Journal of Wildlife Management 74(3):539–548; 2010; DOI: 10.2193/2008-327

Tools and Technology Article

Evaluation of Root-n Bandwidth Selectors for Kernel Density Estimation

TODD D. STEURY,[1,2] *Department of Biology, Washington University, St. Louis, MO 63130, USA*
JOHN E. MCCARTHY, *Department of Mathematics, Washington University, St. Louis, MO 63130, USA*
TIMOTHY C. ROTH, II,[3] *Department of Biology, Indiana State University, Terre Haute, IN 47809, USA*
STEVEN L. LIMA, *Department of Biology, Indiana State University, Terre Haute, IN 47809, USA*
DENNIS L. MURRAY, *Department of Biology, Trent University, Peterborough, ON K9J 7B8, Canada*

ABSTRACT The kernel density estimator is used commonly for estimating animal utilization distributions from location data. This technique requires estimation of a bandwidth, for which ecologists often use least-squares cross-validation (LSCV). However, LSCV has large variance and a tendency to under-smooth data, and it fails to generate a bandwidth estimate in some situations. We compared performance of 2 new bandwidth estimators (root-n) versus that of LSCV using simulated data and location data from sharp-shinned hawks (*Accipter striatus*) and red wolves (*Canis rufus*). With simulated data containing no repeat locations, LSCV often produced a better fit between estimated and true utilization distributions than did root-n estimators on a case-by-case basis. On average, LSCV also provided lower positive relative error in home-range areas with small sample sizes of simulated data. However, root-n estimators tended to produce a better fit than LSCV on average because of extremely poor estimates generated on occasion by LSCV. Furthermore, the relative performance of LSCV decreased substantially as the number of repeat locations in the data increased. Root-n estimators also generally provided a better fit between utilization distributions generated from subsamples of hawk data and the local densities of locations from the full data sets. Least-squares cross-validation generated more unrealistically disjointed estimates of home ranges using real location data from red wolf packs. Most importantly, LSCV failed to generate home-range estimates for >20% of red wolf packs due to presence of repeat locations. We conclude that root-n estimators are superior to LSCV for larger data sets with repeat locations or other extreme clumping of data. In contrast, LSCV may be superior where the primary interest is in generating animal home ranges (rather than the utilization distribution) and data sets are small with limited clumping of locations.

KEY WORDS bandwidth, home range, kernel density, smoothing, utilization distribution.

The kernel density estimator (KDE) is recognized as an accurate and effective method for estimating animal home ranges and utilization distributions (UDs; Worton 1995, Seaman and Powell 1996). The KDE is preferred over other methods such as minimum convex polygon and harmonic mean because the KDE is nonparametric, allows for multiple centers of activity, and is well understood statistically (Silverman 1986, Worton 1989, Kernohan et al. 2001). In addition, the estimator generates a true probability distribution and thus the end result represents an animal's probabilistic use of space (cf. Getz and Wilmers 2004, Getz et al. 2007). Many ecologists regard the KDE as the best available method for estimating UDs (e.g., Powell 2000, Kernohan et al. 2001).

In kernel density estimation, the choice of smoothing parameter (bandwidth) is critical to UD estimate accuracy (Silverman 1986, Worton 1989, Wand and Jones 1995). Currently, bandwidth in home-range and UD studies most often is estimated via least-squares cross-validation (LSCV) because it has low bias and extensive history, and is widely available in spatial analysis software (Silverman 1986, Worton 1989, Gitzen and Millspaugh 2003). The LSCV estimator has been shown to perform well in density estimation, even when compared to newer, second-generation bandwidth estimators such as plug-in and solve-the-equation methods (Seaman and Powell 1996, Loader 1999, Gitzen et al. 2006). However, the LSCV bandwidth estimator is not without its shortcomings; the estimator has large sampling variance (Wand and Jones 1995, Wu and Tsai 2004), and thus bandwidth estimates based on LSCV lack precision. Furthermore, LSCV can under-smooth data, especially in the outer UD isopleths, such that home-range estimates are characterized by small, disjointed polygons (Powell 2000, Getz and Wilmers 2004). Finally, in some cases LSCV can fail to estimate bandwidth altogether (Blundell et al. 2001, Kernohan et al. 2001, Amstrup et al. 2004). Thus, in many cases a suitable alternative to LSCV is needed.

Our objective was to evaluate 2 new bandwidth estimators (root-n and modified root-n) developed for kernel density estimation by Wu and Tsai (2004; also see Chiu 1991, 1992). These bandwidth estimators were derived from, and are closely associated with, the LSCV method but always produce bandwidth estimates, should generate more precise UD estimates (with minimal loss in accuracy), and should be less likely to under-smooth data (Wu and Tsai 2004). We compare the performance of bandwidth estimators using both simulated UDs as well as location data from radiotelemetry studies of sharp-shinned hawks (*Accipter striatus*) and red wolves (*Canis rufus*).

