STOCHASTIC DIFFERENTIAL EQUATIONS IN THE ANALYSIS OF WILDLIFE MOTION

D. R. Brillinger[1], H. K. Preisler[2], A. A. Ager[3], and M. J. Wisdom[3]

The University of California, Berkeley[1]

USDA, Forest Service, Pacific Southwest[2] and Northwest[3] Research Stations

http://www.stat.berkeley.edu/∼brill

Introduction.

This paper presents a method to describe analytically the motion of animals both ranging freely and ranging in a confined region. The motion of the animals is modelled as being affected by their location and, in some cases, the presence of deliberately introduced human "intruders". The interest is the effects of the human intruders on the animals' behavior. The locations of the animals are available at frequent and approximately regular time intervals as are the locations of the intruders.

The data were collected at the Starkey Experimental Forest and Range (Starkey), northeast Oregon, as part of a long term study by the U.S. Forest Service. The study concerned the question of whether recreational uses by humans affect elk (*Cervus elaphus*) and mule deer (*Odocoileus hemionus*), two animal species of keen economic, social, and aesthetic interest to users of National Forests. Further details about Starkey and the recreation experiment may be found in Preisler et al. (2004) and Wisdom et al. (2004). Additional background about the study area and overall research conducted there was described by Rowland et al. (1997) and Brillinger et al. (2001a).

The principal tool employed in the work is a stochastic differential equation, with the drift term depending on location of the elk and the locations of the moving human intruders. The sections of the papers are: the experiment and data, exploratory data analyses, stochastic differential equations, models and results, the case of a boundary, simulation, and discussion.

The experiment and the data.

Data were collected on the movement of animals in the Northeast Study Area of Starkey during the period April to October, 2003. The 3,580-ha study area is enclosed by a 2.4-m high, ungulate-proof fence. The size and shape of the study area, as defined by the ungulate-proof fence, is shown in the figures below.

The study area contains elk and mule deer, but for this analysis only elk are considered. For a given treatment type and period, one of four recreational activities were introduced: hikers, mountain bicyclists, horseback riders, or ATVs (all terrain vehicles). The interest was the effects of the human intruders on the elk. These people followed specific routes. They were GPS equipped and their location estimated about every second. The elk were also GPS equipped and their locations estimated about every 5 min. Each treatment type was applied for 5 day periods, followed by 9-day control periods where no treatments (no human intruders) were applied. There were 8 elk in the study. The order in which each treatment type was implemented was randomly selected, and specifically paired with a control period that immediately followed the randomly-selected treatment.

Thus, the experimental design involved a given 5-day treatment type, such as 5 days of ATV riding, followed by a 9 ATV-control period, followed by another 5 days of a treatment type (e.g. 5 days of hiking) and subsequent 9-day control period, and so on, through the entire study period of April to October. Further details of the experiment may be found in Wisdom et al. (2004).

Exploratory data analyses.

Let $\{\mathbf{r}(t)\}$ denote the path of an elk, giving its location as a function of time t. The components may be thought of as corresponding to latitude and longitude. Preliminary examination of the data included investigation for a circadian rhythm in elk via parallel hourly boxplots of their speed of movement. A purpose was to see if time of day needed to be included in the models of motion. Both control and ATV days were studied. The results are presented in Figure 1 below. In preparing this figure the speed at time t_i was estimated by

$$|\mathbf{r}(t_{i+1}) - \mathbf{r}(t_i)|/(t_{i+1} - t_i)$$

employing the 80% smallest $(t_{i+1} - t_i)$'s. The smallest 80% have been employed because the instantaneous velocity is better estimated when the time difference is small. Here $\mathbf{r}$ is a 2-entry column vector, and $|\mathbf{r}|^2 = r_1^2 + r_2^2$. In order to make the figure more readable, the square root of the speeds has been employed.

The plots differ for the control and ATV cases. There is some dependence on time of day for the control days, they appear to be moving more quickly at 0700 and 1900 hours. The general level of the median of the square-root speed is about .3 $\sqrt{km/hr}$ for the control days. During the hours when the ATV was in action there are many large values with an apparent velocity up to about 9 km/hour. These presumably correspond to animals speeding off once they sensed the approach of the ATV.

It is to be remembered that with the control treatment the elk observation times have a separation of 2 hours while for the ATV treatment the separation is 5 min. This may introduce some bias in the estimate of speed. With a model a correction could be developed.

Stochastic differential equations (SDEs).

