

Ecological Archives E091-020-A1

Toby A. Patterson, Bernie J. McConnell, Mike A. Fedak, Mark V. Bravington, and Mark A. Hindell. 2010. Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error. *Ecology* 91:273–285.

Appendix to: Using GPS data to evaluate the accuracy of state-space methods for correction of Argos satellite telemetry error

Toby A. Patterson, Bernie J. McConnell, Mike A. Fedak, Mark V. Bravington, Mark A. Hindell

Appendix A. A continuous time Kalman Filter/smoothen for Argos locations

During all estimation of positions, locations were converted to eastings and northings using the ‘project’ function from the R-package ‘rgdal’ (see <http://www.gdal.org>). As the seals were tagged on the North Sea Coast of Eastern Scotland, latitude/longitude locations were projected using the Lambert azimuthal equal (LAE) area projection with center position at Longitude 0° E, Latitude 55° N. This avoids the problems associated with working in latitude and longitude coordinates (see Nielsen et al. 2006, Johnson et al. 2008). Because the tags transmit data opportunistically when a seal surfaces, Argos data are irregularly spaced and therefore best handled in a continuous time framework (Johnson et al. 2008). Therefore we must time-index locations by subscripting times t at which observations arrive as t_{τ} where $\tau=1,2,\dots$. So the τ^{th} observation arrived at time t . Differences between observation times are $\Delta t_{\tau} = t_{\tau} - t_{\tau-1}$. The state vector $\mathbf{a}_{\tau} = [x_{\tau}, y_{\tau}]$ is simply the LAE coordinates of location in kilometers from a reference point at 0° East, 55° North.

The state-space form of the movement model is:

$$\mathbf{a}_{\tau} = \mathbf{a}_{\tau-1} + \mathbf{c} \Delta t_{\tau} + \boldsymbol{\varepsilon} \quad (\text{A.1})$$

$$\mathbf{y}_{\tau} = \mathbf{a}_{\tau} + \boldsymbol{\eta} \quad (\text{A.2})$$

Thus the movement process is modeled as a non-isotropic random walk with a drift term $\mathbf{c}_{\tau} = \{u, v\}$ (possibly of magnitude zero) and noise term $\boldsymbol{\varepsilon}$. This model is a simple Normal

Linear Dynamic Model (NLDM) and is therefore amenable to Kalman filtering (Harvey 1992). Non-zero values of u and v imply that the animal has a tendency toward directed movement. Grey seals may be loosely categorized as central place foragers, returning to a set of haulout locations between foraging trips at sea. Thus, over the lifetime of the tags, the net movement is typically zero and directed movement terms are unsuitable. Therefore, we set these to be zero in the estimation process. For generality we have left them in the description of the KF. Here the process noise term has $E(\epsilon)=0$ and $Var(\epsilon) = \mathbf{R} \Delta t$ where

$$\mathbf{R} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}.$$

Similarly, the error covariance has $E(\eta)=0$ and $Var(\eta) = \mathbf{Q}_t$. The \mathbf{Q}_t for a particular location class were calculated from data collected by Vincent et al (2002) - see appendix B. Including GPS locations could be done by specifying an appropriately small \mathbf{Q}_t for these observations.

The KF itself consists of two recursive steps – a prediction step using the process model to predict the next location conditional on the last, and an update step which weights the process prediction by the Argos data and error distributions described by \mathbf{R} . The prediction step for the RW model is given by:

$$\mathbf{a}_{t|t-1} = \mathbf{a}_{t-1} + \mathbf{c}\Delta t_t \quad (\text{A.3})$$

$$\mathbf{P}_{t|t-1} = \mathbf{P}_{t-1} + \mathbf{R}\Delta t_t \quad (\text{A.4})$$

Here \mathbf{a}_t is the estimate of the hidden true state \mathbf{a}_t and $\mathbf{P}_{t|t-1}$ describes the process variance for the t^{th} prediction. Note that we assume, following (Sibert et al. 2003), that the process variance is proportional to the time difference between subsequent locations. The update step equations are given as follows:

$$\mathbf{K}_t = \mathbf{P}_{t|t-1}(\mathbf{P}_{t|t-1} + \mathbf{Q}_t)^{-1} \quad (\text{A.5})$$

$$\mathbf{a}_\tau = \mathbf{a}_{\tau|\tau-1} + \mathbf{K}_\tau(\mathbf{y}_\tau - \mathbf{a}_{\tau|\tau-1}) \quad (\text{A.6})$$

$$\mathbf{P}_\tau = (\mathbf{I} - \mathbf{K}_\tau)\mathbf{P}_{\tau|\tau-1} \quad (\text{A.7})$$

