How heavy-tailed is the distribution of global cargo ship traffic?

Michael T. Gastner*† and César Ducruet‡
*Institute of Technical Physics and Materials Science, Research Centre for Natural Sciences, Hungarian Academy of Sciences, P.O. Box 49, H-1525 Budapest, Hungary, Email: mgastner@gmail.com
†Department of Engineering Mathematics, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, United Kingdom
‡CNRS / University of Paris-I Panthéon Sorbonne, UMR 8504 Géographie-Cités / P.A.R.I.S., 13 rue du Four, F-75006 Paris, France, Email: cdu@parisgeo.cnrs.fr

***Abstract*—Power laws, once believed to be a universal feature of degree distributions in complex networks, have come under attack in recent years. More sophisticated statistical analysis has often revealed other heavy-tailed distributions as more adequate descriptions of real-world data. Here we study degree and strength distributions of the network of worldwide cargo ship movements – the main transport network for world trade – from 14 different years between 1890 and 2008. We compare the Akaike information criterion of various common probabilistic models. In almost all cases, the Akaike weights identify a stretched exponential distribution as the most likely among the investigated models. Simple or truncated power laws, by contrast, do not capture the observations equally well. Cargo ship traffic is thus heavy-tailed with some ports being significantly busier than the average, but the distribution is not scale-free. The maximum-likelihood estimators indicate that the normalized distribution became increasingly shorter-tailed for one century. However, since the start of this millennium this trend appears to be reversing.**

I. Introduction

A great number of recent papers on complex networks have investigated the topological properties of technological networks, many of them falling into the category of so-called "spatial networks" [1]. In particular, scholars have revealed features of spatial networks that differ from the more generic scale-free and small-world network models: a higher clustering coefficient due to the importance of proximity [2], fewer global hubs and more numerous regional hubs (e.g. in air transport [3]), resulting in a lower vulnerability to targeted attacks (see [4] for a review on the relationship between spatial science and network science).

In this study, we focus on the degree and strength distribution of a particularly important spatial network that hauls the majority of world trade [5]: the network of cargo ship movements [6], [7], [8], [9]. In this network, nodes are ports and links are nonstop connections. The degree of a node is defined as the number of ports that the node is connected to by at least one arriving or departing ship. Strength refers to the total number of ship arrivals and departures. The degree and strength distributions are two summary statistics of a weighted network that do not allow a complete reconstruction of traffic on the links. However, these distributions are an important feature of a network's topology and have often been used as circumstantial evidence for mechanistic models of the network's evolution [10], [11], [12]. The strength distribution also plays a crucial role for predicting the full origin-destination matrix (i.e. the traffic between all pairs of ports) because it is an input in transport forecasting (e.g. in the gravity model or intervening opportunities model [13]).

In the early phase of complex network science, many degree distributions of social, technological and biological networks were investigated and characterized as power laws [10]. In this interpretation, the observed distribution is a realization of a probabilistic model that assigns the probability $\Pr(k)$ to the event that an arbitrary node has degree or strength k so that

$$\Pr(k;\tau) = \frac{1}{\zeta(\tau)} k^{-\tau},\ k > 0. \qquad (1)$$

Here $\zeta(\tau)$ is the Riemann zeta function and $\tau > 1$ a fixed parameter that has to be fitted to the data. Because $\Pr(ak;\tau) = a^{-\tau}\Pr(k;\tau)$, the distribution of (1) is also called scale-free. The interest in power law distributions stems mainly from the fact that these are particularly heavy-tailed (i.e. the tail of the distribution decays more slowly than an exponential function). A heavy-tailed distribution causes large fluctuations in degrees and strengths, which is at first glance consistent with many empirical network data.

Most of the time, a power-law degree distribution was inferred from straight-line fits to the log-log diagram of the degree frequency, but this is now generally viewed as an unsatisfactory approach [14]. Identifying a region where the data appear more or less linear is largely arbitrary because most distributions are too noisy and substantially curved on double-logarithmic scales. Straight-line fits based on standard least-squares algorithms can also lead to a bias in the estimated exponents. Furthermore, there is no a priori reason why $\Pr(k)$ has to be a power law. Many other common probability distributions are also heavy-tailed and may fit the data better. Recent studies in fact doubt that power laws are as ubiquitous as once believed [15], [16], [17].

