Brazilian Journal of Probability and Statistics
2021, Vol. 35, No. 1, 50–61
https://doi.org/10.1214/20-BJPS487

Self-organized criticality of aggregated animals attributed to Tweedie convergence

Wayne S. Kendal
University of Ottawa

In memoriam: Prof. Bent Jørgensen (1954–2015), who contributed the fundamentals on which this study was based.

Abstract. Ecologists have had an ongoing interest in a variance to mean power law that governs the clustering of individuals of animal and plant species. This same power law has been reported from disparate biological, physical and mathematical systems, and also characterizes a family of statistical distributions known as the Tweedie exponential dispersion models. Its widespread appearance can be explained by fundamental statistical convergence effects on random data that cause this, and related, power laws to emerge and provide mechanistic insight into its origin, as well as the origin of $1/f$ noise, multifractality and other phenomena attributable to self-organized criticality. A meta-analysis of ecological field data was conducted here to examine how such statistical convergence might affect the power law. These findings provided conjectural insight into a form of self-organized criticality, driven and modulated by the statistical convergence of random data, which could underlie the power law's emergence.

1 Introduction

Ecologists have long recognized that plants and animals tend to cluster together. The mechanism for this clustering, though, remains controversial. The late L. R. Taylor noted that if a habitat were divided into equal-sized regions, and the numbers of a species enumerated within these regions, the mean sample count $\bar{Z}$ and its variance s_Z^2 would obey the power law,

$$s_Z^2 = a \cdot \bar{Z}^p, \tag{1.1}$$

where a and p were constants and p being of fractional value, Taylor (1961). This relationship has come to be called Taylor's power law, Taylor et al. (1983).

Taylor proposed the power law exponent p as a species-specific parameter and found that its values ranged mainly between 1 and 2. This power law has been confirmed for hundreds of animal and plant species, as well as for some microorganisms, Taylor (1961), Taylor, Woiwod and Perry (1978), Taylor and Woiwod (1980), Taylor and Woiwod (1982), and Taylor et al. (1983). On the basis of semi-quantitative arguments Taylor attributed the power law to the balanced non-random repulsive and congregatory behavior of animals, Taylor and Taylor (1977).

Many other explanations for the power law followed: Anderson et al. (1982) proposed a Markovian population model moderated by stochastic demographics; Hanski (1980) suggested a random walk model controlled by reproductive and survival rates; Kemp (1987) advocated the negative binomial, Neyman type A, and Polya–Aeppli distributions as statistical models for Taylor's law; see also, Jørgensen and Kokonendji (2016) and Bonat et al.

Key words and phrases. Emergent phenomena, critical phenomena, phase transitions, Taylor's power law, weak convergence.

Received June 2019; accepted September 2020.

(2018). Perry (1994) argued that chaotic dynamics could yield the law, and Kilpatrick and Ives (2003) showed that simulated species interactions could generate the relation. Despite these (and other) attempts to explain Taylor's law, no consensus towards any particular explanation has emerged.

Variance to mean power laws mathematically identical to Taylor's law have also been implicated in the transmission of human immunodeficiency virus (HIV) by Anderson and May (1988), and recently by Petterle et al. (2019), as well as measles epidemics, Keeling and Grenfell (1999), the geographic clustering of leukemia cases, Philippe (1999), the genomic distribution of single nucleotide polymorphisms (SNPs), Kendal (2003) as well as genes, Kendal (2004a), the distribution of the prime numbers, Kendal and Jørgensen (2015), eigenvalue deviations from random matrices, Kendal and Jørgensen (2011b), $1/f$ noise, Kendal and Jørgensen (2011b), and self-organized criticality (SOC), Kendal (2015). In view of the wide range of manifestations of the power law a more general explanation than that provided by population dynamic models seemed required.

About the same time that Taylor was developing his hypothesis Tweedie (1984), a British medical physicist and statistician, was studying statistical models that expressed an inherent variance to mean power law. These models have come to be called the Tweedie exponential dispersion models (EDMs), Jørgensen (1997). The Tweedie EDMs have a fundamental role in statistical theory as foci of convergence for a wide range of random and pseudorandom data, similar to the role that the Gaussian distribution has in the central limit theorem, Jørgensen, Martínez and Tsao (1994).

The clustering of the potato beetle, *Leptinotarsa decemlineata*, was demonstrated to express Taylor's power law and to correspond to a discrete equivalent of the Tweedie compound Poisson EDM, Kendal (2002). A random distribution of insects would be conventionally represented by the Poisson distribution, which manifests no power law clustering. The compound Poisson EDM, however, provides an alternative random model that inherently expresses the variance to mean power law. An explanation for Taylor's law based on the Tweedie EDMs has met with criticism for reason of its lack of a population dynamic framework and for its inability to account for observations where $0 < p < 1$. The study reported here addresses both of these criticisms.

A meta-analysis of three large ecological reviews, previously reported by Taylor and his colleagues, will be examined. This analysis was focused upon the clustering of invertebrate and bird species in the context of two related Tweedie convergence effects. The first effect, mentioned above, causes certain variance functions of EDMs to converge towards a variance to mean power law; this power law has been traditionally interpreted to represent non-random self-organization amongst individuals of a species. A second convergence effect is presented and conjectured here to shape the distribution of the Tweedie dimensional parameter α, which shares a one- to-one relationship with the exponent p. This latter convergence effect, it will be conjectured, introduces an instability in the values of α, which corresponds to a form of SOC.

