Truncated Lévy walks are expected beyond the scale of data collection when correlated random walks embody observed movement patterns

A. M. Reynolds*

Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, UK

Translating observations taken at small spatio-temporal scales into expected patterns at greater scales is a major challenge in spatial ecology because there is typically insufficient relevant information. Here, it is shown that truncated Lévy walks are the most conservative, maximally non-committal description of movement patterns beyond the scale of data collection when correlated random walks characterize observed movements and when there is partial information about landscape and behavioural heterogeneity. This provides a new conceptual basis for Lévy walks that is divorced from optimal searching theory and free from the difficulties with discerning their presence in empirical data.

Keywords: correlated random walks; Lévy walks; maximum entropy; non-oriented animal movement patterns; landscape and behavioural heterogeneity

1. INTRODUCTION

For many years, correlated random walk (CRW) models have been the dominant conceptual framework for describing non-oriented animal movement patterns ([1] and references therein). In these models, an individual’s trajectory through space is regarded as being made up of a sequence of distinct, independent randomly oriented displacements. To account for the observed tendency of many animals to maintain their direction of travel, a tendency known as ‘directional persistence’, turning angles are typically drawn at random from a unimodal distribution that is peaked around 0°. The width of the distribution determines the angular diffusivity (K) of the walker. Straight-line movements correspond to K = 0. Larger values of K are representative of meandering and circuitous movement patterns.

Patterns of movement do change over time as animals encounter different habitats and engage in different activities. Each mode of movement can be taken to correspond to a different parametrization of the CRW model [2–9], i.e. to a different value of K. Movement patterns at large spatio-temporal scales can then be simulated by occasionally selecting a new value of K from the distribution of modes pK(K). Difficulties arise when the available observational data are not sufficient to parametrize pK(K) accurately. In these cases, the principle of scientific objectivity dictates that we be maximally uncommitted about our knowledge concerning the distribution pK(K). The most conservative, non-committal pK(K) that is consistent with the data (e.g. with estimates for the mean angular diffusivity) is obtained by maximizing Shannon’s differential entropy [10–12]. Any other distribution would assume more information than is known from the data. In this context, Shannon’s differential entropy, \( H = -\int K p_K(K) \log_2 p_K(K) dK \), is a measure of the average surprise of seeing an animal in a particular movement mode (K), given a distribution of modes pK(K). A highly improbable outcome is very surprising. If there are two movement modes (K₁ and K₂), then the entropy is zero when there is no uncertainty, i.e. when pK(K₁) = 1 and pK(K₂) = 0 or when pK(K₁) = 0 and pK(K₂) = 1. It is maximized when pK(K₁) = 1 and pK(K₂) = 1/2 as there is less uncertainty when pK(K₁) ≠ pK(K₂) because then one or other of the modes is more likely to be seen.

Here, using the maximum entropy principle, we show that truncated Lévy walk (LW) movement patterns are the most conservative, maximally non-committal description of movement patterns at large spatio-temporal scales when CRW embody observed movements and when there this is partial information about landscape and behavioural heterogeneity in the form of estimates for some of the lower order moments of K (e.g. when there are estimates the mean or the mean and variance of K). Other models of movement patterns at large spatio-temporal scales would assume more information than is known from the data. This finding broadens considerably the significance of the recent realization that CRW across heterogeneous landscapes can generate LW characteristics [13].

LW arose in a purely mathematical/statistical context in the first half of the last century [14]. They comprise clusters of relatively short step lengths with...
longer steps between them. This pattern is repeated across all scales. The resultant patterns are fractal and have no characteristic scale. Over much iteration, a LW will be distributed much further from its starting position than a Brownian walk of the same length (same sum of step lengths), because there are relatively more long steps and so less meandering. The distribution of step lengths has an inverse power-law tail that is characterized by a power-law exponent (a Lévy exponent, $\mu$) that falls between 1 and 3. LW first entered the literature on animal movement patterns when it was proposed that LW characteristics may be observed in foraging ants [15]. Animal movement patterns cannot, of course, be true LW of indefinite extent. Nonetheless, some movement patterns could be modelled as truncated LW. In truncated LW, power-law scaling does not extend to arbitrarily large scales but instead is either cut off or gives way to exponential scaling. Viswanathan et al. [16] subsequently showed that LW movement patterns can be advantageous in random search scenarios and suggested that animals with LW movement pattern characteristics may be executing an innate, evolved searching strategy. This insight has provided a conceptual basis for interpreting many instances of LW movement patterns [17]. Escalating empirical support for LW movement patterns (reviewed in [18]) subsequently fountained when methodological shortcomings in some approaches to determining the goodness of fits to movement pattern data was identified [19]. A few more recent empirical studies have, however, provided seemingly compelling evidence of truncated LW in the flight patterns of honeybees [20–22] and in the dive patterns of a diverse range of marine predators [23].