This paper proposes a more deductive perspective to understand which type of degree distribution best explains the observed data. We apply information-based model selection [18], an increasingly popular approach for comparing power laws to other heavy-tailed distributions [15], [19], [20], [21]. Another goal of the paper is to test if the best-fitting function for the cargo ship network has changed over time. One hypothesis is that the degree distribution may structurally change alongside major technological transformations of the shipping industry and their consequences on port operations and maritime network configurations.

Data were collected from the Lloyd's Shipping Index, a weekly publication by Lloyd's List, over the period 1890-2008. For 14 selected years, an entire volume of the Index was extracted manually, thereby providing a snapshot of global maritime activity based on the last known voyage of each vessel between two ports, around the months of April and May. The network for each year is weighted by the number of vessel calls by node (i.e. port) and by link (i.e. inter-port movement). A comparison with official data sources for Shanghai and Rotterdam revealed that the Index recorded approximately 0.49% and 0.66% of their respective annual number of total vessel calls. The studied period goes across different dominant ship technologies, such as sail, steam, combustion, specialized vessels (e.g. container, tanker, etc.), and mega-carriers. Such technological evolutions are believed to have been selective, as some ports were dropped from the network and replaced or superseded by new ones better adapted to changing standards, sometimes resulting in an increasing concentration of port activity favouring fewer and larger ports. Containerization is seen as a revolution in itself with profound impacts on network configuration and world trade [22], [23].

II. Probabilistic models

We investigate eight different models that have frequently been used to fit empirical degree distributions in complex networks (Table I). We restrict our study to discrete distributions

(a) whose support are all positive integers and
(b) that depend on maximally two parameters.

Restriction (a) reflects that degree or port calls only have integer values. One might argue that $k = 0$ should also be included and that the maximum degree should have an upper bound because the network is finite. However, from Lloyd's Shipping Index we cannot directly infer which ports were in principle open to traffic, but remained unused. Consequently, distributions with infinite support, but excluding $k = 0$ are more appropriate in the present context. Restriction (b) is primarily to avoid overfitting of the data.

We include four one-parameter models: the Poisson, geometric, zeta and Yule-Simon distribution. The Poisson distribution applies to the node degrees of large sparse Erdős-Rényi random graphs, a common null model in network studies. The tail of a Poisson distribution decays faster than exponentially so that degrees in Erdős-Rényi graphs are effectively limited to values near the mean degree. The geometric distribution decays exponentially, whereas the zeta and Yule-Simon distributions have power law tails. As a mixed case we introduce the exponentially truncated power law as one of our two-parameter models in Table I. The negative binomial is another two-parameter example that is heavy-tailed if its parameter r exceeds 1, but with less weight in the tail than a power law.

Table I
The investigated probabilistic models.

distribution	parameters	$\Pr(k),\ k>0$
Poisson (POIS)	$\lambda > 0$	$\frac{\lambda^k}{(e^\lambda - 1)k!}$
geometric (GEOM)	$p \in (0,1)$	$p(1-p)^{k-1}$
zeta (ZETA) ("power law")	$\tau > 1$	$[\zeta(\tau)k^\tau]^{-1}$
Yule-Simon (YULE)	$\rho > 0$	$\rho \mathrm{B}(k, \rho+1)$
negative binomial (NEGB)	$p \in (0,1)$, $r > 0$	$\frac{\Gamma(k+r)p^k(1-p)^r}{k!\Gamma(r)[1-(1-p)^r]}$
truncated power law (TPOW)	$q \in (0,1)$, $\tau > 1$	$\frac{q^k}{\mathrm{Li}_\tau(q)k^\tau}$
Poisson-lognormal (PLGN) [32]	$\mu \in \mathbb{R}$, $\sigma > 0$	given by (2) and (3)
discrete Weibull (DWEI) [31] ("stretched exponential")	$q \in (0,1)$, $\beta > 0$	$\frac{1}{q}\left[q^{(k^\beta)} - q^{((k+1)^\beta)}\right]$

