Stochastic models for cell motion and taxis

Edward L. Ionides[1], Kathy S. Fang[2], R. Rivkah Isseroff[2]
and George F. Oster[3]

April 18, 2003

Abstract

Certain biological experiments investigating cell motion result in time lapse video microscopy data which may be modeled using stochastic differential equations. These models suggest statistics for quantifying experimental results and testing relevant hypotheses, and carry implications for the qualitative behavior of cells and for underlying biophysical mechanisms. Directional cell motion in response to a stimulus, termed taxis, has previously been modeled at a phenomenological level using the Keller-Segel diffusion equation. The Keller-Segel model cannot distinguish certain modes of taxis, and this motivates the introduction of a richer class of models which is nevertheless still amenable to statistical analysis. A state space model formulation is used to link models proposed for cell velocity to observed data. Sequential Monte Carlo methods enable parameter estimation via maximum likelihood for a range of applicable models. One particular experimental situation, involving the effect of an electric field on cell behavior, is considered in detail. In this case, an Ornstein-Uhlenbeck model for cell velocity is found to compare favorably with a nonlinear diffusion model.

Keywords: Cell migration, chemotaxis, galvanotaxis, nonlinear diffusion, stochastic model.

1 Introduction

Active migration of blood and tissue cells is essential to a number of physiological processes such as inflammation, wound healing, embryogenesis and tumor cell metastasis [6]. It also plays an important role in the functioning of many bioartificial tissues and organs

[1]Department of Statistics, University of Michigan, Ann Arbor, MI 48109
[2]Department of Dermatology, University of California, Davis, CA 95616
[3]Department of Molecular and Cellular Biology, University of California, Berkeley, CA 94720

[37], such as skin equivalents [44] and cartilage repair [41]. Modern techniques in microscopy, genetics and pharmacology have helped to make some progress in unraveling the complex biophysical processes involved in cell motion [39]. Although different cell types show diverse methods of locomotion, there are general principles that are widely applicable for cells moving along a substrate. First a cell extends a protrusion by actin filament polymerization [40], which then attaches to the substrate using integrin adhesion receptors [29]. A contractile force is next generated which moves the cell body. Finally the cell must detach from the substrate at its trailing end.

Various mathematical models incorporating the above principles of cell motion have been proposed. The most ambitious of them attempt to represent all the physical and chemical processes involved in the motion of an entire cell [53, 13, 11]. Others concentrate on a specific process such as extension of a protrusion [40] or receptor dynamics [38]. The primary purpose of these biophysical models is to demonstrate that the proposed mechanisms can in fact produce the forces and behaviors observed experimentally.

Another approach to modeling cell motion is phenomenological in nature. The so-called correlated random walks of Alt [1], Dunn and Brown [19] and Shenderov and Sheetz [49] have been proposed to describe observations of isolated cells locomoting on a substrate. For applications, the behavior of cell populations may be of more direct interest, and here diffusion approximations to population behavior are widely used following Keller and Segel [33]. The theoretical relationships between single cell models and population models have been studied by Alt [1], Dickinson and Tranquillo [14], Ford *et al.* [23, 24]. An empirical comparison between single cell and cell population models is given by Farrell *et al.* [22]. Phenomenological models are used for quantifying experimentally observed cell behavior, and do not require justification in terms of a proposed mechanism. Nevertheless the line dividing biophysical from phenomenological models is in fact only a difference in complexity, and can become blurred as even the simpler phenomenological models can have implications concerning underlying biophysical mechanisms [19].

The questions of scientific and engineering interest about cell motion can be broadly summarized into the following: What biophysical processes are involved in cell motion? How can the speed and direction of the motion be modeled? One approach toward answering these questions is to collect temporal sequences of images of moving cells. Various experimental protocols for collecting such data were discussed in [2] and [3]. Cells may be observed moving separately on a microscope slide, or in a three-dimensional collagen gel. The cells under investigation may be part of a connected tissue, with a stain used to identify and track groups of cells.

Models of cell motion have been categorized by the length and time scales under primary consideration [14]. On the scale of *locomotion* the basic actions of cell motion are apparent: this is the scale on which a cell may extend individual protrusions (broad lamellae, thinner lamellipodia or hairlike filopodia) that can be used to pull itself along. On a longer time and length scale, termed *translocation*, one observes the displacement of the

cell due to one or several motion cycles. As the resolution of detail about the motion decreases, gross tendencies such as directional preferences can become more apparent. On the scale of *migration*, the cumulative effect of many motion cycles is observed. Although on the migration scale one loses the ability to observe directly the mechanisms of cell motion, it is the behavior of cells on this scale which is of primary interest in applications such as development, cancer metastasis, and tissue engineering. Furthermore, some assays such as the Boyden diffusion chamber [10] entail observations of populations of cells for which only the behavior on the migration scale can be directly observed. Dickinson and Tranquillo [14] develop mathematical methods to relate models on different scales, using the method of adiabatic elimination of fast variables [25].

The Keller–Segel model of chemotaxis [33, 10] gives a widely accepted approach for modeling cells on the migration scale, using Fokker–Planck equations. The Fokker-Planck equation models only the density of the cells, however this density also arises from modeling the stochastic paths of individual cells as the solution to a stochastic differential equation (SDE). Fokker-Planck equations apply only to Markov processes [25], whereas SDEs can describe processes with memory. The use of SDEs and techniques of stochastic calculus for biological modeling was pioneered by by Kendall [34] and Brillinger [7, 9]. A contribution of this paper is to show how SDEs can provide a systematic approach for qualitative and quantitative descriptions of cell motion on the scales of translocation and migration.

When the distribution of the velocity process has rotational symmetry about the origin, and does not depend on position, the model is called *isotropic*. Such models, suitable when the cell experiences no directional stimuli, are discussed in Section 2. When the velocity process has directional asymmetry or depends on position then the cell is said to perform *taxis*. The cell must then be picking up some locational or directional cue from its environment. Models for taxis are considered in Section 3. Section 4 addresses the issue of whether a cell can move up a gradient of a chemical attractant while only being aware of its concentration. Section 5 illustrates some uses of the stochastic models via an investigation of cell motion in an electric field.