In what follows it will be supposed that the motion of an elk is described by the following (stochastic differential) equation

$$d\mathbf{r}(t) \;=\; \mu(\mathbf{r}(t),t)dt \;+\; \sigma(\mathbf{r}(t),t)d\mathbf{B}(t) \qquad (1)$$

or in other form

$$\mathbf{r}(t) \;=\; \mathbf{r}(0) \;+\; \int_0^t \mu(\mathbf{r}(s),s)ds \;+\; \int_0^t \sigma(\mathbf{r},s)d\mathbf{B}(s) \qquad (2)$$

where $\mathbf{B}$ is bivariate Brownian, i.e. the increments $\mathbf{B}(t_{i+1}) - \mathbf{B}(t_i)$ are $IN_2(O,(t_{i+1}-t_i)\mathbf{I})$ for all time discretizations $\{t_i\}$ with $t_1 < t_2 < \ldots$. The function μ is called the drift rate and σ the diffusion coefficient. The functions μ, σ depend on location and time and are assumed smooth in the estimation procedures below.

Suppressing the dependence of μ, σ on $\mathbf{r}$ and t solutions to (1) may be approximated by

$$(\mathbf{r}(t_{i+1}) - \mathbf{r}(t_i))/(t_{i+1}-t_i) \;\approx\; \mu \;+\; \sigma\mathbf{Z}_{i+1}/\sqrt{t_{i+1}-t_i} \qquad (3)$$

The $\mathbf{Z}_i$ are $IN_2(0,\mathbf{I})$ and correspond to $(B(t_{i+1}) - B(t_i))/\sqrt{t_{i+1}-t_i})$. Expression (3) suggests how μ and σ might be estimated by least squares procedures.

Models and results.

Figure 2 shows an example of the data available. The points $\mathbf{r}_{ij}(t_i)$ are joined for given j, j indexing the 8 elk. The data of the figure are for the 9-day control periods that were paired with the 5-day periods of ATV riding. The points joined are those no farther apart in time than the 80th percentile of all the time differences. The figure shows a sequence of straight line segments. One sees points of concentration and indications of routes. Two basic models are now considered for the paths, the first does not involve explanatory variables while the second does.

Model I. The control case. Suppose that the drift, μ, depends smoothly on location, $\mathbf{r}$, and one has

$$d\mathbf{r} \;=\; \mu(\mathbf{r})dt \;+\; noise \qquad (4)$$

The model (4) may be fit in the manner of Brillinger et al. (2001a). The resulting estimated velocity field, $(\hat{\mu}_1(\mathbf{r}),\hat{\mu}_2(\mathbf{r}))$, is plotted in Figure 3 for both ATV and control periods. One sees longer arrows, i.e. the animals are moving faster, for the ATV periods. For the control periods one sees the elk moving away from the fences, being led to the lower right, and then swept into the middle. In the case of the ATVs the spacing of the time points is 5 min, while it is 2 hours for the control periods.

Model II. Intruders. Suppose that at time t the animal is influenced by where the ATV is at time t. If the ATV is very close one can to imagine that the behavior of the animal is altered in some fashion, e.g. the animal runs away. Supposing that the ATV is at $\mathbf{x}(t)$ at time t consider the model

$$d\mathbf{r}(t) \;=\; \mu(\mathbf{r})dt \;+\; \nu(|\mathbf{r}(t)-\mathbf{x}(t)|)dt \;+\; noise \qquad (5)$$

Here ν is a 2-vector and $|\mathbf{r}(t)-\mathbf{x}(t)|$ is the distance between the elk and the ATV at time t. The term $\nu(|\mathbf{r}(t)-\mathbf{x}(t)|)$ represents the increment to the velocity of the animal at time t.

The time spacings of measurements differ for the elk and the intruders. Specifically when the ATV is present the spacings are 5 min for the elk and 1 sec for the ATV. To begin the model fitting the values of $\mathbf{x}(t)$ are obtained for the available elk times by interpolation as necessary. The ATV samples are close in time so the interpolation is not difficult.

Assuming that μ and ν in (5) are smooth functions, then the model may be fit by the function *gam* of Splus, see Venables and Ripley (2002). Figure 4 graphs $|\hat{\nu}|(d)$, d distance, for the ATV. (The norm $|\nu| \;=\; \sqrt{\nu_1^2+\nu_2^2}$.) For the ATV one sees an increase in the speed, particularly when elk and humans are close to one another. The increased speed of elk is apparent at distances as far as 1.5 km (Figure 4). An upper 95% null level is indicated in Figure 4 by a dashed line. This is based on standard errors output by *gam*. Also graphed are corresponding figures for the bikers, equestrians and hikers. In these cases distance is to the nearest intruder. The biker and hiker cases show evidence of the intruders disturbing the animals. The equestrian case needs further investigation.

