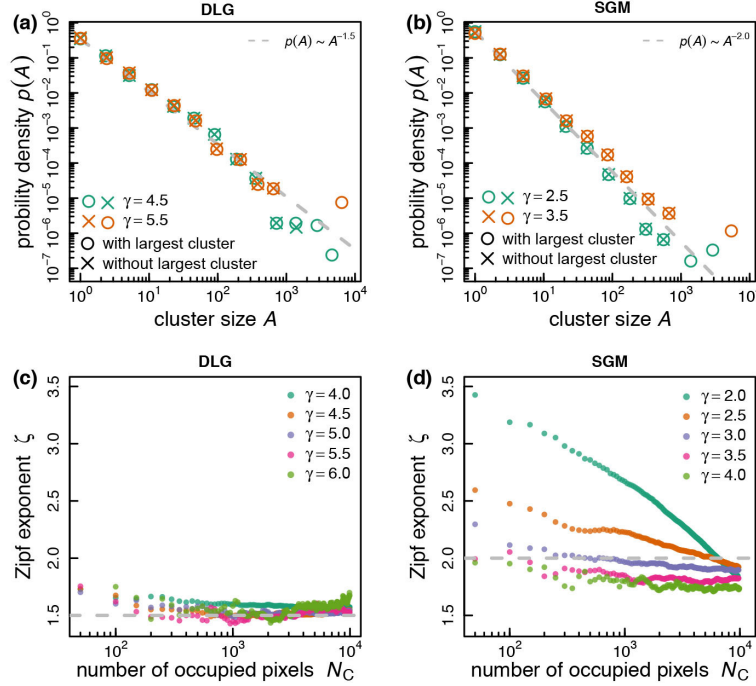


Figure 2. Examples of probability density distributions and Zipf exponent vs. number of urban cells for both models

Panels (a) and (b) show $p(A)$ of the cluster sizes A with (circles) and without (crosses) largest clusters for two γ -values when the number of occupied cells is approximately 7,000. Panels (c) and (d) show how the Zipf exponent ζ depends on the number of urban cells when it is estimated without the largest clusters.

where $p(\cdot)$ is the probability density function (pdf) and A represents the cluster areas (given by their number of urban cells). We use the method proposed by Clauset et al. (2009) to estimate the Zipf exponent ζ (R package “powerlaw” [Gillespie 2015]). Moreover, we fix the lower bound for the estimation of the exponent to 1. In order to have better statistics we combine the output of 10 realizations with the same or similar number of urban pixels. Please note that ζ differs by 1 from the Zipf exponent when it is obtained from a rank-size plot $r \sim A^{-(\zeta-1)}$.

For comparison, we also analyze the distribution of a set where the largest cluster of each realization has been removed. We reason that if the estimated Zipf exponents for the sets with and without largest cluster are similar, then the largest cluster is consistent with being drawn from the same distribution. If the exponents differ, then it is not consistent.

Examples of empirically estimated pdfs are shown in [Figure 2 \(a\) and \(b\)](#). Each panel displays the estimated pdfs for 2 γ -values and with/without largest clusters. Two observations can be made. First, in both cases the largest cluster seems to be dominant for larger γ . Second, DLG leads to a rather small $\zeta \approx 1.5$ and for the SGM the exponent seems to depend on γ .

