Environ Ecol Stat (2009) 16:531–546
DOI 10.1007/s10651-008-0095-7

Inferring linear feature use in the presence of GPS measurement error

Hannah W. McKenzie · Christopher L. Jerde · Darcy R. Visscher · Evelyn H. Merrill · Mark A. Lewis

Received: 13 April 2007 / Revised: 27 February 2008 / Published online: 8 April 2008

Abstract Global Positioning System (GPS) collars are increasingly used to study animal movement and habitat use. Measurement error is defined as the difference between the observed and true value being measured. In GPS data measurement error is referred to as location error and leads to misclassification of observed locations into habitat types. This is particularily true when studying habitats of small spatial extent with large amounts of edge, such as linear features (e.g. roads and seismic lines). However, no consistent framework exists to address the effect of measurement error on habitat classification of observed locations and resulting biological inference. We developed a mechanistic, empirically-based method for buffering linear features that minimizes the underestimation of animal use introduced by GPS measurement error. To do this we quantified the distribution of measurement error and derived an explicit formula for buffer radius which incorporated the error distribution, the width of the linear feature, and a predefined amount of acceptable type I error in location classification. In our empirical study we found the GPS measurement error of the Lotek GPS_3300 collar followed a bivariate Laplace distribution with parameter $\rho = 0.1123$. When we applied our method to a simulated landscape, type I error was reduced by 57%. This study highlights the need to address the effect of GPS measurement error in animal location classification, particularily for habitats of small spatial extent.

H. W. McKenzie (✉) · C. L. Jerde · M. A. Lewis
Centre for Mathematical Biology, Department of Mathematical and Statistical Sciences, University of Alberta, 632 Central Academic Building, Edmonton, AB, Canada T6G 2G1
e-mail: hmckenzie@math.ualberta.ca

D. R. Visscher · E. H. Merrill
Department of Biological Sciences, University of Alberta, CW405, Biological Sciences Building, Edmonton, AB, Canada T6G 2E9

Keywords Bivariate Laplace distribution · Error distribution · Global Positioning System · Habitats of small spatial extent · Location classification · Location error · Rare habitats · Seismic lines

1 Introduction

Global Positioning System (GPS) collars are frequently used by ecologists to collect location data for animals moving across a landscape. The data are used to classify observed animal locations into habitats or to recreate movement paths for the purposes of testing hypotheses about habitat use, movement, and behaviour. However, GPS data are subject to error, including biased fix rates (Frair et al. 2004; Cain et al. 2005; D'Eon and Delparte 2005) and measurement error (D'Eon and Delparte 2005). We focus on measurement error, which is defined as the difference between the observed and true value being measured. Measurement error in location data is often referred to as location error. To ensure correct biological inference from GPS data, it is necessary to evaluate and consider measurement error during analysis. For example, if measurement error is ignored, habitat selection patterns may be misinterpreted (Rettie and McLoughlin 1999; Frair et al. 2004; Visscher 2006), movement distributions miscalculated (Jerde and Visscher 2005), or behaviors misunderstood (Hurford 2005).

Measurement error creates a particularly difficult problem for detecting animal use of habitats of small spatial extent because the area of the habitat is often less than the measurement error (McLoughlin et al. 2002). Consequently, there is increased probability an observed location will be classified outside the habitat when the true location is inside the habitat (type I error), resulting in a bias towards underestimation of habitat use (Rettie and McLoughlin 1999; McLoughlin et al. 2002). Anthropogenic linear features such as roads, seismic lines, and pipelines, are one example of habitats of small spatial extent that are ubiquitous in many North American landscapes (Timoney and Lee 2001). Linear features are known to alter animal distribution, movement, and behaviour (Thurber et al. 1994; James 1999; Dyer et al. 2001, 2002; Whittington et al. 2005). From a management perspective, unbiased detection of animal use of linear features is a crucial first step towards increased understanding of when, where, and how animals use linear features.

One approach for addressing measurement error, commonly used for radio-telemetry locations, is to buffer observed locations by replacing the point location with an area of fixed radius (Samuel and Kenow 1992; Nams and Boutin 1991; but see Saltz 1994). The buffer accounts for the imprecision in the observed location by assuming the animal is located within an area rather than at an exact point location. However, there is no consistent method for choosing the buffer radius, leading to widely varying buffer sizes (e.g., Dickson and Beier 2002; McLoughlin et al. 2002; Dickson et al. 2005). It is not clear how sensitive the biological conclusions in the above studies were to the choice of buffer radius, or whether it is reasonable to compare results from studies where different buffer radii were used because the type I error rate is not available. In addition, the implicit assumption that measurement error follows a uniform distribution results in valuable information provided by the error distribution being discarded. For example, Visscher (2006) showed that differences between true and measured selection coefficients were larger when GPS measurement error was assumed to be uniformly distributed than when it followed either a normal or Laplace distribution. Furthermore, Rettie and McLoughlin (1999) argue that an appropriate buffer should depend on landscape structure (i.e., habitat patch size) as well as distribution of GPS measurement error. This motivated the approach we chose, which was to buffer the habitat using information from the distribution of GPS measurement error.

Here we develop a mechanistic, empirically-based method of buffering linear features addressing the underestimation bias caused by GPS measurement error. We illustrate how to select an appropriate buffer radius that accounts for both the measurement error distribution and the width of the linear feature such that bias introduced by measurement error is minimized. We also show how to test for the robustness of the buffer against observed location misclassification. Using simulated data we demonstrate the effectiveness of the method for reducing the type I error and illustrate how considering measurement error changes our inference of linear feature use. While we focus primarily on the example of linear features, the broad applicability of the method to other types of habitats of small extent is discussed.

