

Fitting Dynamic Models to Animal Movement Data: The Importance of Probes for Model Selection, a Reply to Franz and Caillaud

Benjamin D. Dalziel,^{1,*} Juan M. Morales,² and John M. Fryxell³

1. Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853; 2. Laboratorio Ecotono, Instituto de Investigaciones en Biodiversidad y Medioambiente–Consejo Nacional de Investigaciones Científicas y Técnicas, Universidad Nacional del Comahue Quintral 1250, (8400) Bariloche, Argentina; 3. Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

Submitted August 26, 2009; Accepted February 1, 2010; Electronically published April 21, 2010

Keywords: animal movement, artificial neural networks, model selection, model fitting, probes.

Analyzing animal movement can provide a useful perspective on the interface between landscape patterns and individual behavior (Patterson et al. 2008; Schick et al. 2008). In Dalziel et al. (2008) we proposed a method of fitting dynamic movement models to trajectory and landscape data. The goal was to explore how disparate landscape-behavior processes combine to generate animal movement patterns. As an example of the approach, we used a genetic algorithm (GA) to fit artificial neural network (ANN) models to observations of elk (*Cervus canadensis*) movement.

We formulated seven ANN models that encompassed a factorial combination of three different types of landscape-behavior interaction: the distance from an elk's current location to each point on the landscape, $d(x)$; the resource structure at that point relative to the rest of the home range, $r(x)$; and the estimated memory of previous visits to that point, $m(x)$, where x represents the coordinates of the landscape. The models used this information to estimate a dynamic redistribution kernel that, at each time point, gave the probability that an elk would move to a given location x^* at the next time step. Thus,

$$P(x^* = x) = k[d(x), r(x), m(x)], \quad (1)$$

with P representing probability and k the redistribution kernel. The product of the probabilities for each point in the observed trajectory gives the likelihood of a kernel

given the data. We fitted the models using a GA that iteratively adjusted the weights of the ANNs to produce kernels that maximized this likelihood. Note that rather than modeling step lengths and turn angles, as is often done in random walk theory (e.g., Turchin 1998), this approach focuses on the occurrence of the animal at a particular location x^* at a particular time as a dynamic function of evolving landscape and behavioral variables. This is similar in spirit to the approach proposed in a recent synthesis by Schick et al. (2008).

We compared the models by contrasting statistical probes of their predictions against probes of new trajectory data for the same elk (e.g., Kendall et al. 1999; Morales et al. 2004, 2005). Probes are particularly useful for identifying patterns in stochastic processes such as animal movement where the same process can generate an array of trajectories whose variability belies their common structure. Similar criteria for model assessment have been proposed for complex simulation models by Grimm et al. (2005) and in the posterior predictive simulations used in Bayesian analysis (Gelman et al. 2004). In Dalziel et al. (2008) we used the distribution of daily displacements, a coefficient of resource selection, and the probability that a trajectory would return to a previously visited location as probes to compare actual elk movements with modeled trajectories.

Franz and Caillaud (2010, in this issue) identify two problems with the approach we described in 2008. First, in addition to using probes, we compared the likelihoods of the fitted models. In retrospect we agree with Franz and Caillaud that this was wrong, since the complex models contained the simpler ones as special cases and thus were guaranteed to have at least as high a likelihood score. We also agree with Franz and Caillaud's second point: in the cases where the likelihood of a simpler model exceeded that of more complex one, the GA failed to find exactly the optimal parameter values for the fitted model.

* Corresponding author; e-mail: ben.dalziel@gmail.com.

From these problems, Franz and Caillaud conclude that the approach we described crucially lacks a method for controlling for model complexity and that the fitting algorithm we used was not powerful enough. They further recommend comparing the likelihoods of the models using an information criterion such as Akaike's Information Criterion (AIC; Akaike 1974) to control for model complexity and the use of a more powerful optimization algorithm.

Although we agree with Franz and Caillaud's identification of these problems, we disagree with their analysis and recommendations in several ways. First, assessing model adequacy using probes was a central component of the approach we proposed. Probes have the advantage of focusing on models' ability to predict emergent patterns, rather than focusing on fit and complexity. Second, comparing ANN models using AIC is more difficult than for other classes of models with which AIC is more commonly used because the number of parameters in an ANN model is not necessarily a good proxy for model complexity (Murata et al. 1994; Anders and Korn 1999). As a result, finding the best ANN model using information criteria may require specialized approaches outside the scope of our research (Moody 1994; Murata et al. 1994).

