Learning a Potential Function from a Trajectory

David R. Brillinger

Abstract— **This letter concerns the use of stochastic gradient systems in the modeling of the paths of moving particles and the consequent estimation of a potential function. The work proceeds by setting down a model for the potential function which leads to a stochastic differential equation. The method is simple, direct and flexible being based on a linear model and least squares. The estimated potential function may be used for: simple description, summary, comparison, seeking patterns, simulation, prediction, and model appraisal. Explanatories, attractors and repellors, may be included in the potential function directly. The large sample distribution of the estimated potential function is provided. There is an example analyzing the path of an elk. There are direct extensions to: updating, sliding window, adaptive, robust and real time variants.**

***Index Terms*—Mobility model, monitoring, potential function, stochastic differential equation, stochastic gradient system, surveillance, tracking, waypoint data.**

I. INTRODUCTION

Location signals of moving objects, obtained for example by GPS or LORAN, have become common in practice. Typically one has scattered positions along trajectories of the objects. The questions of how to summarize, how to predict, and how to simulate, such movements arise. This happens particularly when a number of paths are involved or the path of an object is a tangle. (See Figure 1 which shows 1571 locations along the track of an elk in Starkey Reserve, Oregon.) The setup of concern may be viewed as one of state space modeling with its many approaches and methods. One envisages a potential function as a spatial state variable and the paths of objects as determined by the measurement equation. This letter seeks to provide a unified approach for dealing with movement modelling and associated data.

The fields in which movement data have arisen include: animal tracking [1], [2], drifters [3], eyescans [4], [5], soccer, [6], and waypoint data [7]. There are various papers developing the potential approach [1], [2], [6], [14], [15] and [17]. This letter provides the formal background missing in that work and a general discussion.

A. Dynamical Systems and Potential Functions

David R. Brillinger, Statistics Department, University of California, Berkeley, CA 94720

The historical development of the concept of potential function may be described as: Brahe collected data on the paths of planets in the sky, then Kepler analysed that data to produce "laws", next Newton found differential equations consistent with Kepler's laws and produced further laws. Later Lagrange set down the potential function for the gravitational potential case namely $V(\boldsymbol{r}) = -G/|\boldsymbol{r}_0 - \boldsymbol{r}|$ with G the constant of gravitation where $\boldsymbol{r}$ denotes location in R^3, [8]. This function leads to attraction of a particle at the position $\boldsymbol{r}$ towards the position $\boldsymbol{r}_0$, see [9] pages 277-289, [10] pages 13-17. In the mathematical expressions all the vectors appearing are column vectors.

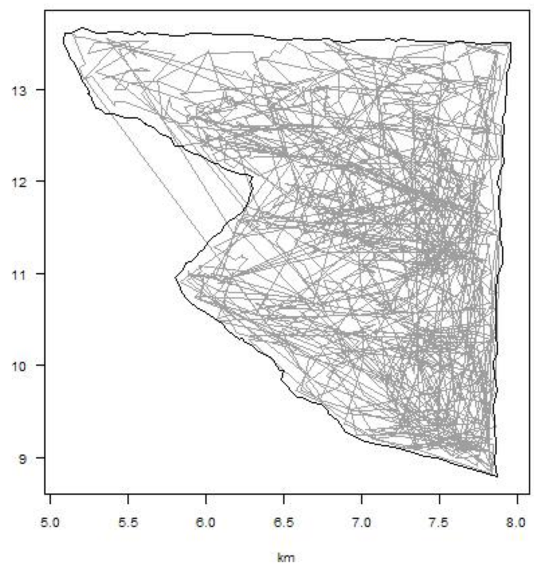

FIG 1. Path of an elk around the NE pasture of the Starkey Experimental Forest in Oregon. Locations were estimated every two hours and are joined consecutively by straight lines.

