

Analyzing Insect Movement as a Correlated Random Walk

P.M. Kareiva¹ and N. Shigesada²

¹ Division of Biology, Brown University, Providence, RI 02912, USA

² Department of Biophysics, Kyoto University, Kyoto, 606 Japan

Summary. This paper develops a procedure for quantifying movement sequences in terms of move length and turning angle probability distributions. By assuming that movement is a correlated random walk, we derive a formula that relates expected square displacements to the number of consecutive moves. We show this displacement formula can be used to highlight the consequences of different searching behaviors (i.e. different probability distributions of turning angles or move lengths). Observations of *Pieris rapae* (cabbage white butterfly) flight and *Battus philenor* (pipe-vine swallowtail) crawling are analyzed as a correlated random walk. The formula that we derive aptly predicts that net displacements of ovipositing cabbage white butterflies. In other circumstances, however, net displacements are not well-described by our correlated random walk formula; in these examples movement must represent a more complicated process than a simple correlated random walk. We suggest that progress might be made by analyzing these more complicated cases in terms of higher order markov processes.

1969). Alternatively, much progress has also been made by describing movement with analytical models that rely on the assumption of random motion (see especially, the pioneering work of Broadbent and Kendall 1953; Skellam 1951). In this paper, we extend an earlier formula concerning discrete random walks (see appendix of Jones 1977) and develop an analytic expression for net displacement in a correlated random walk.

The formula that we develop fills a gap somewhere between the powerful general diffusion models of Skellam (1951) and the realistic detailed simulations of Jones (1977). One of the advantages of our approach is that it produces a summary of behavior that enables comparisons to be made between different organisms, or for the same organism in different situations. In particular, by making the assumption of correlated random movement, we derive a formula for expected net displacement in terms of easily observed distance and directionality parameters. This formula can serve as a means for exploring the consequences of varying rules of movement. We illustrate the approach by applying it to two examples of insect movement.

Introduction

The survival and reproductive rate of insects will depend on their success at locating oviposition sites and food. This success is determined largely by the manner in which individual insects search for these needed items (Jones 1977). Several studies have addressed the general importance of animal movement by either connecting an organism's movement patterns to its population dynamics (cf. Jones 1977; Jones et al. 1980) or examining the adaptive value of different searching behaviors (cf. Smith 1974b). A first step in such studies of movement is the design of a quantitative description of movement behavior (Smith 1974a). When the behavior of individuals is as erratic and irregular as it usually is for insects, it can be extremely difficult to parsimoniously quantify movement patterns. Many investigators have solved this problem by developing simulation models in which probabilistic rules for movement are specified and the computer is used to generate extended movement sequences (e.g. Jones 1977; Kaiser 1976; Siniff and Jessen

A Formula for Net Squared Displacement

Consider an organism in two-dimensional space, moving as shown in Fig. 1a. Although animals typically traverse circuitous routes, we can approximate their displacements by a connected series of straight lines. This approximation allows us to perform comparative analyses using only two simple parameters: move lengths and turning angles between successive moves. But to apply this approach to observations of animals in the field, we need reasonable and consistent criteria for demarcating the endpoints of moves. If we are concerned with flying organisms, landing sites represent natural endpoints for moves. We initially adopted the "move length" and "turning angle" parametrization with butterflies in mind, and it is in such cases that the analyses we describe are best suited. Even when movement is continuous and not punctuated by landing stops (e.g. a crawling beetle), an organism's wanderings may be decomposed into a series of straight line moves. For such continuous travels, we recommend using fixed time intervals to define moves (i.e. a move is the displacement observed at the end of some predetermined constant time interval). Since quan-