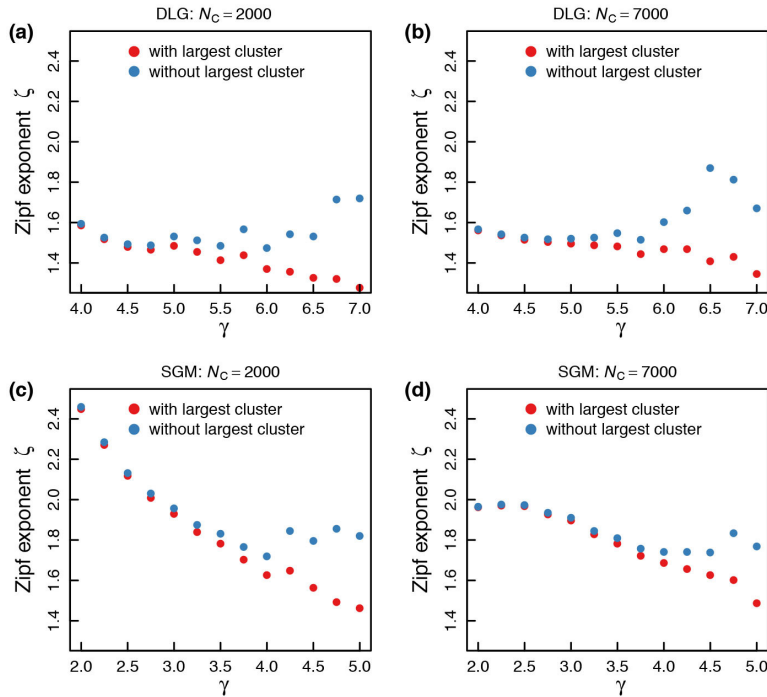


Figure 3. Dependence of the Zipf exponent on the model parameter and influence of in- or excluding the largest clusters

The ζ -exponent is plotted as a function of γ , where the former was estimated with- and without largest clusters. Panels (a) and (b): DLG, panels (c) and (d): SGM. Panels (a) and (c): 2,000 urban cells, panels (b) and (d): 7,000 urban cells. One can see that beyond certain value of the model parameter γ the size of the largest cluster becomes inconsistent with being drawn from the same distribution.

In [Figure 2 \(c\) and \(d\)](#) the resulting ζ -exponents are plotted as a function of the number of urban cells for DLG and the SGM, respectively. We can see that in case of the SGM there can be a strong dependence (as previously reported [Rybski, García Cantú Ros, and Kropp 2013]), which in case of DLG is much less pronounced. Moreover, for the SGM ζ is found around 2 which is consistent with real-world data but for DLG, ζ is close to 1.5 – a more balanced distribution which is rarely found for real cities.

Finally, we study the resulting distributions with and without largest cluster as a function of the model parameter γ . [Figure 3](#) shows the results for 5% and 17.5% urban cells. We find that the estimated exponents – in- or excluding the largest cluster – are similar for small γ but diverge for large γ . In all cases, for small γ the largest cluster is consistent with Eq. (2) but for larger γ it generally outgrows. It is important to note that the ζ estimates for very large γ are less reliable because they are based on critically small sample sizes.

In summary: (i) We propose a new model (DLG) that represents a combination of the well-known DLA with the more recent SGM. DLG is only one possibility and there are other ways of combining DLA and the SGM that can be investigated in future research. Perspectively, SGM could represent a new approach to spatial modeling of human settlements (Strano et al. 2020). (ii) We propose an approach to assess to what extent the largest cluster is consistent with being drawn from the same distribution. From a

methodological point of view, more rigorous statistical testing will be necessary, when e.g. the existence of *primate cities* in real-world data is to be assessed (Linsky 1965). (iii) We find that for both models the largest cluster is compatible with the rest of the distribution if γ is small. For large values it outgrows.

Funding

D. Rybski thanks the Alexander von Humboldt Foundation for financial support under the Feodor Lynen Fellowship. This work was supported by the German Federal Ministry of Education and Research through the Program “International Future Labs for Artificial Intelligence” (Grant number 01DD20002A).

Submitted: March 09, 2021 AEST, Accepted: April 20, 2021 AEST



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for more information.