Here \mathbf{I} is a 2x2 identity matrix. \mathbf{K}_τ is known as the Kalman gain matrix and is used to weight the influence of the data \mathbf{y}_τ on the location prediction \mathbf{a}_τ (Wikle and Berliner 2007). This recursive procedure of prediction (equations A.3 and A.4) and update (equations A.5-A.7) is applied over the time series of Argos locations. Note that missing data are easily handled in this framework (Wikle and Berliner 2007). The estimation of a position when no observation has been taken is simply carried out by application of the prediction step only. The update step is missed out until new data arrives. Interpolation between the gaps between data points is carried out at a regular timestep of 3 hours or at the times a GPS location was received.

A.1.1 Smoothing and interpolation

Applying equations (A.3) to (A.7) over times $\tau = 1 \rightarrow \tau = T$, the final time in the series, gives \mathbf{a}_T , the posterior estimate of the final location conditional on all the previous observations $\mathbf{y}_{1:T}$. In other words, all the data has been used to estimate the posterior on that position. For all the previous data points, only the data $\mathbf{y}_{1:\tau}$ has been incorporated in state (i.e. position) estimation. To obtain the posterior estimates of both the state and its uncertainty, a procedure known as Kalman smoothing is applied using the following equations:

$$\mathbf{P}_{\tau|\tau+1}^* = \mathbf{P}_\tau(\mathbf{P}_{\tau-1} + \mathbf{R})^{-1} \quad (\text{A.8})$$

$$\mathbf{a}_{\tau|\tau+1}^* = \mathbf{a}_\tau + \mathbf{P}^*(\mathbf{a}_{\tau+1}^* - \mathbf{a}_\tau - \mathbf{c}_\tau) \quad (\text{A.9})$$

$$\mathbf{P}_{\tau|\tau+1}^{**} = \mathbf{P}_{\tau-1|\tau} + \mathbf{P}_{\tau-1|\tau}^* \mathbf{P}_{\tau-1|\tau} - \mathbf{P}_{\tau-1|\tau} + \mathbf{R}\Delta t_\tau \mathbf{P}_{\tau-1|\tau}^* \quad (\text{A.10})$$

The final estimates of the path are therefore given by $\mathbf{a}_{1:T}^*$ with uncertainty described by the covariance matrix $\mathbf{P}_{\tau|\tau+1}^{**}$.

A.1.2 Likelihood calculations

As we assume a NLDM, the likelihood of a given point is bivariate Gaussian. The unknown model parameters to be estimated are $\hat{\theta} = \{ \sigma_x^2, \sigma_y^2 \}$. Note that if we were to estimate advective drift terms we would include parameters u and v . The negative log-likelihood is calculated iteratively from the filtering equations and is given as

$$\ln L(\hat{\theta} | \mathbf{y}_\tau) = -\ln \sum_{\tau=1}^T N_2(\mathbf{y}_\tau, \mathbf{a}_\tau, \mathbf{P}_\tau + \mathbf{Q}_\tau) \quad (\text{A.11})$$

Here $N_2(\cdot)$ denotes a bivariate Gaussian distribution. Minimisation of equation A.11 provides the MLEs of the parameters. This was carried out in R using the function `nlminb`.

Numerical calculation of the inverse-Hessian matrix (Fornberg and Sloan 1994) with respect to $\hat{\theta}$ allowed for calculation of approximate standard errors on the parameters. In the results that follow, errors in $\hat{\theta}$ are not considered as they are small compared to errors in the locations. Therefore, we only use point estimates of the parameters to calculate a path.

Appendix References:

- Fornberg, B. and D. M. Sloan. 1994. A review of pseudospectral methods for solving partial differential equations. *Acta Numerica* 3:203–267.
- Harvey, A. C. 1992. Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge.
- Johnson, D. S., J. M. London, M. A. Lea, and J. W. Durban. 2008. Continuous-time correlated random walk model for animal telemetry data. *Ecology* 89:1208-1215.
- Nielsen, A., K. A. Bigelow, M. K. Musyl, and J. R. Sibert. 2006. Improving light-based geolocation by including sea surface temperature. *Fisheries Oceanography* 15:314-325.
- Sibert, J. R., M. K. Musyl, and R. W. Brill. 2003. Horizontal movements of bigeye tuna (*Thunnus obesus*) near Hawaii determined by Kalman filter analysis of archival tagging data. *Fisheries Oceanography* 12:141-151.
- Wikle, C. K. and L. M. Berliner. 2007. A Bayesian tutorial for data assimilation. *Physica D-Nonlinear Phenomena* 230:1-16.