2 Field observations of Taylor's law

To begin, some field observations of Taylor's law from the scientific literature will be introduced as a historical example. These observations were based on the combined study of the aphid (*Aphis fabae* Scop.) from two varieties of bean plants (Mazagan and Claudia Aquadulce), as enumerated and analyzed by Taylor (1970). The mean number of aphids per stem (chosen by Taylor to represent an equitable measurement unit) and its corresponding variance were plotted for $N = 67$ samples (Figure 1). A log-log plot has been typically used

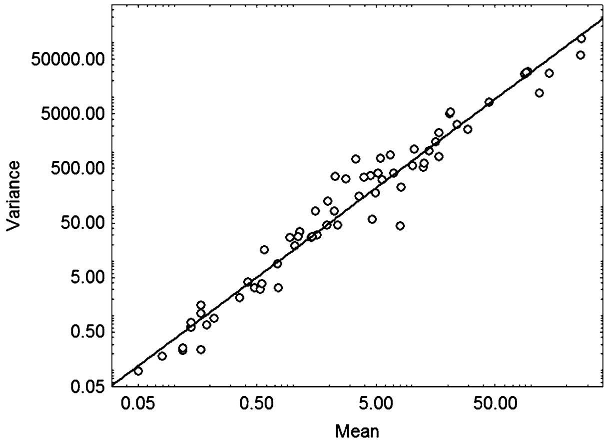

Figure 1 *Taylor's law from infestations of Aphis fabae Scop. on bean plants. These data were previously published by Taylor* (Taylor, 1970). *The means and variances derived from the samples were plotted on a log-log plot. A straight line relationship on this plot* (*solid line*) *suggested an underlying power law. Linear regression yielded the constant* $\log_{10}(a) = 1.20$ (95% *CI*, 1.11 *to* 1.29) *and the exponent* $p = 1.63$ (95% *CI*, 1.54 *to* 1.71) *with* $r^2 = 0.96$.

for this purpose since the power law can be represented by a straight line, and linear regression can provide the power law parameters a and p. Following the methodology employed by Taylor, and numerous other investigators over the last half century, linear regression was used to reproduce the variance to mean power law from Taylor's data. The exponent for these particular data was $p = 1.63$, which was found to have a relatively narrow 95% confidence interval (CI).

In this regression, the logarithm of the sample mean number of counts (the response variable) and the logarithm of the variance of these counts (the explanatory variable) both represented random quantities; consequently care should be exercised in interpreting such regressions. For this reason the analysis was performed here using two different methods: First, following the conventional practice of Taylor (1961) and other investigators, the data were logarithmically transformed and a linear regression performed with the 95% CI determined under the assumption of normality. The second analysis was based on bootstrapping methods (STATA 13, College Station, Texas) as applied to the logarithmically transformed data. Both p and its 95% CI obtained by bootstrapping agreed with the linear regression of the conventional analysis.

The correlation coefficient squared from the regression was $r^2 = 0.96$, indicating that the power law appeared to account for 96% of the variance associated with these data. Inspection of Figure 1 revealed an approximately even distribution of the data points about the linear fit. Field data like these are invariably associated with additional spurious scatter. Consequently, conclusions drawn from the estimated values of p should be considered in light of spurious effects, associated with the unexplained variance.

Other historical assessments of Taylor's law, reviewed in the meta-analysis below, were based on smaller sample numbers N and/or yielded larger proportions of unexplained variance, which opened the possibility for larger spurious influences. The potential that these spurious influences might have caused some assessments for p to fall within the range $0 < p < 1$ will be considered further, particularly in view of criticisms raised regarding inability of the Tweedie EDMs to account for values within this range.

For each of the 67 samples in the analysis of aphids from Figure 1, Taylor provided the mean number of individuals per stem and a variance. The variance is a measure of statistical variability about the mean, and values of the variances Taylor obtained indicated that random

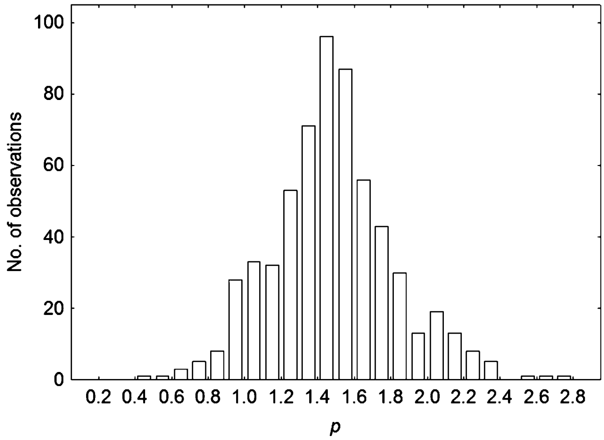

Figure 2 *Histogram of values of Taylor's law exponent p. This figure represents a meta-analysis of data from* 3 *previous publications. Values of p were assessed from linear regression of log-log plots of Taylor's law. In all* 608 *assessments of p were retrieved. Nearly* 90% *of assessments ranged between* 1 *and* 2; 6.4% *were less than* 1 *and* 7.9% *were greater than* 2. *The most frequent values occurred near* $p = 1.4$.

influences had contributed to these assessments beyond the variance to mean power law itself. Taylor considered this variance to mean power law to represent non-random influences upon the spatial distribution of animals. An alternative hypothesis is considered below that the spatial distribution of animals and plants is inherently random, and primarily governed by a Tweedie EDM.