In the cases other than the ATV more than one intruder was active complicating the computations.

A reality check. As a check of their reasonableness of the null line one can proceed as follows. One would expect not much influence if the ATV is far away from the elk. Consider then the model

$$d\mathbf{r}(t) \;=\; \mu(\mathbf{r})dt \;+\; \nu(|\mathbf{r}(t)-\mathbf{x}(t-\tau)|)dt \;+\; noise \qquad (6)$$

where τ is a time lag.

Expression (6) allows the change in speed of an elk is affected by the location of the ATV τ time units earlier. The model (6) is fit for τ = 0; 5; 10; 15 min and the results are given in Figure 5. One sees a diminished ATV effect and less precise confidence intervals at increasing large increments of τ.

Estimation of $|\nu(d)|$ was also done in the absence of the μ terms and the results were very similar. This gave some validity to interpreting the estimate $\hat{\nu}(d)$ on its own.

Wisdom et al. (2004) and Preisler et al. (2004) modelled the probability of elk response to ATVs, for data for the year 2002, and obtained similar results to those here although some differences were apparent.

This completes the data analysis part of the paper. Attention now turns to an analytic problem arising because of the fence.

The case of a boundary.

Figure 2 shows that the motion of the animals is much affected by the boundary of the study area, as defined by the ungulate-proof fence. Effective models need to take note of this, i.e. one needs expressions for SDEs with boundaries. Suppose the animal is restricted to a region D with boundary ∂D. An applicable SDE is

$$d\mathbf{r}(t) \ = \ \mu(\mathbf{r}(t),t)dt \ + \ \sigma(\mathbf{r}(t),t)d\mathbf{B}(t) + d\mathbf{A}(t) \qquad (7)$$

where $\mathbf{A}$ is an increasing process which only increases when $\mathbf{r}(t)$ is on the boundary ∂D.

For more detail on this and related approaches see Brillinger (2003) and references therein.

Simulation.

Simulation finds many uses in probabilistic and statistical work with SDEs. Topics that may be mentioned include: solving an SDE, program checking, estimating functionals, likelihood computation, bootstrapping, and model checking. In the case with no boundaries one can use expression (3).

The following material refers to the constrained case. Suppose the process is restricted to the domain D, e.g. $D \ = \ \{\mathbf{r} : \phi(\mathbf{r}) \ > \ 0\}$ with boundary $\partial D \ = \ \{\mathbf{r} : \phi(\mathbf{r}) \ = \ 0\}$. The idea is to approximate the continuous time process by a discrete time Markov chain. First some notation. Set

$$\mathbf{a}(\mathbf{r},t) \ = \ \frac{1}{2}\sigma(\mathbf{r},t)\sigma(\mathbf{r},t)'$$

Next, for convenience in writing the formulas and carrying out the simulations, the case presented is that of $a_{ij}(\mathbf{r},t) \ = \ 0, \ i \ \neq \ j$. (General formulas may be found in Kushner (1976). Suppose that time is discretized with $t_{k+1} - t_k \ = \ \Delta$. Set $\mathbf{r}_k \ = \ \mathbf{r}(t_k)$. Let D_h refer to the lattice points with separation h in D. Suppose $\mathbf{r}_0$ in D_h. Let $\mathbf{e}_i$ denote the unit vector in the i-th coordinate direction.

Consider the Markov chain with transition probabilities

$$P(\mathbf{r}_k \ = \ \mathbf{r}_0 \ \pm \ \mathbf{e}_i h \mid \mathbf{r}_{k-1} = \ \mathbf{r}_0)$$
$$= \ \frac{\Delta}{h^2}(a_{ii}(\mathbf{r}_0,t_{k-1}) \ + \ h|\mu_i(\mathbf{r}_0,t_k - 1)|^{\pm}) \qquad (8)$$

$$P(\mathbf{r}_k \ = \ \mathbf{r}_0 \mid \mathbf{r}_{k-1} = \ \mathbf{r}_0) \ = \ 1 \ - \ \sum \mathit{preceding} \qquad (9)$$

provided all the possible transition points are in D_h. Here it has been supposed the the probabilities are ≥ 0. By choice of Δ and h this may be arranged.