A potential function, $V(\boldsymbol{r})$, is scalar-valued. This leads to simpler representations of the motion than one based on modeling velocities for example. In the so-called overdamped case, (meaning in a physical situation that the force acting on a particle does not determine the acceleration, but rather its velocity) the equation of motion is

(phone: 510-642-0611, fax: 510-642-3860, email: brill@stat.berkeley.edu)

$$d\mathbf{r}(t) = -\nabla V(\mathbf{r}(t))dt \qquad (1)$$

where $V{:}R^p \rightarrow R$ is a differentiable function and ∇ denotes the gradient. (The negative sign in (1) is traditional.) The entity $d\boldsymbol{r}(t)/dt$ is called a vector field. Reference [9], page 203, defines a gradient system in R^p as a system of differential equations of the form (1). Here $\nabla V = (\partial V/\partial x_1, \ldots, \partial V/\partial x_p)^T$, with $\boldsymbol{r} = (x_1, \ldots, x_p)^T$ and "T" denotes transpose. When $p = 2$ the level surfaces of the potential function are conveniently displayed in contour form and its gradient as arrows on a grid. (See Figures 2 and 3 below.)

B. Stochastic Gradient Systems

The estimation method to be presented can be motivated by stochastic gradient systems that is systems that can be written, in the time invariant case, as

$$d\boldsymbol{r}(t) = -\nabla V(\boldsymbol{r}(t))\, dt + \boldsymbol{\sigma}(\boldsymbol{r}(t))d\boldsymbol{B}(t) \qquad (2)$$

for some differentiable V with $\boldsymbol{B}(t)$ Brownian motion and $\boldsymbol{\sigma}$ a matrix. Brownian motion itself is an example of such a system corresponding to constant V and $\boldsymbol{\sigma}$. Expression (2) is a particular case of the stochastic differential equation (SDE)

$$d\boldsymbol{r}(t) = \boldsymbol{\mu}(\boldsymbol{r}(t))\, dt + \boldsymbol{\sigma}(\boldsymbol{r}(t))d\boldsymbol{B}(t) \qquad (3)$$

but what distinguishes the traditional SDE work from the present study is that the drift term $\boldsymbol{\mu}$ has the special form $-\nabla V$ for some real-valued function V.

II. PROBLEM AND APPROACH

The basic problem supposes the model (2) and seeks to learn $V(\boldsymbol{r})$ given data ($\boldsymbol{r}(t_i)$, $i = 1,\ldots,n$). These data will be viewed as scattered locations of a moving object with locations observed at increasing times t_i along a trajectory of the process (2). One seeks both vector field and potential function estimates.

Supposing $\nabla V(\boldsymbol{r})$ to be a smooth function of $\boldsymbol{r}$, and that the observation times t_i are close together, one can set down the following approximation to (2),

$$\boldsymbol{r}(t_{i+1})-\boldsymbol{r}(t_i)=-\nabla V(\boldsymbol{r}(t_i))(t_{i+1} - t_i) + \boldsymbol{\sigma}\sqrt{(t_{i+1} - t_i)}\boldsymbol{Z}_{i+1} \qquad (4)$$

for $i = 1, 2, 3, \ldots, n$ with, for example, the $\boldsymbol{Z}_i$ zero mean, unit variance independent variates and $\boldsymbol{r}(t_i)$ given.

The reason for the multiplier $\surd(t_{i+1} - t_i)$ is that for real-valued Brownian $Var(dB(t)) = dt$ however other multipliers can be considered. Equation (4) may be set down so the error term has constant covariance matrix, namely

$$(\boldsymbol{r}(t_{i+1})-\boldsymbol{r}(t_i))/\surd(t_{i+1} - t_i)=-\nabla V(\boldsymbol{r}(t_i))\surd(t_{i+1} - t_i)+\boldsymbol{\sigma Z}_{i+1} \quad (5)$$

The approximation

$$(\boldsymbol{r}(t_{i+1})-\boldsymbol{r}(t_i))/(t_{i+1} - t_i) = \boldsymbol{\mu}(\boldsymbol{r}(t_i)) + \boldsymbol{\sigma Z}_{i+1} /\surd(t_{i+1} - t_i) \quad (6)$$

for the SDE (3) was employed in [1], [2] for elk and deer movement. There it was noted that the relationship (5) had the form of a smoothing/nonparametric regression problem as in [11], Chapter 3. In [1] an early attempt was made at estimating a potential function by numerical integration and simulation. The question was also asked of whether the vector field, $\boldsymbol{\mu}$, has the form $-\nabla V(\boldsymbol{r})$? This may be addressed by comparing the unrestricted estimate of $\boldsymbol{\mu}$ with the one assuming the existence of a potential function.