Offprint requests to: P.M. Kareiva

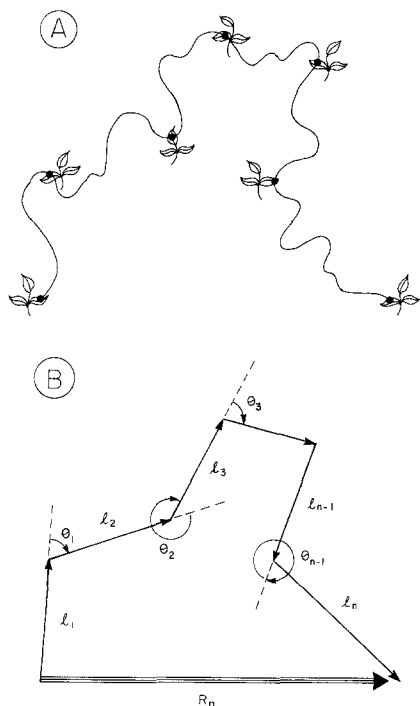


Fig. 1A, B. Method of analyzing movements. **A** Typical movement path of an organism (—), with ● representing move endpoints. **B** The movement parameters recorded for the path shown in **A**

titative analyses of movement may be sensitive to the time interval that defines a move, it is helpful to try out several time intervals before one is settled on. There will often be some intermediate time interval that leads to the most informative characterization of movement. By informative, we mean an interval at which the inter-move variance in turning angles and move lengths is highest, and at which these movement parameters most vividly reflect behavioral responses to varying ecological situations.

Using the convention of straight line moves, we now develop the relationship between an organism's movement behavior and its expected square displacement. We examine square displacement (not linear displacement) because its expectation (and not the expectation of linear displacement) can be conveniently calculated from information about turning angles and move lengths. Designate each m^{th} move as the two dimensional vector (x_m, y_m) , so that the total displacement after n consecutive moves, R_n , is given by

$$R_n = \sum_{m=1}^n (x_m, y_m)$$

We denote θ_m to be the angle (measured clockwise) with which the organism turns at the end of the m^{th} move (consult Fig. 1b). The length of each move,

$$l_m = \sqrt{(x_m^2 + y_m^2)}$$

and the size of each turning angle, θ_m are assumed to be independent random variables with the following probability densities:

$p(l) dl$: the probability that the length of each flight has a value between l and $l + dl$.

$g(\theta) d\theta$: the probability that an angle between two consecutive flights measured clock-wise has a value between θ and $\theta + d\theta$.

Associated with these random variables are the expected values:

$$E(l) = \int_0^{\infty} l p(l) dl$$

$$E(l^2) = \int_0^{\infty} l^2 p(l) dl$$

$$c = E(\cos \theta) = \int_{-\pi}^{\pi} \cos \theta g(\theta) d\theta$$

$$s = E(\sin \theta) = \int_{-\pi}^{\pi} \sin \theta g(\theta) d\theta$$

For our analysis, we assume that a series of moves can be represented by a random draw from $p(l)$ and $g(\theta)$, then another random draw from $p(l)$ and $g(\theta)$, and so on. Because each random draw is independent of the preceding random draws, the random draw process is a first order markov chain. In other words, the derivation that follows assumes that the future pattern of movement is not influenced by the pattern of movement that previously transpired. The result is a correlated random walk in which the distribution $g(\theta)$ provides a measure of the degree to which the direction of movement is correlated. If the direction of movement is strongly correlated, the distribution $g(\theta)$ will be concentrated around a particular value of θ ; if movement can be reduced to a pure random walk with no tendency to continue moving in the same direction, then $g(\theta)$ has a uniform density and both c and s above become zero. The following general formula can be derived (see appendix for details), relating $E(R_n^2)$ to $E(l)$, $E(l^2)$, $E(\cos \theta)$, and $E(\sin \theta)$:

$$E(R_n^2) = nE(l^2) + 2E(l)^2 \frac{(c - c^2 - s^2)n - c}{(1 - c)^2 + s^2} + 2E(l)^2 \frac{2s^2 + (c + s^2)^{\frac{n+1}{2}}}{[1 - c)^2 + s^2]^{\frac{n+1}{2}}} \gamma \quad (1)$$

where

$$\gamma = [1 - c)^2 - s^2] \cos \{(n+1)\alpha\} - 2s(1 - c) \sin \{(n+1)\alpha\}$$