The findings presented here provide a new conceptual basis of LW that is divorced from optimal search scenarios and suggested that animals with LW movement pattern characteristics may be executing an innate, evolved searching strategy. This insight has provided a conceptual basis for interpreting many instances of LW movement patterns [17]. Escalating empirical support for LW movement patterns (reviewed in [18]) subsequently fountained when methodological shortcomings in some approaches to determining the goodness of fits to movement pattern data was identified [19]. A few more recent empirical studies have, however, provided seemingly compelling evidence of truncated LW in the flight patterns of honeybees [20–22] and in the dive patterns of a diverse range of marine predators [23].

The findings presented here provide a new conceptual basis of LW that is divorced from optimal search scenarios and suggested that animals with LW movement pattern characteristics may be executing an innate, evolved searching strategy. This insight has provided a conceptual basis for interpreting many instances of LW movement patterns [17]. Escalating empirical support for LW movement patterns (reviewed in [18]) subsequently fountained when methodological shortcomings in some approaches to determining the goodness of fits to movement pattern data was identified [19]. A few more recent empirical studies have, however, provided seemingly compelling evidence of truncated LW in the flight patterns of honeybees [20–22] and in the dive patterns of a diverse range of marine predators [23].

The findings presented here provide a new conceptual basis of LW that is divorced from optimal search scenarios and suggested that animals with LW movement pattern characteristics may be executing an innate, evolved searching strategy. This insight has provided a conceptual basis for interpreting many instances of LW movement patterns [17]. Escalating empirical support for LW movement patterns (reviewed in [18]) subsequently fountained when methodological shortcomings in some approaches to determining the goodness of fits to movement pattern data was identified [19]. A few more recent empirical studies have, however, provided seemingly compelling evidence of truncated LW in the flight patterns of honeybees [20–22] and in the dive patterns of a diverse range of marine predators [23].

2. METHODS

2.1. Theoretical analysis

In a CRW model considered here, incremental changes in the position $(x, y)$ of a walker made during time-steps of duration, $\Delta t$, are given by $\Delta x = s \cos (\phi(t)) \Delta t$ and $\Delta y = s \sin (\phi(t)) \Delta t$, where $s$ is the speed of the walker and $\phi(t)$ is its direction of travel at time, $t$. A new direction of travel, $\phi(t) = \phi(t - \Delta t) + \theta(t)$, is chosen at the start of each time-step. Here, for the sake of simplicity, turning angles, $\theta(t)$, are drawn at random from a wrapped Gaussian distribution with mean zero and variance $2\pi \Delta t$. This focus on Gaussian statistics is not overly restrictive because the more commonly used von Mises distribution yields analogous results (not shown).

The connection between the CRW modelling and the truncated LW is established in three steps. First, simulated movement patterns for each mode $K$ are represented by sequences of movement bouts or ‘moves’ that are made between successive substantial changes in the direction of travel. Smaller changes in the direction of travel are neglected because they are taken to be representative of responses to environmental ‘micro-cues’ or random meanderings across rough terrain. Second, the distribution of move lengths, $p(l|K)$, for each movement mode $K$ is found. Finally, these different distribution functions are aggregated to form the overall distribution move lengths, $P(l)$, at large spatio-temporal scales. This final step is achieved by combining $p(l|K)$ with the most conservative form of the distribution of modes $p_K(K)$ given reliable estimates for the lower order moments of the angular diffusivity $K$. It will be shown that $P(l)$ has a truncated inverse-square power-law tail. The movement bouts therefore correspond to the steps in a truncated ($\mu = 2$ LW) movement pattern and so the overall movement pattern is truncated ($\mu = 2$ LW). It is subsequently shown that the dispersal characteristics of CRW models with maximally non-committal modes and truncated $\mu = 2$ LW have common dispersal characteristics. This complements the longstanding and widely held belief that CRW movement patterns are diffusive and Gaussian at large spatio-temporal scales [24,25].

2.1.1. Step 1. Movement patterns for a fixed mode $K$ were produced by numerically integrating the CRW model. These movement patterns were then represented by sequences of moves. A move starting at time $t$ is taken to end when the cumulative turning angle $|\phi(t + \tau) - \phi(t)|$ first exceeds some critical value $\Delta \phi$, (figure 1).