III. THE POTENTIAL FUNCTION

A basic issue is how to describe the potential function, $V(\boldsymbol{r})$, $\boldsymbol{r}$ in R^p, supposing one exists. General forms may be considered, but in the development here V is linear in a vector-valued parameter $\boldsymbol{\beta}$ making study and estimation easier. Suppose $V(\boldsymbol{r}) = \boldsymbol{\varphi}(\boldsymbol{r})^T\boldsymbol{\beta}$ with $\boldsymbol{\varphi}$ an L by 1 vector of known functions and $\boldsymbol{\beta}$ an L by 1 unknown parameter. The gradient of V is the p by 1 vector $\nabla \boldsymbol{\varphi}(\boldsymbol{r})^T\boldsymbol{\beta}$. Examples of such a V follow.

Example 1. Polynomial expansion.

Consider $V(\boldsymbol{r}) = \sum \boldsymbol{\beta}_{\boldsymbol{m}}\boldsymbol{r}^{\boldsymbol{m}}$ where $\boldsymbol{m} = (m_1,\ldots, m_p)$, and $\boldsymbol{r}^{\boldsymbol{m}} = \{x_1$ to the power $m_1\}$ times … times $\{x_p$ to the power $m_p\}$ and the sum is over nonnegative integers $m_1,\ldots, m_p$ with $1 \le m_1+\ldots+m_p \le M$.

One could set down a finite trigonometric polynomial instead of the ordinary polynomial here. In a time domain problem, speaker adaptation, [Au-Yeung and Siu] use a polynomial expansion.

Example 2. Node based.

Consider nodal points $\boldsymbol{u}_l$, $l=1,\ldots,L$ in R^p and set $V(\boldsymbol{r}) = \sum \boldsymbol{\beta}_l K(\boldsymbol{r}-\boldsymbol{u}_l)$. for some real-valued differentiable kernel K. As a specific example of K one has the radial basis thin plate splines, [12] and [13], pages 30-34,

$K(\boldsymbol{r})=|\boldsymbol{r}|^{2q-p}\log|\boldsymbol{r}|$ for p even and $=|\boldsymbol{r}|^{2q-p}$ for p odd (7)

Here q denotes the order of differentiability of K , $2q-p>0$, and $|\boldsymbol{r}| = \sqrt{\boldsymbol{r}^T\boldsymbol{r}}$.

Example 3. Attraction and repulsion.

Consider a region A and a point $\boldsymbol{r}$ outside A. Potential functions can be set down allowing attraction or repulsion from A. Specifically let $d_A(\boldsymbol{r})$ denote the minimum distance from $\boldsymbol{r}$ to A and set $V(\boldsymbol{r}) = d_A(\boldsymbol{r})^\alpha$. For $\alpha > 0$ one has attraction to A and repulsion if $\alpha < 0$. One can reverse attraction and repulsion by changing the sign of d_A. It is also convenient to use $V(\boldsymbol{r}) = \beta_1 \log d_A(\boldsymbol{r}) + \beta_2 d_A(\boldsymbol{r})$.

The functional forms of Examples 1-3 may be added together to provide other forms. The gradient may be evaluated analytically for each of these examples and this is necessary for the fitting procedure described in the next section.

Reference [14] considers the observed trajectory of a monk seal near the island of Molokai employing the function

$$V(\boldsymbol{r}) = \gamma_1 x + \gamma_2 y + \gamma_{11} x^2 + \gamma_{12} xy + \gamma_{22} y^2 + C/d(x,y) \quad (8)$$

where $\boldsymbol{r} = (x,y)$, is in R^2, and represents location on the ocean surface. The value $d(x,y)$ is the distance from the seal's location $\boldsymbol{r}$ to the nearest point on the island. The γ 's and C are unknown parameters to be estimated. The final term in (8) keeps the seal off of the island. Another example, [6], studies the motion of a soccer ball in a game. The potential function employed there is

$$V(\boldsymbol{r})=\alpha \log d(\boldsymbol{r})+\beta d(\boldsymbol{r})+\gamma_1 x+\gamma_2 y + \gamma_{11} x^2 +\gamma_{12} xy+\gamma_{22} y^2$$

with $d(\boldsymbol{r})$ the shortest distance to the goal mouth from $\boldsymbol{r} = (x,y)$. The first two terms lead to attraction to the goalmouth. A different monk seal is studied in [15] and a time dependent potential function is employed namely,

$$V(\boldsymbol{r},t) = \alpha \log d(\boldsymbol{r},t) + \beta\, d(\boldsymbol{r},t)$$

with $d(\boldsymbol{r},t)$ the distance from an attractor at time t. In this case the location of attraction switched depending on whether the animal was on an outbound or inbound journey. In each of these examples $V(\boldsymbol{r})$ is linear in the parameter. Potential function and vector field estimates are provided in the papers just referenced. The successful implementation of the proposed method for potential estimation is encouraging.

