Ecology, 89(5), 2008, pp. 1208–1215

Reports

CONTINUOUS-TIME CORRELATED RANDOM WALK MODEL FOR ANIMAL TELEMETRY DATA

Devin S. Johnson,[1] Joshua M. London, Mary-Anne Lea,[2] and John W. Durban

National Marine Mammal Laboratory, Alaska Fisheries Science Center, National Marine Fisheries Service, NOAA, 7600 Sand Point Way NE, Seattle, Washington 98115 USA

Abstract. We propose a continuous-time version of the correlated random walk model for animal telemetry data. The continuous-time formulation allows data that have been nonuniformly collected over time to be modeled without subsampling, interpolation, or aggregation to obtain a set of locations uniformly spaced in time. The model is derived from a continuous-time Ornstein-Uhlenbeck velocity process that is integrated to form a location process. The continuous-time model was placed into a state–space framework to allow parameter estimation and location predictions from observed animal locations. Two previously unpublished marine mammal telemetry data sets were analyzed to illustrate use of the model, by-products available from the analysis, and different modifications which are possible. A harbor seal data set was analyzed with a model that incorporates the proportion of each hour spent on land. Also, a northern fur seal pup data set was analyzed with a random drift component to account for directed travel and ocean currents.

Key words: Argos location; Callorhinus ursinus; *correlated random walk; harbor seal; integrated stochastic process; kalman filter; marine mammals; northern fur seal;* Phoca vitulina; *satellite telemetry; state–space model.*

Introduction

Animal telemetry data are obtained by determining the location of an animal at several points in time. There is often great difficulty in locating an animal and, as such, data are often collected opportunistically. Many movement models for analyzing telemetry data are conceptually based on locations that are collected on regular intervals (Brillinger and Stewart 1998, Jonsen et al. 2003, 2005, Forester et al. 2007). Thus, either a priori data processing or model corrections in the form of subsampling, interpolation, or aggregation must be completed to transform locations to a regular interval time scale. Here, we propose a model that can handle data collected at irregular intervals. By considering movement as a stochastic process on a continuous time scale, irregularly spaced data can simply be thought of as a discrete sample of times. Using this approach, statistical inference can be made using the raw data instead of aggregated, thinned, or interpolated data. The continuous-time model can be placed in a discrete (but still nonuniform) time framework to analyze movement data collected in the field. Practical outputs of this analysis are an estimate of the movement path as well as, estimates of movement rate and travel speeds at each point in the path. Standard errors for each of these quantities are also produced. To illustrate the proposed model, we analyzed marine mammal telemetry data sets from two pinniped species of different families: harbor seals (*Phoca vitulina*) and northern fur seals (*Callorhinus ursinus*). The harbor seal analysis illustrates use of a covariate to modify the movement model to account for haul-out behavior. The fur seal analysis illustrates inclusion of a drift component to model the effect of currents, wind, and directed travel on migration through the North Pacific Ocean.

Continuous-time movement models have been used in the past to model movement with the end goal of estimating the home range of an animal. Blackwell (1997), Dunn and Gibson (1977), and Nations and Anderson-Sprecher (2006) each make use of a bivariate Ornstein-Uhlenbeck process to account for autocorrelation when estimating a home range distribution. Billinger and Stewart (1998) use a continuous process defined on a sphere to model northern elephant seal (*Mirounga angustirostris*) migration. In agreement with Jonsen et al. (2005) and Turchin (1998), we consider the correlated random walk (CRW) model as a more natural way to think about animal movement. The CRW process models correlation in the movement rate under the belief that animals have inertia which keeps them moving at a similar rate over successive times. Therefore, we developed a continuous-time version of the CRW model.

Because the CRW model (both discrete and continuous-time versions) is not Markovian, estimation can be challenging. Correlated movements imply that an

Manuscript received 22 June 2007; revised 3 December 2007; accepted 19 December 2007. Corresponding Editor: A. M. de Roos.

[1] E-mail: devin.johnson@noaa.gov

[2] Present address: Antarctic Wildlife Research Unit, School of Zoology, University of Tasmania, P.O. Box 252-05, Hobart TAS 7001 Australia.

animal location at a given time is dependent on all previous locations, not just the last one. A clever formulation of the continuous-time version into a state–space framework allows use of the Kalman filter (KF) (Durbin and Koopman 2001) to estimate parameters via maximum likelihood and predict locations along the movement path that were not observed. Kalman filtering has been used for many years in the field of wildlife telemetry (Anderson-Sprecher and Ledolter 1991, Anderson-Sprecher 1994, Sibert et al. 2003, Royer et al. 2005, Nations and Anderson-Sprecher 2006, Forester et al. 2007). In addition, the state–space framework allows inclusion of measurement error in telemetry locations.

Methods

A continuous-time model

We begin model development by first considering that movement is a change in location. So, let $\mu(t) = [\mu_1(t), \mu_2(t)]'$ be the location of an animal at time t, with subscript 1 referring to a "latitude" coordinate and subscript 2 referring to a "longitude" coordinate. Then, the difference $\mathbf{d}_\Delta(t) = \boldsymbol{\mu}(t+\Delta) - \boldsymbol{\mu}(t)$ describes movement of the animal over Δ time units. For movement on a uniform time scale (i.e., $\Delta = 1$), Jonsen et al. (2005) created a correlated random walk by applying a first-order autoregressive process (Brockwell and Davis 1991) to the $\mathbf{d}_1(t)$ time series.