2 Isotropic cell translocation

Two characteristics used to describe isotropic cell motion are speed and persistence. The story is that in the short term cells are observed to move with slowly varying direction and speed. After a while they appear to forget their initial orientation. This time scale is termed the *persistence* of the cell. A standard model for isotropic translocation is the Ornstein-Uhlenbeck (O-U) process, where the velocity $\mathbf{v}_t = (v_x(t), v_y(t))$ is supposed to follow the stochastic infinitesimal equation

(M1) $$d\mathbf{v}_t = -a\mathbf{v}_t dt + b d\mathbf{W}_t$$

for positive constants a and b and a two-dimensional Brownian motion $\mathbf{W}_t$. An introduction to SDEs can be found in [43] and an elegant treatment of one dimensional diffusions is given in [31]. The coefficient $-a\mathbf{v}_t$ is called the infinitesimal drift and b^2 the infinitesimal variance. Heuristically, a gives the rate at which the velocity regresses to zero and b gives the magnitude of the random innovations which tend to push the velocity away from zero. The root mean square speed can be calculated as $\sqrt{b^2/a}$, and a measure of persistence is $1/a$. The O-U process was introduced for cell motion in [19], and a thorough presentation is given in [51]. Similar concepts of speed and persistence were developed in earlier probabilistic models [1, 18].

One feature sometimes observed for cells is that their direction of motion changes most rapidly when their speed is small. This behavior is a property of (M1), as is most clearly seen by transforming to polar coordinates (r_t, θ_t) for the velocity by applying Itô's lemma ([43], Theorem 4.2.1). This leads to the infinitesimal equations

$$\begin{aligned} dr_t &= \left(-ar_t + \frac{b^2}{2r_t}\right) dt + b dW_t^r \\ d\theta_t &= \left(\frac{b}{r_t}\right) dW_t^\theta \end{aligned}$$

for two independent one-dimensional Brownian motions W_t^r and W_t^θ.

An extension of the O-U model was introduced by Shenderov and Sheetz [49] to account for the observation that for some cell types the velocity has an oscillatory behavior, with high correlation between velocities at certain time lags. Stochastic infinitesimal equations were not employed in [49], but the model can be written as

$$\text{(M2)} \qquad \begin{aligned} d\mathbf{v}_t &= (\quad -a\mathbf{v}_t \quad -c\mathbf{u}_t \quad) dt + b d\mathbf{W}_t \\ d\mathbf{u}_t &= (\quad \mathbf{v}_t \quad -k\mathbf{u}_t \quad) dt, \end{aligned}$$

where $\mathbf{u}_t$ represents the "memory" of the cell. The the matrix $\begin{pmatrix} -a & -c \\ 1 & -k \end{pmatrix}$ has complex eigenvalues if $(a-k)^2 < 4c$, and in this case the velocity shows oscillatory behavior.

Previously, estimation of parameters in models such as (M1) and (M2) has been carried out by the method of least squares [15, 49] or generalized least squares [12]. Maximum likelihood estimation is the preferred method for efficient inference in similar models arising from econometric and scientific time series analysis [20, 28]. When velocity is modeled as a linear Gaussian Markov process, and position is observed at discrete times, the system is called a linear Gaussian state space space model and the Kalman filter algorithm is available for calculation of the likelihood. The details of applying linear, Gaussian state space methods to our models are not given here, as in practice the general nonlinear method discussed in Section 5.3 were used for the linear models also.

3 Models for taxis

In the non-isotropic case many ways have been proposed by which a cell might respond to a stimulus on the scale of cell translocation, and these are termed *modes of taxis* [14]. *Topotaxis* occurs when a cell turns preferentially toward a stimulus. *Orthotaxis* is said to occur if the magnitude of the velocity of the cell increases when the direction is toward a stimulus. *Klinotaxis* occurs when the rate of turning decreases while traveling toward a stimulus. These three modes depend on the direction of a stimulus, but there are further two modes that depend only on the magnitude of a stimulus. *Orthokinesis* occurs when the magnitude of the velocity decreases with the magnitude of a stimulus. *Klinokinesis* occurs when the rate of turning increases with the magnitude of a stimulus.

The reader may wonder whether these are the only possible modes, and whether the observed motion of a cell toward a stimulus can be uniquely characterized as some combination of these modes. In fact these questions are of basic scientific interest, since modes of taxis are experimentally testable consequences of models at the mechanistic level for the biochemistry and biophysics of cell motion. Unfortunately there has been some confusion in the literature about how to decide empirically upon the modes of taxis, based on observations from a system. Doucet and Dunn [17] discuss this problem and give the example of classifying the mode of taxis of a snake whose head can detect the level of a chemical attractant. By moving its head from side to side this snake detects the gradient of the chemical and moves up it. The whole snake appears to be capable of topotaxis, while mechanistically it can only measure the magnitude of the stimulus and so should be capable only of a kinesis.