Among continuous distributions, two further models whose decay is between an exponential and a power law are the Weibull (also known as stretched exponential if $\beta < 1$) and the lognormal distribution. Previous studies have reported Weibull [24], [25], [26], [27] and lognormal degree distributions [19], [28], [29], [30] in real-world networks, so that we include them as candidates in our study too, albeit in discretized form. We discretize the Weibull distribution following Nakagawa & Osaki [31] so that the complementary cumulative distribution is a discrete stretched exponential. Discretizing the lognormal distribution is a little trickier. Here we apply the method of Bulmer [32] who defined the Poisson-lognormal distribution $f_{\mu,\sigma}(k)$ as the mixing of Poisson distributions whose mean λ is lognormally distributed with parameters μ and σ,

$$f_{\mu,\sigma}(k) = \frac{\int_0^\infty e^{-\lambda}\lambda^{k-1}\exp\left(-\frac{(\ln\lambda-\mu)^2}{2\sigma^2}\right)d\lambda}{\sigma\sqrt{2\pi}k!}. \quad (2)$$

In our calculation we work with the constrained probability that $k > 0$,

$$\Pr(k;\mu,\sigma) = \frac{f_{\mu,\sigma}(k)}{1 - f_{\mu,\sigma}(0)}. \quad (3)$$

The same constraint is used in all other investigated models, which explains why the equations in Table I differ from their textbook form in those cases where $k = 0$ is conventionally included in the distribution's support.

III. Model selection

Assuming that all degrees are independent, the likelihood function for any of the models in Table I has the general form

$$L(\mathbf{v}) = \prod_{i=1}^{n} \Pr(k_i; \mathbf{v}), \quad (4)$$

where $\mathbf{v}$ is the set of the parameters in the second column of the table. The degrees may in reality depend on each other, so that $L(\mathbf{v})$ in (4) is more properly thought of as a composite likelihood. We can justify the use of a composite likelihood in our present context because the degree distribution $\Pr(k)$ that we would like to model is a marginal (rather than the complete joint) distribution of all degrees. Therefore, the full dependence structure is in statistical parlance a “nuisance parameter” which neither matters to us nor is it clear how to specify the full likelihood. In such cases, composite likelihood methods have proved to be a well-behaved alternative [33], [34].

For a specified model i, we determine the parameter $\hat{\mathbf{v}}_i$ that maximizes L and hence also the log-likelihood $\ln(L)$. A comparison between different models can then be performed by ranking their Akaike information criterion (AIC) [35],

$$\text{AIC}_i = -2\ln(L(\hat{\mathbf{v}}_i)) + 2K_i, \quad (5)$$

where K_i is the number of parameters in the respective model. The AIC not only tells us which model is closest to the data in information content, properly taking into account that higher K_i generally allows better fits to the data, but weakens the explanatory power of the model. We can also make quantitative comparisons between different models based on the AIC differences. If AIC_{min} is the minimum AIC over all models, then the differences

$$\Delta_i = \text{AIC}_i - \text{AIC}_{\text{min}} \quad (6)$$

estimate the relative expected information gain between model i and the estimated best model. Because the likelihood of model i given the degrees $k_1, k_2, \dots$ is proportional to $\exp(-\Delta_i/2)$ [18], the relative likelihood is the so-called Akaike weight

$$w_i = \frac{\exp(-\Delta_i/2)}{\sum_j \exp(-\Delta_j/2)}, \quad (7)$$

where the summation in the denominator is over all models included in the comparison. The Akaike weights for the 14 data sets are summarized in Table II for the degree distributions and in Table III for the strength distributions.

IV. Discussion

As a quick glance at Tables II and III reveals, the discrete Weibull distribution always has the largest Akaike weight with only one exception, namely the strength distribution in 1951 when it is a close runner-up behind the Poisson-lognormal. All two-parameter models always perform better than even the best one-parameter model, which is in all cases the Yule-Simon distribution. The added term $2K_i$ in (5) for introducing a second parameter is therefore always more than compensated by an increased likelihood.