In the above expressions the following notation has been employed.

$$|u|^+ \ = \ u \ if \ u \ > \ 0 \ and \ = 0 \ otherwise$$

and

$$|u|^- \ = \ -u \ if \ u \ < \ 0 \ and \ = 0 \ otherwise$$

Next, consider the boundary case. Suppose there is a reflection direction $\gamma(\mathbf{r})$ for $\mathbf{r}$ in ∂D This relates to the setup (6) as follows. The discrete boundary ∂D_h may be defined as

$$\{\mathbf{r} : \ \mathit{line\ connecting}\ \mathbf{r}\ \mathit{to\ neighboring\ grid\ points\ touching}\ \partial D\}$$

$$d\mathbf{A}(t) \ = \ \gamma(\mathbf{r}(t))I_{\partial D}(\mathbf{r}(t))d\mu(t)$$

with

$$d\mu(t) \ = \ lim\{ \ dt \ I_{\partial D_h}(r_t))/h\}$$

I being the indicator function. Now one completes the description of the simulation process by

$$Prob(\mathbf{r} \ = \ \mathbf{r}_0 \pm \mathbf{e_i} h|\gamma, \ \mathbf{r}_0 \ in \ \partial D_h) \ = \ \gamma_i^{\pm}(\mathbf{r})/|\gamma(\mathbf{r})| \qquad (10)$$

Figure 6 provides the results of a simulation taking $a_{ii}(\mathbf{r}) \ = \ \hat{\mu}_i(\mathbf{r})$ and a starting point near the boundary. One sees the "animal" head to the middle of the pasture and then meander around there. Current research concerns the problem of allowing the animal to move away from the middle as Figure 2 indicated happens in the real situation. The grid character of the motion in the figure results from the use of a mesh with spacing h in the simulation.

Discussion.

An SDE has been employed to model the (possible) effects of human disturbance on an elk in a bounded region.

The statistical tools used in this work include: exploratory data analysis, stochastic differential equations, the generalized additive model and simulation.

One might assess the fit of the model via residuals,

$$\frac{\mathbf{r}_j(t_{i+1}) - \mathbf{r}_j(t_i) - \hat{\mu}_j(\mathbf{r}(t_i))(t_{i+1} - t_i)}{\hat{\sigma}_j(\mathbf{r}(t_i))\sqrt{t_{i+1} - t_i}}$$

as was done in Brillinger et al. (2001a).

The statistical tools illustrated here appear useful in future analyses of animal reactions to recreational uses by humans. The use of SDEs and associated tools, as specified here, provides high utility in describing animal reactions under such experiments.

References.

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

Brillinger, D. R., Preisler, H. K., Ager, A. A. and Kie, J. (2001b),"The use of potential functions in modelling animal movement" *Data Analysis From Statistical Foundations*, Nova, New York.

Brillinger, D. R. (2003) "Simulating constrained animal motion using stochastic differential equations," *Lecture Notes in Statistics* 41, Springer, New York.

Kushner, H. J. (1976), *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*, Academic, New York.

Preisler, H. K., Ager, A. A., and Wisdom, M. J. (2004), "Analyzing responses of elk to human disturbances," Submitted.

Rowland, M. M., Bryant, L.D., Johnson, B. K., Noyes, J. H., Wisdom, M. J., and Thomas, J. W. (1997). *"The Starkey Project: History, Facilities, and Data Collection Methods for Ungulate Research"*. USDA Forest Service General Technical Report PNW-GTR-396.

Venables, W. N. and Ripley, B. D. (2002), *Modern Statistics with S*, Springer.

Wisdom, M. J., Ager, A. A., Preisler, H. K., Cimon, N. J., and Johnson, B. K. (2004). "Effects of off-road recreation on mule deer and elk," *Transactions North American Wildlife and Natural Resource Conference* 69, in press.

Acknowledgements.

This research was supported by DRB's NSF grant DMS-02-03921, and Forest Service contract FS 02-JV-11272165-02.