To formalize modes of taxis mathematically one can avoid the snake paradox by defining modes of taxis as properties of models rather than biophysical mechanisms. The time and length scale on which we are modeling a process can determine the characterization of the behavior. Recall the three identified scales of locomotion, translocation and migration. On the scale of locomotion, in which the biophysical properties of a cell result in the extension of pseudopodia and traction along a substrate, modes of taxis lose meaningfulness as a way to characterize the process. On the scale of translocation the position and velocity of a cell, but not its internal processes, are modeled. One can then attempt to define modes of taxis. If there is a scale on which the internal processes regulating the velocity of a cell has negligible memory the velocity process may be modeled by the infinitesimal equations

$$\text{(M3)} \qquad \begin{aligned} dr_t &= \mu_r(r_t,\theta_t,s_t,\phi_t)dt+\sigma_r(r_t,\theta_t,s_t,\phi_t)dW_t^r+\tau_r(r_t,\theta_t,s_t,\phi_t)dW_t^\theta \\ d\theta_t &= \mu_\theta(r_t,\theta_t,s_t,\phi_t)dt+\sigma_\theta(r_t,\theta_t,s_t,\phi_t)dW_t^\theta+\tau_\theta(r_t,\theta_t,s_t,\phi_t)dW_t^r \end{aligned}$$

Here (r_t,θ_t) are the polar coordinates for the velocity $\mathbf{v}_t$, and at location $\mathbf{x}_t$ the stimulus has magnitude $s_t(\mathbf{x}_t)$ and gradient vector $\phi_t(\mathbf{x}_t)$. For multiple stimuli, s_t and ϕ_t take multiple values. W_t^r and W_t^θ are two independent Brownian motions. Assuming the process $\{(\mathbf{v}_t,\mathbf{x}_t)\}$ is continuous, Markov and time homogeneous, it is a small restriction to suppose

it has a representation of the form (M3). If a cell has no preferred direction of turning without a directional cue, (a counter-example was suggested by Alt [3]), $\mu_\theta(r_t, \theta_t, s_t, \phi_t)$ fits the description of a topotaxis term. If $\mu_r(r_t, \theta_t, s_t, \phi_t)$ is written as

$$\mu_r(r_t, \theta_t, s_t, \phi_t) = \mu_r^{(1)}(r_t) + \mu_r^{(2)}(r_t, s_t) + \mu_r^{(3)}(r_t, \theta_t, s_t, \phi_t)$$

then $\mu_r^{(2)}(r_t, s_t)$ has the form of an orthokinesis term and $\mu_r^{(3)}(r_t, \theta_t, s_t, \phi_t)$ has the form of an orthotaxis term. Similarly, if $\sigma_\theta(r_t, \theta_t, s_t, \phi_t)$ is written as

$$\sigma_\theta(r_t, \theta_t, s_t, \phi_t) = \sigma_\theta^{(1)}(r_t) + \sigma_\theta^{(2)}(r_t, s_t) + \sigma_\theta^{(3)}(r_t, \theta_t, s_t, \phi_t)$$

then $\sigma_\theta^{(2)}(r_t, s_t)$ can stake a claim as a klinokinesis term and $\sigma_\theta^{(3)}(r_t, \theta_t, s_t, \phi_t)$ as a klinotaxis term. The remaining terms σ_r, τ_r and τ_θ have no clear roles to play in the recognised modes of taxis, indicating that these modes form an incomplete picture of the possible directional behavior in (M3). For example, a change in the random variation in speed, caused by a varying level of a ligand that interacts with the speed regulation mechanisms of a cell, might cause directional behavior through a term $\sigma_r(r_t, s_t)$.

On the scale of migration when the location $\mathbf{x}_t$ is supposed to have negligible memory one can write down the infinitesimal equation

(M4) $$d\mathbf{x}_t = \mu(s_t, \phi_t)dt + \gamma(s_t, \phi_t)d\mathbf{W}_t.$$

Here $\gamma(s_t, \phi_t)$ is a 2×2 matrix which we may suppose is written $\gamma(s_t, \phi_t) = \gamma^{(1)}(s_t) + \gamma^{(2)}(s_t, \phi_t)$. In this model the two concepts of rate of turning depending on position and of speed depending on position are linked together in the infinitesimal variance term $\gamma(s_t)$. Since the sample paths are not differentiable one has to take a broad-minded view about "speed" and "rate of turning" to recognize $\gamma^{(1)}(s_t)$ as a combined term for both klinokinesis and orthokinesis, and $\gamma^{(2)}(s_t, \phi_t)$ as a term involving klinotaxis and orthotaxis. Similarly $\mu(s_t, \phi_t)$ can be thought of as a taxis term, involving topotaxis, orthotaxis and klinotaxis. It does not generally make sense to include an infinitesimal drift term $\mu(s_t)$ depending only on the magnitude of the stimulus. An alternative interpretation of the parameters in (M4) would come from applying a rescaling argument to (M3). The technique of adiabatic elimination of fast variables [25, 14] can provide such a rescaling for certain particular cases of (M3).

The Fokker-Plank equation corresponding to the infinitesimal equation (M4) is a generalization of the model of Keller and Segel [33] in which μ is written as $\mu = \chi \nabla c$ for a *chemotactic coefficient*, χ, and chemoattractant concentration, $c(\mathbf{x})$. In this context, γ is called the *motility coefficient*.

Choosing between the interpretations of persistence, periodicity, speed and modes of taxis given by different models requires more precise definitions of these concepts than are currently available in the biological literature. The goal here has been to present some options, rather than to come down heavily in favor of any one model.

4 Do kineses work?

There has been some controversy about whether a cell can move up a gradient of a ligand (a small signaling molecule) just by adjusting its speed or rate of turning according to the concentration of the ligand [17]. In other words, do orthokinesis and klinokinesis work as a way of moving up concentration gradients, or must the cell in fact have some memory or ability to detect gradients. A careful theoretical study of klinokinesis where the velocity of a bacterium is treated as a Markov process was undertaken by Stroock [52]. The model considered fails to give convincing evidence that kineses can work. The result obtained in [52] is that if $\{x_t\}$ is the $\mathbb{R}$-valued process giving the position of the bacterium up an increasing gradient of a ligand, and $f : \mathbb{R} \to \mathbb{R}$ is a convex function, then the expected value $E[f(x_t)]$ is monotone increasing with time. This would also be true if $\{x_t\}$ were an unbiased random walk, or a martingale.