If $\boldsymbol{\mu}(t)$ is a smooth and continuous path, then as Δ goes to 0, one obtains the differential equation $d\boldsymbol{\mu}(t) = \mathbf{v}(t)dt$, where $\mathbf{v}(t)$ represents the instantaneous rate of location change (velocity). The Ornstien-Uhlenbeck (OU) process is the continuous-time version of the autoregressive process that Jonsen et al. (2005) used to model location differences. Thus, we consider its use for the instantaneous velocity of an animal. For each coordinate axis, $c = 1, 2$, the OU process $\nu_c(t)$ is defined, for each separation in time, Δ, by the following autoregressive equation:

$$\nu_c(t+\Delta) = \gamma_c + e^{-\beta\Delta}[\nu_c(t) - \gamma_c] + \zeta_c(\Delta) \qquad (1)$$

where γ_c is the mean velocity (can be interpreted as "drift"), β is an autocorrelation parameter, and $\zeta_c(\Delta)$ is a zero mean normal random variable with variance $\sigma^2[1 - \exp(-2\beta\Delta)]/2\beta$. The parameter σ controls the overall variability in velocity. Essentially, the equation states that velocity at time $t+\Delta$ is equal to a random variable whose variance grows with Δ plus an adjustment based on how far away the previous velocity value was from the mean.

The bivariate velocity process, $\mathbf{v}(t) = [\nu_1(t), \nu_2(t)]'$, can be cross-correlated between coordinates (i.e., $\text{cov}[\nu_1(t), \nu_2(t)] \neq 0$); however, an elliptical velocity pattern is more realistic (Anderson-Sprecher and Ledolter 1991). Correlated velocities would produce strange directed travel. For example, positive correlation implies movement predominantly and equally in a northeast and southwest direction. It seems unlikely that, at any given time and current location, an animal would randomly switch, with equal probability, between northeast and southwest travel. This situation might occur if the animal is constrained to an oblong-shaped area in which movements are large relative to the size of the area; a fjord perhaps. In the majority of cases, it would be more realistic to model purposeful travel with the γ_c terms. Therefore, for the remainder of this paper, we consider the velocity processes in each coordinate to be independent.

Using the velocity process, the continuous-time location process $\boldsymbol{\mu}(t)$ can be obtained by integration to give

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}(0) + \int_0^t \mathbf{v}(u)du. \qquad (2)$$

Essentially, the location at time t is the "sum" of the steps plus a starting location. Thus, by modeling velocity, we obtain a model for animal location. Eqs. 1 and 2 define the basic continuous-time correlated random walk model (CTCRW). Because the location process is constructed from a process of correlated velocities, the entire track provides information about the next step (i.e., location process is not Markovian), unlike a simple random walk. This is what gives CTCRW directional persistence. As β tends to ∞ and σ/β tends to a constant, the location process becomes a standard Brownian motion (continuous-time version of a random walk). Small β implies more directional persistence than a simple random walk. In fact, at time separation $\Delta = 3/\beta$, $\nu_c(t+\Delta)$ and $\nu_c(t)$ are roughly independent (correlation approximately 0.05). So, the quantity $3/\beta$ can be considered a measurement of directional persistence in units of time. That is to say, whatever effects are causing the animal to travel in the same direction at the same speed will be independent after $3/\beta$ time units. In addition, $\gamma_c = 0$ usually, but could be modeled to account for drift over time (e.g., see *Examples: Marine mammal movement*). For ease of description in the next section, we have assumed $\gamma_c = 0$ (i.e., no drift).

State–space model formulation

In the previous section, we proposed a continuous-time model for animal movement. The continuous path of an animal, however, can only be observed at sampled times. In addition, there are often measurement errors in the observed locations. In order to model animal movement, while simultaneously accounting for measurement error, we put the CTCRW into a state–space model (SSM) framework. This basic SSM can be modified by the researcher to fit the needs of the data. In the section *Examples: Marine mammal movement* we illustrate some more complex modifications.

The general form of a Gaussian linear SSM for a univariate observation is given by two equations, the observation equation and the state equation:

$$y_i = \mathbf{Z}'_i\boldsymbol{\alpha}_i + \varepsilon_i \qquad \boldsymbol{\alpha}_{i+1} = \mathbf{T}_i\boldsymbol{\alpha}_i + \boldsymbol{\eta}_i \qquad (3)$$

Reports

where $\boldsymbol{\alpha}_i$ is the current state vector, y_i is an observation at time i, $\mathbf{T}_i$ and $\mathbf{Z}_i$ are appropriately sized transformation matrices, ε_i is a normal measurement error with variance H_i, and $\boldsymbol{\eta}_i$ are normal error vectors with covariance matrix $\mathbf{Q}_i$. In the case of animal movement data, y_i represents an observed location at time t_i and α_i is the true location and movement process of the animal at time t_i.