METHODS

Kernel Methods and Bandwidth Estimators

The bivariate KDE is defined as

$$\hat{f}(x,y) = \frac{1}{nh^2}\sum_{j=1}^{n} K\left(\frac{x-X_j}{h}, \frac{y-Y_j}{h}\right),$$

[1] *E-mail: steury@auburn.edu*
[2] *Present address: School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL 36849, USA*
[3] *Present address: Department of Biology, University of Nevada, Reno, NV 89557, USA*

where n is the sample size, h is the bandwidth, K is a chosen kernel function, X_j and Y_j are spatial coordinates from the location data, and x and y are spatial coordinates where the function is evaluated (Silverman 1986, Worton 1989). Whereas we can calculate a different value of h for each dimension (see Sain et al. 1994, Wand and Jones 1995, Wu and Tsai 2004 for such methods), standardizing the dispersion in x and y allows us to apply the same bandwidth in both directions. Such scaling appears to have little effect on accuracy of the UD estimate (Gitzen and Millspaugh 2003). Additionally, although adaptive kernels can use different bandwidth values based on local densities of data, in general the adaptive kernel has not been as well studied statistically and does not appear to provide better estimates of UDs (Worton 1995, Seaman and Powell 1996, Seaman et al. 1999, Powell 2000, but see Discussion). Finally, although numerous kernel functions can be used (e.g., see Silverman 1986 or Worton 1989), we apply the Gaussian function because of its popularity in UD estimation and to simplify calculations (Appendix). Choice of kernel has little impact on accuracy of the UD estimate, especially relative to the effect of the choice of bandwidth (Silverman 1986, Worton 1989).

To estimate bandwidth using LSCV ($\hat{h}_{LSCV}$), we minimized the LSCV score function with respect to h:

$$LSCV(h) = \frac{1}{\pi n h^2} + \frac{1}{4\pi n^2 h^2} \sum_{j,k=1}^{n} e^{\frac{-d_{jk}^2}{4h^2}} - 4e^{\frac{d_{jk}^2}{2h^2}},$$

where

$$d_{jk}^2 = (X_j - X_k)^2 + (Y_j - Y_k)^2$$

(Silverman 1986, Worton 1989). The LSCV score function may have ≥ 1 local minima in addition to the global minimum; the value of h that produces the largest local minimum typically is recommended for use in KDE (Park and Marron 1990, Wand and Jones 1995).

Variance in bandwidth estimates using LSCV tends to be higher and the LSCV method is more likely to fail or under-smooth when data are clustered spatially; such data attributes may occur when animals remain at a site for long periods before moving, return to specific sites (e.g., central place foraging, den- or nest-site attendance), or when data points are discontinuous in space due to rounding (Silverman 1986, Chiu 1991, Wu and Tsai 2004, Hemson et al. 2005, Gitzen et al. 2006). Wu and Tsai (2004) propose a modification to the LSCV score function that eliminates the high-frequency noise in location data caused by repeats or clumping of locations. The value of h that minimizes this score function is a bandwidth estimate that has the optimal convergence rate of $On^{-1/2}$ (i.e., the rate at which variance in the estimate decreases with increasing sample size; Wu and Tsai 2004) and thus has been termed a root-n bandwidth estimate ($\hat{h}_n$). Whereas we used the Gaussian kernel (Appendix), we express the root-n score function as

$$\begin{aligned}\hat{M}(h,T) = {} & \frac{1}{4\pi n h^2} \\ & - \frac{1}{2(\pi n)^2} \sum_{j,k=1}^{n} \int_{-T}^{T} \int_{-T}^{T} e^{-\frac{h^2}{2}(t^2+s^2)} \\ & \cos\left[(X_j - X_k)t\right] \cos\left[(Y_j - Y_k)s\right] dtds \\ & + \frac{n-1}{n^3(2\pi)^2} \sum_{j,k=1}^{n} \int_{-T}^{T} \int_{-T}^{T} e^{-h^2(t^2+s^2)} \\ & \cos\left[(X_j - X_k)t\right] \cos\left[(Y_j - Y_k)s\right] dtds,\end{aligned}$$

where t and s are the Fourier transformations of x and y, and we estimate T as that value ($\hat{T}_\infty$) that (globally) minimizes the following function, which is based on cross-validation:

$$CV^{\infty}(T) = \frac{8T^2}{n+1} - \frac{4}{n^2} \sum_{j,k=1}^{n} \frac{\sin\left[T(X_j - X_k)\right]}{(X_j - X_k)} \frac{\sin\left[T(Y_j - Y_k)\right]}{(Y_j - Y_k)}.$$

Wu and Tsai (2004) note that $\hat{T}_\infty$ may still be an unnecessarily large estimate of T, which truncates the score function to eliminate high-frequency noise (the score function for LSCV is similar to $\hat{M}[h,T]$ with T = ∞). A further modification reduces the estimate based on the aforementioned principle that a local minimum may return superior results. The modified estimate for T is that value ($\hat{T}^*$) that (globally) minimizes

$$CV^*(T) = CV^{\infty}(T) + 1.96$$

$$\sqrt{\left\{\frac{32}{n^4}\left(T - \hat{T}_{loc}\right)^2 \sum_{j,k=1}^{n} \frac{\sin\left[\hat{T}_\infty(X_j - X_k)\right]}{(X_j - X_k)} \frac{\sin\left[\hat{T}_\infty(Y_j - Y_k)\right]}{(Y_j - Y_k)}\right\}},$$

where $\hat{T}_{loc}$ is the smallest local minimizer of $CV^{\infty}(T)$ (Appendix). The modified root-n bandwidth estimate ($\hat{h}^*$) is the (global) minimizer of $\hat{M}(h, \hat{T}^*)$.