Comparing the Itô and Stratonovich solutions [31] to the SDEs introduced in Section 3 gives a new insight into the controversy. Kinesis may be modeled in a simple but instructive way by considering a stochastic process $\{x_t\}$, taking values in $(0,\infty)$, defined by the infinitesimal equation

$$dx_t = \sigma x_t dW_t.$$

This is a particular case of (M4). The Itô solution is $x_t^I = x_0 e^{W_t - t/2}$ and the Stratonovich solution is $x_t^S = x_0 e^{W_t}$. The Itô solution tends to 0 almost surely, though not in expected value, as $E[x_t^I \mid x_0] = x_0$. The Stratonovich solution does not tend to 0, and has increasing expected value $E[x_t^S \mid x_0] = x_0 e^{t/2}$. There is little scientific reason for preferring one solution to the other, and this result suggests that there is equally little reason to decide whether a kinesis results in motion up the gradient or whether there must be an additional taxis for this to occur. It may be valuable to determine whether speed and rate of turning vary with stimulus level but an attempt to assign motion up a gradient to this phenomenon has no scientific basis within the framework of (M4).

5 An investigation of galvanotaxis

Human keratinocytes (skin cells) migrate toward the negative pole in direct current electric fields of physiological strength. This phenomenon is termed galvanotaxis and is of particular interest in wound healing [42], as well as providing a tool for more general study of directional cell motility. One of the challenges in practice of investigating directional response to a stimulus is to set up experimentally the controlled, uniform gradients required for clear and reproducible results. However, such gradients are relatively easy to attain for DC electric fields, making galvanotaxis a convenient model system for investigating basic aspects of directional cell motility [21]. Galvanotaxis is used here to demonstrate the models and methods introduced in this paper. Previous approaches to mathematical modeling

of galvanotaxis may be found in [27] and references therein.

The data analyzed in Section 5.3 were collected by Dr. Fang at University of California, Davis, to investigate the effect of calcium ion (Ca^{2+}) concentration on galvanotaxis. The experimental method follows that used by Fang *et al.* [21] to demonstrate the role of the epidermal growth factor receptor (EGFR) in galvanotaxis. Specifically, normal human keratinocytes from neonatal foreskin epidermis were cultured and plated onto a glass coverslip coated with extra-cellular matrix (collagen). The treated coverslip was placed in a galvanotaxis chamber as described by Nishimura *et al.* [42]. Cells were then observed using phase contrast or differential interference contrast optics with video images being digitally captured. The images were typically captured at intervals of ten minutes, during a one hour observation period, resulting in seven images per experiment.

We consider a nonlinear model for topotaxis in Section 5.1, and compare it with a linear model for taxis introduced in Section 5.2. In Section 5.3 we find maximum likelihood estimates of the parameters, and their errors, and discover that the linear model is in fact more appropriate for these data. This linear model gives some justification for the statistic used to quantify taxis in [21, 42], and also suggests other even more sensitive measures.

5.1 A model for topotaxis in a uniform electric field

Empirically one notices that the speed of the cells is not much affected by the electric field [42]. Theory and observation suggest that changes in cell direction are governed by local behavior around the edges of the leading lamella [19, 4]. Supposing the taxis is generated by the difference in the stimulus (in this case, electric potential) between the two ends of the lamella, the directional sensitivity would be greatest when the cell is moving perpendicular to the electric field. If the electric field is taken to be parallel to the positive x-axis and the velocity $\mathbf{v}_t = (v_x(t), v_y(t))'$ has polar representation (r_t, θ_t) then model consistent with the symmetry and translation invariance of the experiment as well these biological considerations is given by

(M5)
$$d\mathbf{v}_t = \begin{pmatrix} -\alpha & \beta\sin\theta_t \\ -\beta\sin\theta_t & -\alpha \end{pmatrix} \mathbf{v}_t dt + \gamma d\mathbf{W}_t$$

This model is an extension of (M1) and a special case of (M3). An application of Itô's lemma shows the polar representation of the infinitesimal equation defining model (M5) to be

$$\begin{aligned} dr_t &= \left(-\alpha r_t + \frac{2\gamma^2}{r_t}\right) dt + \gamma dW_t^r \\ d\theta_t &= -\beta\sin\theta_t dt + \left(\frac{\gamma}{r_t}\right) W_t^\theta. \end{aligned}$$

The magnitude of the velocity is governed by the same equation as for model (M1). The directional behavior of model (M5) is a rotation of the direction of motion at rate $\beta \sin\theta_t$ toward $\theta = 0$, thus (M5) fits the description of a topotaxis. Although (M5) may be a satisfactory model for taxis generated by selectively increased activation in a pre-existing lamella, it can also happen that the leading lamella is retracted and a new lamella extended. From the representation in polar coordinates we see that at small velocities the taxis term $(-\beta \sin\theta_t dt)$ is dominated by the random variation, which implies a uniform distribution for the orientation of newly formed lamellae. Any attempt to model taxis using a single parameter must make simplifying assumptions, and one interpretation of the data analysis in Section 5.3 is that this particular assumption is inappropriate.

One could quantify galvanotaxis by estimating β in (M5). Another possibility, particularly useful for testing the hypothesis that $\beta = 0$, is to calculate an approximation to the Fisher efficient score. The likelihood function, $L(\alpha,\beta)$, when $\mathbf{v}_t$ is observed for t in the interval $[0,T]$ and γ is known, is taken to be the density of the process (M5) having parameters (α,β,γ) with respect to the process (M5) with parameters $(0,0,\gamma)$, evaluated at $\{\mathbf{v}_t, t \in [0,T]\}$. This density, which in formal probabilistic language is termed a Radon-Nikodym derivative, is given by the Girsanov theorem ([43] theorem 8.6.4) as

$$\begin{aligned} L(\alpha,\beta) &= \exp\left\{ \frac{-\beta}{\gamma^2} \int_0^T r_t^2 \sin\theta_t d\theta_t \right. \\ &\quad \left. -\frac{\beta^2}{2\gamma^2} \int_0^T r_t^2 \sin^2\theta_t dt - \frac{\alpha}{\gamma^2} \int_0^T r_t dr_t - \frac{\alpha^2}{2\gamma^2} \int_0^T r_t^2 dt + 2\alpha T \right\}. \end{aligned}$$