Assume that locations $\mathbf{y}_i = [y_{1i}, y_{2i}]$ are measured at times $t_1, \ldots, t_n$, then, substituting the subscript i for the argument t_i and conditioning on the true location of the animal, $\mu_i = [\mu_{1i}, \mu_{2i}]'$ at time t_i, we have the observation equation for the cth coordinate

$$y_{ci} = \mu_{ci} + \varepsilon_{ci}; \qquad \varepsilon_{ci} \sim N(0, H_{ci}) \tag{4}$$

where H_{ci} is the measurement error variance. The error variance could depend on external location quality covariates (e.g., see *Examples: Marine mammal movement*).

Forming the true location equation from the definition of the CTCRW model is not as obvious because $\boldsymbol{\mu}(t)$ is not Markov. We can form a Markovian state $\boldsymbol{\alpha}_i$, however, by bundling the velocity process (which is Markovian) to the location process into a single state vector. The transition equation for the velocity process at time t_{i+1}, $\nu_{c,i+1}$, is already given in Eq. 1. See Appendix A for the mathematical details of the derivation.

The location, $\mu_{c,i+1}$ can be formulated in terms of the location and velocity at time t_i to obtain the following transition equation:

$$\mu_{c,i+1} = \mu_{ci} + \nu_{ci}\left(\frac{1 - e^{-\beta\Delta_i}}{\beta}\right) + \xi_{ci} \tag{5}$$

where $\Delta_i = t_{i+1} - t_i$ and ξ_{ci} are normal errors with variance

$$V[\xi_{ci}] = \frac{\sigma^2}{\beta^2}\left[\Delta_i - \frac{2}{\beta}\left(1 - e^{-\beta\Delta_i}\right) + \frac{1}{2\beta}\left(1 - e^{-2\beta\Delta_i}\right)\right]. \tag{6}$$

The covariance between ξ_{ci} and ζ_{ci} is also necessary for SMM specification and is given by

$$C[\xi_{ci}, \zeta_{ci}] = \frac{\sigma^2}{2\beta^2}\left(1 - 2e^{-\beta\Delta_i} + e^{-2\beta\Delta_i}\right). \tag{7}$$

Finally, using Eqs. 1, 4, and 5, the CTCRW can be placed into the SSM framework for parameter estimation and prediction from observed locations with the specifications y_{ci} = observed location in the c coordinate at time t_i; $\mathbf{Z}_i = [1\ 0]'$, $\boldsymbol{\alpha}_{ci} = [\mu_{ci}\ \nu_{ci}]'$, $\boldsymbol{\eta}_i = [\xi_{ci}\ \zeta_{ci}]'$, $H_{ci} = H(x_i)$, where x_i is a known location quality covariate; and

$$\mathbf{T}_i = \begin{bmatrix} 1 & (1 - e^{-\beta\Delta_i})/\beta \\ 0 & e^{-\beta\Delta_i} \end{bmatrix};$$

$$\mathbf{Q}_{ci} = \begin{Bmatrix} V[\xi_{ci}] & C[\xi_{ci}, \zeta_{ci}] \\ C[\xi_{ci}, \zeta_{ci}] & V[\zeta_{ci}] \end{Bmatrix}.$$

Statistical inference

Here we focus on the maximum likelihood method of parameter estimation and location prediction, but inference could be made with other methods (see *Discussion*). When using SMMs, the Kalman filter (KF) (Durbin and Koopman 2001) is a fast and efficient computing method for finding maximum likelihood estimates of movement parameters $\hat{\theta} = [\hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}_1, \hat{\sigma}_2]'$ (see Appendix B). In addition, optimum predictions and prediction intervals of unobserved locations, $\hat{\boldsymbol{\mu}}(t)$ (as well as velocity $\hat{\mathbf{v}}(t)$), can be obtained as a by-product.

The likelihood for the CTCRW model is a large, multivariate normal density (based on normal measurement error and movement model). The Kalman filter evaluates this complex likelihood by using matrix manipulations to calculate the conditional log-likelihoods, which are summed to produce the joint log-likelihood. The log-likelihood is maximized to obtain parameter estimates that are then used, with a backwards set of recursions, to produce location predictions. The KF recursions need to be started with an initial value and variance matrix for $\boldsymbol{\alpha}_1$, say $\mathbf{a}_{c1}$ and $\mathbf{P}_{c1}$. We suggest $\mathbf{a}_{c1} = [y_{c1}, 0]$ for an initial state value. The initial covariance matrix $\mathbf{P}_{c1}$ may depend on the type of data. If the initial location is unknown, large variances (relative to the location scale) would be advisable (see harbor seal analysis in the next section). If the first location is known (as in the fur seal example) then $\mathbf{P}_1 = 0$ can be used.

Examples: Marine Mammal Movement

The continuous-time movement model is illustrated with analysis of two pinniped data sets from very different taxa, harbor seals and northern fur seals. Both are expected to have different behaviors which illustrate different modifications of the base model presented in the previous sections. There are, however, some common modeling considerations and methods by which data were collected.

For each species location data were recorded by the Argos system (system description *available online*).[3] Along with locations the Argos system rates the quality of observed location with a score from 1 (lowest measurement error; Argos class 3) to 6 (highest measurement error; Argos class B). Here we used quality as a location error covariate $x_i = 1$ to 5. Locations with $x_i = 6$ were removed due to possibility of outliers and high location error variance. A location error variance model was formulated with a known proportionality constant between the best location quality and the other location qualities, $H_{ci} = \tau_c^2 K_c^2(x_i)$, $c = 1, 2$, where the values of $K_c(\cdot)$ are given in Table 1 and τ_c is a parameter which was estimated. To further reduce the possibility of extreme outliers, both data sets were processed with the swim speed filter of McConnell et al. (1992) using a moving average of 1.5 m/s.