Franz and Caillaud's second point involves the dangers of using a fitting algorithm not powerful enough to find the global maximum in the likelihood surface for the ANN models, which may have had multiple local maxima. We agree that ANNs with more inputs should have at least as high a likelihood as nested ANNs with fewer inputs, since the higher-order models can adopt the structure of the lower-order ones by setting some connections to 0. However, small anomalous differences in likelihood between nested ANN models do not necessarily indicate that the GA has become trapped in a local maximum. An alternate hypothesis is that mutation and recombination of candidate solutions in the GA may prevent the algorithm from reaching the optimum exactly despite being nearby. In Dalziel et al. (2008), we assessed convergence to the global optimum by showing that the parameters selected by the GA did not depend significantly on initial conditions (unpublished results). Another way to test for the existence of multiple local optima would be to compute the likelihood at a series of points on a line in parameter space between two nested ANN models. If the resulting likelihood profile does not show two distinct peaks, then "noise" generated by mutation and recombination rather than a local maximum is likely responsible for nonconvergence. In this case one could run a deterministic hill-climbing algorithm from the best point found by the GA to attain the optimal solution.

Finally, we wish to emphasize that although the problems with the ANNs and the GA identified by Franz and Caillaud are real, they are not part of the general method

we proposed. Estimating dynamic redistribution kernels based on landscape and behavior data in the manner we described requires only flexible statistical models, a method to fit them, and probes to contrast the results. For example, another powerful approach is the use of hierarchical state-based models fitted in a Bayesian framework (e.g., Morales et al. 2004; Schick et al. 2008). This approach may have some advantages over the use of ANNs, such as the ability to mechanistically model landscape and behavioral processes and to incorporate process and observation error (Schick et al. 2008). Time will tell which approaches prove most fruitful in understanding the general mechanisms underlying animal movement patterns. What the approach we described in 2008 can offer these efforts is an example of fitting dynamic models of landscape-behavior interaction using maximum likelihood and contrasting the performance of the models with new data using probes.

Acknowledgments

We thank S. P. Ellner and two anonymous reviewers for their feedback on earlier versions of this manuscript.

Literature Cited

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716–723.
- Anders, U., and O. Korn. 1999. Model selection in neural networks. *Neural Networks* 12:309–323.
- Dalziel, B. D., J. M. Morales, and J. M. Fryxell. 2008. Fitting probability distributions to animal movement trajectories: using artificial neural networks to link distance, resources, and memory. *American Naturalist* 172:248–258.
- Franz, M., and D. Caillaud. 2010. Fitting artificial neural network models to animal movement data: some necessary precautions. *American Naturalist* 175:759–761.
- Gelman, A., J. B. Carlin, and H. S. Stern. 2004. *Bayesian data analysis*. CRC, Boca Raton, FL.
- Grimm, V., E. Revilla, U. Berger, F. Jeltsch, W. M. Mooij, S. F. Railsback, H. H. Thulke, J. Weiner, T. Wiegand, and D. J. DeAngelis. 2005. Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science* 310:987–991.
- Kendall, B. E., C. J. Briggs, W. M. Murdoch, P. Turchin, S. P. Ellner, E. McCauley, R. M. Nisbet, and S. N. Wood. 1999. Why do populations cycle? a synthesis of statistical and mechanistic modeling approaches. *Ecology* 80:1789–1805.
- Moody, J. 1994. Prediction risk and architecture selection for neural networks. In J. H. Friedman and H. Wechsler, eds. *Statistics to neural networks: theory and patterns recognition applications*. Springer, New York.
- Morales, J. M., D. T. Haydon, J. Frair, K. E. Holsinger, and J. M. Fryxell. 2004. Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* 85:2436–2445.
- Morales, J. M., D. Fortin, J. L. Frair, and E. H. Merrill. 2005. Adaptive models for large herbivore movement in heterogeneous landscapes. *Landscape Ecology* 20:301–316.

Murata, N., S. Yoshizawa, and S. Amari. 1994. Network information criterion—determining the number of hidden units for an artificial neural network. *IEEE Transactions on Neural Networks* 5:865–872.

Patterson, T. A., L. Thomas, C. Wilcox, O. Ovaskainen, and J. Matthiopoulos. 2008. State-space models of individual animal movement. *Trends in Ecology & Evolution* 23:87–94.

Schick, R. S., S. R. Loarie, F. Colchero, B. D. Best, A. Boustany, D. A. Conde, P. N. Halpin, L. N. Joppa, C. M. McClellan, and J. S.

Clark. 2008. Understanding movement data and movement processes: current and emerging directions. *Ecology Letters* 11:1338–1350.

Turchin, P. 1998. *Quantitative analysis of movement: measuring and modeling population redistribution in animals and plants*. Sunderland, MA.

Associate Editor: Benjamin Bolker
Editor: Ruth G. Shaw



Sperm Whale. Left, skull of adult *Physeter macrocephalus* seen from above. Right, skull of adult *P. macrocephalus* seen from the side. From “The Sperm Whales, Giant and Pygmy” by Theodore Gill, M.D., Ph.D. (*American Naturalist*, 1871, 4:725–743).