The Fisher efficient score statistic for testing the null hypothesis that $\beta = 0$ is the partial derivative of the logarithm of the likelihood with respect to β evaluated at $\beta = 0$ [46]. Up to an unimportant constant factor, this is

$$Z = \int_0^T r_t^2 \sin\theta_t d\theta_t.$$

This quantity is unobservable, but an approximation can be made in terms of the observed quantities, which are taken to be $\{\mathbf{x}_t, t = 0,1,\ldots,T\}$. Switching back to Cartesian coordinates this becomes

$$Z_1 = \sum_{t=1}^{T-1} \frac{\hat{v}_y(t)}{|\hat{\mathbf{v}}(t)|} (\hat{v}_y(t)\Delta\hat{v}_x(t) - \hat{v}_x(t)\Delta\hat{v}_y(t)).$$

with $\hat{v}_x(t) = x_t - x_{t-1}$, $\Delta\hat{v}_x(t) = \hat{v}_x(t+1) - \hat{v}_x(t)$, *etc*. From symmetry considerations Z_1 has expectation zero when $\{\mathbf{v}_t\}$ is isotropic, and in particular when $\beta = 0$ for any value of α. This follows from the anti-symmetry property that $Z_1(\{\mathbf{v}_t\}) = -Z_1(\{-\mathbf{v}_t\})$. If independent, identically distributed replicates are available, the t-statistic can be used to test the hypothesis that $\beta = 0$.

5.2 A Linear Model for Galvanotaxis

A natural extension of (M1) for motion in a uniform electric field parallel to the x-axis is

(M6) $$d\mathbf{v}_t = -\alpha(\mathbf{v}_t - (\beta,0)')dt + \sigma d\mathbf{W}_t.$$

In polar coordinates this becomes the nonlinear model

$$\begin{aligned} dr_t &= (-\alpha(r_t - \beta\cos\theta_t) + \sigma^2/r_t)dt + \sigma dW_t^r \\ d\theta_t &= -\frac{\beta\sin\theta_t}{r_t}dt + \frac{\sigma}{r_t}dW_t^\theta. \end{aligned}$$

Using the terminology of Section 3, this model includes both an orthotaxis and a topotaxis term. There is no reason, apart perhaps from conceptual simplicity, for preferring a model involving a single mode of taxis. Furthermore, models that have a linear representation are considerably easier to work with. The Fisher score statistic for (M6), up to a constant factor, is

$$Z_2 = x_T - x_0.$$

The cosine statistic employed by [42, 21] can be written as

$$Z_3 = (x_T - x_0)/|\mathbf{x}_T - \mathbf{x}_0|,$$

which caters for variablity between cells by standardising Z_2 by a measure of cell speed. Viewed this way, a more natural modification of Z_2 might be

$$Z_4 = (x_T - x_0)\Big/\sum_{t=0}^{T-1}|\mathbf{x}_{t+1} - \mathbf{x}_t|.$$

These statistics are compared on experimental data in Section 5.3.

5.3 Results

The data analyzed in this section are a treatment group of 24 cells exposed to a 100mV/mm electric field, and a control group of 40 cells with no electric field. Digital images were taken at seven time point separated by 10 minute intervals. Some representative data are displayed in Fig. 1. In Table 1 the statistics $Z_1,\ldots,Z_4$ are compared for testing the presence of an electric field effect. On these data, Z_4 shows the strongest evidence for an effect, and only Z_1 fails to show a statistically significant effect $(p < 0.05)$.

Parameter estimation for models (M5) and (M6) was carried out using maximum likelihood estimation. The likelihood was approximated using sequential importance sampling (also known as particle filtering), a Monte Carlo technique for nonlinear state space models