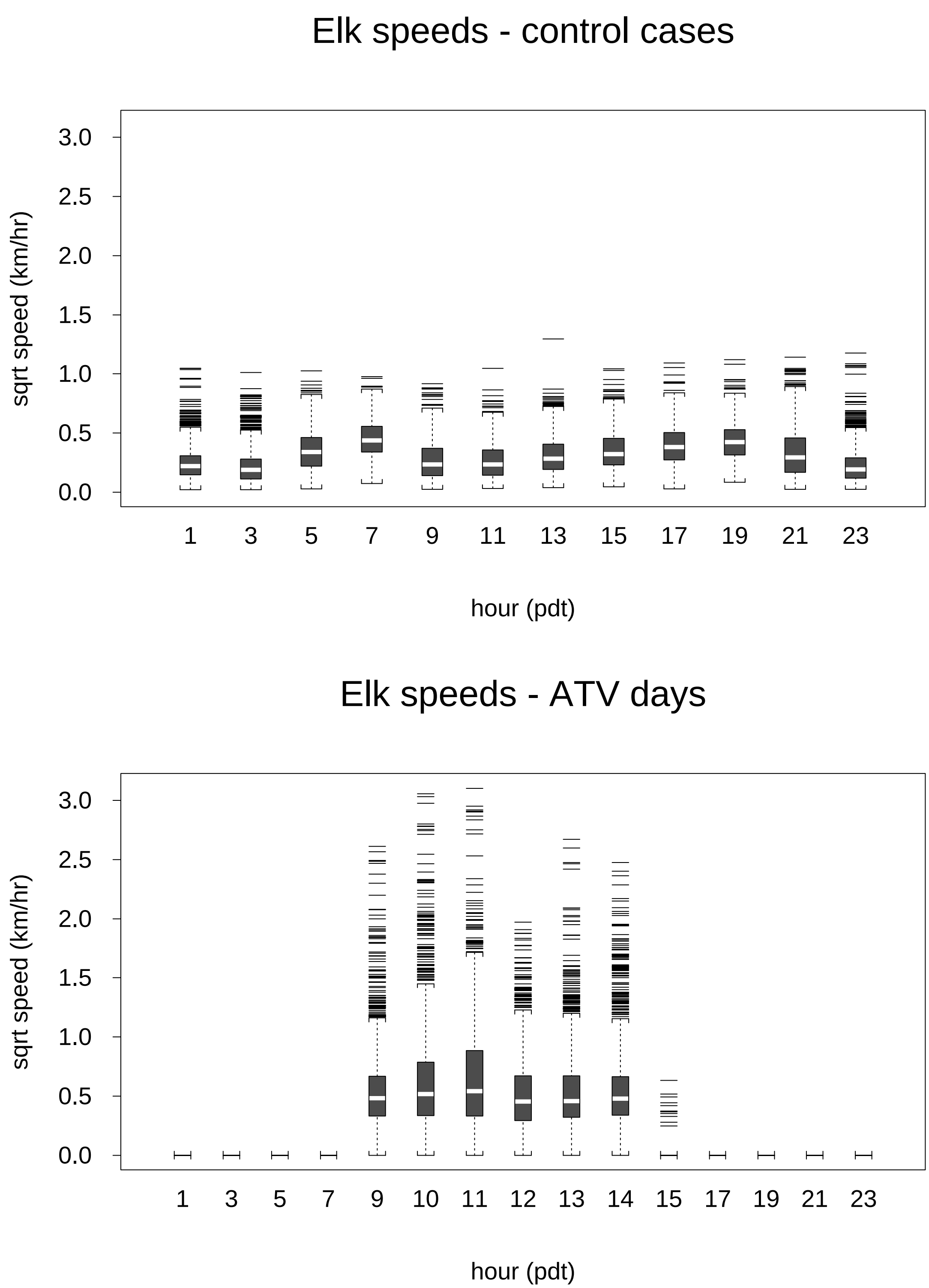


Figure 1: Parallel boxplots of elk speed for control and ATV days respectively.

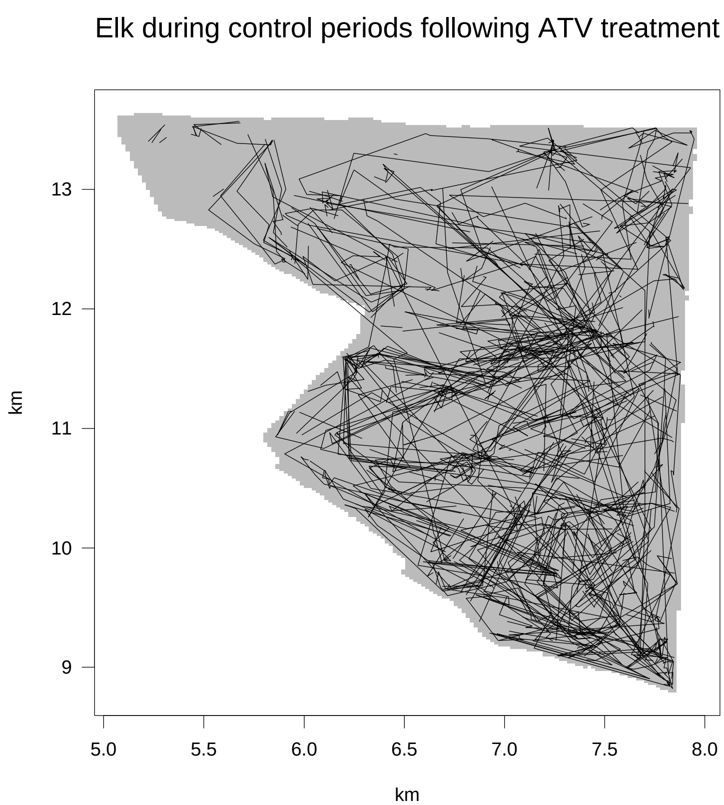
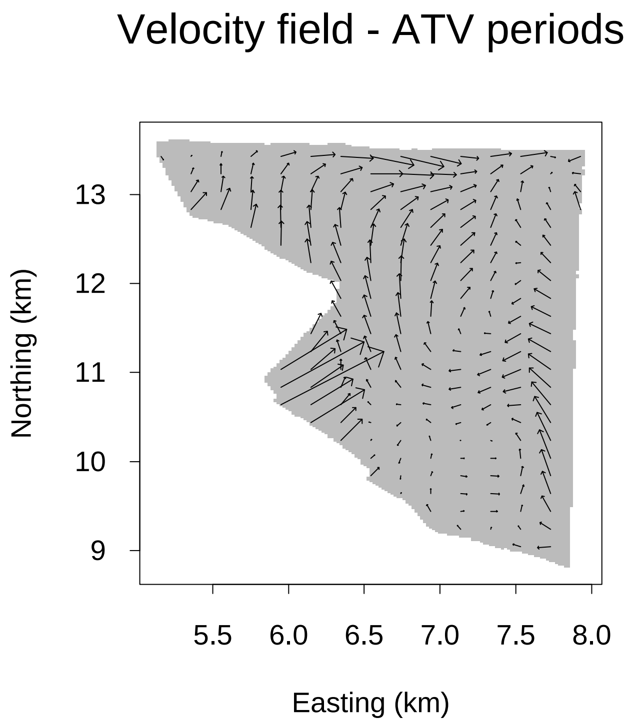


