

Using likelihood to test for Lévy flight search patterns and for general power-law distributions in nature

Andrew M. Edwards*

Pacific Biological Station, Fisheries and Oceans Canada, 3190 Hammond Bay Road, Nanaimo, British Columbia, Canada V9T 6N7, and British Antarctic Survey, High Cross, Madingley Road, Cambridge CB3 0ET, UK

Summary

1. Ecologists are obtaining ever-increasing amounts of data concerning animal movement. A movement strategy that has been concluded for a broad variety of animals is that of Lévy flights, which are random walks whose step lengths come from probability distributions with heavy power-law tails.
2. The exponent that parameterizes the power-law tail, denoted μ , has repeatedly been found to be within the Lévy range of $1 < \mu \leq 3$. Here, we use Monte Carlo simulations to show that the methods used to infer the value of μ are inaccurate.
3. The widely used method of simply logarithmically transforming a standard histogram of movement lengths has been shown elsewhere to be problematic. Here, we further demonstrate how poor it is, and show that it actually biases estimates of μ towards the Lévy range of $1 < \mu \leq 3$, and can bias estimates towards the value of $\mu = 2$. Thus, previous reports of animals undergoing Lévy flights, or of μ being close to the reported optimal value of $\mu = 2$, may simply be a consequence of the bias generated by this method.
4. A technique that has been recently recommended is to logarithmically bin the data and then normalize the resulting histogram. We show that this technique also produces biased results, and suffers from similar problems as those just outlined, although to a lesser extent.
5. The proposed solution is to use likelihood. We find that calculating the maximum likelihood estimate of μ gives the most accurate results (having also tested the rank/frequency method). Likelihood has the further advantages of being the easiest method to implement, and of yielding accurate confidence intervals. Results are applicable to power-law distributions in general, and so are not restricted to inference of Lévy flights.
6. We also re-analyse a data set of grey seal movements that was originally reported to demonstrate Lévy flight behaviour. Using Akaike weights, we test four models, and find no evidence for Lévy flights. Overall, our results suggest that Lévy flights might not be as common as previously thought.

Key-words: animal foraging, animal movement, Akaike weights, Pareto distribution, random walks

Introduction

Huge amounts of quantitative data are obtained from devices attached to animals in their natural environment (Ropert-Coudert & Wilson 2005). Technological advances in miniaturization and memory capacity, driven by consumer demand for new electronic devices, will fuel a continuing increase in the acquisition of such data (Ropert-Coudert & Wilson 2005). Here, we are concerned with data that are used

to ascertain movement of individuals, and how these data are interpreted and analysed to infer search patterns.

A recent approach is to explicitly model the dynamics of individuals using Bayesian state-space models (Jonsen, Myers & Mills Flemming 2003; Jonsen, Mills Flemming & Myers 2005; Jonsen, Myers & James 2007). This approach allows use of explicit mechanistic biological models to determine, for example, the switching of a leatherback turtle between foraging and transiting behavioural modes.

An alternative and easier approach has long been to infer potential movement rules from the shape of the distribution of movement lengths. In particular, this has led to the hypothesis that certain animals' movement patterns can be described by

*Correspondence author. Email: andrew.edwards@dfo-mpo.gc.ca, a.edwards.bas@googlegmail.com

Lévy flights, which are scale-free random walks whose step lengths come from probability distributions that have heavy power-law tails (Shlesinger, Zaslavsky & Frisch 1995; Ben-Avraham & Havlin 2000; for an excellent review of power laws in general see Newman 2005). The resulting pattern is of clusters of short moves being connected by rare long moves, and has fractal properties. Cole (1995) reported the first observation of fractal time variability in animal behaviour, which led him to run simulations suggesting that a Lévy flight might be an efficient search strategy.

The Lévy flight hypothesis then took off with evidence of such a search strategy by wandering albatrosses (Viswanathan *et al.* 1996). This was followed by a claim that a Lévy flight with an exponent of two (defined shortly) represents an optimal foraging strategy under certain conditions, together with further evidence of Lévy flight behaviour by bumblebees and deer (Viswanathan *et al.* 1999, using data from Focardi, Marcellini & Montanaro 1996 and Heinrich 1979). The optimal foraging conclusion prompted other researchers to use the methods of (or similar methods to) Viswanathan *et al.* (1999) to analyse movement data of other animals. This led to conclusions of Lévy flight search patterns for the following (the 'Lévy fleet'): reindeer (Mårell, Ball & Hofgaard 2002), side-striped jackals (Atkinson *et al.* 2002), microzooplankton (Bartumeus *et al.* 2003), grey seals (Austin, Bowen & McMillan 2004), spider monkeys (Ramos-Fernández *et al.* 2004), Peruvian fishing boats (Bertrand *et al.* 2005, 2007), honey bees (Reynolds *et al.* 2007), North Sea fishing boats (Marchal, Poos & Quirijns 2007) and human hunter gatherers (Brown, Liebovitch & Glendon 2007).

However, the conclusions of the original albatross study (Viswanathan *et al.* 1996) have been overturned, because not all of the data used actually represented flight times of albatrosses (Edwards *et al.* 2007), as originally thought (see also Travis 2007). Specifically, the misinterpreted flights turned out to be the longest flights – precisely the 'flights' that were crucial to the Lévy flight hypothesis. Analysis of the corrected data, and of similar, more recent, higher-resolution data, concluded no evidence for Lévy flight search patterns of wandering albatrosses.

The conclusions concerning deer and bumblebees have also been overturned (Edwards *et al.* 2007). In the original study (Viswanathan *et al.* 1999), the data were plotted as histograms on log–log axes, and straight lines drawn to conclude Lévy flight behaviour. Edwards *et al.* (2007) re-analysed these data using an approach involving likelihood (Hilborn & Mangel 1997) and Akaike weights (Burnham & Anderson 2002; Johnson & Omland 2004), and found that the data did not support the Lévy flight hypothesis for either deer or bumblebees (furthermore the deer data were not actually related to movement).

Sims, Righton & Pitchford (2007) helpfully explained how two methods have generally been used to ascertain the Lévy exponent. The first involves constructing a histogram of the movement data, and simply log transforming the axes (termed the LT method by Sims *et al.* 2007; techniques are described here in full in Methods and Results). A straight line

is then fitted to the resulting points, giving a slope of $-\mu$ from which the Lévy exponent μ is calculated. The second method involves constructing a histogram with logarithmic bins, normalizing the count in each bin by the bin width, and then plotting on log–log axes; this is called log-binning with normalization (termed LBN by Sims *et al.* 2007). A straight line is then fitted, again giving a slope of $-\mu$.

Sims *et al.* (2007) tested these methods with simulated data, and also tested a third method known as the cumulative distribution or rank/frequency technique (Newman 2005, here denoted as RF method). Sims *et al.* (2007) concluded that the LBN method was the best method for minimizing the errors when estimating μ .

Here, we further test these three methods, and also test a fourth, that of maximum likelihood estimation (here denoted as the MLE method). We do this by simulating data from a specified power-law distribution (i.e. a Monte Carlo approach), so that we know the true value of μ , and see how well each method recovers the true known value. We explicitly look at a range of values of μ and sample sizes.

We find that the MLE method is clearly the most accurate method of the four, and thus recommend its use. It also has the additional advantage of being the easiest to implement. In particular, we confirm the claim reported in Box 1 of Edwards *et al.* (2007) that the MLE method yields a more accurate estimate for μ than the LBN method.

Another concern is that often the estimate for μ is found, but then no goodness-of-fit tests are performed to see if the power-law conclusion is actually consistent with the data. A straight line can always be drawn, but it may not be meaningful (i.e. is a power law a sensible model in the first place?). Often, no other distributions are tested to see if they might be better models.

To demonstrate this, we apply the likelihood approach, plus the Akaike weight analysis described in Edwards *et al.* (2007), to previously published data on an Atlantic grey seal that were analysed by Austin *et al.* (2004) and then re-analysed by Sims *et al.* (2007). This particular seal was one of those for which Austin *et al.* (2004) had inferred Lévy flight behaviour, by using the LT method. We find that the Lévy flight model is not supported by the data.

The methods shown here to be inaccurate have not been confined just to Lévy flight analysis. This work is therefore applicable to other areas of ecology (and indeed other areas of physics) that are concerned with investigating power-law distributions in nature. Indeed, some of the issues covered here have been studied in other scientific contexts (see Discussion).