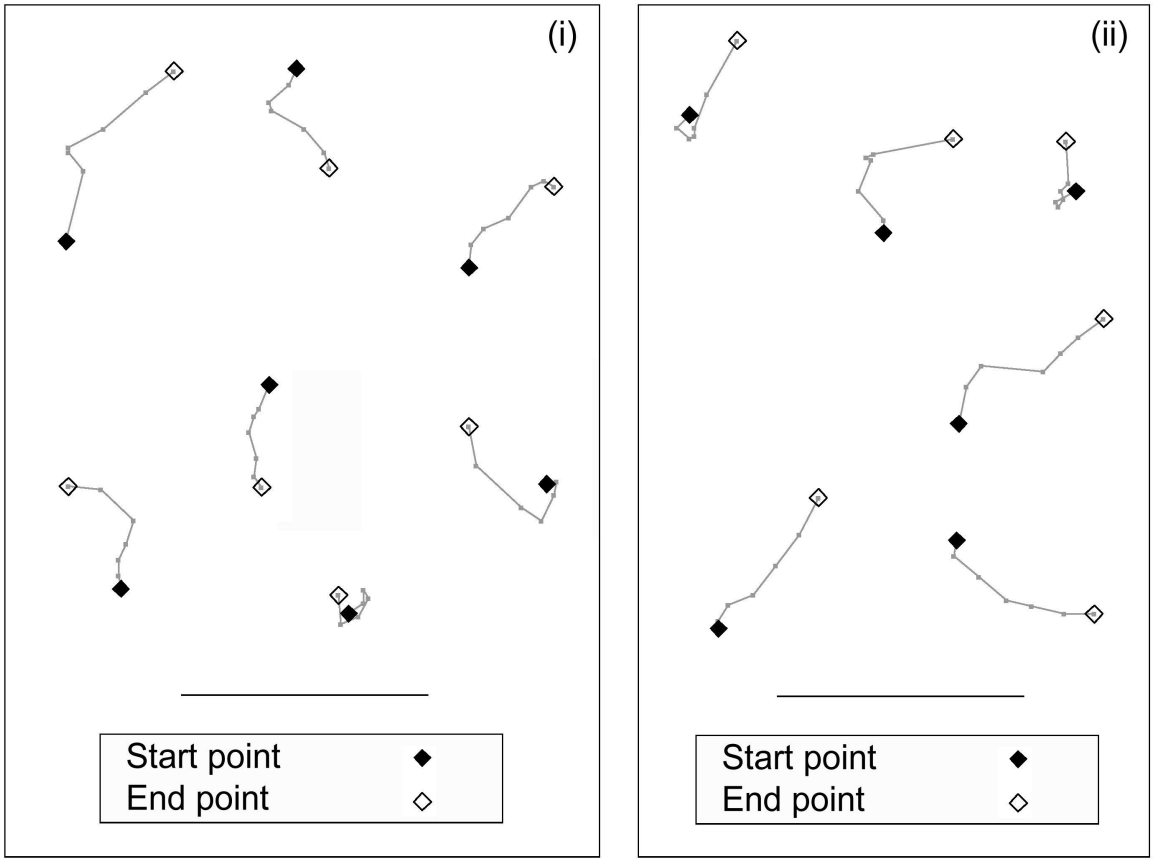

Figure 1: (i) Cell paths resulting from a control experiment, prepared as in Fang *et al.* [21], with no electric field. (ii) Cell paths from a treatment experiment, with an electric field off 100 mV/mm. The cathode is at the top of the page. The bar represents 100 μm.

	treatment group		control group		t-test
	sample mean	SE	sample mean	SE	p-value
Z_1	53.5	18.2	29.8	16.3	0.354
Z_2	21.7	5.2	3.6	4.6	0.015
Z_3	0.490	0.126	0.095	0.120	0.035
Z_4	2.21	0.50	0.22	0.49	0.010

Table 1: The statistics Z_1, Z_2, Z_3, Z_4 were calculated for each of the 24 cells in the treatment group and 40 in the control group. A two sample t-test was carried out to test the hypothesis that the treatment had no effect, against a general alternative hypothesis. Z_1 and Z_2 have units of pixels (1 pixel $\approx 1\mu m$). Z_3 and Z_4 are dimensionless quantities.

[35, 16], as described in [30]. Briefly, sequential importance sampling yields the likelihood, as well as estimates of the unobserved cell velocities, by recursive solution of the filtering and prediction problems. The prediction problem requires simulating from the SDE describing velocity, for which techniques are given in [36]. The filtering problem is then solved by resampling from these simlulated paths with weights proportional to their likelihood given the observed positions. Alternative methods for fitting partially observed diffusion models include Monte Carlo Markov chain [50], Gaussian approximation importance sampling [20], and the EM algorithm [47]. Table 2 presents the resulting estimates, measuring distance in μm, angle in radians, and time in units of 10 minutes.

		treatment		control		p-value for
		estimate	SE_2	estimate	SE_2	difference
M6	α	0.559	0.120	0.609	0.069	0.72
	β	0.455	0.157	0.212	0.116	0.21
	σ	7.94	1.29	7.34	0.78	0.69
M7	α	0.523	0.177	0.592	0.130	0.75
	β	4.93	1.02	0.85	0.78	0.001
	σ	7.22	1.39	7.24	0.93	0.99

Table 2: Parameter estimates for models (M5) and (M6). Estimates are via maximum likelihood, with errors given by the sandwich estimator [54, 30, 32]. A normal approximation gives approximate two-sided p-values for the difference between treatment and control.

The estimates of β divided by their SE's for models (M5) and (M6) are similar to the means of the statistics Z_1 and Z_2 in SE units. This may lead one to suspect that the data are better represented by (M6) than (M5). There are many formal and informal ways of

assessing model fit. Formally one can test model specification using the χ^2 test of [54], or by comparing a nested sequence of models [5]. The clearest evidence found by the author for prefering (M6) to (M5) is a plot of magnitude of cell displacement against direction, displayed in Fig. 2. This plot, shown for the treatment group of cells and for simulations under both models with their fitted parameter values, demonstrates qualitative agreement between the data and model (M6). For both the data and (M6) there are relatively few occasions where cells travel more than five pixels in the direction of the anode. Model (M5), however, acquires its anisotropic behaviour by having fewer occurrences of cells traveling in an anodal direction. Those cells that do travel towards the anode for (M5) do so with almost as large displacements as those traveling towards the cathode.

6 Discussion

This paper has developed models for cell translocation using the tools of stochastic calculus combined with appropriate techniques for statistical inference. Methods for modeling single particles are relevant for other biological systems such as subcellular particle tracking [48, 26] and analysis of telemetry data on animal populations [45, 9]. In principle, the SDE approach allows generalization to interacting particles either through using correlated Brownian motions to drive the SDE [8] or through including an interaction in the infinitesimal drift. Investigating cell-cell interactions using these methods would be an exciting development of our work in this paper.

Modes of taxis can be given concise definitions in the context of models for cell velocity. This can aid interpretation of experimental results. In the case of galvanotaxis it has been previously observed that a DC field, while causing directional behaviour, does not affect cell speed. For our analysis in Section 5.3 we found that the DC field does not affect cell speed or persistence in the sense that estimates of σ and α are similar for treatment and control, but not in the alternative sense (which is a property of the discredited model (M5)) that the distribution of cell speed is independent of cell direction.

In summary, models of cell translocation play a key role in quantitative and qualitative understanding of experimental results. They are also required as input to computer simulations of biological processes. There are many cell types and situations in which one might be interested in studying their motion. It is reassuring to be able to write down, fit and assess a wide class of plausible models, though relatively simple Ornstein-Uhlenbeck type models may still prove to be adequate.

7 Acknowledgement

The authors acknowledge many helpful discussions with David Brillinger. John Ionides and Mitzi Nakatsuka kindly assisted with typing and the preparation of the figures. Several improvements have resulted from helpful suggestions by the reviewers.