2.1.2. Step 2. Move lengths were found to be exponentially distributed (figure 2a). The goodness of fits of the simulation data to exponential distribution functions were tested using an approach advocated by Clauset et al. [26], which is related to the Anderson–Darling test [27]. The simulation data are unlikely to follow an exponential exactly as there will always be some small deviations because of the random nature of the sampling process. The approach of Clauset et al. [26] distinguishes deviations of this type from those that arise because the data are, in actuality, not exponentially distributed. This is achieved by sampling many synthetic datasets from a true exponential distribution, measuring how far they fluctuate from the exponential form, and then comparing the results with similar measurements on the simulation data. If the simulation dataset is much further from the exponential than the typical synthetic ones, then the exponential is not a plausible fit to the data. Implementation of the approach is straightforward. First, the area between the cumulative distribution function of the simulation data and that of the best-fit exponential model is calculated: $A = \sum_{i=1}^{N} |C_{\text{Data}}(x_i) - C_{\text{Model}}(x_i)|$. Synthetic datasets of size $N$ are then created by extracting random numbers from the best-fit model. Each synthetic dataset is then fitted with its own exponential model, and these models are evaluated using the area statistic. The synthetic dataset is therefore compared with their own best-fit models and not with the original simulation data, i.e. the calculation that was performed on the simulation data was also performed on the synthetic data. The proportion of values of these area
later with the aid of numerical simulations. The effect of such truncations is considered not lead to significant numbers of moves being truncated. This constant does not play a significant role in the subsequent analysis. Therefore, without loss of generality and for the sake of simplicity, attention is hereafter focused on $\Delta \phi = 90^\circ$. In this case, $\lambda = K/s$ (figure 2b).

Note that the property of $p_l(l|K)$ of being an exponential distribution would be an explicit model assumption if it is assumed that individual steps of the CRW are exponentially distributed rather than constant lengths.

2.1.3. Step 3. Over sufficiently long periods of time, the overall distribution of move lengths, $P(l)$, will have contributions from all possible modes, each weighted by the likelihood of the animal being in that mode. Mathematically, this aggregation is expressed by the convolution,

$$P(l) = \int_0^\infty p_l(l|K)p_K(K)\,dK. \quad (2.1)$$

This presupposes that switching between modes does not lead to significant numbers of moves being truncated. The effect of such truncations is considered later with the aid of numerical simulations.

Many candidate choices for $p_K(K)$ can be constructed from the observational data but the most conservative, maximally non-committal distribution, is the one that maximizes Shannon’s differential entropy. In the simplest case, observational data may furnish an estimate for the mean angular diffusivity $\langle K \rangle = \eta^{-1}$. The method of Lagrange multipliers can be used to maximize Shannon’s differential entropy under the constraints that probabilities are normalized (sum to unity) and that $\langle K \rangle = \eta^{-1}$ [28]. This amounts to maximizing the functional

$$-\int_K p_K(K) \log_e p_K(K) + \lambda p_K(K) + \beta p_K(K)\,dK, \quad (2.2)$$

where the undetermined multipliers $\lambda$ and $\beta$ are determined from the two constraints. Differentiating equation (2.2) with respect to $p_K(K)$ and equating the derivative to zero, leads to

$$\log_e p_K(K) + 1 + \lambda + \beta K = 0 \quad (2.3)$$

Figure 2. (a) Cumulative frequency distribution of simulated move lengths produced by numerically integrating the CRW model with $K = 1/2$ and $s = 1$ (arbitrary units) over $10^6$ time-steps. The critical angle defining a turn $\Delta \phi = 90^\circ$. A straight-line on this log-linear plot is indicative of exponential scaling and here such scaling is clearly evident. (b) The maximal-likelihood estimates for $\lambda$ as a function of the angular diffusivity $K$ (filled circle). The line is added to guide the eye. Shown for comparison are the maximal-likelihood estimates for $\lambda$ as a function of $K$ when $\Delta \phi = 45^\circ$ (open circle). Each data point was obtained by integrating the CRW model over $10^6$ time-steps.
or
\[ p_\kappa(K) = e^{-(1+\lambda^2)K} \exp(-\eta K). \] 
(2.4)

The normalization and the constraint that \( \langle K \rangle = \eta^{-1} \) together determine \( \beta \) and \( \lambda \) leading to \( p_\kappa(K) = \eta \exp(-\eta K) \). This distribution implies that animals spend a significant proportion of their time moving with arbitrarily small turning angles. Substitution of the entropy maximizer \( p_\kappa(K) = \eta \exp(-\eta K) \) into equation (2.1) gives
\[ P(l) = \frac{l_0}{(l + l_0)^2}, \] 
(2.5)
where \( l_0 = \eta s \). The inverse-square power-law tail of this distribution is indicative of \( \mu = 2 \) LW movement patterns.