2 Statistical model

2.1 Quantifying the error distribution

The distribution of GPS measurement error describes the probability of observing a location $\hat{\mathbf{x}} = (\hat{x}, \hat{y})$ at a given distance from the true location $\mathbf{x} = (x, y)$. The measurement error $\|\hat{\mathbf{x}} - \mathbf{x}\|$, which follows a distribution, reflects the precision of locations obtained by the GPS collar. For example, a leptokurtic measurement error distribution, such as the Laplace, has a larger number of short and long measurements than an equivalent normal distribution with the same variance (Kot et al. 1996). We propose bivariate normal and bivariate Laplace distributions as potential models for the distribution of GPS measurement error (see Kotz et al. 2001 for a description of the properties of the bivariate Laplace distribution),

$$f_\sigma(\hat{\mathbf{x}} \mid \mathbf{x}) = \frac{1}{2\pi\sigma^2} e^{-\frac{1}{2\sigma^2}\left((\hat{x}-x)^2-(\hat{y}-y)^2\right)}, \tag{1a}$$

$$f_\rho(\hat{\mathbf{x}} \mid \mathbf{x}) = \frac{\rho^2}{2\pi} K_0\left(\rho\sqrt{(\hat{x}-x)^2-(\hat{y}-y)^2}\right). \tag{1b}$$

Here σ and ρ are parameters and K_0 is the modified Bessel function of the second kind (Appendix A). To visualize the two-dimensional distributions in one dimension we transform to polar coordinates and find the marginal distribution of the radius r by integrating with respect to θ from 0 to 2π. Following this transformation, the distributions of the radii are given by

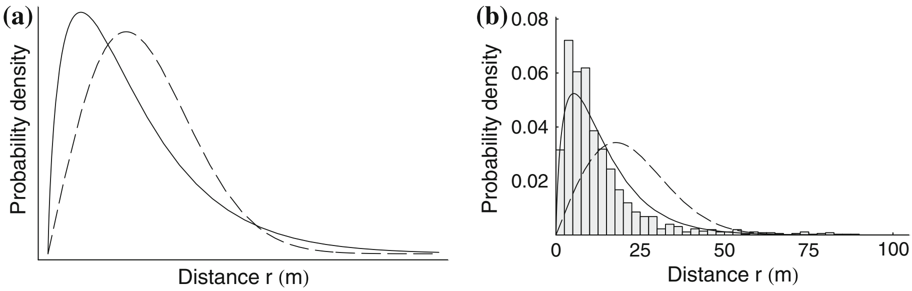

Fig. 1 (**a**) Candidate models for the distribution of GPS measurement error. The marginal distribution of the radii originating from bivariate normal (dashed) and bivariate Laplace (solid) distributions shown with comparable variances. (**b**) Histogram of distances between observed and true locations of a GPS collar in closed conifer forest (Alberta, Canada), including maximum likelihood fits of the candidate models to the data

$$f_\sigma(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}}, \tag{2a}$$

$$f_\rho(r) = \rho^2 r K_0(\rho r). \tag{2b}$$

These models were chosen because they are flexible and can accommodate a range of shapes for the error distribution from mesokurtic to leptokurtic (Fig. 1a). We considered only radially symmetric models because GPS measurement error is not consistently directionally biased (Moen et al. 1996). Model selection techniques are used to determine which of the two models is the best representation of the observed GPS measurement error.

2.2 Computing the buffer radius

The best model of the error distribution is used to compute a buffer radius for the linear features, which reduces the observed location misclassification introduced by the measurement error. We derive a method for choosing the appropriate buffer radius in a hypothesis testing framework. The radius is chosen by finding the rejection region of the test of H_0: the true location is somewhere on the linear feature against H_1: the true location is not on the linear feature, where the test statistic is the observed location $\hat{\mathbf{x}}$. If we consider x to be the distance along a perpendicular line to the original linear feature and y to be the distance along the linear feature, the long, straight nature of linear features allows us to reduce to the problem to one dimension by considering the marginal distribution of $\hat{x}$. The distribution of $\hat{x}$ under the null hypothesis (i.e. when the true location $\mathbf{x}$ is on the linear feature) is given by the general distribution $f_\Theta(\hat{x})$, where Θ is a generalized parameter (Appendix B). The amount of acceptable type I error (α) is specified a priori and corresponds to the proportion of location estimates classified as off the linear feature when the true location is actually on the linear feature. The choice of α fixes the position of the rejection region, and thus the half-width of the buffered linear feature (Fig. 2). The rejection region is found by

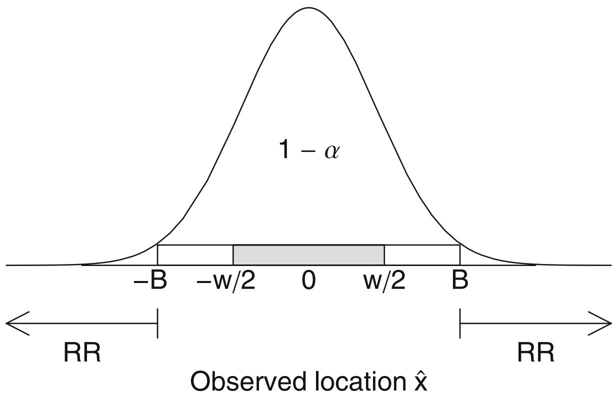

Fig. 2 Rejection regions (RR) for testing the null hypothesis that the true location is on the linear feature against the alternate hypothesis that the true location is not on the linear feature. The test statistic is the observed location $\hat{\mathbf{x}}$, which follows the distribution $f_\Theta(\hat{x}|x)$, where x is on the linear feature (solid line) and Θ is a generalized parameter. The linear feature has width 6.2 m and is shown in grey, while the buffer is shown in white

solving

$$\int_0^B f_\Theta(\hat{x})\, d\hat{x} = \frac{1-\alpha}{2} \tag{3}$$

for the quantity B, which is the half-width of the buffered linear feature.