References

[1] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *J. Math. Biol.*, 9:147–177, 1980.

[2] W. Alt, A. Deutsch, and G. Dunn. *Dynamics of Cell and Tissue Motion*. Birkhäuser, Basel, 1997.

[3] W. Alt and G. Hoffmann. *Biological Motion*. Springer, Berlin, 1990.

[4] M. Bailly, J. S. Condeelis, and J. E. Segall. Chemoattractant-induced lamellipod extension. *Microsc. Res. Tech.*, 43:433–443, 1998.

[5] G. E. P. Box and G. M. Jenkins. *Time Series Analysis: Forecasting and Control*. Holden–Day, San Francisco, 1970.

[6] D. Bray. *Cell Movements*. Garland Publishing, New York, 1992.

[7] D. R. Brillinger. A particle migrating randomly on a sphere. *J. Theoret. Probab.*, 10:429–443, 1997.

[8] D. R. Brillinger, H. K. Preisler, A. A. Ager, and J. G. Kie. An exploratory data analysis (EDA) of the paths of moving animals. *J. Stat. Plan. Infer.*, submitted, 2001.

[9] D. R. Brillinger and B. S. Stewart. Elephant-seal movements: Modelling migration. *Canadian J. Statist.*, 26:431–443, 1998.

[10] H. M. Byrne, G. Cave, and D. McElwain. The effect of chemotaxis and chemokinesis on leukocyte locomotion: A new interpretation of experimental results. *IMA J. Math. Appl. in Medic. And Biol.*, 15:235–256, 1998.

[11] M. Dembo. Field theorems of the cytoplasm. *Comments Theor. Biol.*, 1:159–177, 1989.

[12] R. B. Dickinson and R. J. Tranquillo. Optimal estimation of cell movement indices from the statistical analysis of cell tracking data. *AIChE J.*, 39:1995–2010, 1993.

[13] R. B. Dickinson and R. J. Tranquillo. A stochastic model for adhesion-mediated cell random motility and haptotaxis. *J. Math. Biol.*, 31:563–600, 1993.

[14] R. B. Dickinson and R. T. Tranquillo. Transport equations and indices for random and biased cell migration based on single cell properties. *SIAM J. Appl. Math.*, 55:1419–1454, 1995.

[15] P. A. DiMilla, J. A. Quinn, S. M. Albelda, and D. A. Lauffenburger. Measurement of individual cell migration parameters for human tissue cells. *AIChE J.*, 38:1092–1104, 1992.

[16] A. Doucet, N. de Freitas, and N. J. Gordon. *Sequential Monte Carlo methods in practice*. Springer, New York, 2001.

[17] P. G. Doucet and G. A. Dunn. Distintion between kinesis and taxis in terms of system theory. In W. Alt and G. Hoffmann, editors, *Biological Motion*, pages 498–509. Springer, Berlin, 1990.

[18] G. A. Dunn. Characterising a kinesis response: Time averaged measures of cell speed and directional persistence. *Agents Actions Suppl.*, 12:14–33, 1983.

[19] G. A. Dunn and A. F. Brown. A unified approach to analyzing cell motility. *J. Cell Sci. Suppl.*, 8:81–102, 1987.

[20] J. Durbin and S. J. Koopman. *Time series analysis by state space models*. Oxford University Press, 2001.

[21] K. S. Fang, E. Ionides, G. Oster, R. Nuccitelli, and R. R. Isseroff. Epidermal growth factor receptor relocalization and kinase activity are necessary for directional migration of keratinocytes in dc electric fields. *J. Cell Science*, 112:1967–1978, 1999.

[22] B. E. Farrell, R. P. Daniele, and D. A. Lauffenburger. Quantitative relationships between single-cell and cell-population model parameters for chemosensory migration responses of alveolar macrophages to C5a. *Cell Motil. Cytoskel.*, 16:279–293, 1990.

[23] R. M. Ford and D. A. Lauffenburger. Measurement of bacterial random motility and chemotaxis coefficients: II. Application of single-cell-based mathematical model. *Biotechnology and Bioengineering*, 37:661–672, 1991.

[24] R. M. Ford, B. R. Phillips, J. A. Quinn, and D. A. Lauffenberger. Measurement of bacterial random motility and chemotaxis coefficients: I. Stopped-flow diffusion chamber assay. *Biotechnology and Bioengineering*, 37:647–660, 1991.

[25] C. W. Gardiner. *Handbook of Stochastic Methods*. Springer, New York, second edition, 1985.

[26] P. A. Giardini and J. A. Theriot. Effects of intermediate filaments on actin-based motility of *listeria monocytogenes*. *Biophys. J.*, 81:3193–3203, 2001.

[27] H. Gruler and R. Nuccitelli. The galvanotaxis response mechanism of keratinocytes can be modeled as a proportional controler. *Cell Biochem. Biophys.*, 33:33–51, 2000.

[28] A. C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1989.

[29] A. Huttenlocher, R. R. Sandborg, and A. F. Horwitz. Adhesion in cell migration. *Curr. Opin. Cell Biol.*, 7:697–706, 1995.

[30] E. L. Ionides. *Statistical analysis of cell motion*. PhD thesis, University of California, Berkeley, 2001.

[31] S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, New York, 1981.

[32] G. Kauermann and R. J. Carroll. A note on the efficiency of sandwich covariance matrix estimation. *J. Amer. Statist. Assoc.*, 96:1387–1496, 2001.

[33] E. F. Keller and L. A. Segel. Model for chemotaxis. *J. Theor. Biol.*, 30:225–235, 1971.

[34] D. G. Kendall. Pole-seeking Brownian motion and bird navigation. *J. Roy. Statist. Soc. Ser. B*, 36:365–417, 1974.

[35] G. Kitagawa. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *J. Comput. Graph. Statist.*, 5:1–25, 1996.