[3] ⟨https://www.argos-system.org/manual/⟩

TABLE 1. Values of the known multiplier function $K(\cdot)$ for the location error variance.

Quality	Latitude	Longitude
1	1.00	1.00
2	1.57	1.83
3	3.88	4.71
4	14.17	14.22
5	11.08	5.21

Note: The location quality constant values were obtained from standard deviation ratios presented in Vincent et al. (2002).

In addition to the location quality adjustment, a scale correction was also made in each analysis to adjust for scale differences in latitude and longitude. Namely, for movements near latitude $\dot{y}_1$ there are approximately $1/\cos\{\text{rad}(\dot{y}_1)\}$ degrees longitude per degree latitude, where rad(·) represents conversion to radians. Thus, it follows that if $H_{1i} = \tau_1^2 K_1^2(x_i)$, then $H_{2i} = K_2^2(x_i)/\cos^2\{\text{rad}(\dot{y}_1\}$ implies that τ_1 and τ_2 are approximately equal scale for some appropriately chosen $\dot{y}_1$. Different values of $\dot{y}_1$ can be chosen for different legs of the track to approximate the changes in longitude scale. In addition, there was no reason to believe that movement processes in either coordinate should be different aside from previously mentioned longitude scale changes. Therefore, we set $\beta_2 = \beta_1 = \beta$, $\sigma_1 = \sigma$, and $\sigma_2 = \sigma/\cos\{\text{rad}(\dot{y}_1\}$. All computations were performed with the R statistical package (*available online*).[4]

Harbor seal movement

Between September 2004 and May 2006 satellite-linked time-depth recorders were attached to harbor seals in lower Cook Inlet, Alaska, USA. Harbor seals routinely haul out on land during the course of day-to-day travels. Haul-out data are derived from a conductivity sensor in the tag, and this behavior was summarized into hourly bins indicating the percentage of each hour the tag was dry.

It would be unwise to use the basic CTCRW model as it assumes the animal is in continuous motion for the length of the track. The haul-out behavior of harbor seals needs to be taken into account to provide accurate estimates of location. Other authors have proposed inclusion of characteristics into movement models that account for behavioral shifts in movement. (Blackwell 1997, 2003) proposes a switching model for inclusion of a discrete behavior covariate that controls movement parameters. Other authors (Morales et al. 2004, Jonsen et al. 2005) have proposed the same type of model, but treated the discrete behavior variable as a latent random effect to create a mixture movement model. Here we consider inclusion of a continuously valued covariate which produces a model in which a smooth range of haul-out behavior is allowed to act on the movement of the seal. This avoids subjective loss of covariate information by discretization. However, the covariate must be available, something not all telemetry devices can record.

To develop a haul-out model, first, let $u_1, \ldots, u_m$ be the cut points for which the proportion of time the instrument was dry, D_i, is measured. For the present analysis, u_i and u_{i+1} are an hour apart. The dry time D_i is associated with the interval $(u_i, u_{i+1}]$. If dry time equals 1, the model should, in essence, slow down so that $\hat{\boldsymbol{\mu}}(t)$ is equal to $\hat{\boldsymbol{\mu}}(u_i)$ for any t in the interval $(u_i, u_{i+1}]$. Upon examining specification of the $\nu_c(t)$ process in Eq. 1, one can see that letting β tend to infinity (while σ remains constant) will give the desired result that velocity tends to zero. If all location observations occurred at $u_1, \ldots, u_m$, the CTCRW state-space model could be used by replacing β with $\beta_i = \beta/(1 - D_i)^{\phi}$, where ϕ is positive, in the matrices $\mathbf{T}_i$ and $\mathbf{Q}_i$. This is not, however, the case. It would be impossible to achieve a perfect correspondence.

In order to overcome the fact that locations and dry time values are measured at different times, we used the built-in missing data handling properties of the KF (see Appendix B). Times for which an estimated location is desired can be included in the data set as a missing value and the filter automatically will process the entire augmented data set at the both the prediction and estimation stages. Thus, to implement the haul-out model we augmented the location times $t_1, \ldots, t_n$ with the cut points $u_1, \ldots, u_{m-1}$, here we let $u_m = t_n$. Now, let t_i^*, $i = 1, \ldots, n + m - 1$, be the augmented time set. For the t_i^* where a location is measured, D_i was set to the dry time measured at the closest cut point previous to t_i^* (see Appendix C: Fig. C1 for illustration). For hours where D_i was missing, $D_i = 0$ was used as missingness was likely due to the animal being in the water. This represented a small fraction of the total number of hours. Harbor seals do not range over large latitudinal gradients; therefore, $\dot{y}_{1i}$ was set to the geometric mean latitude.

To obtain parameter MLEs, the CTCRW KF is run over the entire augmented data set. The likelihood is maximized to obtain estimates $\hat{\beta}$, $\hat{\sigma}$, $\hat{\tau}_1$, $\hat{\tau}_2$, and $\hat{\phi}$. Because the animals were allowed to range freely to reduce capture effects, initial locations were unknown. So, we used $\mathbf{a}_{c1} = [y_{c1}, 0]$ and $\mathbf{P}_{c1} = \text{diag}\{[1, 1]\}$. We felt that this was sufficiently large to represent an unknown initial state when recording locations in degrees latitude and longitude.