2.3 Assessing robustness to type I and type II error

The robustness of the observed location classification is evaluated using the power function $\beta(\mathbf{x})$ (Appendix C). The power function represents the probability that the null hypothesis will be rejected for any given true location $\mathbf{x}$, and is particularly useful because it graphically represents both type I (α) and type II (β) errors simultaneously. The type I error given by $\beta(\mathbf{x})$ is different from the type I error specified by α in Eq. 3 because the hypotheses under consideration are slightly different. Previously we knew only that the true location was on the linear feature, whereas for $\beta(\mathbf{x})$ we know the true location. The types of error can be calculated directly from the graph of the power function. For a location $\mathbf{x}$ on the linear feature, the type I error is $\beta(\mathbf{x})$. If $\mathbf{x}$ is off the linear feature the type II error is $1 - \beta(\mathbf{x})$. Ideally the power function would be 0 for any $\mathbf{x}$ that is on the linear feature and 1 for any $\mathbf{x}$ that is off the linear feature. In other words, for a true location on the linear feature we should never observe a location estimate off the linear feature, and vice versa. The presence of measurement error prevents achievement of this ideal, but an appropriate buffer will have a power function near 0 for all $\mathbf{x}$ on the linear feature and close to 1 for all $\mathbf{x}$ off the linear feature (Casella and Berger 2002).

3 Methods

3.1 Data

Two sets of GPS data were collected in the central east slopes of the Rocky Mountains, Alberta, Canada (52°27′ N, 115°45′ W) using a Lotek GPS_3300 collar (Lotek Wireless, Ontario, Canada). We focused solely on the Lotek GPS_3300 collar. The variability between collar brands, as demonstrated by Hebblewhite et al. (2007), is not addressed here. To select between candidate models for the distribution of GPS measurement error, location data were collected from a stationary GPS collar placed in closed conifer forest 1 m off the ground recording location estimates in UTM coordinates at 5-min intervals over 4 days in February 2005 ($n = 1{,}422$). The law of large numbers states that if observations are independent, then the sample mean is a consistent estimator for the population mean in the absence of bias (Casella and Berger 2002). We assume there is no directional bias in location estimates (Moen et al. 1996) and this assumption is supported by visual inspection of the data (Fig. 3), so given the large sample size we use the mean to estimate the true location.

To assess the performance of the best model for the GPS error distribution, data were collected from a collar placed at nine consecutive locations along a transect in closed conifer forest perpendicular to a 6.2 m wide linear feature. Location estimates were recorded at 5-min intervals for 24 h at the center and edges, as well as 25, 50, and 75 m on either side of the linear feature. These distances were chosen from estimated standard deviations of GPS collars (D'Eon and Delparte 2005) so as to vary the amount of overlap between the error distribution and the linear feature.

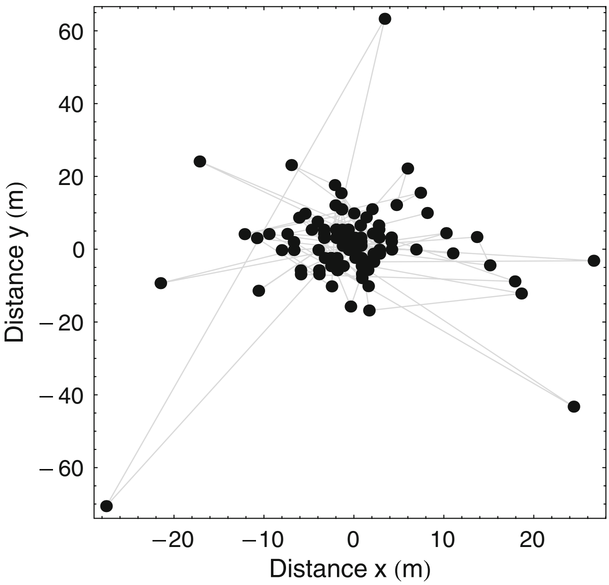

Fig. 3 Distribution of the gps measurement error for the Lotek_3500 collar in closed conifer forest. Points shown are an 8-h subset of the observed locations, with consecutive locations connected by the grey lines

3.2 Model selection and validation

To determine the best model for the error distribution of the Lotek GPS_3300 collar, the two candidate models for measurement error distribution were fit to the observed location data using maximum likelihood. Maximum likelihood estimate for the parameter of the bivariate normal distribution is

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n}((\hat{x}_i - x)^2 + (\hat{y}_i - y)^2)}{2n}}. \tag{4}$$

The maximum likelihood estimate for the bivariate Laplace distribution parameter was found numerically by maximizing the log-likelihood function,

$$LL(\rho \mid \hat{\mathbf{x}}) = \sum_{i=1}^{n} \log\left(\frac{\rho^2}{2\pi} K_0\left(\rho\sqrt{(\hat{x}_i - x)^2 + (\hat{y}_i - y)^2}\right)\right), \tag{5}$$

using the BFGS quasi-Newton (Mathematica 5.1, Wolfram Research, Inc.). In all likelihood estimates, $(\hat{x}_i, \hat{y}_i)$ is the ith observed location, (x, y) is the true location, and n is the sample size. Confidence intervals for the parameter estimates were constructed using the parametric bootstrap (Efron and Tibshirani 1993). The best model was selected using Akaike Information Criterion (AIC) (Burham and Anderson 1998).