The full meta-analysis began with the construction of a frequency histogram for the exponent p derived from 3 large reviews of field data derived mainly from insect and avian species, Taylor, Woiwod and Perry (1978) and Taylor and Woiwod (1980, 1982). The resultant distribution appeared more or less symmetrically peaked with a mode near $p = 1.4$ (Figure 2).

About 6.4% of the estimates for p were less than 1. From this initial examination it was not clear as to the degree that spurious effects might have influenced the estimates in this range. Before the data are examined more closely it would be useful to review some further background regarding the Tweedie EDMs.

3 Foundations of the Tweedie exponential dispersion models

EDMs are statistical models created to describe the error distributions of generalized linear models. See Jørgensen (1997) for further detail of the theory. These EDMs are characterized by unique variance functions $V(\mu)$, expressed in terms of the position parameter, μ, Nelder and Wedderburn (1972). The Tweedie EDMs correspond to a subclass of these models distinguished by the mathematical properties of closure under additive and reproductive convolution, as well as under scale transformation, Jørgensen (1997), and Bonat and Kokonendji (2017). This latter property leads the Tweedie EDMs to have variance functions of the form $V(\mu) = \mu^p$ and thus an inherent variance to mean power law because if $Y \sim \mathrm{ED}(\mu, \lambda)$ then $\mathrm{E}(Y) = \lambda\mu$ and $\mathrm{Var}(Y) = V(\mu)$.

The Tweedie EDMs are classified by their power law exponent p: $p < 0$ gives the extreme stable distributions; $p = 0$, the Gaussian distribution; $p = 1$, the Poisson distribution; $1 < p < 2$, the compound Poisson distribution; $p = 2$, the gamma distribution; $2 < p < 3$, the positive stable distributions; $p = 3$, the inverse Gaussian distribution; $p > 3$, the positive stable distributions; $p = \infty$, the extreme stable distributions; and (as noted above) for $0 < p < 1$ the Tweedie models do not exist.

Additive Tweedie EDMs have cumulant generating functions of the form, Jørgensen (1997),

$$K_p^*(s;\theta,\lambda)=\begin{cases}\lambda e^{\theta}(e^{s}-1), & p=1,\\ \lambda\kappa_p(\theta)[(1+s/\theta)^{\alpha}-1], & p\neq 1,2,\\ -\lambda\log(1+s/\theta), & p=2,\end{cases}\tag{3.1}$$

where s is the generating function variable and θ and λ are the canonical and index parameters, respectively. The cumulant function $\kappa_p(\theta)$ is

$$\kappa_p(\theta)=\begin{cases}e^{\theta}, & p=1,\\ \dfrac{\alpha-1}{\alpha}\left(\dfrac{\theta}{\alpha-1}\right)^{\alpha}, & p\neq 1,2,\\ -\log(-\theta), & p=2.\end{cases}\tag{3.2}$$

The compound Poisson distribution (as indicated by the second CGF in Eq. (3.1)) represents the sum of a random (Poisson-distributed) number of gamma-distributed variables. It has its values for p restricted within the range $1<p<2$, which corresponded to the majority of assessments for p from the field data provided here.

The constant α in Eq. (3.1) is referred to here as the Tweedie dimensional exponent; it is related to the power law exponent p through the equations

$$\alpha=(p-2)/(p-1)\quad\text{and}\tag{3.3}$$

$$p=(\alpha-2)/(\alpha-1).\tag{3.4}$$

The Tweedie convergence theorem, used here to justify the application of the Tweedie compound Poisson distribution to Taylor's law, is directed towards the variance functions of EDMs rather than the distributions, themselves. We have for independent and identically distributed random variables, described by EDMs with variance functions that approximate a power law at some point, these variance functions are mathematically required to converge towards the variance functions of the Tweedie EDMs, Jørgensen, Martínez and Tsao (1994). In this regard a theorem by Mora (1990) has shown that variance functions behave this way provided uniform convergence on compact sets. Moreover, because the variance functions of numerous probability distributions will approach $V(\mu)=\mu^p$ for marginal values of μ, the Tweedie EDMs would be expected to act as foci of convergence for a fairly comprehensive range of random and pseudorandom data, Kendal and Jørgensen (2011a). The term pseudorandom is applied here to indicate data that are generated by deterministic means yet appear to manifest with random properties such as a distribution function.

A population dynamic model based on the Tweedie compound Poisson distribution has been proposed to explain the manifestation of Taylor's law amongst individuals of a species, Kendal (2004b): Clusters of individuals are assumed to be randomly distributed throughout their habitat, in accordance with a Poisson distribution. The number of individuals within each cluster is assumed governed by a stochastic birth, death and immigration process, and consequently obeys a negative binomial distribution, Kendall (1949). The mean size of the clusters, overall, would then be represented by a gamma distribution, Boswell and Patil (1970). Indeed, the gamma distribution maximizes the entropy for positive valued distributions of a certain mean, providing additional insight into the role that this distribution may have with cluster size. The number of individuals per quadrat then would obey a compound Poisson distribution and Taylor's law would result.