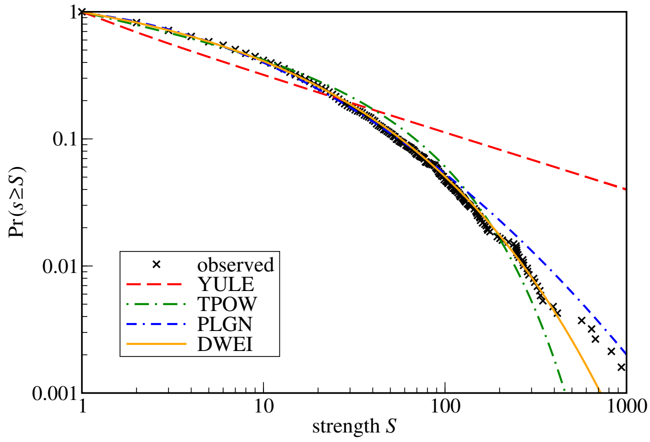

Figure 1. The complementary cumulative strength distribution function in the year 1995 together with the maximum-likelihood Yule-Simon, truncated power law, Poisson-lognormal and discrete Weibull distributions. The one-parameter Yule-Simon distribution fits the data far worse than any of the two-parameter alternatives. Among these the discrete Weibull approximates most data points best.

The effect of including a second parameter can be seen in Fig. 1 where we compare the observed strength distribution in 1995 with the maximum-likelihood estimates for the Yule-Simon, truncated power law, Poisson-lognormal and discrete Weibull distribution. (There is nothing peculiar about the year 1995; we have chosen it as a representative example to demonstrate general features that repeat in most other years too.) The cumulative distribution of the data is substantially curved on a log-log scale and thus difficult to fit by an asymptotic power law such as the Yule-Simon distribution. All other plotted distributions have the flexibility to follow the curvature more accurately. Still, the truncated power law decays too rapidly on the right, whereas the Poisson-lognormal does not decay quickly enough. The discrete Weibull distribution with its intermediate asymptotic decay is a better compromise.

It is striking that this stretched exponential has remained the best fit for almost all data sets between 1890 and 2008 despite the technological transformations and regional shifts that took place during this period. For instance, Asia as a whole concentrated 6% and 38% of world port traffic in 1890 and 2008 respectively, while North America dropped from 17% to 6%. This observation points towards resilience in the overall distribution of ship traffic: ports that declined in importance or completely disappeared from the network were replaced by others so that the shape of the distribution remained approximately a stretched exponential. Such results tend to question the commonly believed impacts

Table II
AKAIKE WEIGHTS FOR THE DEGREE DISTRIBUTION. VALUES BELOW 10^{-200} ARE ROUNDED TO ZERO. THE SMALLEST AKAIKE WEIGHT IN EACH YEAR IS HIGHLIGHTED IN BOLD FONT.