To validate the model, for each transect location we found the proportion of observed locations on the linear feature and compared this with the proportion predicted by the model. The relevant measure for calculating the proportion of observed locations predicted by the model to be on the linear feature, $\hat{p}$, is the marginal distribution $f_\Theta(\hat{x}|x)$ of the model $f_\Theta(\hat{\mathbf{x}}|\mathbf{x})$. The marginal distribution describes the univariate distribution of $\hat{x}$ for all values of $\hat{y}$. Thus, the proportion of observed locations predicted by the model to be on the linear feature, for each transect location, is the integral of $f_\Theta(\hat{x}|x)$ from $-w/2$ to $w/2$, where w is the width of the linear feature and x is the transect location. We compared the model predicted proportions to CIs created using a non-parametric bootstrap of the observed data (Efron and Tibshirani 1993). It was possible to do this only for the three central transect locations because all others had fewer than 25 locations observed on the linear feature (Efron and Tibshirani 1993). We used Bonferroni adjusted 90% confidence intervals to protect experiment-wide error.

4 Results

4.1 The distribution of measurement error

All collars had fix rates (i.e. proportion of total possible GPS locations successfully obtained) of greater than 97%. The observed distribution of GPS error was unimodal with several long distance outliers and a mean of 14 m (Fig. 1b). Using AIC, the bivariate Laplace model was the best representation of the empirical distribution of GPS measurement error (Table 1). The normal model was not supported by the data,

Table 1 Results of maximum likelihood parameter estimation and model selection for the three candidate models (Eq. 1a and b). Confidence intervals (95%) for the parameters are shown in brackets

Model	Parameter estimates (95% C.I.)	ΔAIC
Bivariate Laplace	$\hat{\rho} \rightarrow 0.1123$ (0.1072, 0.1175)	0
Bivariate normal	$\hat{\sigma} \rightarrow 17.7213$ (17.2694, 18.1979)	2617

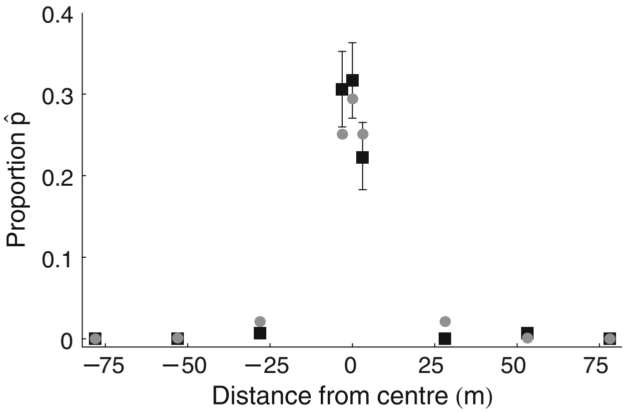

Fig. 4 Validation of the bivariate Laplace model for the distribution of GPS measurement error using a transect across a 6.2 m wide linear feature. Circles represent the probability of the observed location being on the linear feature predicted by the model. Boxes represent the probability of the observed location being on the linear feature calculated from the data, shown with 90% Bonferroni adjusted C.I. where possible. The three central points correspond to true locations on the linear feature

with ΔAIC of 2617. The marginal distribution of the bivariate Laplace model is a symmetric Laplace distribution (Kotz et al. 2001). Therefore, the predicted proportion of locations on the linear feature given the transect location x, is

$$\hat{p} = \int_{-w/2}^{w/2} \frac{\rho}{2} e^{-\rho|\hat{x}-x|}\, d\hat{x}. \tag{6}$$

The closer the true location was to the center of the linear feature, the greater the proportion of location estimates observed and predicted to be on the linear feature (Fig. 4). In all cases the predicted proportion was either within or near ($<$1%) the Bonferroni adjusted 90% confidence intervals, indicating that the bivariate Laplace model is a good representation of the observed GPS error distribution.

4.2 Buffer selection and assessment

Based on the bivariate Laplace model of error distribution, we derived the formula for the half-width of the buffered linear feature using Eq. 3 and replacing the general distribution $f_\Theta(\hat{x})$ with $f_\rho(\hat{x})$ (see Appendix B). For the case where $\int_0^{w/2} f_\rho(\hat{x})\, d\hat{x} < (1-\alpha)/2$, meaning less than $(1-\alpha)/2$ of the density of the error distribution occurs between the center and edge of the linear feature, the half-width of the buffered linear feature is given by

$$B = \frac{1}{\rho} \log \left(\frac{2 \sinh(wp/2)}{\alpha w \rho} \right). \tag{7}$$

Recall ρ is the parameter of the bivariate Laplace distribution and α is the specified amount of type I error. We give a brief discussion of the behavior or Eq. 7 as each of the parameters are varied independently. For ρ and w fixed, B decreases logarithmically to zero as α goes to 1. For a particular error distribution and choice of α, B increases as the width of the linear feature increases. For small w the rate of increase is quadratic, but for large w the rate of increase is linear. As ρ increases, meaning the variance of the error distribution decreases, B decreases exponentially.