Other statistical models can be proposed that yield Taylor's law, either approximately or exactly. Arguably, on the basis of the Tweedie convergence theorem, many of these other models would themselves be required to converge towards the form of the Tweedie EDMs.

Indeed, the impact inhomogeneity model proposed by Eisler, Bartos and Kertesz (2008) for Taylor's law, and the maximum entropy model for Taylor's law proposed by Fronczak and Fronczak (2010), have been shown to mathematically yield the Tweedie EDMs, Eisler, Bartos and Kertesz (2008) and Kendal and Jørgensen (2011a).

4 Outlying assessments for the power law exponent

As noted above, 39 (6.4%) of 608 assessments of p in this meta-analysis came within the range $0 < p < 1$. Table 1 provides detail regarding these 39 outlying cases. They were ranked here in accordance with the upper bound of their 95% CIs for p. Only 8 (1.3%) of these outlying cases had an upper bound of less than 1. (The other 31 cases could thus be considered potentially within the range of values of p for the Tweedie distributions.) Of these 8 remaining cases one was based on 5 measurements; the other cases were based on from 10 to 92 measurements. The highest value for r^2 determined from these 8 cases was 0.81, as reported for *Tribolum confusum*. In this particular case, 19% of the variance remained unexplained by the regression model. The other 7 cases had anywhere from 43% to 65% unexplained variance, indicating major spurious effects, which provided reason to question the assessed values for p.

Figure 3 provides a log-log plot of the data from *Tribolum confusum*, Naylor (1959), which arguably presented the strongest case for p falling within the range $0 < p < 1$. The solid line represents the best regression fit for Taylor's power law; the dotted line represents the equation $s_Z^2 = \bar{Z}$, an alternative model based on the Poisson distribution. There was qualitatively not a large difference between both models; the Poisson model ($p = 1$) had no adjustable parameters whereas Taylor's law had two parameters. For these reasons a Poisson model might have been more appropriate to consider rather than application of Taylor's law. Of the remaining 8 studies reviewed above, Naylor's data provided a relatively large number of cases; the remaining members of this group possessed larger degrees of unexplained variance. The meta-analysis presented so far thus appeared to indicate that spurious effects may have had a major role in causing the assessments of p to range between 0 and 1.

5 Field assessments of the Tweedie exponent

The exponent α from the Tweedie compound Poisson distribution represents one of two parameters that characterized its constituent gamma distribution (Eq. (3.1)). Equation (3.3) can be used to estimate α from the values of p within the meta-analysis, provided that $p \neq 1$. Figure 4 gives a frequency histogram for α derived from the evaluable data.

A sharply peaked distribution was evident with a modal value near $\alpha = -0.7$, which corresponded to $p = 1.6$. The peak was located close to the transition point ($\alpha = 0$, $p = 2$) where the Tweedie compound Poisson distribution is replaced by the gamma distribution. This peak also represents a mathematical critical point, since at this point the derivative with respect to α does not exist. The solid line from Figure 4 represents a fitted asymmetric Laplace distribution with the probability density function,

$$f(\alpha; \hat{\mu}, \hat{\gamma}, \hat{m}) = \frac{1}{\sqrt{2\hat{\gamma} + \hat{\mu}^2}} \exp\left\{\frac{1}{\hat{\gamma}}\left[(\alpha - \hat{m})\hat{\mu} - |\alpha - \hat{m}|\sqrt{2\hat{\gamma} + \hat{\mu}^2}\right]\right\}, \tag{5.1}$$

where $\hat{m}$ is the mode and $\hat{\mu}$ and $\hat{\gamma}$ are the mean and variance, Jørgensen and Kokonendji (2011). The values corresponding to $0 < p < 1$ appeared only as infrequent outliers in this histogram.