Fig. 1 illustrates the estimated hourly locations and haul-out locations for a single harbor seal. These locations correspond to many known haul outs from extensive aerial surveys in Cook Inlet. Parameter estimates and standard errors are given in Appendix C: Table C1. Knowing that the animal remains in a fixed location during haul-out allows more rigorous estimation of location error. The estimates and 95% confidence intervals of τ_1 and τ_2 imply estimates of measurement

[4] ⟨http://www.r-project.org/⟩

Reports

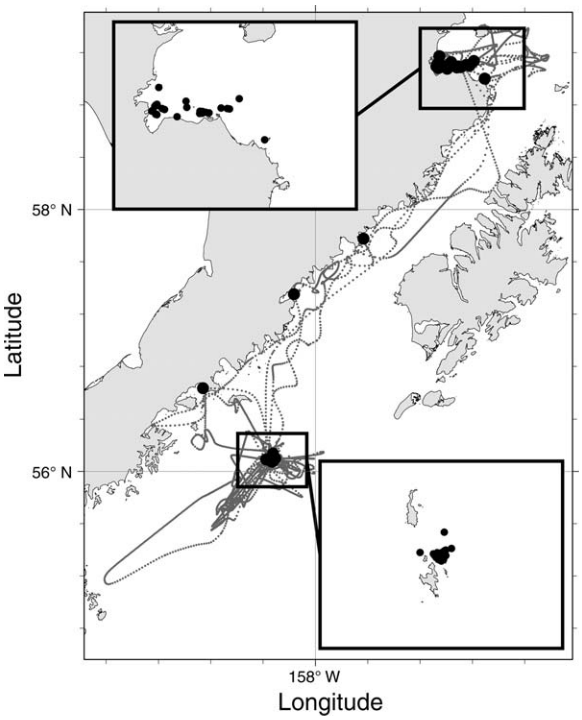

FIG. 1. Haul-out adjusted continuous-time model for harbor seals (*Phoca vitulina*) in Cook Inlet, Alaska, USA. Light gray points represent predicted hourly locations when the animal was swimming for some portion of the hour (i.e., dry time <1). Large black points represent predicted locations when the animal is hauled out (i.e., dry time = 1).

error standard deviations for the highest quality locations of 81.5 m (95% CI: 74.1–89.7 m) in the latitude coordinate and 179.9 m (95% CI: 167.0–193.8 m) in the longitude coordinate. This is consistent with Vincent et al. (2002) who also found larger Argos error in the longitude coordinate. The small value $\hat{\beta} = 0.67$ (95% CI: 0.53–0.85) show that this animal exhibited persistence in direction while swimming; a simple random walk would not have been appropriate. From $\hat{\beta}$ the estimated directional persistence is 4.5 h (95% CI: 3.5–5.6 h) when the seal is swimming, significantly far from zero.

Northern fur seal pup migration

In November 2005, northern fur seal pups (NFS) from St. Paul Island, Alaska were equipped with satellite tags prior to departure from the breeding colonies. Pup departure times from rookeries were calculated as the midpoint between the time of the last land location and the first location at sea.

The NFS location data presents another challenge to the basic CTCRW model such that another modification was necessary. Due to the fact that this is a migration path where pups are traveling long distances to feeding grounds in the North Pacific Ocean, it is conceivable that they may exhibit some directed travel or large-scale ocean current effects. Therefore, we will include a slowly varying drift model for the mean velocity.

A random drift model can be fit by thinking of movement as the sum of two independent zero mean OU velocity models:

$$\nu_c(t) = \gamma_c(t) + \vartheta_c(t) \tag{8}$$

where $\vartheta_c(t)$ has parameters β and σ, as before, and $\gamma_c(t)$ has parameters β/ψ and σ_γ. The parameter $\psi > 1$ is a scale-multiplying factor for the drift process. This forces $\boldsymbol{\gamma}(t)$ to vary more slowly (longer directional persistence) than $\boldsymbol{\vartheta}(t)$. Thus, $\boldsymbol{\gamma}(t)$ represents the effects of slowly changing conditions and $\boldsymbol{\vartheta}(t)$ represents small-scale adjustments in velocity.

The drift model was fit to the NFS data by making the following changes to the SMM matrices:

$$\mathbf{Z}_{ci} = [1\ 0\ 0]' \qquad \alpha_{ci} = [\mu_{ci}\ \nu_{ci}\ \gamma_{ci}]'$$

$$\mathbf{T}_{ci} = \begin{bmatrix} 1 & (1-e^{-\beta\Delta_i})/\beta & \psi(1-e^{-\beta\Delta_i/\psi})/\beta \\ 0 & e^{-\beta\Delta_i} & 0 \\ 0 & 0 & e^{-\beta\Delta_i/\psi} \end{bmatrix}$$

and

$$\mathbf{Q}_{ci} = \left\{ \begin{matrix} V[\xi_{ci}] & C[\xi_{ci}, \zeta_{ci}] & C[\xi_{ci}, \omega_{ci}] \\ C[\xi_{ci}, \zeta_{ci}] & V[\zeta_{ci}] & 0 \\ C[\xi_{ci}, \omega_{ci}] & 0 & V[\omega_{ci}] \end{matrix} \right\}.$$