For the case when $\int_0^{w/2} f_\rho(\hat{x})\, d\hat{x} \geq (1 - \alpha)/2$, less than $\alpha/2$ of the density of the error distribution occurs beyond the edge of the linear feature. Therefore, the probability of misclassifying an observed location as off the linear feature when the true location is on the linear feature is already less than or equal to α and buffering is no longer necessary. This result highlights why GPS measurement error is of particular importance in the context of narrow habitats, such as linear features. The power function for assessing the type I and type II errors associated with B is (Appendix C)

$$\beta(x) = \begin{cases} 1 - e^{\rho x} \sinh(\rho B) & \text{if } x < -B, \\ e^{-\rho b} \cosh(\rho x) & \text{if } -B < x < B, \\ 1 - e^{-\rho x} \sinh(\rho B) & \text{if } x > B. \end{cases} \tag{8}$$

5 Example data analysis

In this section we apply our approach using location data simulated with ArcGIS 9 (ESRI) and Mathematica 5.1 (Wolfram Research, Inc.) (Fig. 5). "True" locations were placed in a 10 km × 10 km landscape containing 9 m wide linear features according to a Poisson process, and constrained so that 250 points fell on the linear features and 750 points fell off the linear features. For each true location, an "observed" location was generated using the bivariate Laplace error distribution for the Lotek GPS_3300 collar. Buffer selection included four steps. (1) Choice of a priori type I error rate. For this analysis we selected an α-level of 0.05, which means we are willing to accept a misclassification of an observed location off the linear feature five times out of 100. The choice of α will vary depending on the biological question under consideration and knowledge of the study system. The implications of the choice of α are further considered in the discussion. (2) Quantification of the error distribution. For this example, we used the error distribution obtained from the Lotek GPS_3300 collar, which followed the bivariate Laplace distribution with parameter $\rho = 0.1123$. (3) Calculation of the buffer. From Eq. 7, $B = 27$ m for a linear feature width of 9 m. B represents the half-width of the buffered linear feature, so the total width of the buffered linear feature is 54 m. The buffer calculation must be repeated for each linear feature of

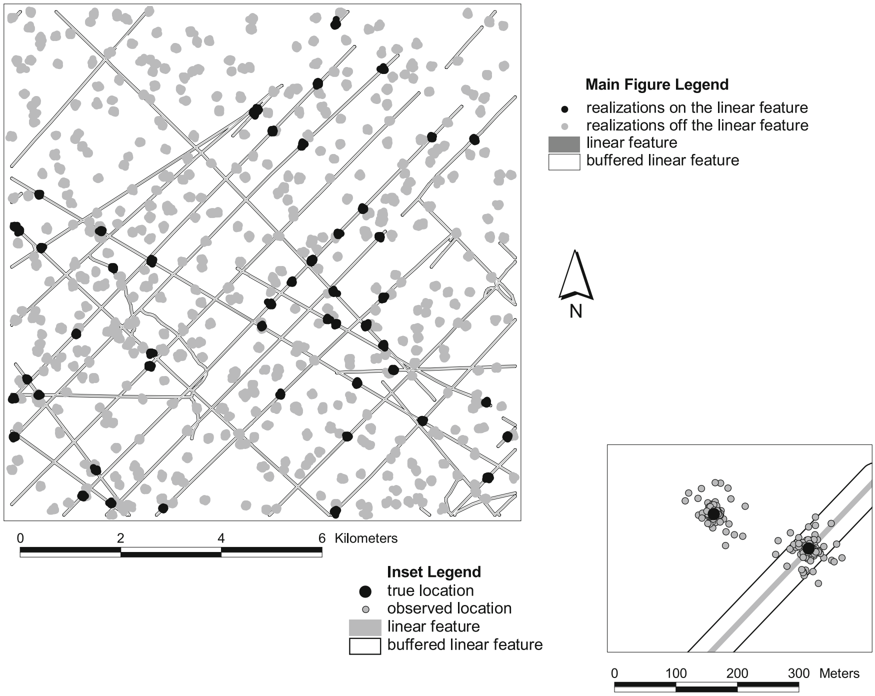

Fig. 5 Simulated location data for the example analysis. The landscape is 10 km × 10 km. "True" locations were placed in the landscape according to a Poisson process. "Observed" locations were generated under the assumption of a bivariate Laplace error distribution with parameter $\rho = 0.1123$

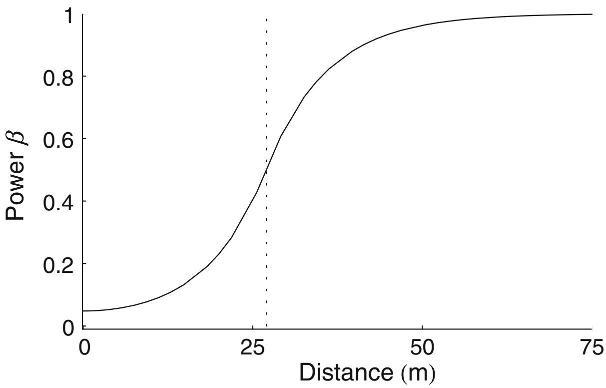

Fig. 6 The power function β (solid line) of the buffer (location shown by dotted line) chosen for the example data analysis

different width. (4) Error Assessment. Once the buffer is selected, we graphically assess the robustness of the buffer to classification error by computing the power function (Eq. 8, Fig. 6). The probability of type I error ranged from 4.8% for true locations at the centre to 5.5% for true locations near the edge of the linear feature. For true locations off the linear feature the probability of type II error was high near the edge

Table 2 Location classification of the example data. Columns 1–4 are the number of points classified as habitat type i, given that they are truly located in habitat j. Type I error is (off|on)/250 and type II error is (off|on)/750. Proportion of use is (on|on + on|off)/1000

Classification	on|on	off|on	on|off	off|off	Type I error	Type II error	Proportion of use
Truth	250	–	–	750	–	–	0.25
No buffer	101	149	8	742	0.60	0.01	0.11
Buffer	243	7	56	694	0.03	0.07	0.30

of the linear feature (94.5%), but dropped to 50% at the edge of the buffer (22.5 m from the edge of the linear feature), and was trivial ($< 1\%$) at 60 m from the edge of the linear feature.