$$\alpha = \arctan s/c$$

This equation often reduces to a much simpler expression for special cases. For instance, if $g(\theta)$ has a uniform density, we end up with $E(R_n^2) = nE(l^2)$. A more interesting and biologically realistic special case arises when organisms exhibit equal probabilities of turning left or right. In that situation, $g(\theta)$ is symmetric about $\theta = 0^\circ$ and $E(\sin \theta) \equiv 0$, reducing Eq. 1 to the convenient form:

$$E(R_n^2) = nE(l^2) + 2E(l)^2 \frac{c}{1 - c} \left(n - \frac{1 - c^n}{1 - c} \right) \quad (2)$$

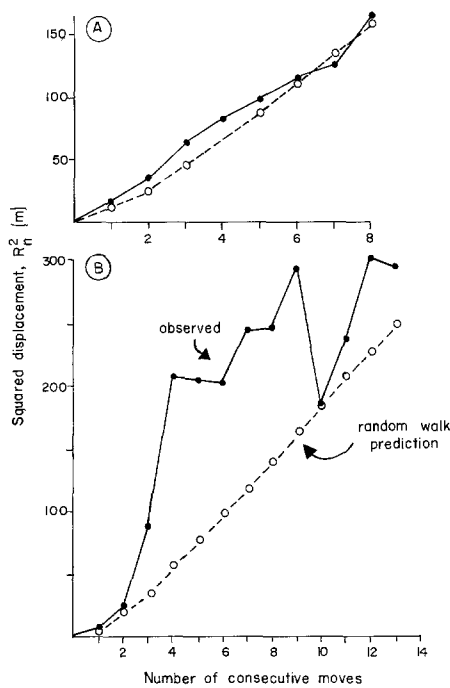


Fig. 2A, B. The relationship between the mean squared distance *Pieris rapae* butterflies are displaced and the number of consecutive moves they have made: **A** represents flight while ovipositing in a collard garden and **B** represents flight while nectar-feeding in a goldenrod field. The correlated random walk prediction is calculated, as explained in the text, using Eq. 2

One of the merits of this analysis is that it allows us to relate behavioral changes that influence $g(\theta)$ or $p(l)$ to consequent changes in the expected net displacements. Turning angles are a feature of searching behavior that have received much attention (cf. Smith 1974a, b). The effect of turning angles on net displacement is clearly seen for large n , where Eq. 2 tends to:

$$E(R_n^2) \rightarrow nE(l^2) + 2nE(l)^2 \frac{c}{1-c} - 2E(l)^2 \frac{c}{(1-c)^2}$$

$$E(R_n^2) \sim_{n \gg 1} nE(l^2) + 2nE(l)^2 \frac{c}{1-c}$$

Turning angles determine $E(R_n^2)$ through the ratio $c/1-c$ above, with small turning angles producing the largest $c/1-c$ ratios and thus the largest displacements.

In general, if movement sequences represent the outcome of a correlated random walk, then Eqs. 1 and 2 can be used to compare dispersal behaviors in different circumstances. In these situations the probability distributions for θ and l adequately summarize an animal's movement behavior. Changes in the distributions of θ and l will alter rates of net displacement according to Eqs. 1 and 2.

Given data describing the straight line moves of several individuals, the appropriateness of Eqs. 1 or 2 is easily tested. From observations of movement we can obtain averages for l , l^2 , $\cos \theta$ and $\sin \theta$ and then use these values to calculate the expected curve for R_n^2 (applying either Eqs. 1 or 2 and varying n). If the theoretical curve generated in this fashion for R_n^2 matches the observed plot of mean R_n^2 values, then the movement

would appear to be a correlated random walk. If observed R_n^2 averages do not correspond to the theoretical curve (see examples), then specific violations of the random walk assumptions can be pursued. More complicated behaviors might require the consideration of cross-correlations between θ_i and l_i , possible violations of the markov property, or orientation to particular landmarks.