Figure 2: Elk paths during control periods following the ATV treatments. The points joined are 2 hours apart.



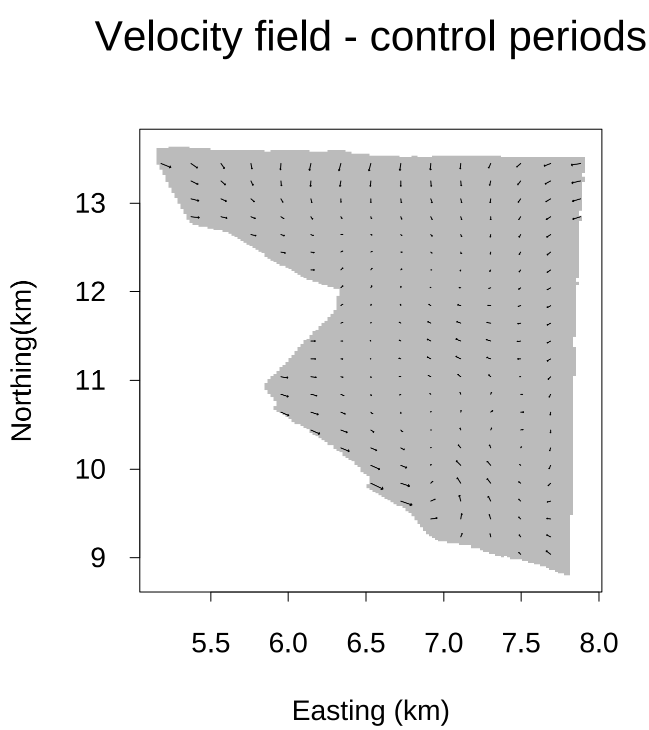


Figure 3: Top:Estimated vector field of the elk movement during ATV periods. Bottom: For the control periods at the beginning and end of the experiment.