All calculations for the work here, plus production of the figures, were performed using R Version 2.2.0 (R Development Core Team 2007).

Methods and results

DESCRIPTION OF METHODS

If an organism performs a Lévy flight, then the probability density function (pdf), $f(x)$, of its movement lengths, x , will have a power-law tail, such that

$$f(x) \sim x^{-\mu}, \tag{1}$$

where μ is the exponent and $1 < \mu \leq 3$. The symbol \sim is used in physics to mean that $f(x)$ ‘goes like’ $x^{-\mu}$ in the tail. So the tail of the pdf decays as a power law (rather than, say, an exponential e^{-x}), and, furthermore, \sim also implies that one is not concerned with precisely defining the constant (or proportionality) term that appears in front of $x^{-\mu}$. This last, subtle, point, may have created some confusion in the ecological literature: \sim cannot simply be replaced by $=$ without explicitly defining the start of the tail (see below), yet it has been in the literature (e.g. Brown *et al.* 2007). It also precludes any test of the goodness-of-fit of the proposed pdf to the data, because the pdf is not precisely defined.

The Lévy flight hypothesis is only concerned with the tail of the pdf (i.e. the long movements); the distribution of short movements is not relevant. We define the start of the tail by a , such that we are only concerned with the distribution of movements that are $\geq a$. In empirical applications, it is often the case that the full pdf of measured movements has been considered to be a power law (e.g. Viswanathan *et al.* 1996; Mårell *et al.* 2002; Bartumeus *et al.* 2003; Ramos-Fernández *et al.* 2004; Brown *et al.* 2007; Marchal *et al.* 2007). Thus, the ‘tail’ is actually the full pdf of the data. Then a would be given by the shortest measured or measurable movement.

So, just considering movements that are $\geq a$ (which may be all movements in the data set), the pdf for x is

$$f(x) = Cx^{-\mu}, \quad x \geq a, \tag{2}$$

where C is the normalization constant given by $C = (\mu - 1)a^{\mu-1}$. The \sim has now been replaced by $=$, and the constant C explicitly calculated by solving $\int_a^\infty f(x)dx = 1$; $f(x)$ now defines a valid pdf.

It is required that $\mu > 1$; otherwise $\int_a^\infty f(x)dx = \infty$ such that $f(x)$ would

not integrate to 1 (whatever the value of C) and so $f(x)$ would not represent a pdf. [For $\mu \leq 1$, $f(x)$ decays too slowly – at the extreme of $\mu = 0$ we would have a uniform distribution, which obviously requires specification of an upper cut-off point that cannot be defined as $x \rightarrow \infty$]. For the power-law tail to represent a Lévy flight requires that $1 < \mu \leq 3$, such that the distribution has infinite variance:

$$\int_a^\infty Cx^{-\mu}x^2dx = \infty \text{ only for } 1 < \mu \leq 3. \text{ For } \mu > 3 \text{ the pdf is still a power}$$

law, but without the special properties of a Lévy flight because the variance is finite. Equation (2) is a Pareto distribution of the first kind (Johnson, Kotz & Balakrishnan 1994).

Note that μ does not depend on the unit of measurement of x ; that is, it is a dimensionless parameter. This is one reason why relatively similar values of μ are theoretically possible for the diverse range of scales associated with the aforementioned Lévy fleet. For the exponential distribution discussed later, the parameter λ does depend on the scale of measurement (and would thus change if one measured distances in metres or kilometres, unlike μ). Thus, a universal value of λ would not be possible for diverse animals.

The practical problem when analysing a particular data set is: given the data, what is the value of μ ? To be more explicit, consider a data set of values $\mathbf{x} = \{x_i\}$, with $i = 1, 2, 3, \dots, n$. Suppose we know that these values come from (2) with $a = 1$, and are each independent from one another. We wish to find the value of μ that is most supported by the data. Determination of μ has been carried out by the aforementioned four methods that are now described in detail.

Figure 1a shows a standard histogram for an example simulated data set of 1000 independent random numbers drawn from the pdf

$$f(x) = x^{-2}, \quad x \geq 1; \tag{3}$$

that is, eqn (2) with $a = 1, \mu = 2$. This is the distribution concluded in the original wandering albatross study (Viswanathan *et al.* 1996).

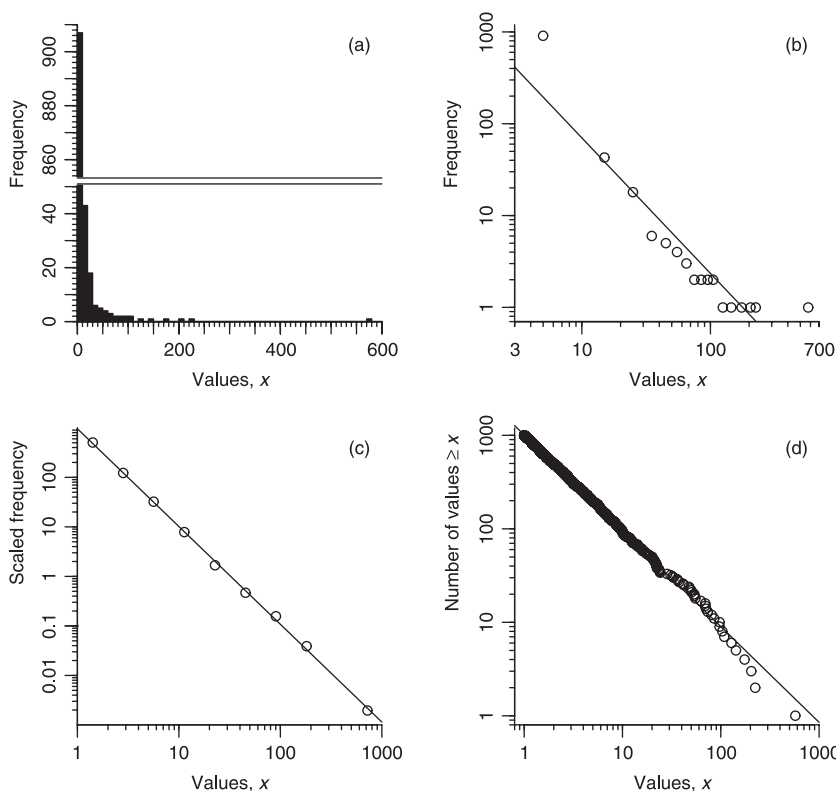


Fig. 1. Example of 1000 values drawn from the power law in eqn (3), for which $a = 1, \mu = 2$. (a) Simulated data plotted as a standard histogram, with bin breaks at 0, 10, 20, ... , as determined by the Scott (1979) algorithm. Note the break in the y-axis because 907 of the 1000 values fall in the first bin (the expected proportion of such values is $\int_1^{10} x^{-2}dx = 0.9$). The few large values of x are indicative of the heavy power-law tail. (b) The LT (log transform) method for determining μ essentially plots the points from (a) on logarithmic axes, and fits a straight line, yielding the poor estimate of $\mu = 1.476$. (c) The LBN (log-binning with normalization) method yields $\mu = 1.975$. (d) The RF (rank/frequency) method yields $\mu = 2.025$. The non-graphical MLE method yields $\mu = 2.019$.

Random numbers were obtained using the inverse method (also called the transformation method, Newman 2005): if u is a random number drawn from the uniform distribution over the range $[0, 1]$, then $x = a(1 - u)^{1/(1-\mu)}$ is a random number from the power-law distribution (2). We now explain how each of the four methods was applied to this data set, to estimate the value of μ from the 1000 data points.

The LT method (Fig. 1b) involves computing a standard histogram, whereby each bin has the same width. The choice of bin width is often not stated. Here, we followed Bertrand *et al.* (2005, 2007) and Marchal *et al.* (2007) in using the Scott (1979) algorithm to obtain the bin width (the same algorithm was used for Fig. 1a). The resulting midpoints of each bin and the counts for each bin are then plotted on logarithmic (base 10) axes. So $\log_{10}(\text{count in bin } i)$ was plotted against $\log_{10}(\text{midpoint of bin } i)$, such that Fig. 1b is essentially Fig. 1a on logarithmic axes. A straight line is then fitted by linear regression, for which the slope is assumed to be $-\mu$, yielding an estimate for μ . The LT method was originally used by Viswanathan *et al.* (1999) to analyse the deer data, and a 'smoothed' version used for the bumblebee data (although lines were sketched rather than fitted). Using the LT method for our sample of 1000 simulated values, we obtained $\mu = 1.476$, clearly a very poor estimate for the true known value of $\mu = 2$.