When the true locations are known we can directly assess the performance of the buffering method (Table 2). For example, in these example data significant reduction in type I error caused only a small increase in type II error. The type I error decrease by 57% while the type II error increased by only 7%. The total number of correctly classified observed locations (on|on + off|off) increased with the addition of the buffer from 843 to 937 locations. Therefore, the estimate of the proportion of observed locations on linear features changed from 0.11 to 0.30, while the true proportion was 0.25. This particular example highlights both the effectiveness and the limitations of the buffering method, which are further addressed in the discussion.

6 Discussion

The example data analysis demonstrated that measurement error leads to underestimation of linear feature use. Failure to evaluate and consider GPS measurement error often interferes with our ability to detect ecological mechanisms (Rettie and McLoughlin 1999), resulting in poorly informed management decisions. In the introduction we identified several weaknesses of the current buffering approach. The method for buffering linear features presented here increases our ability to correct for the bias introduced by GPS measurement error by addressing these concerns.

The method explicitly includes information about the distribution of GPS measurement error. We found the bivariate Laplace distribution best represented the measurement error of the Lotek GPS_3300 collar in a closed conifer forest (Table 1). This result differs from previous studies where uniform (Dickson and Beier 2002; McLoughlin et al. 2002; Conner et al. 2003; Dickson et al. 2005) or normal (Samuel and Kenow 1992; Jerde and Visscher 2005; Visscher 2006) distributions were assumed to describe measurement error, but were not validated. Future investigations should assess if there are differences in kernel shape between different brands of collars to determine if a particular brand of collar would be better suited for assessing linear feature use. Because habitat variables affect the distribution of GPS measurement error (Moen et al. 1996; Frair et al. 2004; Cain et al. 2005), researchers should quantify error distributions for each collar and habitat type. If habitat type is heterogeneous within the study area, it may be necessary to use GPS measurement error distributions specific to each habitat type. For example, consider a linear feature running partly through closed conifer (cc) and partly through open deciduous (od). If previous studies quantifying the measurement error distribution between these habitats found them to be significantly different (i.e. $\rho_{cc} \neq \rho_{od}$), each section of the linear feature would have a different buffer size. However, habitat-specific distributions will increase the complexity of applying the method and the effect of habitat type on the GPS measurement error distribution should be investigated before adopting this approach.

The buffer, which directly incorporates habitat structure via the width of the linear feature and gives the a priori level of type I error, is amenable to error analysis. The ability to control the type I error permits flexibility in the choice of buffer size depending on the research question. Recall type I error corresponds to classifying an observed location outside a habitat when in truth it is inside the habitat. For example, it is often important for conservation to understand the value of specific corridors (Haddad et al. 2003), animal behaviours associated only with linear features (Dyer et al. 2001), or predator–prey interactions on linear features (James 1999). Therefore, a researcher is likely to choose a conservative type I error rate to avoid underestimating linear feature use.

While our method ensures that observed location classification achieves a specified level of type I error, there is no direct control of type II error (i.e. classifying a location as in a habitat, when in truth it is outside of a habitat). Because the level of type II is not constrained in the method, it may remain constant or increase after the application of the buffers. The probability of making a type II error depends on the distribution of animal locations relative to the linear features. For example, if animals are found either on linear features or quite far away from linear features, buffering is unlikely to cause a large increase in type II error. However, if the animals are often found off linear features, but near the edges, then type II error will increase with buffering. In the latter case, the importance of the edge habitat will be missed and the importance of the linear feature overestimated. Therefore, buffers may not always lead to better estimates of habitat use since total error (i.e. the sum of the type I and type II errors) may increase or decrease.

Two approaches can be used to gain insight into the trade off between type I and type II error. First, the distribution of animal locations can be used to determine whether type II error is likely to remain constant or increase significantly after buffering. If there are relatively few animal locations between the edge of the linear feature and the edge of the buffer, as compared to the total number of locations, it is unlikely type II error will increase significantly with buffering. Therefore, it is the local density of animal locations near the edge of linear features, and not the overall density of locations in the landscape, that will affect type II error. In the example data analysis, 8% of the locations off linear features were between the edge of the linear feature and the edge of the buffer. Therefore, we could expect a similar percentage increase in the type II, which we saw (Table 2). This suggests animal location data should be assessed before applying the buffering method in order to determine if type II error is likely to be a significant problem. Second, by inspecting the graph of the power function (Fig. 6), it is possible to evaluate where the probability of making a type II error (i.e. $1 - \beta$) becomes small. By visually examining the data and the power function researchers can trade off between type I and type II error by varying the chosen value of α and comparing the corresponding power functions.

To compare our buffering method to traditional approaches we reanalysed the example data. We assumed a location error of 31 m 95% of the time (D'Eon et al. 2002) and applied this as a buffer on the linear features. This buffer had 0 type I error (correctly classified 250 points on linear features) and a type II error of 0.1 (incorrectly classified 56 points on linear features). The traditional buffer resulted in no type I error, but there was increased type II error compared to the buffer we proposed. Although the traditional buffer performed similarly to the buffer we proposed for these example data, its performance is not guaranteed to remain consistent over different sample sizes. In addition, because it does not depend on the width of the linear feature, the traditional buffer remains constant for linear features of different sizes. The buffer proposed here would adjust, becoming smaller as linear features got wider (see Eq. 7), in order to appropriately correct for bias introduced by GPS measurement error, which depends to a large part on the spatial extent of the habitat in which you are trying to detect the location.

We considered linear features as one example of habitats of small spatial extent, but the method is more broadly applicable. The generally straight nature of the linear feature allowed us to reduce the problem from two to one dimensions. For non-linear habitats of small spatial extent this simplification is inappropriate. In these cases it is possible to compute B by numerically integrating the error distribution in two spatial dimensions. In addition, we assumed the location of the linear feature was known exactly. This is not often the case for real landscapes. Future work should focus on identifying the effect of measurement error in landscape features, as it is potentially a significant source of classification error that may compound or dominate the effect of GPS measurement error in the animal locations.