It is interesting to note that in general $E(R_n^2)$ tends to increase linearly with n , as n becomes large. This can be seen directly for the special cases represented by Eqs. 1 and 2. More importantly, even if we consider k^{th} order markovian movement (the i^{th} movement vector is correlated to the vectors of the preceding k moves), $E(R_n^2)$ will tend to increase linearly with n , for $n \gg 1$ (Saito 1967). Thus, even if we do not know the details of movement, we can often extrapolate expected square displacements as linear functions of n . But if expected square displacements are linear functions of n , expected *unsquared* displacements will not be (an exception is directed or oriented movement towards a fixed point, Nossal and Weiss 1974). Assuming a constant velocity of random movement, longer observation bouts will yield lower displacements per unit time than will shorter observation bouts. Because it is square displacement, rather than displacement, that is expected to be a linear function of time spent moving (see also Skellam 1951 and Crumpacker and Williams 1973), square displacement represents the most convenient parameter with which to quantify dispersal or movement rates.

Examples of Analysis

Cabbage Butterfly Flight

Root and Kareiva (1982) observed the flight sequences of over 200 *Pieris rapae* females near Ithaca, New York. A move was defined as the transition from one landing site to the next. The straight line paths between landing sites were mapped and the used to generate frequency distributions for l_i and θ_i (following the conventions shown in Fig. 1b). Since $g(\theta)$ was symmetric about $\theta=0^\circ$ for these data, theoretical R_n^2 curves were calculated by substituting into Eq. 2 the averages for l , l^2 and $\cos \theta$. To determine if butterfly movement conformed to the expectations of a correlated random walk, we compared the theoretical curves calculated from Eq. 2 to plots of observed mean squared displacements (generated by averaging R_n^2 values for each different n). Data from two different situations were analyzed in this manner: 20 females ovipositing in a homogeneous garden of collards and 18 females nectar-feeding in a goldenrod field. The butterflies ovipositing on collards conformed remarkably well to the expectations of the correlated random walk model (see Fig. 2a). Because ovipositing butterflies appeared to engage in a simple correlated random walk, Root and Kareiva (1982) were able to use $g(\theta)$, $p(l)$ and Eq. 2 as quantitative descriptors of butterfly flight – in particular of how ovipositing females responded to changes in the dispersion of their food plants. In contrast, the mean square displacements of nectar-feeding butterflies departing widely from the predictions of Eq. 2 (see Fig. 2b); this indicates that when they are nectar-feeding, the flight patterns of butterflies cannot be simply described as correlated random walks. We do not know

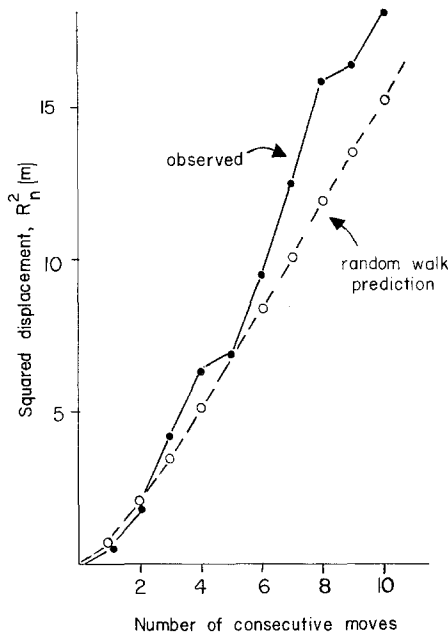


Fig. 3. The relationship between the mean squared distance *Battus philenor* larvae are displaced and the number of consecutive moves they have made (the # of 2 min intervals that have elapsed). The correlated random walk prediction is calculated, as explained in the text, using Eq. 2

what complications might be involved in nectar-feeding that violate the simplifying assumptions leading to Eq. 2.