Now,

$$\begin{aligned} V[\xi_{ci}] = {} & \frac{\sigma^2}{\beta^2}\left[\Delta_i - \frac{2}{\beta}\left(1-e^{-\beta\Delta_i}\right) + \frac{1}{2\beta}\left(1-e^{-2\beta\Delta_i}\right)\right] \\ & + \frac{\psi^2\sigma_\gamma^2}{\beta^2}\left[\Delta_i - \frac{2\psi}{\beta}\left(1-e^{-\beta\Delta_i/\psi}\right)\right. \\ & \left. + \frac{\psi}{2\beta}\left(1-e^{-2\beta\Delta_i/\psi}\right)\right] \end{aligned} \tag{9}$$

and

$$V[\omega_{ci}] = \frac{\psi\sigma_\gamma^2}{2\beta}\left[1-e^{-2\beta\Delta_i/\psi}\right]$$

$$C[\xi_{ci}, \omega_{ci}] = \frac{\psi^2\sigma_\gamma^2}{2\beta^2}\left(1-2e^{-\beta\Delta_i/\psi} + e^{-2\beta\Delta_i/\psi}\right).$$

Appendix A provides mathematical details of the drift process state-space model derivation.

Due to the fact that the pups were instrumented on land and rookery locations are known, the initial state and state variance were set to $\mathbf{P}_{c1} = 0$ and $\mathbf{a}_{c1} = [R_c, 0, 0]$, where R_c is the coordinate c value of the rookery location.

Because a large range of latitudes was traversed during migration (see Fig. 2 for a single pup), the longitude scale correction factor becomes important. For each location, a value of $\dot{y}_1$ is necessary. To provide values, each observed latitude was truncated its whole degree value, $\dot{y}_{1i} = [y_{1i}]$. For each t in the interval $[t_i, t_{i+1})$, where a prediction was desired, $\dot{y}_1$ was set to $\dot{y}_{1i}$. Admittedly, this is somewhat ad hoc, however, various alternatives produced nearly identical results, so, the CTCRW model seems rather robust to this choice.

One quantity of interest for NFS pups is speed. The ground speed of the pups is likely to be an indicator of when the animals are foraging vs. traveling during their trip. The instantaneous speed, $\hat{S}(t)$, of the pup at time t (in km/h) is approximately

$$\hat{S}(t) = 111.325\sqrt{\hat{v}_1(t)^2 + [\hat{v}_2(t)\cos\{\text{rad}[\hat{\mu}_1(t)]\}]^2} \quad (10)$$

where $\hat{v}_c(t) = \hat{\gamma}_c(t) + \hat{\vartheta}_c(t)$. For the small distances traveled (<10 km) by the animal in one hour, this approximation is within 0.02% error of the true rate when calculated using the great circle distance. The standard error of $\hat{S}(t)$ can be estimated via the delta method (Casella and Berger 2002). The CTCRW KF provides all the variances and covariances necessary for the delta method calculation using the generic *delta-method*( ) function in the *msm* library for R (see Appendix B).

Fig. 2 illustrates hourly speed estimates and location predictions for a single pup. Very fast speed is frequently attained by pups through the Aleutian Island passes (see Fig. 2 left inset). The degree to which pups actively swim or ride the swift pass currents is unknown and the subject of further study. The right inset map in Fig. 2 illustrates the slow speeds around possible foraging areas of the instrumented pup.

Fig. 3 illustrates the estimated drift and small-scale velocity trace. Examination of the drift process trace in Fig. 3a shows directed southerly travel for approximately the first two weeks. This is followed by generally eastern progression (Fig. 3b). The small-scale autocorrelation MLE $\hat{\beta} = 0.57$ (95% CI: 0.53–0.85), hence the estimated directional persistence for $\boldsymbol{\vartheta}(t)$ is 5.3 h (95% CI: 3.7–7.6 h). Interestingly, quite similar to the harbor seal in the previous section. The directional persistence ratio $\hat{\psi} = 58.92$ h gives an approximately 13-day persistence effect for the random drift component $\boldsymbol{\gamma}(t)$.

Discussion

In this paper, we presented a continuous-time model for animal movement. The continuous-time formulation allows the data to be used without subsampling or aggregating data to fit into a regularly spaced time scale. By further placing the model into a state-space framework the fast and computationally efficient Kalman filter can be used to estimate locations at a set of desirable times. This represents a significant improvement over discrete time scale models. First, and foremost, because it does not require a researcher to select what the modeled movement time scale will be (e.g., daily or weekly). With movement models such as Jonsen et al. (2005), the assumption that true location at