Year	number of ports	POIS	GEOM	ZETA	YULE	NEGB	TPOW	PLGN	DWEI
1890	872	0	3.67×10^{-120}	2.64×10^{-32}	1.22×10^{-21}	5.65×10^{-7}	6.86×10^{-3}	2.12×10^{-1}	$\mathbf{7.81 \times 10^{-1}}$
1925	1140	0	1.17×10^{-166}	6.62×10^{-70}	5.01×10^{-55}	1.04×10^{-3}	1.73×10^{-3}	7.47×10^{-3}	$\mathbf{9.90 \times 10^{-1}}$
1946	1184	0	5.29×10^{-175}	9.04×10^{-57}	7.18×10^{-42}	3.35×10^{-4}	3.62×10^{-2}	3.56×10^{-2}	$\mathbf{9.28 \times 10^{-1}}$
1951	1280	0	1.19×10^{-184}	1.49×10^{-69}	1.25×10^{-52}	4.11×10^{-5}	4.88×10^{-4}	4.75×10^{-2}	$\mathbf{9.52 \times 10^{-1}}$
1960	1463	0	0	1.73×10^{-88}	2.90×10^{-69}	4.21×10^{-5}	1.01×10^{-4}	3.35×10^{-3}	$\mathbf{9.97 \times 10^{-1}}$
1965	1506	0	0	2.66×10^{-105}	7.41×10^{-85}	2.95×10^{-5}	2.29×10^{-5}	1.20×10^{-3}	$\mathbf{9.99 \times 10^{-1}}$
1970	1467	0	0	4.02×10^{-97}	2.82×10^{-78}	1.58×10^{-2}	1.53×10^{-2}	5.85×10^{-5}	$\mathbf{9.69 \times 10^{-1}}$
1975	1565	0	0	7.00×10^{-91}	6.35×10^{-71}	6.27×10^{-4}	2.71×10^{-3}	7.40×10^{-4}	$\mathbf{9.96 \times 10^{-1}}$
1980	1581	0	7.85×10^{-190}	5.61×10^{-115}	3.08×10^{-93}	1.42×10^{-4}	8.92×10^{-5}	5.26×10^{-5}	$> \mathbf{9.99 \times 10^{-1}}$
1985	1844	0	0	5.14×10^{-136}	3.25×10^{-110}	2.50×10^{-6}	1.33×10^{-6}	2.27×10^{-5}	$> \mathbf{9.99 \times 10^{-1}}$
1990	1849	0	0	6.49×10^{-158}	2.03×10^{-132}	1.07×10^{-4}	4.13×10^{-5}	3.34×10^{-8}	$> \mathbf{9.99 \times 10^{-1}}$
1995	1880	0	2.37×10^{-190}	1.03×10^{-158}	2.28×10^{-132}	1.07×10^{-3}	4.53×10^{-4}	7.91×10^{-9}	$\mathbf{9.98 \times 10^{-1}}$
2000	1916	0	0	4.81×10^{-126}	5.17×10^{-100}	9.65×10^{-3}	7.72×10^{-3}	3.08×10^{-7}	$\mathbf{9.83 \times 10^{-1}}$
2008	1963	0	7.70×10^{-191}	4.34×10^{-102}	2.72×10^{-76}	1.46×10^{-1}	3.37×10^{-1}	1.81×10^{-7}	$\mathbf{5.16 \times 10^{-1}}$

Table III
AKAIKE WEIGHTS FOR THE STRENGTH DISTRIBUTION. VALUES BELOW 10^{-200} ARE ROUNDED TO ZERO. THE SMALLEST AKAIKE WEIGHT IN EACH YEAR IS HIGHLIGHTED IN BOLD FONT.

Year	POIS	GEOM	ZETA	YULE	NEGB	TPOW	PLGN	DWEI
1890	0	0	1.78×10^{-25}	3.69×10^{-17}	9.43×10^{-23}	6.52×10^{-5}	4.20×10^{-1}	$\mathbf{5.80 \times 10^{-1}}$
1925	0	0	8.17×10^{-51}	4.64×10^{-39}	5.14×10^{-23}	1.53×10^{-9}	4.20×10^{-1}	$\mathbf{5.80 \times 10^{-1}}$
1946	0	0	1.78×10^{-46}	1.91×10^{-34}	9.66×10^{-20}	9.85×10^{-6}	2.94×10^{-1}	$\mathbf{7.06 \times 10^{-1}}$
1951	0	0	1.51×10^{-56}	2.40×10^{-42}	3.64×10^{-20}	8.84×10^{-9}	$\mathbf{5.04 \times 10^{-1}}$	4.96×10^{-1}
1960	0	0	1.87×10^{-73}	3.62×10^{-57}	3.39×10^{-21}	1.69×10^{-11}	2.93×10^{-1}	$\mathbf{7.07 \times 10^{-1}}$
1965	0	0	1.29×10^{-78}	2.12×10^{-62}	2.93×10^{-28}	8.59×10^{-16}	3.89×10^{-1}	$\mathbf{6.11 \times 10^{-1}}$
1970	0	0	7.00×10^{-79}	4.45×10^{-63}	2.46×10^{-14}	3.40×10^{-8}	1.95×10^{-2}	$\mathbf{9.80 \times 10^{-1}}$
1975	0	0	2.23×10^{-77}	3.35×10^{-60}	3.00×10^{-20}	2.22×10^{-10}	1.25×10^{-1}	$\mathbf{8.75 \times 10^{-1}}$
1980	0	0	3.41×10^{-95}	1.86×10^{-76}	6.12×10^{-16}	1.80×10^{-12}	4.71×10^{-2}	$\mathbf{9.53 \times 10^{-1}}$
1985	0	0	3.90×10^{-119}	3.95×10^{-96}	1.52×10^{-20}	6.42×10^{-18}	1.82×10^{-1}	$\mathbf{8.18 \times 10^{-1}}$
1990	0	0	1.34×10^{-146}	8.66×10^{-124}	2.58×10^{-20}	3.37×10^{-20}	4.77×10^{-3}	$\mathbf{9.95 \times 10^{-1}}$
1995	0	0	7.04×10^{-143}	4.46×10^{-119}	1.34×10^{-17}	1.61×10^{-17}	7.76×10^{-4}	$\mathbf{9.99 \times 10^{-1}}$
2000	0	0	4.26×10^{-119}	3.29×10^{-95}	3.06×10^{-18}	8.34×10^{-16}	4.47×10^{-3}	$\mathbf{9.96 \times 10^{-1}}$
2008	0	0	7.67×10^{-94}	3.11×10^{-71}	1.21×10^{-15}	5.94×10^{-9}	4.37×10^{-5}	$> \mathbf{9.99 \times 10^{-1}}$