Table 1 *Meta-analysis*

Data sources	N	p	SE*	Upper 95% CI†	r^2
Taylor, Woiwod and Perry (1978) Total number of cases, 156	13 cases with $0 < p < 1$				
Tribolum confusum, Naylor (1959)	80	0.75	0.04	0.83	0.81
Tellina tenuis, Holme (1950)	5	0.56	0.16	0.88	N/A
Wheat straws, Bardner and Lofty (1971)	92	0.75	0.11	0.96	0.35
Dacus tyroni, Monro (1967)	10	0.47	0.22	0.98	0.4
Asphondylia, Shibuya and Ouchi (1955)	24	0.98	0.01	1	N/A
Arnica montana, Blackman (1942)	14	0.88	0.06	1.01	0.94
Agriotes, Burrage and Gyrisco (1956)	25	0.99	0.04	1.07	0.54
Wheat attacked by virus/insects, Bardner and Lofty (1971)	96	0.92	0.1	1.12	0.47
Ranuculus bulbosa, Blackman (1942)	20	0.99	0.08	1.16	0.99
Poterium sanguisorba, Blackman (1942)	20	0.99	0.1	1.2	0.85
I. grandicollis Mason (1970)	37	0.91	0.15	1.21	N/A
Sahlbergella singularis, Youdeowei (1965)	10	0.85	0.21	1.33	0.68
Surviving wheat plants, Bardner and Lofty (1971)	7	0.84	0.25	1.48	0.7
Taylor and Woiwod (1980) Total no. of cases, 360	0 cases with $0 < p < 1$				
Taylor and Woiwod (1982) Total number of cases, 92	26 cases with $0 < p < 1$				
Athene nocua	52	0.61	0.11	0.83	0.37
Picus viridis	35	0.73	0.072	0.88	0.57
Gallinago gallinago	27	0.75	0.081	0.92	0.74
Cuculus canorus	84	0.83	0.081	0.99	0.44
Turdus viscivorus	89	0.87	0.064	1	0.55
Buteo buteo	11	0.7	0.147	1.03	0.58
Numenius arquata	20	0.79	0.113	1.03	0.69
Motacilla alba	83	0.91	0.061	1.03	0.67
Sylvia atricapilla	71	0.94	0.051	1.04	0.83
Alectoris rufa	46	0.92	0.064	1.05	0.82
Garrulus glandarius	50	0.92	0.067	1.05	0.64
Pica pica	76	0.93	0.065	1.06	0.62
Emberiza citrinella	98	0.93	0.065	1.06	0.61
Perdix perdix	95	0.91	0.08	1.07	0.65
Streptopelia turtur	53	0.96	0.066	1.09	0.72
Strix aluco	48	0.93	0.081	1.09	0.59
Parus palustris	47	0.96	0.065	1.09	0.7
Pyrrhyla pyrrhula	84	0.99	0.049	1.09	0.74
Scolopax rusticola	10	0.86	0.104	1.1	0.65
Vanellus vanellus	87	0.99	0.059	1.11	0.77
Parus montanus	36	0.97	0.069	1.11	0.75
Miliaria calandra	47	0.93	0.1	1.13	0.86
Tyto alba	17	0.84	0.142	1.14	0.68
Saxicola rubetra	13	0.86	0.128	1.14	0.75
Corvus corone	95	0.98	0.11	1.2	0.56
Haematopus ostralegus	12	0.66	0.26	1.24	0.33

N, number of cases in the regression; p, Taylor's power law exponent; N/A, not available; *SE, standard error; †95% CI, 95% confidence interval

A chi-squared goodness of fit test revealed a statistically significant difference between the histogram and the theoretical asymmetric Laplace model (probability value < 0.001). This poor fit was not surprising considering that from a sample of 92 regression estimates of p from one of the component studies used in this meta-analysis, Taylor and Woiwod (1982), the percentage of unaccounted for variance was on average 25% (standard error 1%). There

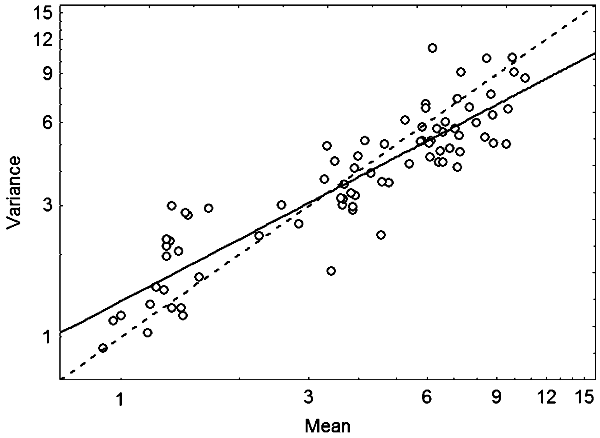

Figure 3 *Taylor's power law from Tribolum confusum,* Naylor (1959)*. Naylor provided* 80 *assessments of the variance and mean from laboratory experiments of this beetle's distribution in flour meal. The solid line represents the best fit of Taylor's law. The broken line represents the relationship expected from a Poisson distribution where the variance equaled the mean.*

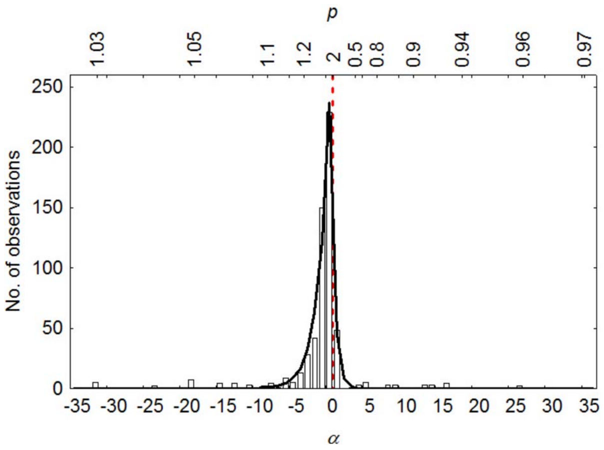

Figure 4 *Frequency histogram for the Tweedie exponent α. Meta-analysis of data reported by* Taylor, Woiwod and Perry (1978), Taylor and Woiwod (1980), *and* Taylor and Woiwod (1982) *provided* 601 *evaluable assessments for α. Their frequency histogram appeared sharply peaked and roughly corresponded to a normalized asymmetric Laplace distribution (solid line, $\hat{m} = -0.5$, $\hat{\gamma} = 1.8$, $\hat{\mu} = -0.9$). The top horizontal axis provided values for p corresponding to those of α (bottom horizontal axis). The broken line denotes the point ($\alpha = 0$, $p = 2$) where the compound Poisson distribution would be replaced by the gamma distribution.*

seemed to be a major component of statistical error associated with such assessments that might have interfered with the fit histogram.