All 3 bandwidth estimators seek to minimize mean integrated squared error (MISE) of the UD estimate, or the mean of the squared difference between estimated and true UDs, and the score functions are approximations of the MISE. Thus, minimizing the score functions with respect to h should provide good estimates of the h that minimize differences between estimated and true UDs ($h_{\text{opt;}}$ Silverman 1986, Wand and Jones 1995).

Simulations

We used simulations of simple and complex UDs to compare performance of the LSCV and root-n bandwidth estimators. We used methods similar to those of Seaman and Powell (1996; see also Worton 1995, Seaman et al. 1999, Gitzen et al. 2006). We generated our basic simulated UDs using mixtures of 1, 4, 8, and 16 bivariate normal distributions. For each level of complexity, we generated 20 UDs, whereby we randomly selected attributes for each normal distribution in the mixture from uniform distributions with the following ranges: means from 0.0 to 12.0, standard deviations from 0.5 to 7.5, and x, y covariances from −1 to 1; each x and y had its own randomly

determined mean and standard deviation (Seaman and Powell 1996). We constrained mixing proportions to sum to 1. Mixed distributions such as these can generate varied and highly irregular UDs, including multimodal UDs and UDs with considerable skew or kurtosis (Seaman and Powell 1996, Gitzen et al. 2006). From each of these 80 simulated UDs, we generated 50 random data sets of 50 and 150 locations each. For each of the 8,000 data sets, we calculated values for $\hat{h}_{LSCV}$, $\hat{T}_{\infty}$, $\hat{T}_{loc}$, $\hat{T}^*$, $\hat{h}_n$, and $\hat{h}^*$ as described above.

Although LSCV typically performs well with simulations, such simulations generally do not contain repeat locations or extreme clumping of data. Thus, we also compared bandwidth estimator performance with simulated UDs containing repeated points. We generated repeat-point simulated UDs using mixtures of 4 or 16 bivariate normal distributions, whereby we generated attributes for distributions in the mixture as for our basic simulated UDs. However, in generating simulated location data, 4%, 8%, 12%, 16%, or 20% of locations were repeats, drawn (with replacement) from the randomly generated location data. For the repeat-point UD simulations, we generated 10 UDs at each level of complexity and proportion of repeats, from which we generated 20 random datasets of 50 and 150 locations each for a total of 4,000 generated data sets. We applied a small random jitter (<0.001 units in both x and y directions) to repeat locations to prevent LSCV failure and thereby facilitate comparison between estimators; thus, the relative performance of LSCV may be positively biased.

To evaluate and compare performance of different bandwidth estimators via simulations, we used MISE generated by the estimator relative to the minimum possible MISE for that data set (adapted from Wu and Tsai 2004):

$$R = \mathrm{MISE}_{\hat{h}} / \mathrm{MISE}_{h_{opt}} - 1.$$

For our simulations, we determined MISE using discrete approximation:

$$\mathrm{MISE} = \frac{1}{m} \sum_{i=1}^{m} \left[\hat{f}(x,y) - f(x,y) \right]^2,$$

where $\hat{f}(x, y)$ is estimated density of the UD, evaluated as defined previously, and $f(x, y)$ is true density of the UD (Silverman 1986). We determined true UD density $f(x, y)$ at any x, y coordinate by summing bivariate densities from each distribution in the mixture, multiplied by their respective mixing proportions. In our calculations we evaluated $\hat{f}(x, y)$ and $f(x, y)$ over a fine grid of m coordinates, whereby we chose the size of the grid such that it was sufficiently large to completely encompass all distributions in the mixture to 4 standard deviations and sufficiently fine that the calculated area of the true UD = 1.00 (to 2 decimal places; adapted from Seaman and Powell 1996, Seaman et al. 1999). We solved for $\mathrm{MISE}_{h_{opt}}$ using standard minimization routines.

Often, ecologists are interested in calculating an animal's home range, which frequently is the 95% isopleth of the UD (e.g., Gitzen et al. 2006). Thus, for each estimator we also calculated relative error in home-range area:

$$\mathrm{REA} = \frac{\hat{A} - A}{A},$$

where $\hat{A}$ and A are the area contained in the 95% isopleths of the estimated and true UDs, respectively (adapted from Seaman and Powell 1996, Seaman et al. 1999). We calculated variance and other indices of dispersion in performance measures over the random data sets at a given sample size for each simulated UD (Seaman and Powell 1996, Seaman et al. 1999).

Animal Location Data

In addition to using simulated UDs to evaluate root-n and LSCV estimators, we compared their performance using actual location data from radiocollared sharp-shinned hawks and red wolves. These 2 species are characterized by different patterns of space use; hawks can move quickly, are not territorial, return to a central roost each day, and their large-scale movement patterns are not different from random (Roth and Lima 2007), whereas wolves are slower, territorial, and move more systematically through their home range.