of technological changes on shipping networks, such as increasing hierarchical linkages favoured by containerization and related hub-and-spoke systems. Modern shipping operations are path-dependent as they rely upon port areas and functions that were already in place [36], and so the port hierarchies of the late nineteenth and early twenty-first centuries differ more in scale than in nature [37]. Another possible explanation is the growing trade and economic integration within regions that favoured more transversal, short-sea and coastal shipping linkages that are not always routed through large hubs. Hierarchical and nonhierarchical tendencies thus overlap and mitigate one another at the global level.

A closer look at the maximum-likelihood estimators for the Weibull parameters, however, reveals that the distribution has not remained static. In Fig. 2 we plot the tail exponent β for all years included in this study. For the degree distribution, β increased from 0.304 ± 0.010 to 0.453 ± 0.008 between 1890 and 1995. The strength distribution shows a similar behaviour with β increasing from 0.200 ± 0.008 to 0.369 ± 0.007 during the same period. In other words, the distributions became less heavy-tailed for more than one century, suggesting a flattening hierarchy among the world's ports. Generally, β is smaller for the strength distribution indicating that its tail is more stretched than for the degree. This result is consistent with a previous observation in [8]

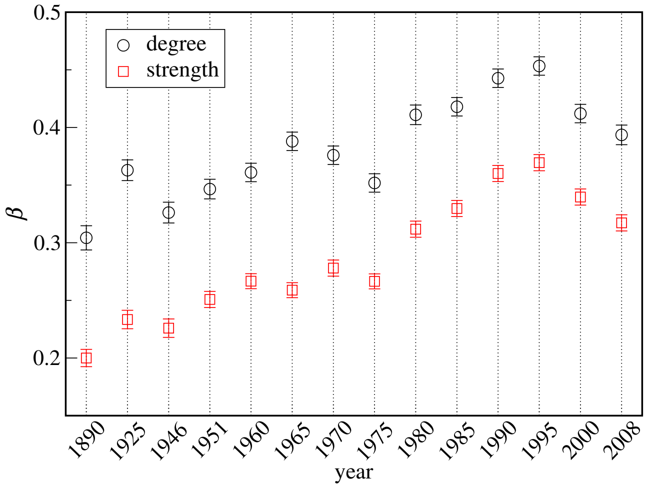

Figure 2. The maximum-likelihood estimate for the tail exponent β in the discrete Weibull distribution defined in Table I. Circles refer to the degree, squares to the strength distributions. The data show an overall increase in β between 1890 and 1995 (i.e. the distributions became shorter-tailed). Since then, β has decreased again.

that the strength of a port scales on average superlinearly with its degree.