We will see from the next section some theoretical justification for the choice of an asymmetric Laplace distribution in this application. Further theoretical justification can be derived from the placement of the mathematical critical point at $\alpha = -0.7$, just prior to the transition point where $\alpha = 0$ (Section 7). As well, previously published analyses derived from the eigenvalue deviations of the Gaussian Unitary Ensemble (GUE), the deviations of the Chebyshev prime counting function Kendal (2014), and the growth fluctuations of bristlecone pine trees Kendal (2017) revealed similar correspondences and placements of the asymmetric Laplace distribution to histograms derived from the Tweedie parameter α. For these reasons, and in spite of the poor statistical fit, this distribution could be considered on an exploratory basis to conjecture a process analogous to SOC that was associated with the aggregation of these animals.

6 The geometric Tweedie dispersion models

The asymmetric Laplace distribution is a member of a second family of Tweedie models called the geometric Tweedie dispersion models, for which geometric sums of random variables possess properties analogous to the Tweedie EDMs, Jørgensen and Kokonendji (2011). These second models are characterized by their v-function, much as EDMs are characterized by their variance function, $V(\mu)$, Jørgensen (1997) and Jørgensen and Kokonendji (2011). The geometric Tweedie dispersion models are further identified by the power law $v(\mu) = \lambda\mu^p$, with $\lambda > 0$. The asymmetric Laplace distribution is the geometric Tweedie dispersion model generated from the Laplace distribution. It has a role analogous to that of the Gaussian distribution, given its power law exponent $p = 0$, Jørgensen and Kokonendji (2011). (See also Abid, Kokonendji and Masmoudi (2019) and Abid, Kokonendji and Masmoudi (2020) for recent properties and regression models pertaining to this matter.)

Geometric dispersion models with v-functions asymptotic to $v(\mu) = \mu^p$ are mathematically required to converge towards the geometric Tweedie models in a second (and related) convergence effect, Jørgensen and Kokonendji (2011). The correspondence of an asymmetric Laplace distribution to the data from Figure 4 provided reason to conjecture that the latter convergence effect may have influenced the variation of α amongst the assessments drawn from the meta-analysis.

7 Discussion and conclusion

Taylor, Woiwod and Perry (1978) have argued that the behavior of living organisms causes their spatial distribution to be non-random. They chose to restrict their definition of randomness to spatial distributions governed by the Poisson distribution, and to consider the variance to mean power law as being non-random. As noted above, though, randomness can manifest in other ways. In particular, the Tweedie EDMs provide a formal explanation for how a variance to mean power law can be represented by a random distribution. A more general view of randomness was adopted here. Traditional count models, like those described by Kemp (1987), can also be shown to have a random origin within a discrete Poisson-Tweedie convergence theorem, Jørgensen and Kokonendji (2016): Here, the variance functions $V(\mu) = \mu^p$ and v-functions $v(\mu) = \mu^p$ are replaced by dispersion functions $v(\mu) = \mu^p$ that have the Hermite, Neyman Type A, Pólya–Aeppli, binomial, negative binomial and Poisson-Inverse Gaussian distributions as foci of convergence, analogous to the Tweedie EDMs, Jørgensen and Kokonendji (2016).

Emergent phenomena is a term applied to patterns that arise within complex systems that are characterized by scale-free or power law behaviour, Barabási and Albert (1999). SOC is a hypothesis used to explain the emergence of certain power laws like $1/f$ noise, Bak, Tang and Wiesenfeld (1988). The aggregation of plants and animals also could similarly be considered to reflect an emergent phenomenon by virtue of its apparent variance to mean power law. It is conjectured here that such emergence could reflect a form of self-organization, driven by a statistical convergence effects that govern non-Gaussian random data, which are focused primarily on the Tweedie compound Poisson and asymmetric Laplace distributions.

The meta-analysis of the exponent p, from collected studies of animals presented here, revealed that most of these values ranged between 1 and 2, and were within the domain of p associated with the compound Poisson distribution. The amount of unexplained variance with the typical field evaluations of the power law exponent p makes it difficult to exclude the application of the Tweedie EDMs because of apparent observations being in the range $0 < p < 1$.

The hypothesis of SOC postulates that the power law relationship associated with $1/f$ noise (and other power laws) represent the response of deterministic dynamical systems to the self-organized critical state, Bak, Tang and Wiesenfeld (1988). This state was based by analogy to phase transitions where at the critical point between phases dynamical variables become unstable and manifest power law behavior. Such critical effects can be associated with sharp peaks in their dynamical variables (for example, see Shan et al. (2014)).

Analysis of the exponent α in the meta-analysis provided a frequency histogram that was sharply peaked and had qualitative resemblance to an asymmetric Laplace distribution, a distribution that acts as focus of a second convergence effect. This peak was closely located to the transition point between the domains of the compound Poisson distribution and the gamma distribution, much as how similar sharp peaks in dynamical variables can be associated with phase transitions.

The evolution of such systems towards critical states could be considered an identifying feature of SOC, Bak, Tang and Wiesenfeld (1988). The hypothesis of SOC, though, was proposed for dynamical systems; that is, systems whose evolution is described by deterministic trajectories. The evidence drawn from the ecological systems presented here instead indicated that the manifestation of Taylor's law, and its associated criticality, might be more appropriately attributed to random processes.