To clarify a potential source of confusion, note that the slope $-\mu$ is calculated on logarithmic axes for which the axes are labelled by the log of frequency or log of x [i.e. the ordinate of Fig. 1b would be labelled $\log_{10}(\text{frequency})$ with numbers 0, 1, 2 and 3, instead of 1, 10, 100 and 1000]. But axes are labelled here using unlogged numbers (1, 10, 100 and 1000) to aid interpretation and make it easier to relate the diagrams in Fig. 1 to each other. The same applies for the LBN and RF methods.

The LBN method (Fig. 1c) involves setting the bin breaks such that each bin is twice the width of the preceding bin (although, if desired, bin widths could be chosen to, say, triple in size). Here, since $a = 1$, we used bin breaks of 1, 2, 4, 8, 16, 32, 64, ... , as used by Viswanathan *et al.* (1996). The count in each bin is then divided by the width of that bin (e.g. Pueyo 2006), to yield a frequency density for each bin, which can be intuitively thought of as a count per unit interval of x (Newman 2005). Viswanathan *et al.* (1996) used such a normalization (Edwards *et al.* 2007), although this was not explicitly stated. The frequency density for each bin is then plotted against the geometric midpoint of that bin (so for the bin covering 512–1024, this is $\sqrt{(512 \times 1024)} \approx 724$). A linear regression on the plotted points then gives an estimate of μ (as the negative of the slope). The LBN method for the simulated data set here gave $\mu = 1.975$, performing far better than the LT method. The LBN method is recommended by Sims *et al.* (2007).

Note that in Fig. 1c, the right-most point represents the bin 512–1024, and thus the count in this bin is 1, as only the value 574 falls in this bin (see Fig. 1a). Yet in the linear regression, this point is as important as the other points in determining the slope of the line. Also, the penultimate bin is empty, covering the range 256–512, but cannot be plotted on Fig. 1c as zero cannot be plotted on logarithmic axes.

The RF method (Fig. 1d) involves ranking the values in order, with 1 being the largest. The number of data points $\geq x_i$ (the rank of x_i) is then plotted against x_i , for each data value x_i . Logarithmic axes are used, and a straight line fitted to the points. The slope of the straight line is then assumed to be $1 - \mu$, from which the estimate for μ is obtained. As Newman (2005) pointed out, such plots are useful as they show all the data points, and avoid any (often arbitrary) determination of bins. Here, we ask whether fitting lines to such plots is of use when determining μ . For the simulated data, we

obtained $\mu = 2.025$, which, for this data set, is as close to the true $\mu = 2$ as the estimate from the LBN is.

Finally, the MLE method first requires determination of the likelihood function (Hilborn & Mangel 1997; Burnham & Anderson 2002; Crawley 2002). Given known data \mathbf{x} , and assuming a to be known, the log-likelihood function for the power-law tail (Newman 2005; Edwards *et al.* 2007) is

$$\log[L(\mu | \text{data } \mathbf{x})] = n \log(\mu - 1) + n(\mu - 1) \log a - \mu \sum_{j=1}^n \log x_j, \quad (4)$$

where $L(\mu | \text{data } \mathbf{x})$ is the likelihood of a particular value of the unknown parameter μ given the known data \mathbf{x} (and log is natural log). The MLE for μ , denoted $\hat{\mu}$, is the most likely value for μ given the data \mathbf{x} , and is given by

$$\hat{\mu} = 1 - \frac{n}{n \log a - \sum_{j=1}^n \log x_j}. \quad (5)$$

Note that if $a = 1$, this simplifies further to

$$\hat{\mu} = 1 + \frac{1}{\text{mean}(\log \mathbf{x})}. \quad (6)$$

Presented with a particular data set \mathbf{x} , the MLE $\hat{\mu}$ is far easier to compute than any of the graphical methods that involve fitting straight lines to the data plotted on log-log axes – just insert the data values into eqn (5) or eqn (6). If the full data set is assumed to come from a power-law distribution but a is not known a priori, then the MLE for a is the minimum value of the data.

For the simulated data set, the MLE method yielded $\hat{\mu} = 2.019$, with a 95% confidence interval (computed using the profile likelihood-ratio test, Hilborn & Mangel 1997) of (1.957, 2.083). Thus, based on the particular simulated data set shown in Fig. 1a, the MLE method appears best, but only by a small margin. Newman (2005) also found the MLE method to be accurate for a single simulated data set.

ACCURACY OF METHODS

The four estimates of μ just discussed were calculated for the single simulated data set of 1000 random numbers drawn from eqn (3). The results obviously depend on the particular 1000 random numbers. So we repeated the above calculations on 10 000 independent simulated samples, each of 1000 random numbers (10 000 is the number recommended by Crawley 2002, and used by Sims *et al.* 2007). Thus, for each of the four methods, we obtained 10 000 estimates for μ , each of which depended on the particular 1000 random numbers drawn each time from the power law with known exponent $\mu = 2$. Thus, if observational data truly come from a power law with the theoretically optimal exponent of $\mu = 2$, we asked how well would each of the four methods estimate the value of μ ?

Figure 2 shows the resulting histograms of 10 000 estimates for μ , one histogram for each method. Summary statistics are shown in Table 1.

The LT method is clearly very poor (Fig. 2a). All 10 000 estimates of μ fell below the true value of $\mu = 2$ (Table 1). Goldstein, Morris & Yen (2004) found similar inaccuracies for the discrete equivalent of our continuous power law (known as the zeta distribution or discrete Pareto distribution). Sampling 10 000 random numbers from the discrete Pareto distribution with known exponent $\gamma = 2.5$, and repeating this 50 times, they found the LT method to give a poor estimate of $\gamma = 1.59 \pm 0.184$ (standard error).

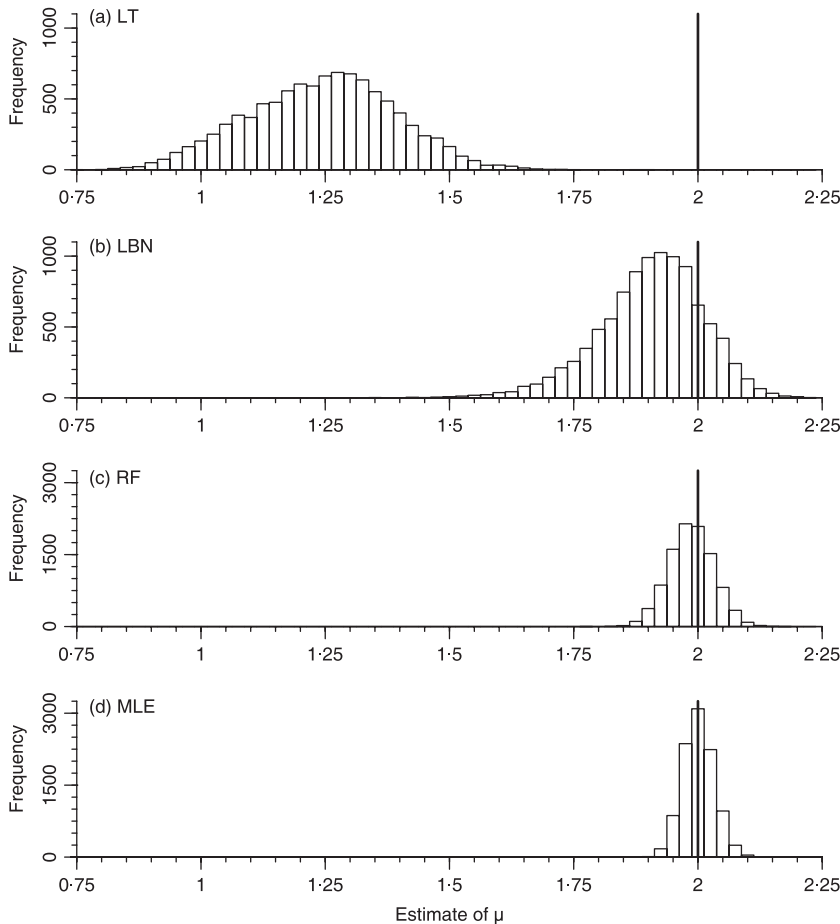


Fig. 2. Histograms of estimates for μ from 10 000 simulated data sets, each data set containing 1000 independent random numbers drawn from the power law (3) that has known $\mu = 2$. The true value of $\mu = 2$ is shown as a thick vertical line. Results are for the four methods: (a) LT, (b) LBN, (c) RF and (d) MLE. Note the change in vertical scales between (a), (b) and (c), (d). Summary statistics are shown in Table 1. The MLE method clearly gives the best estimates for μ , given the true value $\mu = 2$.