REFERENCES

- Auerbach, F. 1913. "Das Gesetz der Bevölkerungskonzentration." *Petermanns Geogr. Mitteilungen* 59: 73.
- Batty, M. 1991. "Generating Urban Forms from Diffusive Growth." *Environment and Planning A: Economy and Space* 23 (4): 511–44. <https://doi.org/10.1068/a230511>.
- Batty, M, and P. A. Longley. 1988. "The Morphology of Urban Land Use." *Environment and Planning B: Planning and Design* 15 (4): 461–88. <https://doi.org/10.1068/b150461>.
- Batty, M., P. Longley, and S. Fotheringham. 1989. "Urban Growth and Form: Scaling, Fractal Geometry, and Diffusion-Limited Aggregation." *Environ. Plan. A* 21: 1447.
- Berry, B. J. L., and A. Okulicz-Kozaryn. 2012. "The City Size Distribution Debate: Resolution for US Urban Regions and Megalopolitan Areas." *Cities* 29: S17.
- Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman. 2009. "Power-Law Distributions in Empirical Data." *SIAM Review* 51 (4): 661–703. <https://doi.org/10.1137/070710111>.
- Cliff, A. D., R. L. Martin, and J. K. Ord. 1974. "Evaluating the Friction of Distance Parameter in Gravity Models." *Regional Studies* 8: 281.
- Cottineau, Clémentine. 2017. "MetaZipf. A Dynamic Meta-Analysis of City Size Distributions." Edited by Wei-Xing Zhou. *PLoS One* 12 (8): e0183919. <https://doi.org/10.1371/journal.pone.0183919>.
- Couclelis, H. 1996. "The Death of Distance." *Environment and Planning B: Planning and Design* 23 (4): 387–89. <https://doi.org/10.1068/b230387>.
- Fotheringham, A. S., M. Batty, and P. A. Longley. 1989. "Diffusion-Limited Aggregation and the Fractal Nature of Urban Growth." *Papers in Regional Science* 67: 55.
- Friedmann, J. R. P. 1956. "Locational Aspects of Economic Development." *Land Econ* 32: 213.
- Gillespie, C. S. 2015. "Fitting Heavy Tailed Distributions: The Powerlaw Package." *Journal of Statistical Software* 64 (2): 1–16.
- Li, Yunfei, Diego Rybski, and Jürgen P. Kropp. 2021. "Singularity Cities." *Environment and Planning B: Urban Analytics and City Science* 48 (1): 43–59. <https://doi.org/10.1177/2399808319843534>.
- Linsky, A. S. 1965. "Some Generalizations Concerning Primate Cities." *Annals of the American Association of Geographers* 55: 506.
- Makse, Hernán A., Shlomo Havlin, and H. Eugene Stanley. 1995. "Modeling Urban Growth Patterns." *Nature* 377 (6550): 608–12. <https://doi.org/10.1038/377608a0>.
- Niemeyer, L., L. Pietronero, and H. J. Wiesmann. 1984. "Fractal Dimension of Dielectric Breakdown." *Physical Review Letters* 52: 1033.
- Nitsch, V. 2005. "Zipf Zipped." *Journal of Urban Economics* 57: 86–100.
- Ribeiro, Haroldo V., Milena Oehlers, Ana I. Moreno-Monroy, Jürgen P. Kropp, and Diego Rybski. 2021. "Association between Population Distribution and Urban Gdp Scaling." *PLoS One* 16 (1): e0245771. <https://doi.org/10.1371/journal.pone.0245771>.

Rozenfeld, H. D., D. Rybski, X. Gabaix, and H. A. Makse. 2011. "The Area and Population of Cities: New Insights from a Different Perspective on Cities." *American Economic Review* 101: 2205.

Rybski, Diego. 2013. "Auerbach's Legacy." *Environment and Planning A: Economy and Space* 45 (6): 1266–68. <https://doi.org/10.1068/a4678>.

Rybski, Diego, Anselmo García Cantú Ros, and Jürgen P. Kropp. 2013. "Distance-Weighted City Growth." *Physical Review E* 87 (4): 042114. <https://doi.org/10.1103/physreve.87.042114>.

Rybski, Diego, and Y. Li. 2021. "Comparing Power-Laws and Exponentials in Simulations of Gravitational Growth." In *Handbook on Entropy, Complexity and Spatial Dynamics. the Rebirth of Theory? (Forth-Coming)*. Edward Elgar.

Schweitzer, F., and J. Steinbrink. 1998. "Estimation of Megacity Growth: Simple Rules versus Complex Phenomena." *Applied Geography* 18: 69.

Soo, K.T. 2005. "Zipf's Law for Cities: A Cross-Country Investigation." *Regional Science and Urban Economics* 35: 239.

Strano, E., F. Simini, M. De Nadai, T. Esch, and M. Marconcini. 2020. "Precise Mapping, Spatial Structure and Classification of All the Human Settlements on Earth." arXiv:2006.06584 [physics.soc-ph].

Witten Jr., T. A., and L. M. Sander. 1981. "Diffusion-Limited Aggregation, a Kinetic Critical Phenomenon." *Physical Review Letters* 47: 1400.

Zipf, G. K. (1949) 2012. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology*. Manfield Centre, CT: Martino Publishing.