Data from 10 sharp-shinned hawks were collected from late November to late January during 2000–2004 in Indiana (Roth and Lima 2007). Hawks were located via telemetry between 2 hours and 10 hours per day, in 10-minute intervals. Number of locations averaged 1,938.1 (±255.9 SE) and ranged from 1,360 to 3,264. For the hawk data, we evaluated estimator performance by comparing UDs estimated from subsamples to the ranked density of the original data. Specifically, we ranked points from each original data set according to an index, equal to the sum of distances to the 10 nearest neighboring points. We deemed the X% of points with the smallest index values to belong inside the X isopleth of the true UD. Previous research into similar nearest-neighbor methods have shown that such methods are accurate at determining the probability density at sample locations (Mack and Rosenblatt 1979). From each hawk data set, we then randomly generated 100 subsamples (without replacement) of sizes 50 and 150. For each subsample, we generated UDs using LSCV and both root-n estimators. We then evaluated each estimated UD against the ranked, original data set; we calculated number of errors of commission or omission (data points inside or outside an isopleth that should have been outside or inside, respectively) at the 50%, 60%, 70%, 80%, 90%, and 95% isopleths. We added a small random jitter to the data to prevent LSCV failure; thus the relative performance of LSCV may be positively biased.

Radiotelemetry data from 47 red wolf packs was gathered in the early 1990s (n = 15) and early 2000s (n = 32) in eastern North Carolina (A. Beyer, United States Fish and Wildlife Service, unpublished data). Locations of individual wolves were obtained via aerial telemetry on roughly a weekly basis (A Beyer, unpublished data). In calculating pack UDs, we avoided pseudo-replication by only including one location per day for all pack members in close proximity (<500 m), plus the location(s) of any isolated pack

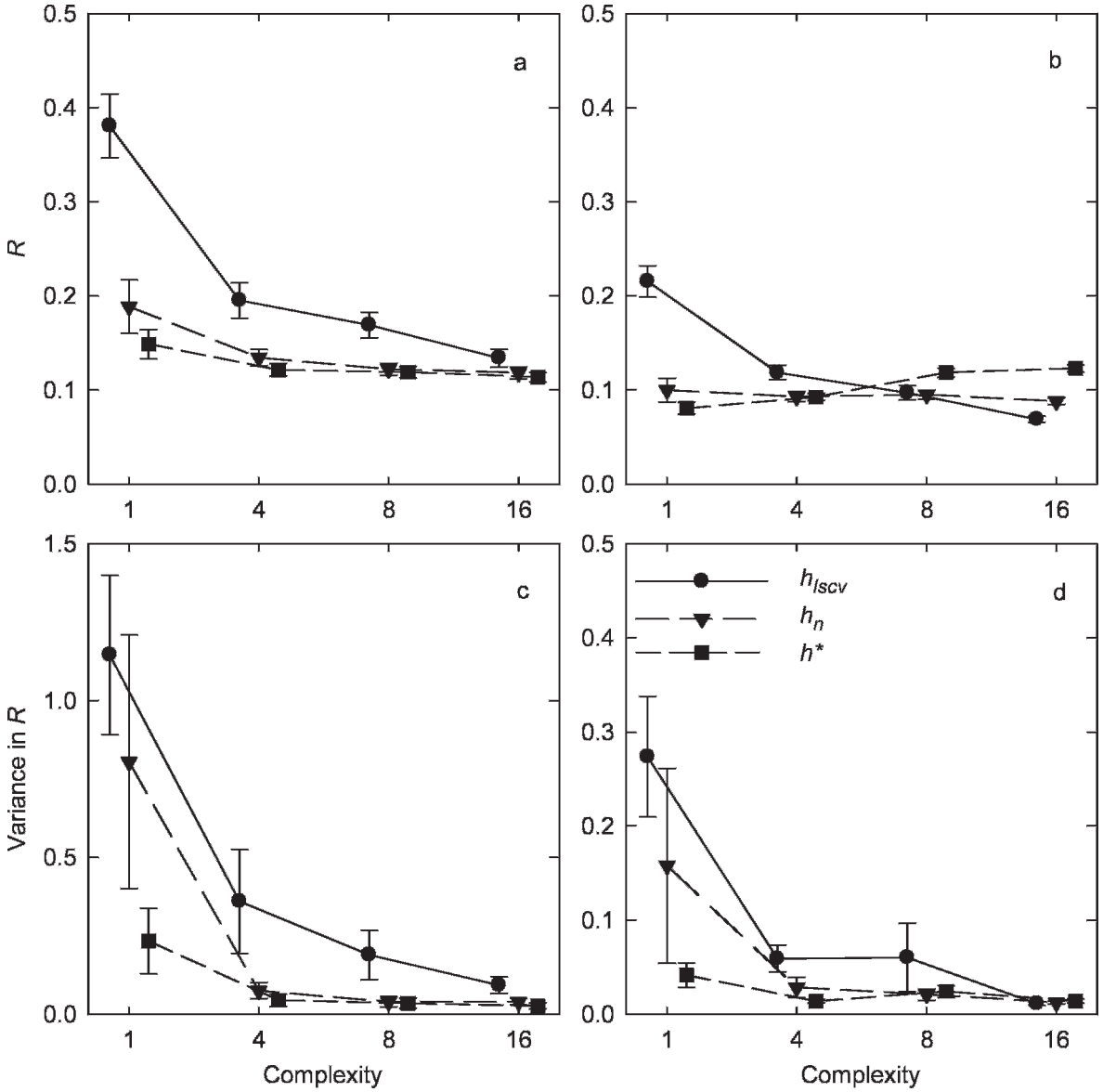

Figure 1. Relationship between the complexity of the utilization distribution (UD) and bandwidth performance. Complexity is the number of mixed distributions used to create the simulated UD. Measures of performance are the relative mean integrated square error (R) of the UD generated by each of the estimators (a, b) and the variance in R (c, d). Each point in a and b is the average of 1,000 random simulated data sets, whereas each point in c and d is the average of 20 randomly generated UDs. Error bars are standard error of the mean. Frames a and c are for $n = 50$; b and d are for $n = 150$.