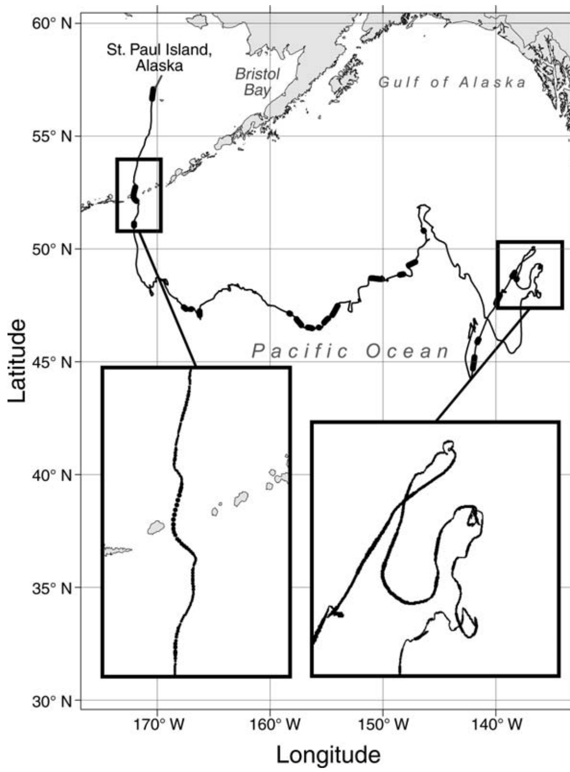

Fig. 2. Estimated hourly locations for a northern fur seal pup (*Callorhinus ursinus*) tracked from St. Paul Island, Alaska, USA, November 2006 to June 2007. In the full-scale track, only two speed classes are plotted, <4 km/h (small dots) and ≥4 km/h (large dots). In the inset maps, speed is discretized to 1 km/h blocks for more detail. The left inset shows predicted hourly movement through Amukta Pass; the right inset shows tortuosity of northern fur seal pups highlighting rapid (travel/current) and slow (foraging) movement regions.

the time of measurement is a linear interpolation of the bracketing modeled locations means that small scale movements are assumed to be linear. If there are multiple observations between two modeled times all deviation from linear travel is assumed to be measurement error. This implies the need for careful consideration of the modeled time scale. Using the CTCRW model, once parameters are estimated, the KF can make predictions at any chosen time scale and the small scale-movements are retained.

The basic CTCRW model can be generalized to include movement covariates and large- and small-scale movement rate modeling. Something difficult to do with a discrete time scale, unless the covariate times are aligned with the location time scale. Both of these modifications were illustrated with example analysis of marine mammal telemetry data. First by including a dry time covariate model to model harbor seal movement in the presence of haul-out behavior. Secondly, a random drift modification was used to model northern fur seal migration in the presence of long-term directed travel and ocean currents. Here, environmental covariates

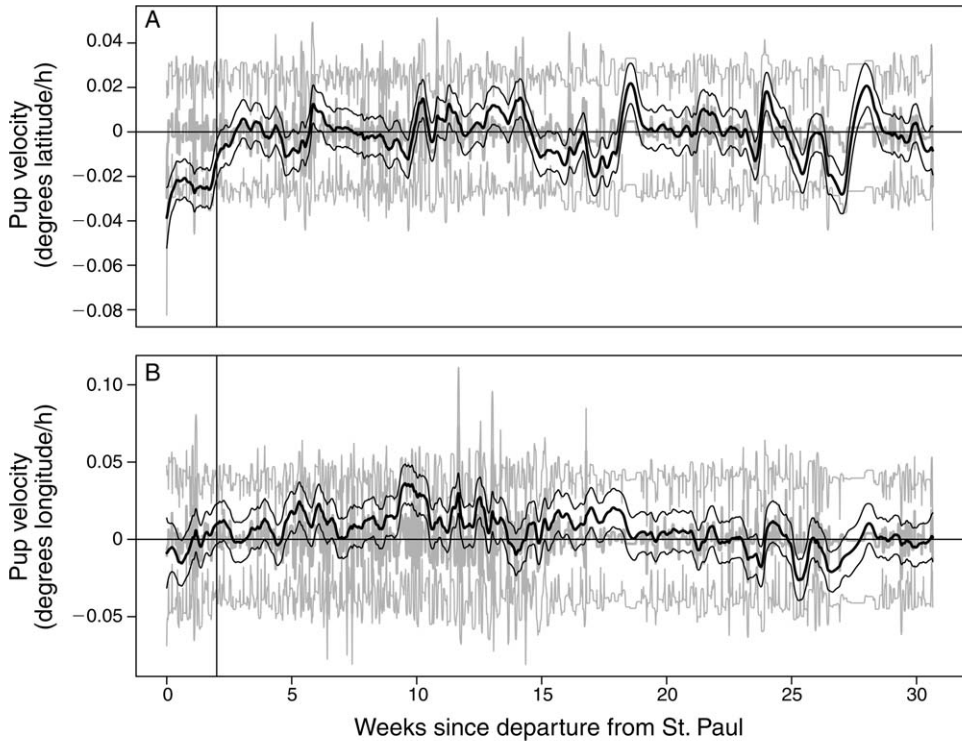

FIG. 3. Estimated hourly velocity processes for a northern fur seal pup tagged on St. Paul Island, Alaska, USA. Heavy black lines represent the drift process $\gamma_c(t)$ and the heavy gray lines represent the small-scale $\vartheta_c(t)$ process ($c = 1$ for latitude and $c = 2$ for longitude). Finer lines represent 95% confidence bands for each process. Negative values for latitude processes represent movement to the south. Negative values for longitude represent travel to the west. A vertical line is drawn at two weeks of migration to illustrate what appears to be a change in overall direction and speed by the animal.

were not included. But it is easy to envision a model, similar to the haul-out model, where the velocity parameters are a function of these variables (Forester et al. 2007).