IV. ESTIMATION

The representation (4) with $\boldsymbol{r}$ in R^p and $\nabla V(\boldsymbol{r}) = \nabla \boldsymbol{\varphi}(\boldsymbol{r}_i)^T\boldsymbol{\beta}$ will be employed. The values $\boldsymbol{r}(t_i)$ will be written $\boldsymbol{r}_i$. Define the p by 1 vector $\boldsymbol{y}_{i+1} = (\boldsymbol{r}_{i+1}-\boldsymbol{r}_i)/\sqrt{(t_{i+1}-t_i)}$. Following expression (4) one has the form

$$\boldsymbol{y}_{i+1} = -\nabla \boldsymbol{\varphi}(\boldsymbol{r}_i)^T\boldsymbol{\beta}\sqrt{(t_{i+1} - t_i)} + \boldsymbol{\sigma Z}_{i+1},\ i = 1,\ldots,n \qquad (9)$$

involving the L by 1 vector $\boldsymbol{\beta}$, the L by p matrix $\nabla \boldsymbol{\varphi}(\boldsymbol{r}_i)$, the p by p matrix $\boldsymbol{\sigma}$, and the p by 1 vector $\boldsymbol{Z}_{i+1}$. Suppose, for simplicity, $\boldsymbol{\sigma} = \sigma \boldsymbol{I}$ with σ real-valued and with $\boldsymbol{I}$ the p by p identity matrix. Stack the n values $\boldsymbol{y}_{i+1}$, i=1,…,n vertically to form an np by 1 vector $\boldsymbol{Y_n}$. Stack the n matrices $-\nabla \boldsymbol{\varphi}(\boldsymbol{r}_i)^T \sqrt{(t_{i+1} - t_i)}$ to form the np by L matrix $\boldsymbol{X}_n$. Let $\boldsymbol{x}_i$ denote the i-th row of $\boldsymbol{X}_n$ so $\boldsymbol{X}_n^T\boldsymbol{X}_n = \sum \boldsymbol{x}_i\boldsymbol{x}_i^T$. Stack the n values $\sigma\boldsymbol{Z}_{i+1}$ to form $\boldsymbol{\varepsilon}_n$. Then one has the regression model $\boldsymbol{Y_n} = \boldsymbol{X_n\beta} + \boldsymbol{\varepsilon}_n$ with the difference from ordinary regression that $\boldsymbol{X_n}$ is random. Supposing the matrix $\boldsymbol{X}_n^T\boldsymbol{X}_n$ to be nonsingular one can compute the ordinary least squares estimate $\boldsymbol{b} = (\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}\boldsymbol{X}_n^T\boldsymbol{Y}_n$ of $\boldsymbol{\beta}$ and then $\boldsymbol{\varphi}(\boldsymbol{r})^T\boldsymbol{b}$ as an estimate of $V(\boldsymbol{r})$.

By setting $\boldsymbol{x}_i = -\nabla \boldsymbol{\varphi}(\boldsymbol{r}_i)$ and supposing the ε_{ij} of $\boldsymbol{\varepsilon}_n = [\varepsilon_{ij}]$ to be independent, zero mean, variance σ^2 variates asymptotic properties of $\boldsymbol{\varphi}(\boldsymbol{r})^T\boldsymbol{b}$ may be obtained from Theorem A.1 below. The theorem, and some additional assumptions needed, is found in [16].