Method	1st quartile	Median	Mean	3rd quartile	Percentage of estimates below true value of $\mu = 2$
LT	1.14	1.25	1.24	1.35	100.00%
LBN	1.84	1.91	1.90	1.98	82.62%
RF	1.956	1.986	1.986	2.017	61.87%
MLE	1.979	2.000	2.001	2.022	49.91%

Table 1. Statistics for the 10 000 simulations with known value $\mu = 2$ for each of the four methods, as shown in Fig. 2. Results are given to two or three decimal places as appropriate depending on closeness of the mean and median to 2

For the LBN method (Fig. 2b), μ was generally underestimated, with 83% of the estimates below the true value of $\mu = 2$ (Table 1). There was also a large spread of estimated values. Thus, although the mean of the μ estimates was 1.90, which is only 5% below the true value, this method still performed poorly and is severely biased.

Sims *et al.* (2007) stated that the imposition of a maximum movement length in their simulated data introduced a slight bias for all methods. However, no maximum movement length was imposed here, and the LBN method is still biased.

The RF method performed better (Fig. 2c), with a narrow spread of estimates around a mean of 1.986 (Table 1). However, 62% of values still fell below the true value of $\mu = 2$.

The MLE method had an even narrower spread of estimated values (Fig. 2d), around a mean of 2.001 (Table 1). And 49.9% of estimates fell below the true value, and 50.1% above. Thus, the MLE method here yielded an unbiased estimate for μ , and also has the

minimum spread of the four estimation techniques. The MLE method clearly outperforms the others.

Furthermore, 95% confidence intervals were obtained for the MLE of μ using the likelihood profile method (Hilborn & Mangel 1997). Of the 10 000 simulations, 95.06% of the 95% confidence intervals contained the true value of $\mu = 2$. Thus, the 95% confidence intervals were also accurate.

Goldstein *et al.* (2004) also used an MLE method for their aforementioned discrete Pareto distribution, although for the discrete distribution an analytical solution equivalent to (5) cannot be found. They found the true power-law exponent of $\gamma = 2.5$ to be estimated as $\gamma = 2.500 \pm 0.017$ (SD), also showing excellent accuracy.

The results presented in Fig. 1 suggested that the LBN, RF and MLE methods gave estimates of similar accuracy. But the results of Table 1 and Fig. 2 show that this was a particular characteristic of the specific simulated data set of 1000 numbers in Fig. 1.

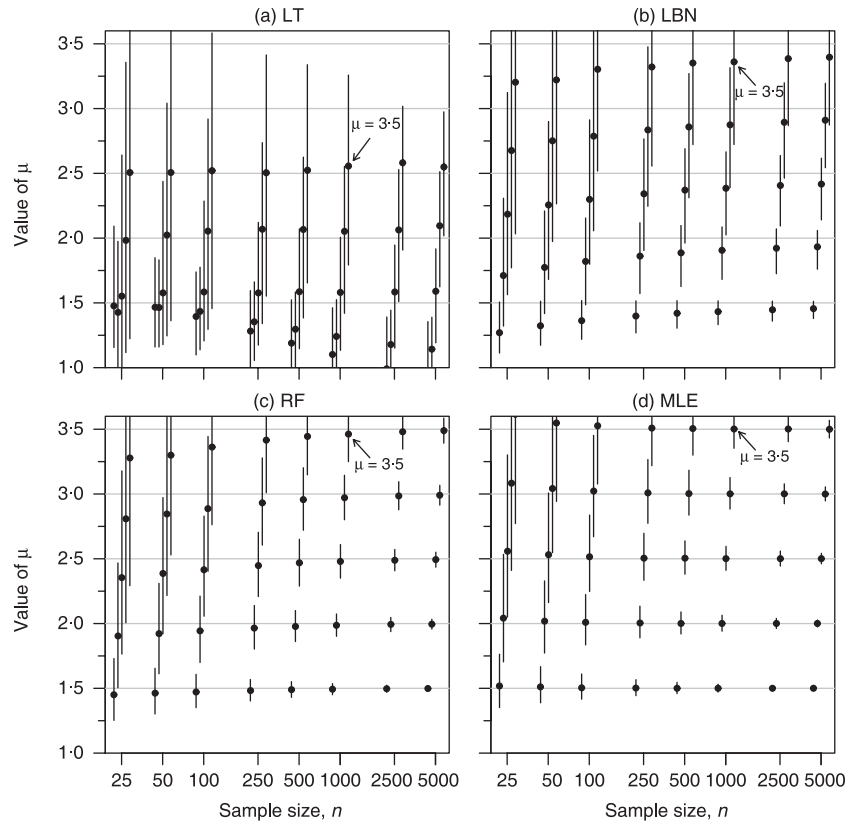


Fig. 3. Further simulations similar to Fig. 2 for different combinations of μ and n , using each method: (a) LT, (b) LBN, (c) RF and (d) MLE. Horizontal grey lines represent the known μ values used (namely 1.5, 2, 2.5, 3 and 3.5), and the tick marks on the logarithmic x -axis indicate the sample sizes n . For a given combination of μ and n , n random numbers were sampled from the power law of exponent μ , and μ was estimated from the n random numbers. This was repeated 10 000 times (as in Fig. 2 for $n = 1000$, $\mu = 2$), with the mean of the 10 000 estimates indicated by a black solid circle on the diagram for the method used. The vertical bars indicate the range within which lay 95% of the estimated values. Thus, there is one black circle with a vertical line for each of the 40 combinations of μ and n (for clarity, circles and lines are offset slightly from their prescribed value of n). For illustration, the arrows indicate the mean value of μ estimated by each method for $n = 1000$ and a true $\mu = 3.5$. Only for the MLE method does the circle lie on the $\mu = 3.5$ grey line. Overall, the MLE method is clearly best.

SENSITIVITY TO EXPONENT μ AND SAMPLE SIZE n

Just as the results of Fig. 1 were shown to be dependent on the particular simulated data sets of 1000 random numbers, the results of Table 1 and Fig. 2 may be dependent on the particular values of $\mu = 2$ and $n = 1000$. So we repeated the above analyses for a range of μ and n values.

Figure 3 shows the dependence of the results on the prescribed values of μ and n , one diagram for each method. Results are shown for five values of μ : 1.5, 2, 2.5, 3 and 3.5. In the Supplementary material, results for further values of μ are shown.

For clarity of explanation (as the results are the best and so the diagram is easiest to interpret), first consider Fig. 3d for the MLE method. The five μ values tested are indicated by the five horizontal grey lines. Eight sample sizes, n , were used: 25, 50, 100, 250, 500, 1000, 2500 and 5000 – a sample size of only 25 was the case for the high-food bumblebee data set analysed by Viswanathan *et al.* (1999). These eight sample sizes are indicated by the values shown on the logarithmic x -axis.

Thus, there are 40 combinations of μ and n values. For a given combination, n random numbers were sampled from the power law with exponent μ , and the MLE method used to then estimate μ from the generated data (assuming $a = 1$ to be known). This was repeated 10 000 times, yielding 10 000 estimates of μ for the given combination of μ and n . At the corresponding value of n , the mean value of the 10 000 estimates of μ was then plotted as a black circle in Fig. 3d. A vertical bar was then added, to indicate the range within which 95% of the estimated μ values fell (the endpoints of the bars represent the 2.75% and 97.25% quantiles). For each sample size n , this was then repeated for each of the other values of μ . For visual clarity (to stop the bars overlapping each other), the circles and bars are slightly horizontally offset from the precise n value, with the

left-most (or bottom) circle representing a true value of $\mu = 1.5$, the second-from-left circle representing $\mu = 2$, etc. Thus, there are eight values of n , and each uses five values for μ , yielding 40 combinations in all giving 40 black circles and vertical bars (although the circle for $n = 25$ with $\mu = 3.5$ lies just outside of the plotted range).