After the peak in 1995, β has begun to decline again. As possible reasons, China's economic rise or the increasing size of container ships come to mind. We are planning to carry out further analysis to determine if the changing tails are indeed caused by regional shifts affecting the distribution globally. We will also search for possible impacts of technological changes and enhance the data base with more regular snapshots in the period between 1890 and 1946. In each selected year, more movements will be included as data extraction from Lloyd's Shipping Index proceeds.

V. Conclusion

As in all model selection problems, one should bear in mind that reality is almost certainly more complex than any of the candidate models. In our case, it might be possible to reduce the AIC further by allowing more than two parameters. With additional data, it might also become possible to statistically analyze the dynamics of the network with full likelihood methods [38] rather than resorting to summary statistics such as degree and strength distributions. Despite these caveats, model selection based on Akaike weights is a statistically rigorous approach based on information theory [18]. Therefore, we can firmly conclude that – among all investigated models – the Weibull distribution is clearly the overall most likely candidate to explain the degree and strength distributions in the global network of cargo shipping during the past 125 years.

Acknowledgment

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013) / ERC Grant Agreement n. [313847] "World Seastems". M. T. G. thanks the European Commission for financial support (project number FP7-PEOPLE-2012-IEF 6-456412013).

References

[1] M. Barthélemy, "Spatial networks," *Physics Reports*, vol. 499, pp. 1–101, 2010.

[2] M. Barthélemy, "Crossover from scale-free to spatial networks," *Europhysics Letters*, vol. 63, pp. 915–921, 2003.

[3] A. Barrat, M. Barthélemy, and A. Vespignani, "The effects of spatial constraints on the evolution of weighted complex networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2005, p. P05003, 2005.

[4] C. Ducruet and L. Beauguitte, "Spatial science and network science: review and outcomes of a complex relationship," *Networks and Spatial Economics*, 2014. doi: 10.1007/s11067-013-9222-6.

[5] United Nations Conference on Trade and Development, *Review of Maritime Transport 2013*, Available: unctad.org/en/publicationslibrary/rmt2013_en.pdf

[6] D. Wei-Bing, G. Long, L. Wei, and C. Xu, "Worldwide marine transportation network: efficiency and container throughput," *Chinese Physics Letters*, vol. 26, p. 118901, 2009.

[7] Y. Hu and D. Zhu, "Empirical analysis of the worldwide maritime transportation network," *Physica A*, vol. 388, pp. 2061–2071, 2009.

[8] P. Kaluza, A. Kölzsch, M. T. Gastner, and B. Blasius "The complex network of global cargo ship movements," *J. Roy. Soc. Interface*, vol. 7, pp. 1093–1103, 2010.

[9] C. Ducruet, "Network diversity and maritime flows," *J. Transp. Geogr.*, vol. 30, pp. 77–88, 2013.

[10] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, 1999.

[11] B. Tadić, "Temporal fractal structures: origin of power laws in the world-wide Web," *Physica A*, vol. 314, pp. 278–283, 2002.

[12] C. Moore, G. Ghoshal, and M. E. J. Newman, "Exact solutions for models of evolving networks with addition and deletion of nodes," *Physical Review E*, vol. 74, p. 036121, 2006.

[13] A. G. Wilson, "A statistical theory of spatial distribution models," *Transportation Research*, vol. 1, pp. 253–269, 1967.

[14] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, pp. 661–703, 2009.

[15] A. M. Edwards *et al.*, "Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer," *Nature*, vol. 449, pp. 1044–1047, 2007.

[16] W. Willinger, D. Alderson, and J. C. Doyle, "Mathematics and the Internet: a source of enormous confusion and great potential," *Notices of the AMS*, vol. 56, pp. 586–599, 2009.

[17] M. P. H. Stumpf and M. A. Porter, "Critical truths about power laws," *Science*, vol. 335, pp. 665–666, 2012.

[18] K. P. Burnham and D. R. Anderson, *Model selection and multimodel inference*. New York: Springer, 1998.