One can compute $s_n^2 = n^{-1}\sum(\boldsymbol{y}_i-\boldsymbol{x}_i^T\boldsymbol{b})^T(\boldsymbol{y}_i-\boldsymbol{x}_i^T\boldsymbol{b})$ as an estimate of σ^2 and, for example, then set down a confidence interval for $\boldsymbol{\varphi}(\boldsymbol{r})^T\boldsymbol{\beta}$ using the results of Theorem A.2. Specifically, provided lim log $\lambda_{\max}(\boldsymbol{X}_n^T\boldsymbol{X}_n)/n \to 0$ a.s., one has

$$(\boldsymbol{\varphi}(\boldsymbol{r})^T(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{-1}\boldsymbol{\varphi}(\boldsymbol{r}))^{1/2}\boldsymbol{\varphi}(\boldsymbol{r})^T(\boldsymbol{b}-\boldsymbol{\beta})/s_n \to N(0,1),$$

the standard normal, in distribution as $n\to\infty$. This last leads to the approximate 100(1-α)% confidence interval

$$\boldsymbol{\varphi}(\mathbf{r})^T \boldsymbol{\beta} = \boldsymbol{\varphi}(\mathbf{r})^T \mathbf{b} \pm z_{\alpha/2} (\boldsymbol{\varphi}(\mathbf{r})(\mathbf{X}_n^T\mathbf{X}_n)^{-1}\boldsymbol{\varphi}(\mathbf{r})^T)^{1/2}$$

where $z_{\alpha/2}$ denotes the 100α/2 percent point of the standard normal.

V. AN EXAMPLE

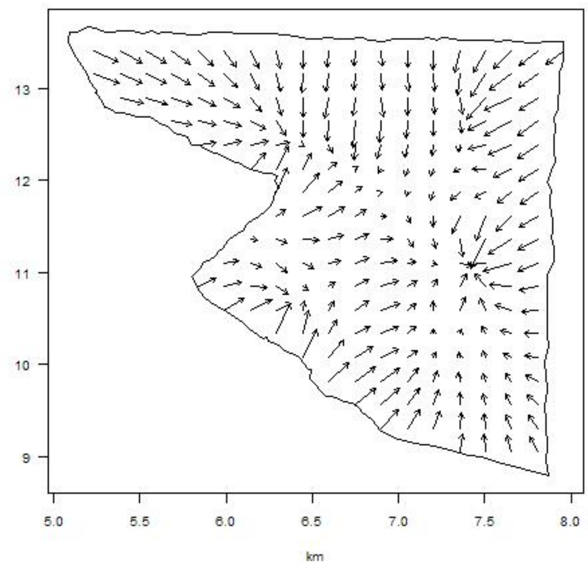

FIG 2. The estimated vector field for the path of Figure1.

The Starkey Reserve is a large area in Oregon set aside to study the interactions of elk, deer, cows and man sharing an environment. There is a high fence around the Reserve.Figure 1 shows a sampled trajectory of one of the elk.There were 1571 GPS locations taken with a time interval of approximately 2

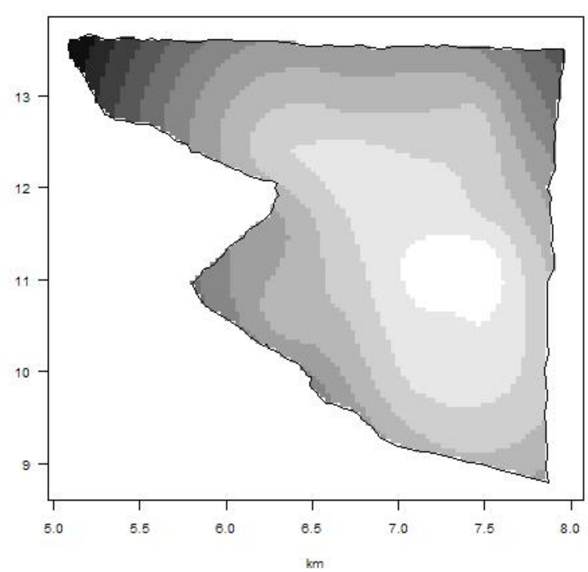

FIG 3. The estimated potential for the path of Figure 1. The lighter shading corresponds to smaller values.

hours between successive locations [17].

A potential function $V(\boldsymbol{r})$ was approximated by thin plate radial basis splines employing the kernel function of (7) with p and $q = 2$. The coefficients $\boldsymbol{\beta}_l$ were estimated by ordinary least squares employing the model (5). The results are provided in Figures 2 and 3. One sees the confusion of Figure 1 much reduced. A point of attraction appears near (7.5,11.0). Figure 3 provides an image plot of the potential function. Now one sees the point of attraction with but a glance.