The previous results in Fig. 2d showed that the MLE method was accurate for $\mu = 2$, $n = 1000$. This can be seen in Fig. 3d at $n = 1000$, by the second black dot from the bottom. The black dot represents the mean of the 10 000 simulations, and clearly lies on the horizontal line at $\mu = 2$. And the shortness of the vertical bar indicates that 95% of the estimated μ values indeed lie close the true value of $\mu = 2$.

It can be seen that as sample size n is increased, the MLE value remains unbiased (the black circles remain on the horizontal $\mu = 2$ line), and the 95% range of estimates becomes narrower. This is expected – a larger sample size yields a more accurate estimate for μ . As n decreases, the 95% range becomes larger (as expected), and the mean estimate for μ becomes slightly positively biased. This general result is seen across all values of μ .

The same 400 000 simulated data sets from Fig. 3d were also used to estimate μ by each of the other three methods (there were 10 000 data sets for each of the 40 combinations of μ and n). The next most accurate method is the RF method shown in Fig. 3c. The mean estimates (black circles) are not as accurate as for the MLE method. Even at large sample sizes n , the higher μ values are slightly underestimated. For sample sizes $n = 25$ and 50, the estimates when $\mu = 3.5$ (the highest black circle for each value of n) are close to $\mu = 3.25$. Across all combinations of μ and n , the 95% ranges are larger than those for the MLE method (as was suggested in Figs 2c and d for $\mu = 2$ with $n = 1000$).

The LBN method (Fig. 3b) underestimates the true value of μ across the full ranges of μ and n . The underestimation originally

shown in Fig. 2b and Table 1, for $\mu = 2$ and $n = 1000$, is not markedly improved with an increase in sample size.

For the LT method (Fig. 3a), the estimated μ values are so poor (as expected from Fig. 2a), that most of the black circles lie a long way from the corresponding horizontal line that represents the true μ value that the circles should be estimating. This is true even for large sample sizes. The LT method should clearly never be used.

The inaccurate methods directly impact the biological conclusions. For the LBN method and non-Lévy value of $\mu = 3.5$ (top circle for each value of n in Fig. 3b), the 95% range overlaps the Lévy range $\mu \leq 3$, even for $n = 5000$. And for the borderline value of $\mu = 3$, the estimated μ value, and a large majority of the 95% range, lie in the Lévy range $\mu \leq 3$. The LBN method therefore biases the estimate of μ towards the Lévy range.

Thus, the LBN method, the preferred method of Sims *et al.* (2007), will often falsely classify data generated from a non-Lévy power law ($\mu > 3$) as being a Lévy flight ($1 < \mu \leq 3$). Therefore, an organism whose movelengths really come from a pdf with a non-Lévy power-law tail with $\mu > 3$, can be incorrectly identified by the LBN method as undergoing Lévy flights.

This problem is clearly more pronounced for the LT method in Fig. 3a. For a sample size of $n = 5000$, and $\mu = 3$ or 3.5, the entire 95% range of estimated μ values lies within the Lévy range. For the RF method, with $\mu = 3.5$ the 95% range reaches well into the Lévy range ($\mu \leq 3$) for the smaller sample sizes.

The downward biasing of estimates of μ from the LT and LBN methods also means that estimated values of μ that come from power laws for which the true $\mu > 2$, will be biased towards the theoretically optimal value $\mu = 2$. Thus, the empirical evidence for μ values close to 2 may actually be a consequence of biasing due to the methods. Alternatively, for a true value of $\mu = 2$ the inferred reported value will generally be < 1.5 for the LT method and < 2 for the LBN method. But given that many reported values are close to 2, it seems that the overall consequence of the downward biasing will have been to exaggerate reports of the number of data sets with $\mu = 2$ (if indeed the data are from power laws).

Empirical re-analysis

The above analysis was concerned with correctly estimating the power-law exponent μ . But this does not tell us if the power law is a good description of the data in the first place. Here we apply modern methods of model selection (Burnham & Anderson 2002) to an empirical data set.

Austin *et al.* (2004) used a variety of techniques to analyse movement data that they had collected concerning grey seals [*Halichoerus grypus* (Fabricius, 1791)]. The seals were fitted with satellite-relay data loggers when they were on Sable Island, a vegetated sand bar 300 km from the coast of Nova Scotia, Canada. A total of 15 333 useable locations from 52 seals were obtained – this gives an example of the aforementioned huge data sets that are emerging from animal monitoring studies, and demonstrates the need for mathematical and statistical tools to analyse such data.

As a small part of their study, Austin *et al.* (2004) tested data for individual seals to see if the seals might exhibit Lévy flights. Following the example of Viswanathan *et al.* (1999), Austin *et al.* (2004) analysed their data using the LT method. Only 9 of the 27 daily sampled seals appeared to fit the Lévy

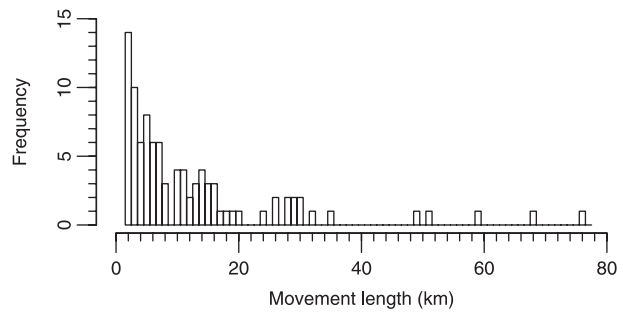


Fig. 4. True raw data from seal 6118 (from D. Austin), with bin breaks at 1.5, 2.5, 3.5, ... km. Move lengths < 1.5 km are not shown, as Austin *et al.* (2004) considered only move lengths ≥ 1.5 km in their analysis (only 10 moves were < 1.5 km).

flight model. In particular, for their seal 6118, Austin *et al.* (2004) concluded a Lévy flight with $\mu = 1.26$.

Sims *et al.* (2007) noted that the erroneous LT method had been used. Focussing on seal 6118, they also noted that the relationship between $\log(\text{frequency})$ and $\log(\text{movement length})$ – see Austin *et al.*'s Fig. 3a or Sims *et al.*'s Fig. 4b – is 'not linear, indicating the distribution of steps is unlikely to follow a power law'. The approach described by Edwards *et al.* (2007) provides a more formal way to judge quantitatively whether the data do indeed follow a power law or not, by comparing the strength of evidence for the power law with the evidence for other distributions.

Sims *et al.* (2007) digitized the data from Austin *et al.*'s Fig. 3a and re-analysed the data using the LBN method, obtaining $\mu = 0.80$. They found that their re-analysis 'provides no evidence for the presence of Lévy flights in the data (e.g. $\mu < 1$) and confirms the absence of a power-law distribution.'

Here, we also present a re-analysis of the data for seal 6118, by first using the binned form of the data that was shown in Fig. 3a of Austin *et al.* (2004). Thus, we used the same information as used by Sims *et al.* (2007). (Actual binned values were obtained from D. Austin, as were the raw unbinned data values that are analysed later). We tested the evidence for four models – PL, the standard Lévy flight model of a power-law distribution over the range $[a, \infty)$; Exp, a standard exponential distribution over $[a, \infty)$; PLB, a power law bounded over the range $[a, b]$, also known as the truncated Pareto distribution (Johnson *et al.* 1994); and ExpB, an exponential distribution bounded over the range $[a, b]$. These four models were tested for the deer and bumblebee data by Edwards *et al.* (2007) (see that reference for the four equations), who included the bounded distributions because of the negligible support for the conventional PL model. A standard exponential distribution describes the time intervals between events of a simple Poisson random process (Grimmett & Stirzaker 1990), and so represents an obvious model to test (as also mentioned by Cole 1995). Following Austin *et al.* (2004), the value $a = 1.5$ was used for all models, and $b = 15.5$ used for the bounded models because it represents the maximum value of the highest bin in Austin *et al.*'s Fig. 3a.

Maximum likelihood estimates for μ (PL and PLB) or λ (Exp and ExpB) were calculated numerically, using the likelihood functions for binned data derived in Supplementary methods 4 and 5 of Edwards *et al.* (2007). The 95% confidence interval for each MLE was calculated using the profile likelihood-ratio test (Hilborn & Mangel 1997).