members. Number of locations we used in UD estimation averaged 144.6 (±19.4) but ranged from 17 to 608. For the red wolf data sets, we assessed failure rates of different estimators. If neither estimator failed, we identified disjointed polygons or lacunae (holes) in red wolf home ranges (95% UD isopleths), which are potentially unrealistic and may indicate under-smoothing and thereby an underestimate of home-range size. We note, however, that we have no way of knowing what the true wolf home ranges were, and thus disjointed polygons may be accurate representations of wolf space use.

RESULTS

Simulations

With data from our basic simulated UDs, the 2 root-n estimators tended to produce UD estimates with lower relative MISE (R), on average, as well as lower variance in R, than that generated by LSCV, except at higher sample sizes and more complex UDs (Fig. 1). Differences in R and variance in R between the UDs generated by LSCV and those generated by root-n estimators were greatest at low complexity, because these measures of performance tended to converge among estimators at higher complexity (Fig. 1). Increasing sample size tended to decrease R and variance in R, but had little influence on relationships among R, variance in R, and the estimators (Fig. 1). The modified root-n estimator tended to generate lower average values of R and variance in R than the original root-n estimator, except at higher sample sizes and more complex UDs (Fig. 1). On a sample-by-sample basis, the root-n bandwidth estimators were more likely to provide superior fit than LSCV at the lowest UD complexity (proportion of

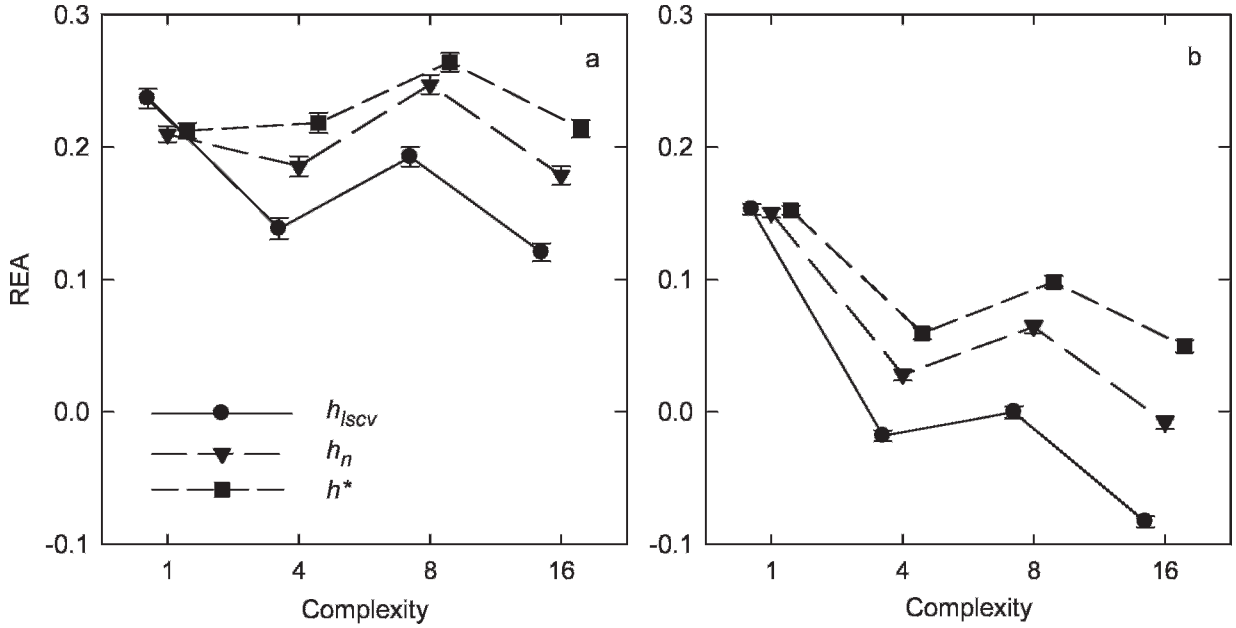

Figure 2. Relationship between the complexity of the utilization distribution (UD) and relative error in home-range area (REA). Complexity is the number of mixed distributions used to create the UD. Frames a and b depict REA area for $n = 50$ and $n = 150$, respectively. Each point is the average of 1,000 random simulated data sets. Error bars are standard error of the mean.

times that $R_n < R_{lscv} = 0.70 \pm 0.03$ [± 95% CI] for $n = 50$ and 0.69 ± 0.03 for $n = 150$). However, at highest UD complexity the root-n bandwidth estimators were less likely to provide superior fit than LSCV (proportion of times that $R_n < R_{lscv} = 0.44 \pm 0.03$ for $n = 50$ and 0.43 ± 0.03 for $n = 150$). At moderate levels of complexity (4 and 8 mixed normals), the 2 estimators had approximately equal probability of generating a superior fit, regardless of sample size (proportion of times that $R_n < R_{lscv}$ at 4 mixed normals $= 0.53 \pm 0.03$ for $n = 50$ and 0.48 ± 0.03 for $n = 150$; at 8 mixed normals $= 0.48 \pm 0.03$ for $n = 50$ and 0.5 ± 0.03 for $n = 150$). Differences between average and sample-by-sample results at higher UD complexity were due to high variance in LSCV generating a number of poor individual estimates, which skewed the average. The modified root-n estimator was less likely than the original root-n estimator to provide a better UD estimate on a sample-by-sample basis, except at lowest complexity (proportion of times that $R^* < R_n = 1.00$ for $n = 50$ and 150 at complexity = 1 mixed normal; proportion of times that $R^* < R_n$ at >1 mixed normal <0.482 and decreases with increasing sample size and complexity).