Crawling Caterpillars

As part of a detailed study concerning the population ecology of pipe-vine swallowtails, Rausher (1979) quantified the searching abilities of *Battus philenor* larvae by following their movements in the field. He released larvae singly in the middle of a 15.4×10.8 m grid in natural vegetation. At 2 min intervals Rausher marked the position of larvae with small flags and was thus able to retrace their movement patterns at the end of each observation bout. A total of 29 larvae were followed until they wandered out of the study grid. In this case, a "move" ends and a new one begins at the end of each 2 min interval. Using Rausher's data, we calculated the values expected for R_n^2 (assuming a correlated random walk) by substituting observed means for l , l^2 and $\cos \theta$ into Eq. 2 [again, $g(\theta)$ was symmetric about $\theta=0^\circ$]. Both the expected R_n^2 curve and the realized curve from Rausher's experiment are plotted in Fig. 3. While there is modest agreement between theory and observation, Eq. 2 consistently underestimates R_n^2 for move numbers greater than $n=5$ (by an average of 23%). Detailed analysis of the turning angle data indicates that *Battus* larvae alternate turning angles so that right turns follow left turns and left turns follow right turns in a significantly non-random pattern (G -test of independence; $P < 0.001$). By alternating turns in this nonrandom fashion, *Battus* larvae reduce the probability of turning back on themselves and returning to a previous position. As a result the larvae's net displacement is greater than would be expected if their movement were a simple correlated random walk.

Discussion

The analysis of animal movement as a correlated random walk is best viewed as a null hypothesis. When observations of movement sequences generate R_n^2 curves that conform to Eqs. 1 or 2, then the null hypothesis may be acceptable. Moreover, the formulae we derived which relate net displacements to movement parameters are especially useful for examining the consequences of changes in $g(\theta)$ or $p(l)$ behavior. Although in each particular situation an animal may behave as though engaged in a correlated random walk, the probability distributions governing that process may change as a result of varying ecological circumstances. For that reason, movement that can be represented as a random walk should not be confused with the notion of an animal that behaves randomly with respect to the environment. It may be especially fruitful to investigate the foraging or searching movements of animals as correlated random walks whose parameters (turning angle or move length) depend on local ecological conditions. Thus, for example, an organism might always move randomly, but with sharper turning angles in the presence of food resources.

When previous experience alters an animal's movement behavior, we are left with the vexing problem of analyzing k^{th} order markov chains, or constructing elaborate simulations of movement. Although several investigators have faced the problem of complicated searching behavior by developing computer simulations of movement (cf. Siniff and Jessen 1969; Kaiser 1976; Jones 1977; Jones et al. 1980), much less effort has been directed towards formal analyses of movement in terms of higher-order markov chains. We recommend the exploration of markov chain analyses of movement because such approaches may provide greater generality than simulation studies.

Appendix

Derivation of Equation (1). The mean square distance reached after the n^{th} move is obtained by averaging the square of R_n with respect to l_m and θ_m ($m=1, 2, \dots, n$). For more general cases, this problem has previously been solved by Nossal and Weiss (1974) in their analysis of the surface locomotion of biological cells. The following derivation is essentially similar, but somewhat simpler than that given by Nossal and Weiss (1974). Since movement is assumed to be a random walk, the cross-correlations among l_m , l_j , θ_m and θ_j are taken to be zero, and we have the following equation

$$E(R_n^2) = \sum_{m=1}^n E(W_m \cdot W_m) + 2 \sum_{m>j} E(W_m \cdot W_j) \\ = nE(l^2) + 2[E(l)]^2 \sum_{m>j} E\left(\cos \sum_{k=j}^{m-1} \theta_k\right)$$

where $E(\cdot)$ denotes the expectation of the associated random variable, and $(W_m \cdot W_j)$ is the inner product of the move vectors (x_m, y_m) and (x_j, y_j) . The difficult term in this expression is

$$\sum_{m>j} E\left(\cos \sum_{k=j}^{m-1} \theta_k\right),$$

which we can calculate by using Euler's formulae to express cosines as complex numbers. In particular, note that

$$\cos \sum_{k=j}^{m-1} \theta_k = \frac{1}{2} \left(\exp i \sum_{k=j}^{m-1} \theta_k + \exp -i \sum_{k=j}^{m-1} \theta_k \right)$$

Next, since the turning angles $\theta_k (k=1, 2, \dots)$ are assumed to be independent of each other, we obtain

$$E \left(\cos \sum_{k=j}^{m-1} \theta_k \right) = \frac{1}{2} \{ (c + is)^{m-j} + (c - is)^{m-j} \}$$

where $c = E(\cos \theta)$ and $s = E(\sin \theta)$.