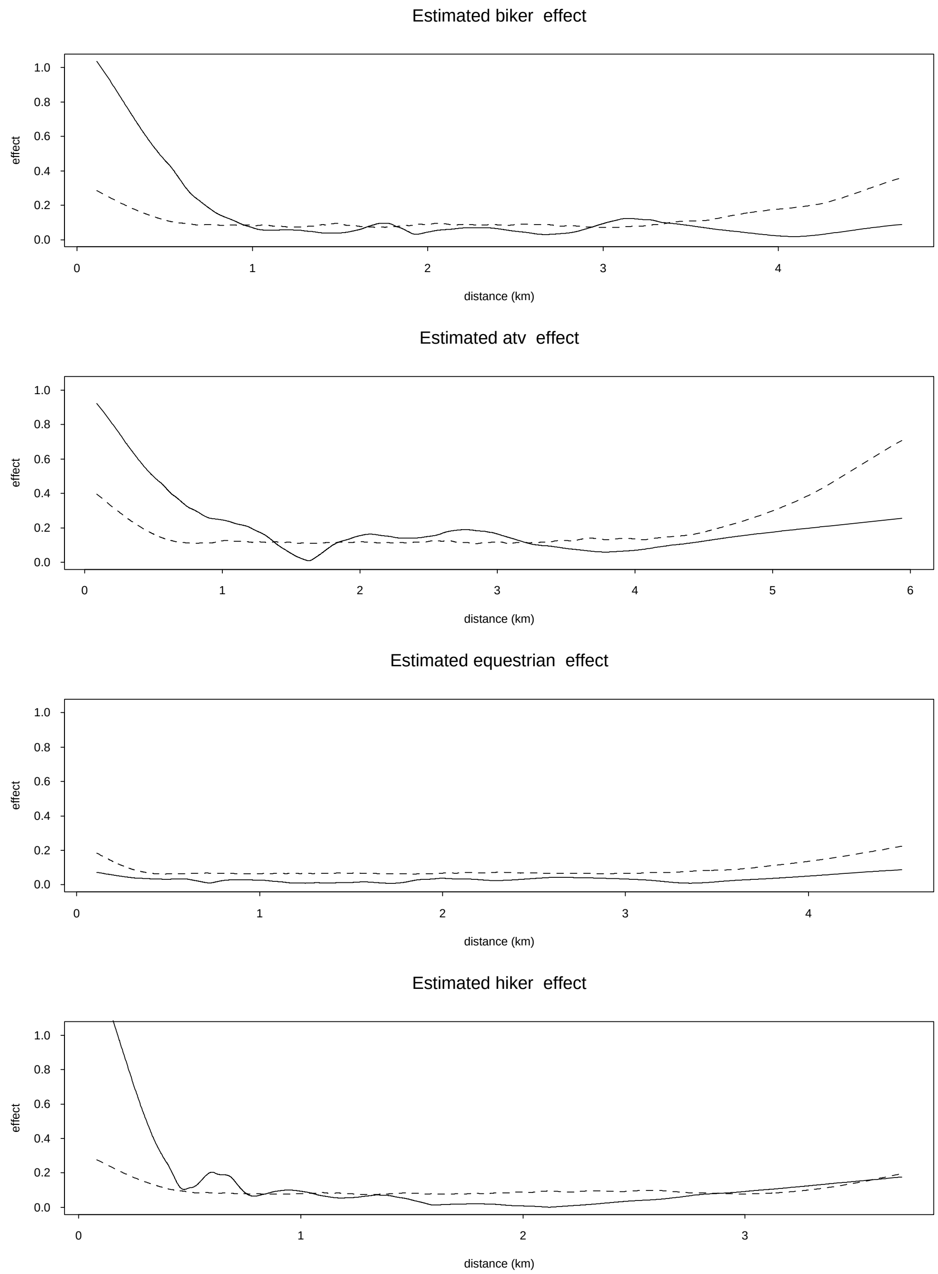


Figure 4: The estimated effect, or $\nu(d)$, per equation (5), which is the incremental increase in the speed of elk at time t, in relation to distance from humans engaged in riding bicycles, an ATV, horses and on foot. The dashed line represents the upper 95% null level, below which the human effect may be deemed not different than control periods.