The main assumption of the CTCRW model and associated state–space modeling methods is one of normality. The measurement error and velocity were both assumed to be normally distributed. For any given data set it is certainly possible to violate this assumption on both counts. Argos locations are known to have large outliers (Vincent et al. 2002, Jonsen et al. 2005). The filtering we performed and use of higher quality locations alleviated this problem to a visually detectable degree.

If preprocessing of the data is not possible there are some slight modifications of the CTCRW model that can be used. If the measurement error assumption is significantly violated the CTCRW model could be placed in the robust framework of Jonsen et al. (2005) by assuming t distributed location errors. Inference could then make via Markov Chain Monte Carlo (MCMC). We believe, however, that it is better to invest research effort into outlier detection in Argos locations rather than use of a robust black box. The robustness of the t errors comes at a cost. If every location is potentially an outlier, any model will tend to produce predicted paths which are more linear as velocity shrinks toward the mean. By detecting outliers, removing them, and using the normally distributed CTCRW model, small-scale movements will be better preserved.

A nonnormal movement process, such as a heavier tailed process, has the potential to be more challenging to handle. A heavy tailed process implies that the animal exhibits "flight" behavior where there are occasionally very large movements mixed with mostly small movements. The mixture model framework of Morales et al. (2004), however, could be employed. A latent variable could indicate "flight" sections where velocity variance would be large. And, conditional on this latent indicator, the movement could be assumed to be normal (albeit, with different variances). The resulting mixture of normal movements will be heavier tailed than a single normal movement model. MCMC would be the most straightforward method of inference for this type of model.

In the presence of possible normality assumption violations the normal CTCRW model can be slightly modified to overcome these problems. Even with the normality assumptions it is a very flexible model for animal movement data.

Acknowledgments

We thank J. Ver Hoef, J. Laake, G. Duker, and two anonymous reviewers for comments which improved the manuscript.

Literature Cited

Anderson-Sprecher, R. 1994. Robust estimates of wildlife location using telemetry data. Biometrics 50:406–416.

Anderson-Sprecher, R., and J. Ledolter. 1991. State-space analysis of wildlife telemetry data. Journal of the American Statistical Association 86:596–602.

Blackwell, P. G. 1997. Random diffusion models for animal movement. Ecological Modelling 100:87–102.

Blackwell, P. G. 2003. Bayesian inference for Markov processes with diffusion and discrete components. Biometrika 90:613–627.

Brillinger, D., and B. Stewart. 1998. Elephant seal movements: modelling migration. Canadian Journal of Statistics 26:431–443.

Brockwell, P. J., and R. Davis. 1991. Time series: theory and methods. Second edition. Springer-Verlag, New York, New York, USA.

Casella, G., and R. L. Berger. 2002. Statistical inference. Second edition. Duxbury, Pacific Grove, California, USA.

Dunn, J., and P. Gibson. 1977. Analysis of radio-telemetry data in studies of home range. Biometrics 33:85–101.

Durbin, J., and S. Koopman. 2001. Time series analysis by state space methods. Oxford University Press, Oxford, UK.

Forester, J. D., A. R. Ives, M. G. Turner, D. P. Anderson, D. Fortin, H. L. Beyer, D. W. Smith, and M. S. Boyce. 2007. State-space models link elk movement patterns to landscape characteristics in Yellowstone National Park. Ecological Monographs 77:285–299.

Jonsen, I., J. Flemming, and R. Myers. 2005. Robust state-space modeling of animal movement data. Ecology 86:2874–2880.

Jonsen, I. D., R. A. Myers, and J. M. Flemming. 2003. Meta-analysis of animal movement using state-space models. Ecology 84:3005–3063.

McConnell, B. J., C. Chambers, and M. A. Fedak. 1992. Foraging ecology of southern elephant seals in relation to the bethymetry and productivity of the southern ocean. Anarctic Science 4:393–398.

Morales, J. M., D. T. Haydon, J. Friar, K. E. Holsinger, and J. M. Fryxell. 2004. Extracting more out of location data: building movement models as mixtures of random walks. Ecology 85:2436–2445.

Nations, C., and R. Anderson-Sprecher. 2006. Estimation of animal location from radio telemetry data with temporal dependencies. Journal of Agriculture, Biological, and Environmental Statistics 11:87–105.

Royer, F., J.-M. Fromentin, and P. Gaspar. 2005. A state-space model to derive bluefin tuna movement and habitat from archival tags. Oikos 109:473–484.

Sibert, J., M. Musyl, and R. Brill. 2003. Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. Ecology 84: 3005–3063.

Turchin, P. 1998. Quantitative analysis of movement. Sinauer Associates, Sunderland, Massachusetts, USA.

Vincent, C., B. J. McConnell, V. Ridoux, and M. A. Fedak. 2002. Assessment of Argos location accuracy from satellite tags deployed on captive grey seals. Marine Mammal Science 18:156–166.

APPENDIX A

Mathematical details of the continuous-time correlated random walk (CTCRW) model (*Ecological Archives* E089-074-A1).

APPENDIX B

Details for application of the Kalman filter to the CTCRW model (*Ecological Archives* E089-074-A2).

APPENDIX C

Additional details for the analysis of marine mammal data (*Ecological Archives* E089-074-A3).