The power law aggregation of animals and plants, as indicated by Taylor's law, is conjectured here to represent the emergence of a form of SOC from random systems governed by Tweedie convergence. Similar power law clustering found amongst other biological, physical and mathematical systems would seem to indicate a general underlying mechanism, which relates the variance to mean power law to phenomena like $1/f$ noise and multifractality, Kendal (2017).

References

Abid, R., Kokonendji, C. C. and Masmoudi, A. (2019). Geometric dispersion models with real quadratic v-functions. *Statistics & Probability Letters* **145**, 197–204. MR3873907 https://doi.org/10.1016/j.spl.2018.09.010

Abid, R., Kokonendji, C. C. and Masmoudi, A. (2020). Geometric Tweedie regression models for continuous and semicontinuous data with variation phenomenon [in reliability]. *AStA Advances in Statistical Analysis* **104**, 33–58. MR4067219 https://doi.org/10.1007/s10182-019-00350-8

Anderson, R., Gordon, D., Crawley, M. and Hassell, M. (1982). Variability in the abundance of animal and plant species. *Nature* **296**, 245–248.

Anderson, R. M. and May, R. M. (1988). Epidemiological parameters of HIV transmission. *Nature* **333**, 514–519.

Bak, P., Tang, C. and Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A* **38**, 364–374. MR0949160 https://doi.org/10.1103/PhysRevA.38.364

Barabási, A.-L. and Albert, R. (1999). Emergence of scaling in random networks. *Science* **286**, 509–512. MR2091634 https://doi.org/10.1126/science.286.5439.509

Bardner, R. and Lofty, J. (1971). The distribution of eggs, larvae and plants within crops attacked by wheat bulb fly *Leptohylemyia coarctata* (Fall.). *Journal of Applied Ecology* **8**, 683–686.

Blackman, G. E. (1942). Statistical and ecological studies in the distribution of species in plant communities 1. Dispersion as a factor in the study of changes in plant populations. *Annals of Botany London* **6**, 351–370.

Bonat, W. H., Jørgensen, B., Kokonendji, C. C., Hinde, J. and Demetrio, C. G. B. (2018). Extended Poisson–Tweedie: Properties and regression models for count data. *Statistical Modelling* **18**, 24–49. MR3746503 https://doi.org/10.1177/1471082X17715718

Bonat, W. H. and Kokonendji, C. C. (2017). Flexible Tweedie regression models for continuous data. *Journal of Statistical Computation and Simulation* **87**, 2138–2152. MR3656096 https://doi.org/10.1080/00949655.2017.1318876

Boswell, M. T. and Patil, G. P. (1970). Chance mechanisms generating negative binomial distributions. In *Random Counts in Scientific Work* (G. P. Patil, ed.), *Random Counts in Scientific Work: Expanded from the Proceedings of the Biometric Society Symposium*, *Dallas, Texas*, *December 1968*, **1**, 3–22. University Park, PA: Pennsylvania State University Press. MR0287599

Burrage, R. and Gyrisco, G. (1956). The transformation of counts of European chafer larvae for analysis of variance. *Journal of Economic Entomology* **49**, 179–182.

Eisler, Z., Bartos, I. and Kertesz, J. (2008). Fluctuation scaling in complex systems: Taylor's law and beyond. *Advances in Physics* **57**, 89–142.

Fronczak, A. and Fronczak, P. (2010). Origins of Taylor's power law for fluctuation scaling in complex systems. *Physical Review E* **81**, 066112.

Hanski, I. (1980). Spatial patterns and movements in coprophagous beetles. *Oikos* **34**, 293–310.

Holme, N. (1950). Population-dispersion in *Tellina tenuis* da Costa. *Journal of The Marine Biological Association Plymouth* **29**, 267–280.

Jørgensen, B. (1997). *The Theory of Dispersion Models*. London: Chapman & Hall. MR1462891

Jørgensen, B. and Kokonendji, C. C. (2011). Dispersion models for geometric sums. *Brazilian Journal of Probability and Statistics* **25**, 263–293. MR2832887 https://doi.org/10.1214/10-BJPS136

Jørgensen, B. and Kokonendji, C. C. (2016). Discrete dispersion models and their Tweedie asymptopics. *AStA Advances in Statistical Analysis* **100**, 43–78. MR3452285 https://doi.org/10.1007/s10182-015-0250-z

Jørgensen, B., Martínez, J. R. and Tsao, M. (1994). Asymptotic behaviour of the variance function. *Scandinavian Journal of Statistics* **213**, 223–243. MR1292637

Keeling, M. and Grenfell, B. (1999). Stochastic dynamics and a power law for measles variability. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* **354**, 769–776.

Kemp, A. W. (1987). Families of discrete distributions satisfying Taylor's power law. *Biometrics* **43**, 693–699.

Kendal, W. (2017). $1/f$ noise and multifractality from bristlecone pine growth explained by the statistical convergence of random data. *Proceedings of the Royal Society A* **473**, 20160586. MR3621489 https://doi.org/10.1098/rspa.2016.0586

Kendal, W. S. (2002). Spatial aggregation of the Colorado potato beetle described by an exponential dispersion model. *Ecological Modelling* **151**, 261–269.