For each model and MLE, Akaike's information criterion (Hilborn & Mangel 1997; Burnham & Anderson 2002; Johnson & Omland 2004) was calculated, as was the Akaike weight (Burnham & Anderson 2002; Edwards *et al.* 2007). The number of parameters being estimated for each model (K_i in Box 1 of Edwards *et al.* 2007) was one (just μ) for the unbounded PL and Exp models, and two (μ and b) for the bounded PLB and ExpB models.

The Akaike weight, w_i , for model i is considered to be the weight of evidence in favour of model i being the best model for the data, out of the four models being considered. Thus, the model with the highest weight is most supported by the data. Here it is clearer, since we have four models, to show the evidence ratio for each model, which is defined (Burnham & Anderson 2002) as w_{best}/w_i , where w_{best} is the Akaike weight for the best model (highest weight). Thus, the best model has an evidence ratio of 1, and the other models have evidence ratios > 1 .

Our re-analysis ('Reported data' in Table 2) finds the MLE of $\mu = 1.82$ (95% CI: 1.66–1.99) for the unbounded PL model, compared to Austin *et al.*'s reported value of 1.26, and Sims *et al.*'s LBN-computed value of 0.80. Thus, Austin *et al.*'s value of 1.26, using the erroneous LT method, is actually closer to the MLE than Sims *et al.*'s recalculation using the LBN method.

But for the PLB model, the MLE is $\mu = 0.87$ (95% CI: 0.57–1.17), which is closer to Sims *et al.*'s LBN-computed value. This apparent paradox arises because the graphical LBN method does not distinguish between bounded and unbounded distributions (i.e. between PLB and PL). For a bounded distribution, we assume that the data are restricted to being less than the largest observed value (a conservative assumption that will favour Akaike weights for a bounded power law over a bounded exponential – see Edwards *et al.* 2007). For an unbounded distribution, the assumption is that larger values are possible but were not observed in this particular data set. For this latter case, we essentially have

counts of zero in the higher bins. But such zero counts cannot be plotted on log-log axes and so the bounded and unbounded cases cannot be distinguished by the LBN method.

The evidence ratios show that the ExpB model is most supported by the data, and is fairly convincingly supported over the next best model (PLB). The unbounded PL and Exp models are not supported at all, with the PL model by far the worst model of the four.

As Goldstein *et al.* (2004) noted, 'the parameter estimate of a power-law exponent has very limited meaning without some assessment of its goodness-of-fit'. A goodness-of-fit test (G -test, or likelihood-ratio test, with Williams's correction; Sokal & Rohlf 1995) for the ExpB model with $\lambda = 0.16$ shows that the reported data are indeed consistent with coming from this model ($n = 96$, degrees of freedom = 7, $G = 8.6$, $p = 0.28$).

Thus, we conclude that the data do not appear to follow a power law (either bounded or unbounded), but that the bounded exponential model is an acceptable model.

We analysed the data as plotted in Austin *et al.* (2004) in order to use the same information as Sims *et al.* (2007) in their analysis. However, during the re-analysis it became apparent that the original grey seal data were plotted incorrectly in Fig. 3a of Austin *et al.* (2004) (D. Austin, personal communication). The numbers on the x -axis actually indicated bin number, not movement length as reported. For example, bin number 6 represents movement lengths in the range 7.5–10.0 km rather than the range 6.5–7.5 km indicated in the figure. Only the histograms and log-log plots in Fig. 3 of Austin *et al.* (2004), and their associated calculations, are incorrect – the rest of the figures and results are not affected (D. Austin, personal communication).

So now we analyse the true raw data from Austin *et al.* (2004) corresponding to the correct movement lengths for seal 6118, as shown in the spatial plot of Austin *et al.*'s Fig. 3b, and as a standard histogram in Fig. 4.

First, in Fig. 5a, we use the LBN method to estimate μ , thus following Sims *et al.*'s preferred approach but using the true data. Some curvature is apparent in Fig. 5a, suggesting a straight line (i.e. a power law) may not be an excellent fit. The slope gives a value of $\mu = 1.59$, with a high r^2 of 0.94.

Using the likelihood approach (Table 2), we find that the LBN value of $\mu = 1.59$ just discussed is actually very close to

Table 2. Re-analysis of movement length data of grey seal 6118. 'Reported data' are the data as plotted in Fig. 3a of Austin *et al.* (2004), for which $\mu = 1.26$ was calculated by Austin *et al.* using the LT method, and $\mu = 0.80$ was calculated by Sims *et al.* (2007) using the LBN method. 'True data' are the corrected raw data set (see text), with the value $\mu = 1.59$ calculated here using the LBN method. Four models were tested here using the likelihood approach: PL (power law unbounded), Exp (exponential unbounded), PLB (power law bounded) and ExpB (exponential bounded), and the maximum likelihood estimate (MLE) for μ or λ calculated as appropriate. Bold represents the model with the evidence ratio of 1.0 (i.e. the model which is most supported by the data).

Data set	Reported μ	PL: MLE for μ (95% CI)	Exp: MLE for λ (95% CI)	PLB: MLE for μ (95% CI)	ExpB: MLE for λ (95% CI)	Evidence ratios			
						PL	Exp	PLB	ExpB
Reported data	1.26, 0.80	1.82 (1.66, 1.99)	0.22 (0.18, 0.27)	0.87 (0.57, 1.17)	0.16 (0.11, 0.22)	$> 10^{15}$	240	17	1.0
True data	(1.59)	1.60 (1.49, 1.73)	0.090 (0.073, 0.109)	1.25 (1.07, 1.43)	0.089 (0.072, 0.108)	$> 10^6$	2.2	1.0	5.3

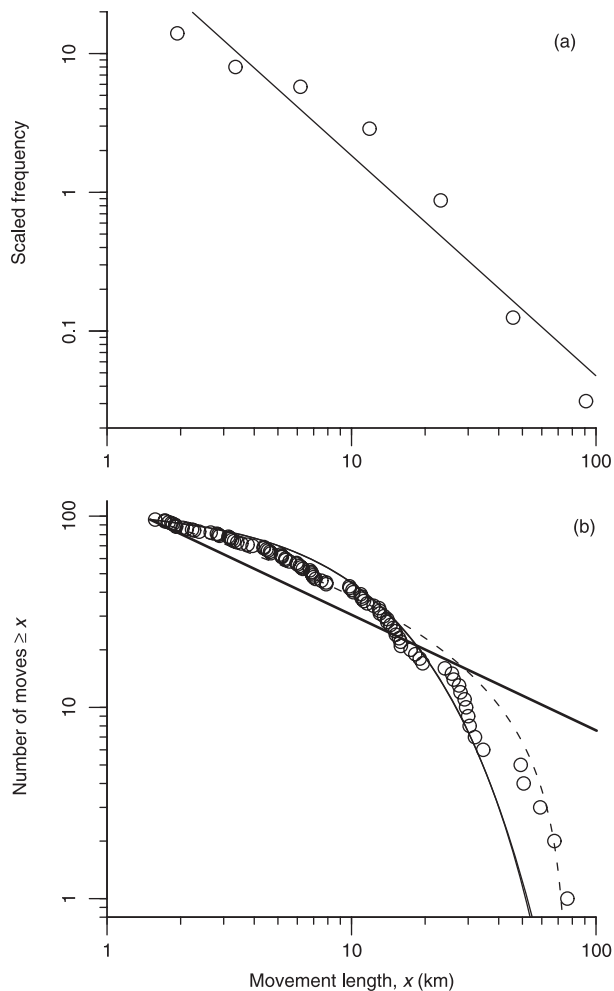


Fig. 5. (a) True move lengths from Fig. 4 analysed using the LBN method. Bin breaks are at 1.5, 2.5, 4.5, 8.5, 16.5, 32.5, 64.5 and 128.5. The linear regression fit to the points yields the estimate $\mu = 1.59$. (b) Rank/frequency plot of the true move lengths. The four models (using their respective MLEs) from Table 2 are: PL (thick straight line), Exp (thin curved line), PLB (dashed curve line), ExpB (thin curved line, which overlaps almost perfectly with the Exp model). In agreement with the evidence ratios, the PL model is a very poor model, despite appearing to be not too bad in (a).

the MLE value of 1.60 for the unbounded PL model. However, for the 'Reported data', the LBN value was close to the MLE for the PLB model. It may not be so surprising that such an inconsistency arises because the LBN method does not distinguish between the PL and PLB models.