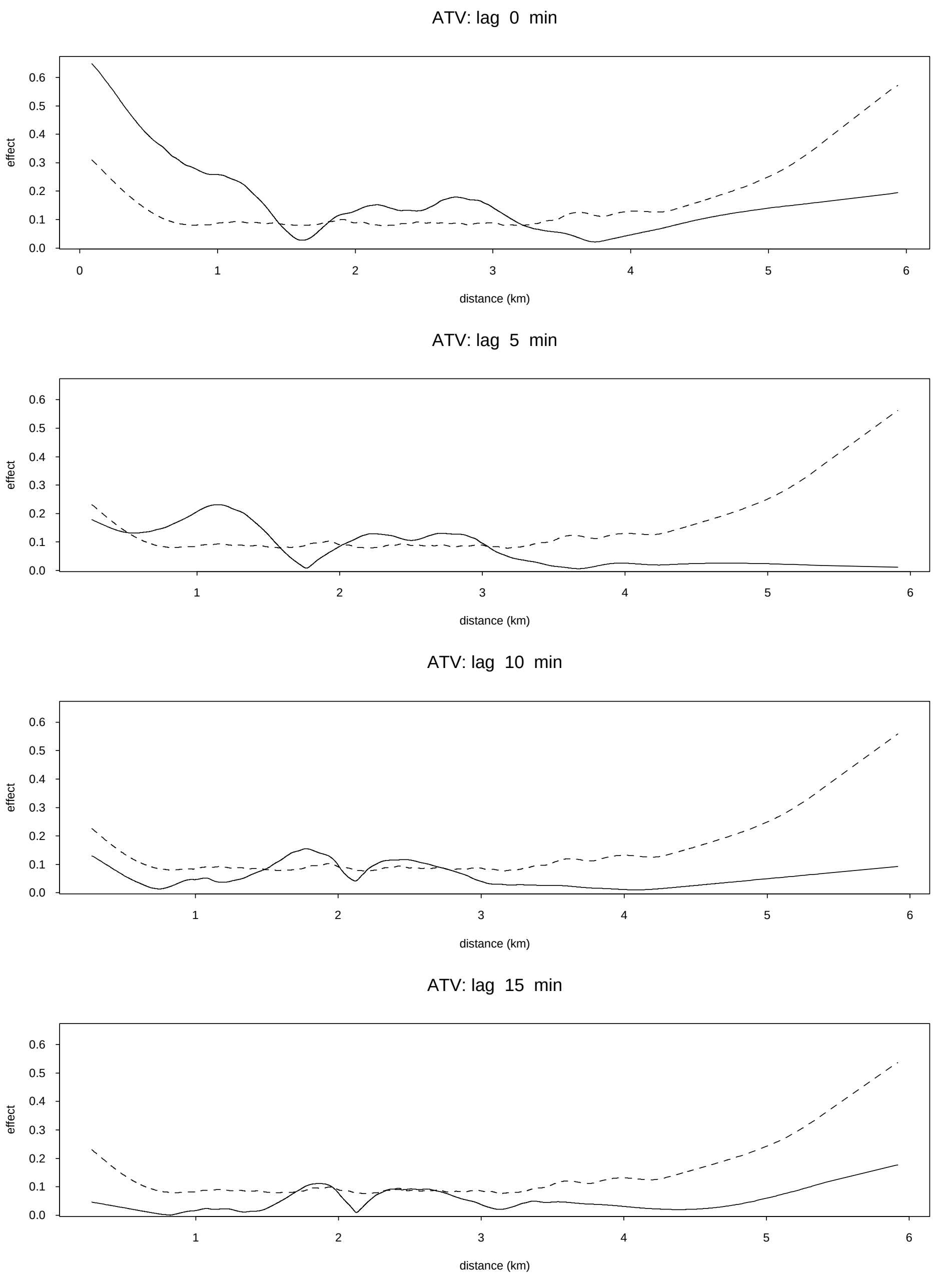


Figure 5: The estimated $|\nu|(d)$ for the model (6) with τ = 0, 5, 10, 15 min.

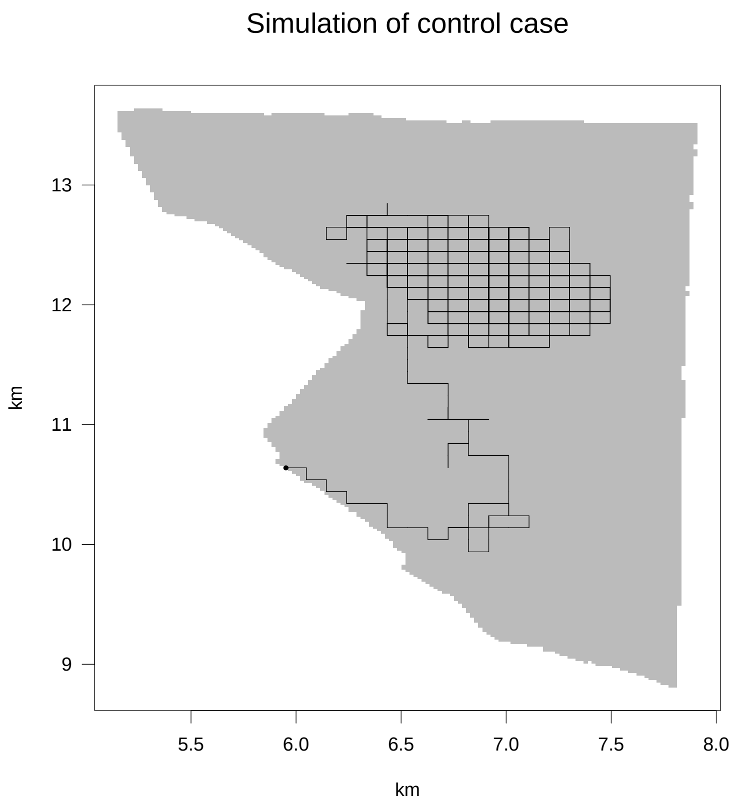


Figure 6: A simulation, by the equations (8-10), of the fitted SDE using the data from the control days.