VI. EXTENSIONS

Various generalizations of the letter's results may be mentioned. The references listed contain some mention to the ideas. One could set down an expansion for V employing wavelet functions, [18]. One could consider updating methods for real-time work and also robust estimation, [19]. If the potential function is changing slowly one could consider a sliding window estimate., [20]. In video analysis one might consider the model $I(\boldsymbol{r},t) = I_0(\boldsymbol{r}) + \delta(\boldsymbol{r}(t)\text{-}\boldsymbol{r})$ with t indexing frame, I_0 representing background, and δ the Dirac delta [21], [22]. Lastly once could consider adaptive estimates. [23].

Because of the simplicity of the approach obtaining results for these cases appears quite direct.

VII. CONCLUSION

This letter presents a novel estimation method for handling moving objects. the computations may be implemented by least squares algorithms.

APPENDIX

*Theorem A.*1. [16] Consider the regression model $\boldsymbol{y}_i = \boldsymbol{x}_i^T\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i$, i=1,2,… with $\{\boldsymbol{\varepsilon}_i\}$ martingale differences with respect to an increasing sequence of σ-fields $\{F_n\}$. Suppose that $\sup_n \mathrm{E}(\|\boldsymbol{\varepsilon}_n\|^\alpha|F_{n-1}) < \infty$ a.s. for some $\alpha > 2$. Suppose further that $\lim_{n\to\infty} \mathrm{var}\{\boldsymbol{\varepsilon}_n|F_{n-1}) = \boldsymbol{\sigma}^2$ almost surely for some nonstochastic $\boldsymbol{\sigma}$. Assume that $\boldsymbol{x}_n$ is a F_{n-1}-measurable random variable and that there exists a non-random positive definite symmetric L by L matrix $\boldsymbol{B}_n$ for which $\boldsymbol{B}_n^{-1(}\boldsymbol{X}_n^T\boldsymbol{X}_n)^{1/2}\rightarrow\boldsymbol{I}$ and $\sup_{1\le i\le n}\| \boldsymbol{B}_n^{-1}\boldsymbol{x}_n\|\rightarrow 0$ in probability. Then

$$(\boldsymbol{X}_n^T\boldsymbol{X}_n)^{1/2}(\boldsymbol{b}\text{-}\boldsymbol{\beta}) \rightarrow N(0, \sigma^2\boldsymbol{I})$$

in distribution as $n\rightarrow\infty$.

Note that 0 mean independent observations like the $\boldsymbol{\sigma Z}_{i+1}$ of (5) form a martingale difference sequence with respect to the σ-field F_i generated by $\{\boldsymbol{r}(t_1),\ldots,\boldsymbol{r}(t_i)\}$.

Theorem A.2. Under the assumptions of Theorem A.1 and lim log $\lambda_{\max}(\boldsymbol{X}_n{}^T\boldsymbol{X}_n)/n \to 0$ almost surely, one has

$$((\boldsymbol{\varphi}(\boldsymbol{r})(\boldsymbol{X}_n{}^T\boldsymbol{X}_n)^{-1}\boldsymbol{\varphi}(\boldsymbol{r})^T)^{-1/2}\boldsymbol{\varphi}(\boldsymbol{r})^T\,(\boldsymbol{b}\text{-}\boldsymbol{\beta})/s_n \to N(0,1)$$

in distribution as $n \to \infty$.

The additional assumption in A.2 is to have the almost sure convergence of s_n to σ.

ACKNOWLEDGEMENT

The author thanks A. A. Ager and H. K. Preisler for involving him with the Starkey data set that motivated the potential function approach to track problems.