A focus on GPS measurement error and animal use of linear features is timely. GPS technology is now commonly used to acquire animal location data, linear feature densities are likely to increase in the future (Timoney and Lee 2001), and linear features affect several aspects of animal ecology including movement and survival (Thurber et al. 1994; James 1999; Dyer et al. 2001, 2002; Whittington et al. 2005). Although GPS telemetry is more accurate than traditional radio telemetry, access to more precise location data has stimulated biological inquiry at increasingly finer scales (Deutsch et al. 1998). As technology advances, it is necessary to acknowledge that limitations still exist. Rigorous methods for quantifying and addressing measurement error are needed to ensure biological investigation occurs at an appropriate scale given the measurement error inherent in the data (Ryan et al. 2004). Empirically justified buffers correct for GPS measurement error and lead to more accurate, consistent, and informed inference about animal use of habitats of small spatial extent.

Acknowledgements The authors thank N. Webb, J. Burger, J. Boyd, and Alberta Sustainable Resource Development for assistance with the GPS data collection. W. Nelson and M. Wonham provided useful discussions and comments. HWM was supported by an NSERC CGS-M, an Alberta Ingenuity Studentship, and the University of Alberta. CLJ was supported by a NSERC CRO Grant to MAL. DRV was supported by the University of Alberta and NSERC (Industrial to Weyerhaeuser). EHM was supported by an NSERC CRO. MAL was supported by a NSERC Discovery Grant and a Canada Research Chair in Mathematical Biology.

Appendix A: The modified Bessel function of the second kind

After Abramowitz and Stegun (1972), the modified Bessel function of the second kind, $K_n(x)$, is one of the solutions to the modified Bessel differential equation. For the special case where $n = 0$ it can be written as

$$K_0(x) = \int_0^\infty \frac{\cos(xt)}{\sqrt{t^2 + 1)}}\, dt. \tag{A-1}$$

Appendix B: The distribution of the observed locations

We derive the distribution of $\hat{x}$ where x is on the linear feature. Given no prior information regarding the distribution of true locations across the linear feature, we assume a uniform distribution for x such that

$$\phi(x) = \begin{cases} \frac{1}{w} & x \in [-w/2, w/2], \\ 0 & \text{elsewhere.} \end{cases} \tag{B-1}$$

If prior information is available indicating that animals prefer using certain regions of the linear features, such as the edges, the procedure could be adjusted to account for this by assuming an alternate distribution for x. Using Bayes' Rule, the marginal distribution of $\hat{x}$ is given by

$$f_\Theta(\hat{x}) = \frac{1}{w} \int_{-w/2}^{w/2} f_\Theta(\hat{x} \mid x)\, dx. \tag{B-2}$$

Replacing the general error distribution with the Bessel model,

$$\begin{aligned} f_\rho(\hat{x}) &= \int_{-w/2}^{w/2} \frac{\rho}{2w} e^{-\rho|\hat{x}-x|}\, dx \\ &= \begin{cases} \frac{e^{\rho\hat{x}} \sinh(\frac{\rho w}{2})}{w} & \text{if } \hat{x} < -w/2, \\ \frac{1-e^{-\frac{\rho w}{2}} \cosh(\rho\hat{x})}{w} & \text{if } -w/2 < \hat{x} < w/2, \\ \frac{e^{-\rho\hat{x}} \sinh(\frac{\rho w}{2})}{w} & \text{if } \hat{x} > w/2, \end{cases} \end{aligned} \tag{B-3}$$

where the solution is found using a change of variable as in Kot et al. (1996).

Appendix C: Derivation of the power function

Given a buffered linear feature of half-width B, the rejection region (RR) for the hypothesis test is $(-\infty, -B] \cup [B, \infty)$. The power function $\beta(x)$ is defined to be

$$\begin{aligned}\beta(x) &= P_x(\hat{x} \in RR) \\ &= \begin{cases} P(\text{type I Error}) & \text{if } x \in [-w/2, w/2], \\ P(1 - \text{type II Error}) & \text{otherwise} \end{cases} \\ &= 1 - \int_{-B}^{B} f_{\Theta}(\hat{x} \mid x)\, d\hat{x}. \end{aligned} \tag{C-1}$$

Replacing the general error distribution with the Bessel model and following Kot et al. (1996),

$$\begin{aligned}\beta(x) &= 1 - \int_{-B}^{B} \frac{\rho}{2} e^{-\rho|\hat{x}-x|}\, d\hat{x} \\ &= \begin{cases} 1 - e^{\rho x}\sinh(\rho B) & \text{if } x < -B \\ e^{-\rho b}\cosh(\rho x) & \text{if } -B < x < B \\ 1 - e^{-\rho x}\sinh(\rho B) & \text{if } x > B. \end{cases}\end{aligned} \tag{C-2}$$

References

Abramowitz M, Stegun IA (eds) (1972) Handbook of mathematical functions with formulas, graphs, and mathematical tables, 9th printing, chapter Modified Bessel Functions I and K, Dover, pp 374–377

Burham KP, Anderson DR (1998) Model selection and multinomial inference, 2nd edn. Springer

Cain JW, Krausman PR, Jansen BD, Morgart JR (2005) Influence of topography and GPS fix interval on GPS collar performance. Wildlife Soc Bull 33(3):926–934

Casella G, Berger RL (2002) Statistical inference, 2 edn. Duxbury

Conner LM, Smith MD, Burger LW (2003) A comparison of distance-based and classification-based analyses of habitat use. Ecology 84(2):526–531