If there are reliable estimates for the first \( n \) moments of \( K \), then the entropy maximizer is given by \( p_\kappa(K) = \exp(\sum_{n=0}^\infty a_n K^n) \), where the constants \( a_0, a_1, \ldots, a_n \) are determined by the requirements that the distribution sums to unity and has the known moments [28]. Given such a suite of information about habitat structure and behavioural heterogeneity, the most conservative, maximally non-committal distribution of move lengths will once again have an inverse-square power-law tail. Similarly, if angular diffusivities are known to be bounded (i.e. are known to range from \( k_{\text{min}} \) to \( k_{\text{max}} \)), but otherwise unknown, then the maximally non-committal move length distributions will have inverse-square power-law tails. In this case, the entropy maximizer, \( p_\kappa(K) = 1/(K_{\text{max}} - K_{\text{min}}) \) [28].

2.2. Dispersal characteristics

Long-time net displacements, \( x \), of walkers in CRW in mode \( K \) will be Gaussian-distributed with mean zero and variance \( \sigma_x^2 = (\bar{s}^2/K)t \) [29]. This, of course, is not representative of ballistic (straight-line) movements associated with \( K = 0 \). Ballistic movements are characterized by \( \sigma_x^2 = \bar{s}^2 t^2 \). Nonetheless, the distribution of long-time net displacements made by CRW with maximally non-committal models can be approximated by the convolution:
\[ P(x, t) = \int_0^\infty \frac{1}{\sqrt{2\pi} \sigma_x} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) p_\kappa(K) dK, \] 
(2.6)
which is just the weighted average of Gaussian distribution functions: one Gaussian distribution function for each mode \( K \), each weighted by the expected likelihood of the animal being in that mode. Equation (2.6) is approximate because it has been assumed that switching between modes does not lead to significant numbers of moves being truncated and because the Gaussian distributions wrongly imply that some walkers have travelled distances in excess of \( st \) in a time \( t \).

The validity of this approximation is examined in §2.3, where the results of numerical simulations are reported. For the simplest entropy maximizer, \( P_\kappa(K) = \eta \exp(-\eta K) \), equation (2.6) yields
\[ P(x, t) = \frac{\eta}{2\sqrt{2\pi} \bar{s} t} \left(\frac{x^2}{2\bar{s}^2 t + \eta}\right)^{-3/2}. \] 
(2.7)

In §2.3, it will be shown that equation (2.7) provides a very good representation of simulation data for long-time displacements that are less than \( st \).

2.3. Numerical simulations

In the foregoing theoretical analysis, switching between movement modes during the execution of a move rather than only after completing a move was not accounted for. Such switching will inevitably lead to truncation of the distribution of move lengths. As a consequence the net displacement distribution will eventually transition from equation (2.7) to a Gaussian by virtue of the central limit theorem. Gaussian net displacements are indicative of diffusion. Such a movement pattern is illustrated in figure 3 together with the associated distribution of move lengths. The truncated inverse power-law distribution of move lengths (figure 3c) indicates that this movement pattern can be modelled as a truncated LW.

Simulated movement patterns are initially super-diffusive but do eventually become Gaussian and diffusive in character (figure 4). The effective, long-time, diffusion rate depends on the distribution of angular diffusivities as well as on the switching rate. This is analogous to the results obtained by Skalski & Gilliam [30] for switching between two CRW models.

The dispersal characteristics of CRW models with exponentially distributed modes (i.e. with maximally non-committal modes; figure 4) closely resemble that of truncated LW per se shown in figure 5, and this resemblance serves to further illustrate the close correspondence between these two patterns of movement. Figure 5 shows that truncated \( \mu = 2 \) LW movement patterns are superdiffusive up to scales comparable with the truncation scale but at larger scales are diffusive and Gaussian.

3. DISCUSSION

Extrapolating from observational scales to much greater scales is a major challenge of spatial ecology because animal behaviour changes over time as animals encounter different habitats and engage in different activities [5,31]. To scale from limited observations to the landscape, we must understand how to aggregate and simplify, retaining essential information without getting encumbered by unnecessary detail. In such an analysis, natural scales and frequencies may emerge, and in these rests the essential nature of the system dynamics [32]. Even though there is no single natural scale at which movement patterns should be studied there could be scaling laws that describe how pattern changes across scale [31]. This is the major lesson of the theory of fractals [33].