REFERENCES

[1] Brillinger, D. R., Preisler, H. K., Ager, A. A. and Kie, J. G., (2001a). The use of potential functions in modelling animal movement. Pp. 369-386 in *Data Analysis from Statistical Foundations* (Ed. A. K. M. E. Saleh). Nova Science, Huntington. www.stat.berkeley.edu/~brill/Papers/fraser8.pdf

[2] Brillinger, D. R., Preisler, H. K., Ager, A. A. Kie, J., and Stewart, B. S. (2001b). Modelling movements of free-ranging animals, *University of California Berkeley Statistics Technical Report 610*. www.stat.berkeley.edu/~brill/Preprints/610.pdf

[3] Etre (2006). Eyetracking. www.etre.com/usability/eyetracking/

[4] Lumpkin, R. and Pazos, M. (2007). Measuring surface currents with Surface Velocity Program drifters: the instrument, its data, and some recent results. *Lagrangian Analysis and Prediction of Coastal and Ocean Dynamics,* Eds. Griffa,, A., Kirwan, A. D., Mariano, A. J., Özgökmen, T., and Rossby, H. T., *Cambridge University Press.*

[5] Tourangeau, R., Couper, M. P., Galesic, M. (2005). Use of eye-tracking for studying survey response processes. *ESF Workshop on Internet Survey Methodology*, Dubrovnik, 26-28 September 2005.

[6] Brillinger, D. R. (2007). A potential function approach to the flow of play in soccer. *Journal of Quantitative Analysis of Sports*, to appear, January. www.stat.berkeley.edu/~brill/Papers/goal21.pdf

[7] Navidi, W. and Camp, T. (2004). Stationary distributions for the random waypoint mobility model. *IEEE Trans. Mobile Comp.* 3, 99-108.

[8] Hirsch, M. W. (1984). The dynamical systems approach to differential equations. *Bull. Amer. Math. Soc.* 12, 1-64.

[9] Hirsch, M. W., Smale, S. and Devaney, R. L. (2004). *Differential Equations, Dynamical Systems and An Introduction to Chaos*, Second Edition. Elsevier, Amsterdam.

[10] Taylor, J. R. (2005). *Classical Mechanics*. University Science, Sausalito.

[11] Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*. Chapman and Hall, London.

[12] Duchon, J. (1977). Splines minimizing rotation-invariant semi-norms in Sobolev spaces. *Constructive Theory of Functions of Several Variables*. Lecture Notes in Mathematics 571. Springer, Berlin.

[13] Wahba, G. (1980). *Spline Models for Observational Data*. SIAM, Philadelphia.

[14] Brillinger, D. R., Stewart, B., and Littnan, C. (2006b). A meandering hylje, *Festschrift for Tarmo Pukkila on His 60th Birthday*, Eds. E. P. Liski, J. Isotalo, S. Puntanen, and G.P.H. Styan Dept. of Mathematics, Statistics and Philosophy, Univ. of Tampere. www.stat.berkeley.edu/~brill/Papers/hylje.pdf

[15] Brillinger, D. R., Stewart, B., and Littnan, C. (2006a). Three months journeying of a Hawaiian monk seal. *Lecture Notes in Statistics*. www.stat.berkeley.edu/~brill/Papers/monk32.pdf

[16] Lai,T. L. and Wei, C. Z. (1982). Least squares estimation in stochastic regression models with applications to identification and control of dynamic systems. *Annals of Statistics* 10, 154-166.

[17] Brillinger, D. R., Preisler, H. K., Ager, A. A. and Wisdom, M. J. (2004). Stochastic differential equations in the analysis of wildlife motion. *2004 Proceedings of the American Statistical Association, Statistics & the Environment Section*. http://www.stat.berkeley.edu/~papers/jsm2004j.pdf

[18] Blu, T. and Unser, M. (2002). Wavelets, fractals, and radial basis functions. *IEEE Trans. Signal Processing* 50, 543-553.

[19] Yang, B., Zhang, Z. and Sun, Z. (2006). Computing nonlinear τ-estimation based on dynamic differential evolution strategy. *IEEE Signal Processing Letters* 13, 756-759.

[20] Belge, M. and Miller, E. L. (2000). A sliding window RLS-like adaptive algorithm for filtering alpha-stable noise. *IEEE Signal Processing Letters* 13, 86-89.

[21] Zlokolica, V., Pizurica, A. and Philips, W. (2006). Noise estimation for video processing based on spatio-temporal gradients. *IEEE Signal Processing Letters* 13, 337-340.

[22] Piroddi, R. and Vlachos, T. (2006). A simple framework for spatio-temporal video segmentation and delayering using dense motion fields. *IEEE Signal Processing Letters* 13, 421-424.

[23] Au-Yeung, S-K. and Siu, M-H. (2006). Maximum likelihood linear regression adaptation for polynomial segment models. *IEEE Signal Processing Letters* 13, 644-647.