With data from our basic simulated UDs, LSCV tended to generate smaller home ranges than did root-n estimators. These differences resulted in lower positive relative error in home-range area (REA) with LSCV than with root-n estimators at low sample sizes (Fig. 2). However, at the lowest level of complexity all 3 estimators tended to generate similar REA (Fig. 2). Furthermore, at high sample sizes, LSCV tended to cause underestimation of home-range area, and the negative REA with LSCV was 67% larger (in absolute terms) than the REA generated by root-n estimators at highest UD complexity (Fig. 2). In general, REA decreased an average of 80% for all 3 estimators with larger sample size (Fig. 2). On a sample-by-sample basis, the LSCV and root-n bandwidth estimators were about equally likely to provide superior home range estimates at higher UD complexity (proportion of times that absolute value [abs] of $[REA_n] <$ abs$[REA_{lscv}]$ ranged from 0.474 to 0.536 [all 95% CI = 0.03] at 4, 8, and 16 mixed normals for $n = 50$ and 150). However, the root-n estimators were more likely to provide superior home-range estimates at the lowest UD complexity (proportion of times that abs$[REA_n]$ < abs$[REA_{lscv}] = 0.66 \pm 0.03$ for $n = 50$ and 0.64 ± 0.03 for $n = 150$). The original root-n estimator was more likely than the modified root-n estimator to generate a superior home-range estimate at higher levels of UD complexity and both sample sizes (proportion of times that abs$[REA_n] <$ abs$[REA_{lscv}] = 0.60 \pm 0.30$ for $n = 50$ and 150, complexity = 1; 0.37 ± 0.10 for $n = 50$, complexity = 4; 0.29 ± 0.08 for $n = 150$, complexity = 4; 0.42 ± 0.10 for $n = 50$ and $n = 150$; 0.40 ± 0.08 for $n = 50$, complexity = 16; 0.42 ± 0.07 for $n = 150$, complexity = 16).

Our repeat-point simulations revealed that repeats in location data can have drastic effects of the performance of the LSCV estimator. As proportion of repeats in the data increased, relative error in UD estimates and negative error in home range area increased dramatically with LSCV (Fig. 3). At 20% repeat locations, R was 50-fold greater with LSCV than with either root-n estimator, and the estimated home range area using LSCV was zero for virtually all data sets. Proportion of repeats had virtually no effect on performance with either root-n estimator, although we did detect a slight decrease in home-range area with increasing repeats (Fig. 3). On a sample-by-sample basis, the proportion of times that either root-n estimator returned superior results compared to LSCV increased from approximately 0.5 with no location repeats to