[19] M. P. H. Stumpf, P. J. Ingram, I. Nouvel, and C. Wiuf, "Statistical model selection methods applied to biological networks," in *Lecture notes in computer science*, vol. 3737, *Transactions on computational systems biology III*, C. Priami, E. Merelli, P. Gonzalez, and A. Omicini, Eds. Berlin: Springer, 2005, pp. 65–77.

[20] D. T. Hamilton, M. S. Handcock, and M. Morris, "Degree distributions in sexual networks: a framework for evaluating evidence," *Sexually Transmitted Diseases*, vol. 35, pp. 30–40, 2008.

[21] F. Prieto and J. M. Sarabia, "Fitting the degree distribution of real-world networks," *Int. J. Complex Systems in Science*, vol. 1, pp. 129–133, 2011.

[22] K. P. B. Cullinane and M. Khanna, "Economies of scale in large containerships: optimal size and geographical implications," *J. Transp. Geogr.*, vol. 8, pp. 181–195, 1999.

[23] D. M. Bernhofen, Z. El-Sahli, and R. Kneller, "Estimating the effects of the container revolution on world trade," Lund University Working Paper, Department of Economics, School of Economics and Management, 2013.

[24] J. Laherrere and D. Sornette, "Stretched exponential distributions in nature and economy: fat tails with characteristic scales," *European Physical Journal B*, vol. 2, pp. 525–539, 1998.

[25] A. Broido and kc claffy, "Internet topology: connectivity of IP graphs," *Proc. SPIE*, vol. 4526, pp. 172–187, 2001.

[26] Y. He, G. Siganos, M. Faloutsos, and S. Krishnamurthy, "A systematic framework for unearthing the missing links: measurements and impact," *Proc. 4th USENIX Conf. Networked Systems Design & Implementation*, p. 14, 2007.

[27] L. E. C. Rocha, F. Liljeros, and P. Holme, "Information dynamics shape the sexual networks of Internet-mediated prostitution," *Proc. Nat. Acad. Sci.*, vol. 107, pp. 5706–5711, 2010.

[28] K. Bhattacharya, G. Mukherjee, J. Saramaki, K. Kaski, and S. S. Manna, "The international trade network: weighted network analysis and modelling," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, p. P02002, 2008.

[29] V. Gómez, A. Kaltenbrunner, and V. López, "Statistical analysis of the social network and discussion threads in Slashdot," *Proc. 17th Int. Conf. World Wide Web*, pp. 645–654, 2008.

[30] A. Todor, A. Dobra, and T. Kahveci, "Uncertain interactions affect degree distribution of biological networks," *IEEE Int. Conf. Bioinformatics and Biomedicine*, pp. 457–461, 2012.

[31] T. Nakagawa and S. Osaki, "The discrete Weibull distribution," *IEEE Trans. Reliability*, vol. R-24, pp. 300–301, 1975.

[32] M. G. Bulmer, "On fitting the Poisson lognormal distribution to species-abundance data," *Biometrics*, vol. 30, pp. 101–110, 1974.

[33] D. R. Cox and N. Reid, "A note on pseudolikelihood constructed from marginal densities," *Biometrika*, vol. 91, pp. 729–737, 2004.

[34] C. Varin, N. Reid, and D. Firth, "An overview of composite likelihood methods," *Statistica Sinica*, vol. 21, pp. 5–42, 2011.

[35] H. Akaike, "A new look at the statistical model identification," *IEEE Transactions on Automatic Control*, vol. 19, pp. 716–723, 1974.

[36] D. K. Fleming and Y. Hayuth, "Concepts of strategic commercial location: the case of container ports," *Maritime Policy and Management*, vol. 21, pp. 187–193, 1994.

[37] B. Marnot, "Interconnexion et reclassements: l'insertion des ports français dans la chaîne multimodale au XIXe siècle," *Flux*, vol. 1, pp. 10–21, 2005.

[38] C. Wiuf, M. Brameier, O. Hagberg, and M. P. H. Stumpf, "A likelihood approach to analysis of network data," *Proc. Nat. Acad. Sci*, vol. 103, pp. 7566–7570, 2006.