Many candidate models of movement patterns at large spatio-temporal scales will be consistent with observations. However, the most conservative one will
Figure 3. (a) An example movement pattern produced by the CRW model with exponentially distributed modes. The mean angular diffusivity $K = 0.1$, $s = 1$ and $\Delta t = 0.01$ (arbitrary units). New values of $K$ are selected at random at the end of each move or after executing a move for 500 time-steps. (b) Representation of the simulated movement pattern in terms of straight-line ‘moves’ ($\Delta \phi = 90^\circ$). (c) Survival function (the complement of the cumulative distribution function) of move lengths for the aforementioned movement pattern simulated over $10^6$ time-steps (solid line). Shown for comparison is the predicted cumulative frequency distribution of move lengths in the absence of switching (dashed line). The survival function associated with the predicted probability distribution, equation (2.5), is given by $l_0/(l + l_0)$.

Figure 4. (a) Simulation data for mean square displacements, $R^2$, produced by the CRW model with exponentially distributed modes. The mean angular diffusivity $K = 1$, $s = 1$ and $\Delta t = 0.01$ (arbitrary units; solid line). The data were produced by ensemble-averaging $10^4$ model runs each with a constant angular diffusivity. Switching between modes is taken to occur on much longer timescales. The asymptotic ballistic, $R^2 \propto t^2$, and diffusive, $R^2 \propto t$, scaling behaviours are indicated (dashed lines). The temporal extent of the superdiffusive movements does not change significantly when the mean angular diffusivity is decreased or increased by a factor of 10. (b) Simulation data for the distribution of displacements after $10^4$ time-steps (jagged line) together with the analytical prediction (equation (2.7); smooth line). (c) Dispersion at $10^5$ time-steps (jagged line) together with a Gaussian with the same variance (smooth line). In this case, a new angular diffusivity is chosen at random from an exponential distribution with $K = 1$ (arbitrary units) at the end of each move or after executing a move for $10^3$ time-steps. Comparable results (not shown) were obtained for switching times drawn at random from an exponential distribution.
R models embody observed movement patterns and patterns beyond the scale of data collection when CRW are maximally non-committal models of movement patterns. (J. R. Soc. Interface)

Here, it was shown that truncated LW, produced by truncated $\mu = 2$ LW. Move lengths are drawn at random from an inverse-square power-law distribution ranging from 0.01 to 1000 (arbitrary units). Turning angles are selected at random and are uniformly distributed between $0^\circ$ and $360^\circ$. The walkers have speed $s = 1$ (arbitrary units). Diffusive $R^2 \propto t$ scaling behaviour is indicated (dashed lines). (b) Dispersal at $10^4$ time-steps (jagged line) together with the analytical prediction (smooth line), equation (2.7), for CRW with multiple modes. The effective value of $\eta$ was chosen to maximize the goodness of fit. (c) Dispersal at $10^5$ time-steps (jagged line) together with a Gaussian with the same variance (smooth line). These displacements are nearly Gaussian.

be the one that uses the available information and nothing more. Here, it was shown that truncated $\mu = 2$ LW are maximally non-committal models of movement patterns beyond the scale of data collection when CRW models embody observed movement patterns and when there is partial information about landscape/behavioural heterogeneity, in the form of reliable estimates for some of the lower order moments of the angular diffusivity. Any other model would assume more information than is known from the data. Truncated $\mu = 2$ LW thereby provides a robust link between local and emergent dynamics which is largely insensitive to the description of heterogeneity and so have the potential to become a valuable tool in ecological modelling when scaling up from observational scales to assess the potential effects of landscape heterogeneity and changes in behaviour. This shows that with landscape and behavioural heterogeneity, the unusual thing is not truncated LW movement patterns but their absence, as foreshadowed by Reynolds [13]. Large-scale CRW movement patterns, if they arise at all, would be an emergent phenomenon, not a mathematically self-evident state from which any deviation is a worrisome anomaly. This has significant ecological consequences because LW and Gaussian diffusive movement patterns lead to different expectations for population dispersal and for encounter rates between individuals that regulate population dynamics by setting bounds on the spread of communicable diseases, predation and mating. These findings should not be confused with the modelling of the observed movement patterns as truncated LW, which would be applicable if foragers select move lengths directly from a truncated power-law distribution functions ([18] and references therein).

The analysis also indicates that the hallmark signature of $\mu = 2$ LW movement patterns, namely inverse-square power-law distribution functions of move length, could arise at the population level if the movement patterns of non-identical individuals are characterized by fixed but different angular diffusivities. It follows from this that inverse-square power-law distributions of move lengths at the population level are not necessarily indicative of individuals having $\mu = 2$ LW movement patterns.

Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council.

REFERENCES
Truncated Lévy walks A. M. Reynolds