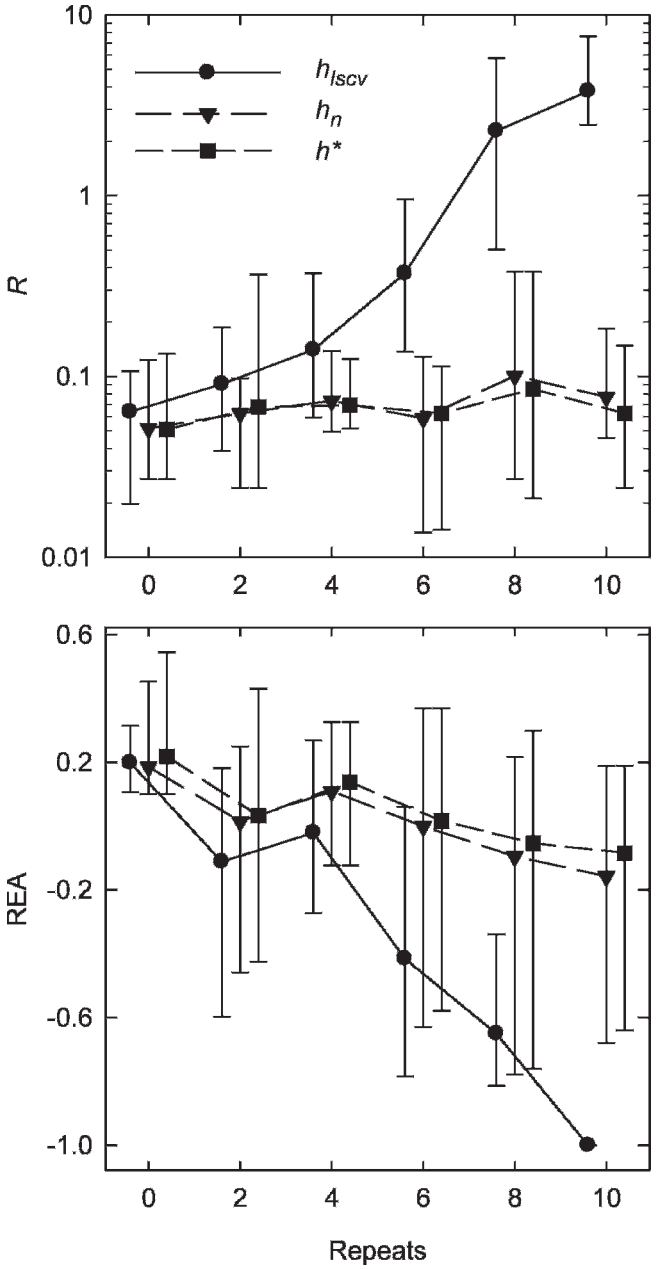

Figure 3. Relationship between the number of repeat locations in the data and bandwidth performance. Measures of performance are the relative mean integrated square error (R) of the utilization distribution (UD) generated by each of the estimators and the relative error in home range area (REA). Each point is the median of 1,000 random simulated data sets. Error bars are the minimum and maximum observed medians across 10 simulated UDs; we calculated medians from 20 data sets per UD. In both frames, complexity = 4, n = 50.

approximately 0.9 with 20% repeat locations. All of these results held across sample sizes and both levels of UD complexity.

Animal Location Data

Root-n methods generally performed superiorly to LSCV when generating UDs with hawk data. Total proportions of data points correctly identified as in or out of the 6 examined isopleths were 0.61 (SE = 0.0021), 0.70 (0.0015), and 0.71 (0.0012) for UDs generated using LSCV, original root-n, and modified root-n, respectively, for subsamples of 50 locations. For subsamples of 150 locations, the proportions of correctly identified data points were 0.64 (0.0008), 0.73 (0.0011), and 0.74 (0.0011) using LSCV, original root-n, and modified root-n estimators, respectively. On a sample-by-sample basis, the root-n and modified root-n estimators generated more accurate UDs than did LSCV in 98.6% and 98.5% of cases, respectively, for subsamples of size 50; root-n estimators were always more accurate at subsamples of size 150. The modified root-n estimator generated more accurate UDs than the original root-n in 93.6% and 96.7% of cases for subsamples of size 50 and 150, respectively. The LSCV estimator was more likely to under-smooth the data (omission error); 32.0%, 21.1%, and 18.7% of errors were due to under-smoothing for LSCV, root-n, and modified root-n UDs, respectively. However, the root-n estimators were more likely to over-smooth as 3.2%, 5.9%, and 6.9% of errors were due to over-smoothing. All of the above results held for individual isopleths.

The LSCV method failed to provide an estimate for the bandwidth for 11 (23.4%) of the 47 red wolf home ranges; all data sets with >5% repeat locations invoked LSCV failure, whereas LSCV did not fail for any datasets with <5% repeats. The root-n method provided an estimate for all wolf home ranges. For most pack home ranges, LSCV generated home-range estimates that were more disjunct, with more polygons and lacunae (e.g., Fig. 4). The average number of polygons and lacunae per home range produced under LSCV was 9.0 ± 1.4 and 1.5 ± 0.3, respectively. These numbers decreased to 4.8 ± 0.8 and 0.6 ± 0.2 polygons and lacunae, respectively, for root-n and 4.1 ± 0.6 and 0.5 ± 0.2, respectively, for modified root-n. Differences between estimators were significant for both number of polygons (repeated measures analysis of variance, $F_{2,72}$ = 18.82, $P < 0.001$; although root-n was not different from modified root-n, $P = 0.72$, Tukey's post-hoc test) and lacunae ($F_{2,72} = 7.28$, $P = 0.001$; root-n not different from modified root-n, $P = 0.83$).

DISCUSSION

Wu and Tsai (2004) suggested that performance of the root-n bandwidth estimators is superior to that of LSCV, with the modified root-n ($\hat{h}^*$) typically outperforming the original root-n ($\hat{h}_n$; also see Chiu 1992). However, Wu and Tsai (2004) evaluated performance of these estimators for simple bivariate probability distributions; performances of the root-n estimators have not been tested for more complex bivariate probability distributions, such as those that might describe animal UDs (e.g., Seaman and Powell 1996, Gitzen et al. 2006), or with real location data.

Without repeat locations, our simulations revealed that LSCV was more likely to generate the best fit between UD estimates and true UDs at high UD complexity in individual samples. Furthermore, on average, LSCV tended to have lower positive relative error in estimates of home-range area at low sample size or moderate complexity. However, the