[36] P. E. Kloeden and E. Platen. *Numerical Soluion of Stochastic Differential Equations*. Springer, New York, 1992.

[37] R. S. Langer and J. P. Vacanti. Tissue engineering: The challenges ahead. *Scientific American*, 280(4):86–89, 1999.

[38] D. A. Lauffenburger and J. J. Linderman. *Receptors: Models for Binding, Tracking and Signalling*. Oxford University Press, New York, 1993.

[39] G. Maheshwari and D. A. Lauffenburger. Deconstructing (and reconstructing) cell migration. *Microscopy Research and Technique*, 43:358–368, 1998.

[40] A. Mogilner and G. Oster. The physics of lamellipodial protrusion. *European Biophysical J.*, 25:47–53, 1996.

[41] D. J. Mooney and A. G. Mikos. Growing new organs. *Scientific American*, 280(4):60–65, 1999.

[42] K. Y. Nishimura, R. R. Isseroff, and R. Nuccitelli. Human keratinocytes migrate to the negative pole in DC electric fields comparable to those measured in mammalian wounds. *J. Cell Sci.*, 109:199–207, 1996.

[43] B. Oksendal. *Stochastic Differential Equations*. Springer, New York, 1998.

[44] N. Parenteau. Skin: The first tissue-engineered products. *Scientific American*, 280(4):83–84, 1999.

[45] H. K. Preisler, D. R. Brillinger, A. A. Ager, and J. Kie. Analysis of animal movement using telemetry and gis data. In *Proc. ASA Section on Statistics and the Environment*, pages 100–105, 1999.

[46] C. R. Rao. *Linear Statistical Inference and Its Applications*. Wiley, New York, 2nd edition, 1973.

[47] G. O. Roberts and O. Stramer. On inference for partially observed nonlinear diffusion models using the Metropolis-Hastings algorithm. *Biometrika*, 88:603–621, 2001.

[48] M. J. Saxton. Single-particle tracking: The distribution of diffusion coefficients. *Biophys. J.*, 72:1744–1753, 1997.

[49] A. D. Shenderov and M. P. Sheetz. Inversely correlated cycles in speed and turning in an amoeba: An oscillatory model of cell locomotion. *Biophysical J.*, 72:2382–2389, 1997.

[50] N. Shephard and M. K. Pitt. Likelihood analysis of non-Gaussian measurement time series. *Biometrika*, 84:653–667, 1997.

[51] C. L. Stokes, D. A. Lauffenbuger, and S. K. Williams. Migration of individual microvessel endothelial cells; stochastic model and parmeter measurement. *J. Cell Sci.*, 99:419–430, 1991.

[52] D. W. Stroock. Some stochastic processes which arise from a model of the motion of a bacterium. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, 28:305–315, 1974.

[53] R. T. Tranquillo and W. Alt. Stochastic model of receptor-mediated cytomechanics and dynamic morphology of leukocytes. *J. Math. Biol.*, 54:361–412, 1996.

[54] H. White. Maximum likelihood estimation of mis-specified models. *Econometrica*, 50:1–26, 1982.

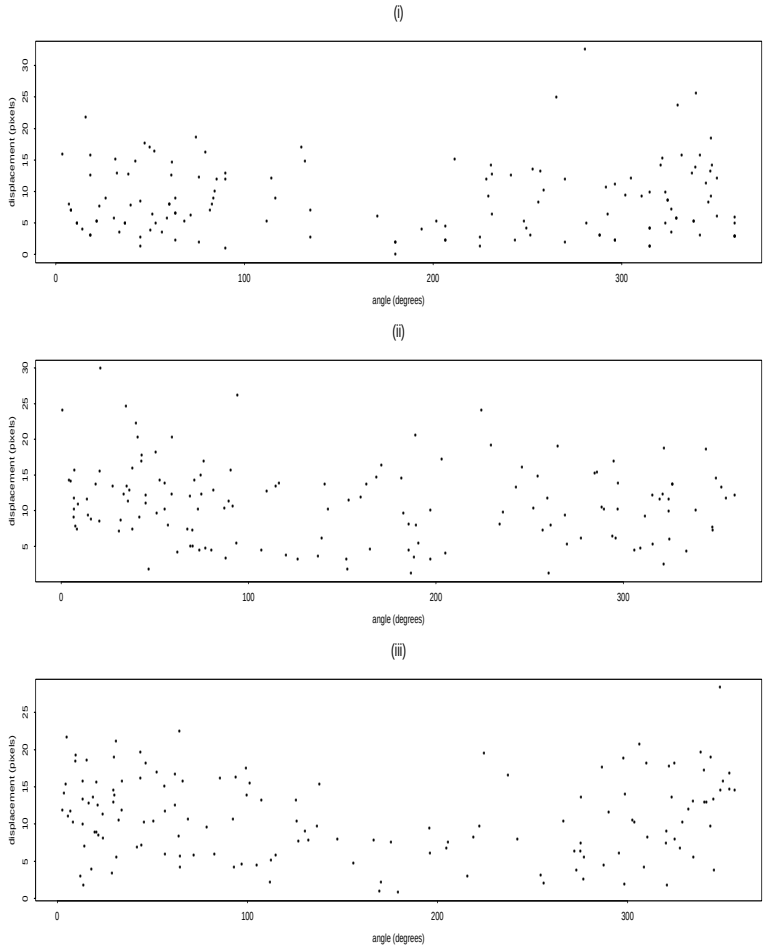

Figure 2: Setting $\mathbf{x}_{t+1}^{(i)} - \mathbf{x}_t^{(i)} = (r_t^{(i)} \cos\theta_t^{(i)}, r_t^{(i)} \sin\theta_t^{(i)})^T$, the displacement, $r_t^{(i)}$, is plotted against the angle, $\theta_t^{(i)}$, for each cell i and each time point t. (i) The treatment group. (ii) Simulated data for model (M5), using the fitted parameter values. (iii) Simulated data for model (M6), using the fitted parameter values.