For the true data, the evidence ratios show that the PLB model has the most support (Table 2). But it is not convincingly the best (Burnham & Anderson 2002), because the Exp model has an evidence ratio of only 2.2, and the ExpB model of only 5.3. Again, the standard Lévy flight (PL model) is completely unsupported by the data. This is despite the LBN plot in Fig. 5a giving an apparently good fit.

Figure 5b shows the true data plotted as a rank/frequency plot, together with the curves for the four models of Table 2 as fitted by their respective MLEs. Note that the PLB model on

the rank/frequency plot is not a straight line, demonstrating that a straight-line fit would have no chance of working for this model. Figure 5b shows that the PL model is clearly very poor. The other models do not look so bad, with the PLB model looking the best, as ascertained (but not conclusively) by the evidence ratios.

However, the goodness-of-fit test (G -test, as above) for the PLB model finds that the data are *not* consistent with this model ($n = 96$, degrees of freedom = 9, $G = 19.7$, $p = 0.020$). So, while the PLB model is indeed the best of the four models tested, it is not a satisfactory model. Further distributions or hypotheses would need to be considered. Hence, we conclude that none of the four models tested are suitable for the data, and, in particular, the grey seals are not exhibiting Lévy flight movement patterns. The breaks in Fig. 5b suggest that no single distribution will describe the data, and thus that more detailed models and biological interpretation would be required.

Discussion

Using simulated data generated from known power-law distributions, we have tested four methods that have previously been used to estimate the power-law exponent μ . The MLE method is clearly superior to the other methods, and we therefore recommend its use. It has the added benefits of being easy to compute, without having to worry about the binning issues that plague the LT and LBN methods, and of providing proper construction of confidence intervals.

The LT method has been used in many studies (see Sims *et al.* 2007; Edwards *et al.* 2007). Pueyo & Jovani (2006) and Sims *et al.* (2007) showed that this method is erroneous, and the extensive analysis here shows that it never gives the correct answer, even for large sample sizes. In fact, it is unclear, from a theoretical perspective, why it was ever used in the first place. Sometimes, in other scientific areas, the bins containing the scarce large values in the tail have been removed to improve the estimate of μ , but this is an *ad hoc* approach with no guarantee of accuracy.

Unfortunately, the LT method has been shown here to severely bias estimates of μ towards the Lévy flight range (and often towards the reportedly optimal $\mu = 2$). The consequence of such bias is that researchers have been misled as to the biological meaning of their analyses.

The LBN method seems more warranted from a theoretical viewpoint, yet the fact that some points represent many more data values than other points remains problematic. The LBN method for the true grey seal data in Fig. 5a resulted in a high r^2 of 0.94, yet it was then shown that the power law (both unbounded and bounded) was not a suitable model for the data. Similar r^2 values have been reported elsewhere (e.g. Bartumeus *et al.* 2003). But r^2 is related to the amount of variance explained by the regression line (e.g. Crawley 2002), and is not a statistical goodness-of-fit test of the power-law distribution.

Sims *et al.* (2007) did conclude by advocating the use of likelihood methods in addition to plotting approaches. Here,

we have shown that likelihood methods are the most accurate for ascertaining power-law exponents. Regarding plotting, it is indeed useful to plot the data in the different ways described by Sims *et al.* (2007), as in Fig. 1 (and Edwards *et al.* 2007). One of the reasons for the different plotting methods is that if data do come from a power law over several orders of magnitude, then a standard histogram (linear axes) poorly represents the data. But we do not recommend using any subsequent regression fits to determine the power-law exponent.

Benhamou (2007) discussed further issues concerning Lévy flights that have not been addressed in this paper. For example, there is the general problem of being careful about inferring a process from a pattern. Also see Travis (2007) for Benhamou's and others' thoughts concerning Lévy flights.

The MLE and Akaike weight approaches used here are applicable to the determination of power laws in other areas of ecology, and even other areas of science (e.g. understanding human memory; Rhodes & Turvey 2007). White, Enquist & Green (2008) have also recently looked at the issue of estimating power-law exponents, giving examples of power-law distributions being inferred in areas of biology as diverse as sizes of islands and fluctuations in metabolic rate. White *et al.* (2008) tested the performance of fitting methods on simulated data (our complementary studies were performed simultaneously but independently). Their conclusions agree with those derived here, namely that MLE is the most accurate method and therefore should be used. Furthermore, White *et al.* (2008) looked at the effect of changing the bin sizes for the binning methods, and tested the binning methods with the lowest counts excluded (as is sometimes carried out). Such refinements could slightly improve the accuracy of the methods as reported here, although we agree with White *et al.*'s comment that there are many conceivable rules that could be applied, yet 'no amount of tweaking' will produce a consistent binning method.

Beyond ecology, similar and related issues to those examined here have been investigated previously, also leading to recommendations to use likelihood over graphical methods (e.g., Clark, Cox & Laslett 1999 in geology). White *et al.* (2008) noted that the bias of the MLE for μ seen at small sample sizes (see their Fig. 2a and Fig. 3d here) can be eliminated – see Johnson *et al.* (1994) and Clark *et al.* (1999). Clauset, Shalizi & Newman (unpublished, see <http://arxiv.org/abs/0706.1062v1>) present a wide-ranging overview, including a discussion of why the methods involving regression on log–log plots violate the assumptions of linear regression, and give further technical properties of the MLE for μ (see also Newman 2005 and Johnson *et al.* 1994). They also re-analysed 24 data sets using likelihood, and suggest that some power-law behaviour might be 'merely a figment of the researcher's imagination'.

Many ecological studies track animals moving vertically (e.g. Watwood *et al.* 2006) or horizontally (e.g. Pinaud & Weimerskirch 2007; Phalan *et al.* 2007). Manipulation experiments are also performed (e.g. Luschi *et al.* 2001) to try and reveal the navigational cues used by animals. A complicating factor is that movement will be a consequence of an animal's

own movement plus any drift due to, say, ocean currents (e.g. Luschi, Hays & Papi 2003) or wind (e.g. Green *et al.* 2004). Deducing the underlying rules driving animal movement will continue to require novel quantitative approaches.

This purposefully methodological paper has shown some of the pitfalls that can arise when analysing movement data. Testing has shown the likelihood method to perform very well, and Akaike weights allow for proper selection between models, as advocated by Burnham & Anderson (2002). Robust conclusions regarding animal movement and behaviour rely on correctly analysing data. Such data are expensive, difficult and time-consuming to collect. It is hoped that the work presented here will help ecologists reach meaningful and accurate conclusions about the animals that they study.

Acknowledgements

I am very grateful to Debbie Austin for providing grey seal data and thank her plus Nicholas Watkins, Mervyn Freeman, Gary Abel and Ryan Woodard for many useful discussions. I appreciate Ransom Myers' interest in the early stages of this work, and regret that he was not able to see its fruition – his influence on those that knew and worked with him will continue for many years. I thank Ethan White, Graeme Hays and an anonymous referee for insightful constructive comments and additional references. This work commenced while I was part of the Natural Complexity Programme at the British Antarctic Survey, funded by the UK Natural Environment Research Council.