Thus the sum, $\sum_{m>j}^n E \left(\cos \sum_{k=j}^{m-1} \theta_k \right)$, is given by

$$\begin{aligned} \sum_{m>j}^n E \left(\cos \sum_{k=j}^{m-1} \theta_k \right) &= \frac{1}{2} \sum_{p=1}^{n-1} \sum_{j=1}^{n-p} \{ (c + is)^p + (c - is)^p \} \\ &= \frac{c + is}{2(1 - c - is)} \left(n - \frac{1 - (c + is)^n}{1 - c - is} \right) \\ &\quad + \frac{c - is}{2(1 - c + is)} \left(n - \frac{1 - (c - is)^n}{1 - c + is} \right) \\ &= \frac{(c - c^2 - s^2) n - c}{(1 - c)^2 + s^2} \\ &\quad + \frac{2s^2 + (c^2 + s^2)^{\frac{n+1}{2}} \left[\{(1 - c)^2 - s^2\} \right.}{\cos \{(n+1)\alpha\} - 2s(1 - c) \sin \{(n+1)\alpha\}} \\ &\quad \left. + \frac{\{(1 - c)^2 + s^2\}^2}{\cos \{(n+1)\alpha\} - 2s(1 - c) \sin \{(n+1)\alpha\}} \right]}{\{(1 - c)^2 + s^2\}^2} \end{aligned}$$

where we put $\alpha = \arctan s/c$.

When we substitute this in our equation above for $E(R_n^2)$, we arrive at Eq. 1 of the text.

Acknowledgements. We would like to thank Howard Taylor and Akira Oktubo for stimulating and encouraging this study,

and Mark Rausher for providing maps of *Battus philenor* movement. The comments of anonymous reviewers, and especially of Jim McNair, greatly improved the notation and the interpretation of the displacement formula derived in this paper. This work was completed while N. Shigesada was on sabbatic leave in the Biology Department of Stanford University.

References

- Broadbent SR, Kendall DG (1953) The random walk of *Trichostrogylus retortaeformis*. *Biometrics* 9:460-466
- Crumpacker DW, Williams JS (1973) Density, dispersion, and population structure in *Drosophila pseudoobscura*. *Ecol Mon* 43:498-538
- Jones R (1977) Movement patterns and egg distribution in cabbage butterflies. *J Anim Ecol* 46:195-212
- Jones R, Gilbert N, Guppy M, Neals V (1980) Long-distance movement of *Pieris rapae*. *J Anim Ecol* 49:629-642
- Kaiser H (1976) Quantitative description and simulation of stochastic behavior in dragonflies (*Aeschna cyanea*, Odonata). *Acta Biotheor* 25:163-210
- Nossal RJ, Weiss GH (1974) A descriptive theory of cell migration on surfaces. *J Theor Biol* 47:103-113
- Rausher M (1979) Coevolution in a simple plant-herbivore system. Dissertation. Cornell University, Ithaca, New York, USA
- Root RB, Kareiva PM (1983) The search for resources by cabbage butterflies (*Pieris rapae*): ecological consequences and adaptive significance of markovian movements in a patchy environment. *Ecol Monog* (in press)
- Saito N (1967) *Introduction to Polymer Physics*. Syokabo Press, Tokyo
- Siniff D, Jessen C (1969) A simulation model of animal movement patterns. *Adv Ecol Res* 6:185-219
- Skellam JG (1951) Random dispersal in theoretical populations. *Biometrika* 38:196-218
- Smith JNM (1974a) The food searching behavior of two European thrushes. I. Description and analysis of search paths. *Behaviour* 48:276-302
- Smith JNM (1974b) The food searching behavior of two European thrushes. II. The adaptiveness of the search patterns. *Behaviour* 49:1-61

Received August 8, 1982