Kendal, W. S. (2003). An exponential dispersion model for the distribution of human single nucleotide polymorphisms. *Molecular Biology and Evolution* **20**, 579–590.

Kendal, W. S. (2004a). A scale invariant clustering of genes on human chromosome 7. *BMC Evolutionary Biology* **4**, 3.

Kendal, W. S. (2004b). Taylor's ecological power law as a consequence of scale invariant exponential dispersion models. *Ecological Complexity* **1**, 193–209.

Kendal, W. S. (2014). Multifractality attributed to dual central limit-like convergence effects. *Physica A* **401**, 22–33. MR3294810 https://doi.org/10.1016/j.physa.2014.11.035

Kendal, W. S. (2015). Self-organized criticality attributed to a central limit-like convergence effect. *Physica A* **421**, 141–150. MR3294810 https://doi.org/10.1016/j.physa.2014.11.035

Kendal, W. S. and Jørgensen, B. (2011a). Taylor's power law and fluctuation scaling explained by a central-limit-like convergence. *Physical Review E* **83**, 066115.

Kendal, W. S. and Jørgensen, B. (2011b). Tweedie convergence: A mathematical basis for Taylor's power law, $1/f$ noise and multifractality. *Physical Review E* **84**, 066120.

Kendal, W. S. and Jørgensen, B. (2015). A scale invariant distribution of the prime numbers. *Computation* **3**, 528–540.

Kendall, D. G. (1949). Stochastic processes and population growth. *Journal of the Royal Statistical Society, Series B* **11**, 230–264. MR0034977

Kilpatrick, A. M. and Ives, A. R. (2003). Species interactions can explain Taylor's power law for ecological time series. *Nature* **422**, 65–68.

Mason, R. (1970). Comparison of flight aggregation in two species of southern ips (Coleoptera: Scolytidae). *Canadian Entomologist* **102**, 1036–1041.

Monro, J. (1967). The exploitation and conservation of resources by populations of insects. *Journal of Animal Ecology* **36**, 531–547.

Mora, M. (1990). Convergence of the variance functions of natural exponential families. *Annales de la Faculté des Sciences de Toulouse, Serie 5* **11**, 105–120. MR1191714

Naylor, A. (1959). An experimental analysis of dispersal in the flour beetle, *Tribolum confusum*. *Ecology* **40**, 453–465.

Nelder, J. A. and Wedderburn, R. W. M. (1972). General linearized models. *Journal of the Royal Statistical Society, Series A* **135**, 370–384.

Perry, J. N. (1994). Chaotic dynamics can generate Taylor's power law. *Proceedings of the Royal Society of London Series B, Biological Sciences* **257**, 221–226.

Petterle, R. R., Bonat, W. H., Kokonendji, C. C., Seganfredo, J. C., Morares, A. and Gomes da Silva, M. M. (2019). Double Poisson–Tweedie regression models. *The International Journal of Biostatistics* **15**, 20180119. MR3962660 https://doi.org/10.1515/ijb-2018-0119

Philippe, P. (1999). The scale-invariant spatial clustering of leukemia in San Francisco. *Journal of Theoretical Biology* **199**, 371–381.

Shan, C.-J., Cheng, W.-W., Liu, J.-B., Cheng, Y.-S. and Liu, T.-K. (2014). Scaling of geometric quantum discord close to a topological phase transition. *Scientific Reports* **4**, 1–6.

Shibuya, M. and Ouchi, Y. (1955). Pattern of spatial distribution of the soy bean pod gall midge in a soy bean field. *Oyo-Kontyu* **11**, 91–97.

Taylor, L. R. (1961). Aggregation, variance and the mean. *Nature* **189**, 732–735.

Taylor, L. R. (1970). Aggregation and the transformation of counts of *Aphis fabae* Scop. on beans. *Annals of Applied Biology* **65**, 181–189.

Taylor, L. R. and Taylor, R. A. J. (1977). Aggregation, migration and population mechanics. *Nature* **265**, 415–421.

Taylor, L. R., Taylor, R. A. J., Woiwod, I. P. and Perry, J. N. (1983). Behavioural dynamics. *Nature* **303**, 801–804.

Taylor, L. R. and Woiwod, I. P. (1980). Temporal stability as a density-dependent species characteristic. *Journal of Animal Ecology* **49**, 209–224.

Taylor, L. R. and Woiwod, I. P. (1982). Comparative synoptic dynamics. I. Relationships between inter- and intraspecific spatial and temporal variance/mean population parameters. *Journal of Animal Ecology* **51**, 879–906.

Taylor, L. R., Woiwod, I. P. and Perry, J. N. (1978). The density-dependence of spatial behaviour and the rarity of randomness. *Journal of Animal Ecology* **47**, 383–406. MR0600375 https://doi.org/10.2307/4114

Tweedie, M. C. K. (1984). An index which distinguishes between some important exponential families. In *Statistics: Applications and New Directions* (J. K. Ghosh and J. Roy, eds.), *Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, 579–604. Calcutta, India: Indian Statistical Institute. MR0786162

Youdeowei, A. (1965). A note on the spatial distribution of the cocoa mirid *Sahlbergella singularis* Hagl. in a cocoa farm in western Nigeria. *Nigerian Agricultural Journal* **2**, 66–67.

Department of Physics
University of Ottawa
Ottawa
Canada
E-mail: wkend037@uottawa.ca