References

- Atkinson, R.P.D., Rhodes, C.J., Macdonald, D.W. & Anderson, R.M. (2002) Scale-free dynamics in the movement patterns of jackals. *Oikos*, **98**, 134–140.
- Austin, D., Bowen, W.D. & McMillan, J.I. (2004) Intraspecific variation in movement patterns: modeling individual behaviour in a large marine predator. *Oikos*, **105**, 15–30.
- ben-Avraham, D. & Havlin, S. (2000) *Diffusion and Reactions in Fractals and Disordered Systems*. Cambridge University Press, Cambridge, UK.
- Bartumeus, F., Peters, F., Pueyo, S., Marrasé, C. & Catalan, J. (2003) Helical Lévy walks: adjusting searching statistics to resource availability in microzooplankton. *Proceedings of the National Academy of Sciences USA*, **100**, 12771–12775.
- Benhamou, S. (2007) How many animals really do the Lévy walk? *Ecology*, **88**, 1962–1969.
- Bertrand, S., Bertrand, A., Guevara-Carrasco, R. & Gerlotto, F. (2007) Scale-invariant movements of fishermen: the same foraging strategy as natural predators. *Ecological Applications*, **17**, 331–337.
- Bertrand, S., Burgos, J.M., Gerlotto, F. & Atiquipa, J. (2005) Lévy trajectories of Peruvian purse-seiners as an indicator of the spatial distribution of anchovy (*Engraulis ringens*). *ICES Journal of Marine Science*, **62**, 477–482.
- Brown, C.T., Liebovitch, L.S. & Glendon, R. (2007) Lévy flights in Dobe Ju/'hoansi foraging patterns. *Human Ecology*, **35**, 129–138.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer, New York.
- Clark, R.M., Cox, S.J.D. & Laslett, G.M. (1999) Generalizations of power-law distributions applicable to sampled fault-trace lengths: model choice, parameter estimation and caveats. *Geophysical Journal International*, **136**, 357–372.
- Cole, B.J. (1995) Fractal time in animal behaviour: the movement activity of *Drosophila*. *Animal Behaviour*, **50**, 1317–1324.
- Crawley, M.J. (2002) *Statistical Computing: An Introduction to Data Analysis using S-Plus*. John Wiley & Sons Ltd., Chichester, UK.
- Edwards, A.M., Phillips, R.A., Watkins, N.W., Freeman, M.P., Murphy, E.J., Afanasyev, V., Buldyrev, S.V., da Luz, M.G.E., Raposo, E.P., Stanley, H.E. & Viswanathan, G.M. (2007) Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer. *Nature*, **449**, 1044–1048.
- Focardi, S., Marcellini, P. & Montanaro, P. (1996) Do ungulates exhibit a food density threshold? A field study of optimal foraging and movement patterns. *Journal of Animal Ecology*, **65**, 606–620.

- Goldstein, M.L., Morris, S.A. & Yen, G.G. (2004) Problems with fitting to the power-law distribution. *European Physical Journal B*, **41**, 255–258.
- Green, M., Alerstam, T., Gudmundsson, G.A., Hedenström, A. & Piersma, T. (2004) Do Arctic waders use adaptive wind drift? *Journal of Avian Biology*, **35**, 305–315.
- Grimmett, G.R. & Stirzaker, D.R. (1990) *Probability and Random Processes*. Oxford University Press, Oxford, UK.
- Heinrich, B. (1979) Resource heterogeneity and patterns of movement in foraging bumblebees. *Oecologia*, **40**, 235–245.
- Hilborn, R. & Mangel, M. (1997) *The Ecological Detective: Confronting Models with Data*, Vol. 28, Monographs in Population Biology, Princeton University Press, Princeton, New Jersey.
- Johnson, N.L., Kotz, S., & Balakrishnan, N. (1994) *Continuous Univariate Distributions*, Vol. 1, 2nd edn. John Wiley and Sons, New York.
- Johnson, J.B. & Omland, K.S. (2004) Model selection in ecology and evolution. *Trends in Ecology & Evolution*, **19**, 101–108.
- Jonsen, I.D., Mills Flemming, J. & Myers, R.A. (2005). Robust state-space modeling of animal movement data. *Ecology*, **86**, 2874–2880.
- Jonsen, I.D., Myers, R.A. & James, M.C. (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Marine Ecology Progress Series*, **337**, 255–264.
- Jonsen, I.D., Myers, R.A., & Mills Flemming, J. (2003) Meta-analysis of animal movement using state-space models. *Ecology*, **84**, 3055–3063.
- Luschi, P., Åkesson, S., Broderick, A.C., Glen, F., Godley, B.J., Papi, F., & Hays, G.C. (2001) Testing the navigational abilities of ocean migrants: displacement experiments on green sea turtles (*Chelonia mydas*). *Behavioral Ecology and Sociobiology*, **50**, 528–534.
- Luschi, P., Hays, G.C., & Papi, F. (2003) A review of long-distance movements by marine turtles, and the possible role of ocean currents. *Oikos*, **103**, 293–302.
- Marchal, P., Poos, J.-J., & Quirijns, F. (2007) Linkage between fishers' foraging, market and fish stocks density: Examples from some North Sea fisheries. *Fisheries Research*, **83**, 33–43.
- Mårell, A., Ball, J.P., & Hofgaard, A. (2002) Foraging and movement paths of female reindeer: insights from fractal analysis, correlated random walks, and Lévy flights. *Canadian Journal of Zoology*, **80**, 854–865.
- Newman, M.E.J. (2005) Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, **46**, 323–351.
- Phalan, B., Phillips, R.A., Silk, J.R.D., Afanasyev, V., Fukuda, A., Fox, J., Catry, P., Higuchi, H. & Croxall, J.P. (2007). Foraging behaviour of four albatross species by night and day. *Marine Ecology Progress Series*, **340**, 271–286.
- Pinaud, D. & Weimerskirch, H. (2007) At-sea distribution and scale-dependent foraging behaviour of petrels and albatrosses: a comparative study. *Journal of Animal Ecology*, **76**, 9–19.
- Pueyo, S. (2006) Diversity: between neutrality and structure. *Oikos*, **112**, 392–405.
- Pueyo, S. & Jovani, R. (2006) Comment on 'A keystone mutualism drives pattern in a power function'. *Science*, **313**, 1739 c.
- R Development Core Team (2007) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0
- Ramos-Fernández, G., Mateos, J.L., Miramontes, O., Cocho, G., Larralde, H. & Ayala-Orozco, B. (2004) Lévy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). *Behavioral Ecology and Sociobiology*, **55**, 223–230.
- Reynolds, A.M., Smith, A.D., Menzel, R., Greggers, U., Reynolds, D.R., & Riley, J.R. (2007). Displaced honey bees perform optimal scale-free search flights. *Ecology*, **88**, 1955–1961.
- Rhodes, T. & Turvey, M.T. (2007). Human memory retrieval as Lévy foraging. *Physica A*, **385**, 255–260.
- Ropert-Coudert, Y. & Wilson, R.P. (2005) Trends and perspectives in animal-attached remote sensing. *Frontiers in Ecology and the Environment*, **3**, 437–444.
- Scott, D.W. (1979) On optimal and data-based histograms. *Biometrika*, **66**, 605–610.
- Shlesinger, M.F., Zaslavsky, G.M. & Frisch, U., eds (1995) *Lévy Flights and Related Topics in Physics*. Springer-Verlag, Berlin and Heidelberg, Germany.
- Sims, D.W., Righton, D.R., & Pitchford, J.W. (2007) Minimizing errors in identifying Lévy flight behaviour of organisms. *Journal of Animal Ecology*, **76**, 222–229.
- Sokal, R.R. & Rohlf, F.J. (1995) *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn, W. H. Freeman and Company, New York.
- Travis, J. (2007) Do wandering albatrosses care about math? *Science*, **318**, 742–743.
- Viswanathan, G.M., Afanasyev, V., Buldyrev, S.V., Murphy, E.J., Prince, P.A., & Stanley, H.E. (1996) Lévy flight search patterns of wandering albatrosses. *Nature*, **381**, 413–415.
- Viswanathan, G.M., Buldyrev, S.V., Havlin, S., da Luz, M.G.E., Raposo, E.P., & Stanley, H.E. (1999) Optimizing the success of random searches. *Nature*, **401**, 911–914.
- Watwood, S.L., Miller, P.J.O., Johnson, M., Madsen, P.T., & Tyack, P.L. (2006) Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*). *Journal of Animal Ecology*, **75**, 814–825.
- White, E.P., Enquist, B.J., & Green, J.L. (2008) On estimating the exponent of power-law frequency distributions. *Ecology*, **89**, 905–912.

Received 12 December 2007; accepted 18 April 2008
Handling Editor: Graeme Hays

Supplementary material

The following supplementary material is available for this article online:

Appendix S1. Sensitivity to exponent μ and sample size n – full results

Fig. S1. Full results for the LT method.

Fig. S2. Full results for the LBN method.

Fig. S3. Full results for the RF method.

Fig. S4. Full results for the MLE method.

This material is available as part of the online article from:
<http://www.blackwell-synergy.com/doi/full/10.1111/j.1365-2656.2008.01428.x>

(This link will take you to the article abstract).

Please note: Blackwell Publishing is not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.