D'Eon RG, Delparte D (2005) Effects of radio-collar position and orientation on GPS radio-collar performance, and the implications of PDOP in data screening. J Appl Ecol 42(2):383–388

D'Eon RG, Serrouya R, Smith G, Kochanny CO (2002) GPS radiotelemetry error and bias in mountainous terrain. Wildlife Soc Bull 30:430–439

Deutsch CJ, Bonde RK, Reid JP (1998) Radio-tracking manatees from land and space: tag design, implementation, and lessons learned from long-term study. Mar Technol Soc J 32(1):18–29

Dickson BG, Beier P (2002) Home-range and habitat selection by adult cougars in southern California. J Wildlife Manage 66(4):1235–1245

Dickson BG, Jenness JS, Beier P (2005) Influence of vegetation, topography, and roads on cougar movement in southern California. J Wildlife Manage 69(1):264–276

Dyer SJ, O'neill JP, Wasel SM, Boutin S (2001) Avoidance of industrial development by woodland caribou. J Wildlife Manage 65(3):531–542

Dyer SJ, O'neill JP, Wasel SM, Boutin S (2002) Quantifying barrier effects of roads and seismic lines on movements of female woodland caribou in northeastern Alberta. Can J Zool 80(5):839–845

Efron B, Tibshirani RJ (1993) An introduction to the bootstrap. Chapman and Hall/CRC

Frair JL, Nielsen SE, Merrill EH, Lele SR, Boyce MS, Munro RHM, Stenhouse GB, Beyer HL (2004) Removing GPS collar bias in habitat selection studies. J Appl Ecol 41(2):201–212

Hebblewhite M, Percy M, Merrill EH (2007) Are all global positioning system collars created equal? Correcting habitat-induced bias using three brands of collars in the Central Canadian Rockies. J Wildlife Manage 71(6):2026–2033

Haddad NM, Bowne DR, Cunningham A, Danielson BJ, Levey DJ, Sargent S, Spira T (2003) Corridor use by diverse taxa. Ecology 84(3):609–615

Hurford AL (2005) Wolf movement within and beyond territory boundary. Master's thesis, University of Alberta

James ARC (1999) Effects of industrial development on the predator–prey relationship between Wolves and Caribou in Northeastern Alberta. Ph.D. thesis, Universtiy of Alberta

Jerde CL, Visscher DR (2005) GPS measurement error influences on movement model parameterization. Ecol Appl 15(3):806–810

Kot M, Lewis MA, vandendriessche P (1996) Dispersal data and the spread of invading organisms. Ecology 77(7):2027–2042

Kotz S, Kozubowski T, Podgorski K (2001) The Laplace distribution and generalizations, a revisit with applications to communications, economics, engineering, and finance. Birkhauser

McLoughlin PD, Case RL, Gau RJ, Cluff HD, Mulders R, Messier F (2002) Hierarchical habitat selection by barren-ground grizzly bears in the central Canadian Arctic. Oecologia 132(1):102–108

Moen R, Pastor J, Cohen Y, Schwartz CC (1996) Effects of moose movement and habitat use on GPS collar performance. J Wildlife Manage 60(3):659–668

Nams VO, Boutin S (1991) What is wrong with error polygons. J Wildlife Manage 55(1):172–176

Rettie WJ, McLoughlin PD (1999) Overcoming radiotelemetry bias in habitat-selection studies. Can J Zool 77(8):1175–1184

Ryan PG, Petersen SL, Peters G, Gremillet D (2004) GPS tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of African Penguins. Mar Biol 145(2):215–223

Saltz D (1994) Reporting error measures in radio location by triangulation – a review. J Wildlife Manage 58(1):181–184

Samuel MD, Kenow KP (1992) Evaluating habitat selection with radiotelemetry triangulation error. J Wildlife Manage 56(4):725–734

Thurber JM, Peterson RO, Drummer TD, Thomasma SA (1994) Gray wolf response to refuge boundaries and roads in Alaska. Wildlife Soc Bull 22(1):61–68

Timoney K, Lee P (2001) Environmental management in resource-rich Alberta, Canada: first world jurisdiction, third world analogue?. J Environ Manage 63(4):387–405

Visscher DR (2006) GPS measurement error and ressource selection functions in a fragmented landsscape. Ecography 29(3):458–464

Whittington J, St. Clair CC, Mercer G (2005) Spatial responses of wolves to roads and trails in mountain valleys. Ecol Appl 15(2):543–553

Author Biographies

Hannah W. McKenzie is a Ph.D. student at the University of Alberta, in Edmonton, Alberta. She obtained her M.Sc. in Mathematical and Statistical Biology at the University of Alberta, where she researched the effect of industrial development on wolf movement and predator–prey interactions. She currently works on population dynamics in streams.

Christopher L. Jerde received his Ph.D. at University of Alberta and is currently a postdoctoral research associate at the University of Notre Dame's Center for Aquatic Conservation where he works on applied probability models for predicting biological invasions.

Darcy R. Visscher is a Ph.D. student at the University of Alberta, in Edmonton, Alberta. He obtained his M.Sc. in African Zoology at the University of Pretoria (South Africa), where he researched the population dynamics of African buffalo in Kruger National Park. He currently works on models of adaptive behaviour in the movement and patch use of elk.

Dr. Evelyn H. Merrill is professor in the Department of Biological Sciences at the University of Alberta. She completed her Ph.D. at the University of Washington and her M.Sc. at the University of Idaho. She works on how landscape patterns influence trophic interactions of large mammals and the spread of diseases.

Mark A. Lewis is Professor and Canada Research Chair in Mathematical Biology at the University of Alberta. He received his D.Phil. from Oxford, and works on the application of quantitative methods to spatial ecology.