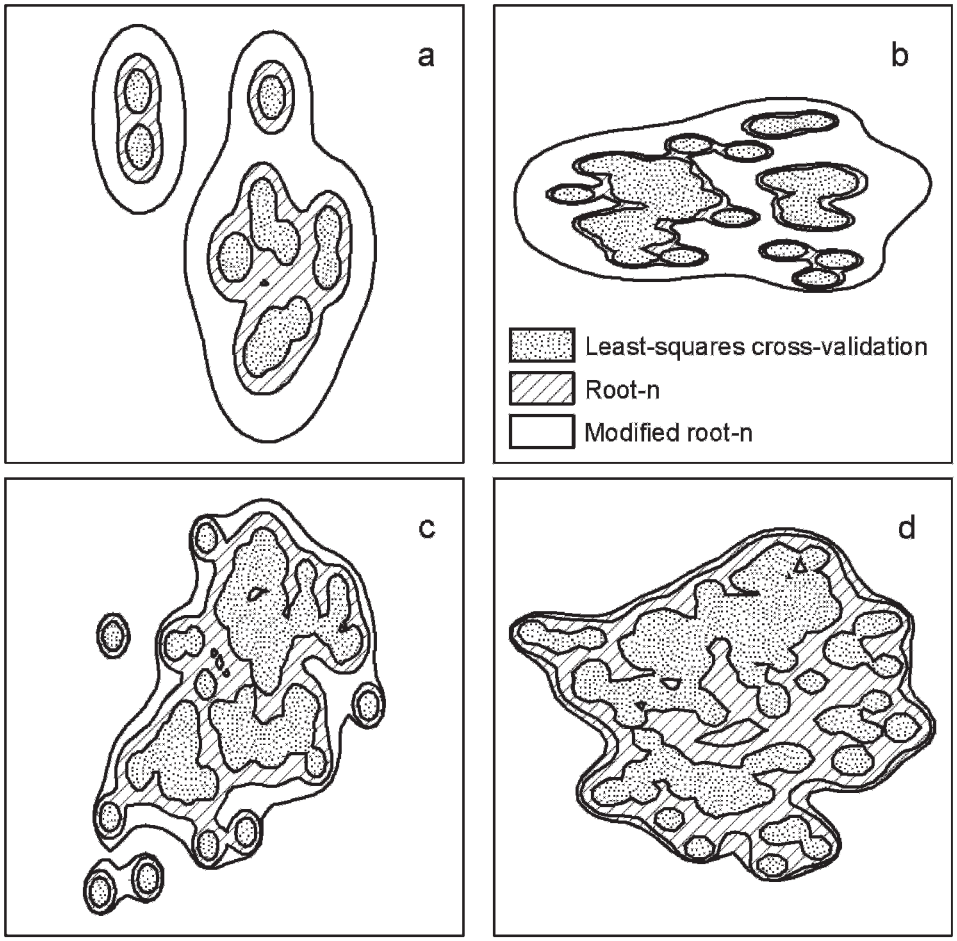

Figure 4. Examples of red wolf home ranges (95% isopleths) generated using the 3 bandwidth estimators. a) $n = 17$, $\dot{h}_{LSCV} = 0.11$, $\dot{h}_n = 0.22$, $\dot{h}^* = 0.47$. b) $n = 58$, $\dot{h}_{LSCV} = 0.13$, $\dot{h}_n = 0.17$, $\dot{h}^* = 0.57$. c) $n = 100$, $\dot{h}_{LSCV} = 0.09$, $\dot{h}_n = 0.20$, $\dot{h}^* = 0.30$. d) $n = 145$, $\dot{h}_{LSCV} = 0.07$, $\dot{h}_n = 0.17$, $\dot{h}^* = 0.22$.

root-n estimators generated UD estimates that were more accurate, on average, and also more precise. For individual samples, root-n estimators also were more likely to generate the best UD and home-range estimates at low UD complexity. Finally, the root-n estimators were less likely than LSCV to underestimate home-range size at high sample size and complexity.

Presence of repeat locations had a considerable negative effect on LSCV performance. Root-n estimators were more likely to generate the most accurate UD estimates with as few as 5–10% repeat locations, regardless of data sample size or UD complexity. Furthermore, as proportion of repeats increased, LSCV tended to generate UDs that drastically underestimated home-range size and had extremely poor fit to true UDs. Conversely, the performance of neither root-n estimator was substantially affected by proportion of repeats. Finally, we note that repeat locations were not true repeats, because we applied a random jitter to the data to prevent LSCV failure. Thus, such work-arounds do not eliminate poor performance in LSCV, although true repeats are likely to generate even worse performance.

The superiority of the root-n methods also was clear with real data. Root-n estimators nearly always generated UD estimates from subsampled hawk data that were a better fit to the original data than did LSCV. Similarly, LSCV often generated red wolf home-range estimates that were disjunct and comprised of many polygons and lacunae, whereas the root-n estimators generated more realistic-looking home ranges. Differences in the relative performances of LSCV and root-n estimators between our basic simulations and our repeat-point simulations or real data are due to differences in characteristics of the data. Our basic simulated data did not have discretization error, repeat locations, or clumping in the data. Thus, only minimal high-frequency noise existed, and consequently root-n estimators could not improve on LSCV for estimates lacking such constraints. Least-squares cross-validation may work well for data sets (real or simulated) that have little noise, especially those with small sample sizes or if researchers are only interested in the home range instead of the full UD (Gitzen et al. 2006); LSCV is less likely to over-smooth data in such situations. However, UD estimates generated using root-n generally had higher accuracy and precision on average even in our basic simulations. Furthermore, the performance of root-n was clearly superior with our repeat-point simulations and real animal location data, even when generating home ranges with small sample sizes (e.g., Fig. 4). Note that these data probably represent a worst-case scenario for the

LSCV estimator because of extreme clumping in the data in the form of repeats (Gitzen et al. 2006). Indeed, preliminary analyses indicated that LSCV and root-n estimators tend to perform similarly with the various clumped distributions used by Gitzen et al. (2006; T. Steury, Washington University, unpublished data), which are characterized by only moderate levels of clumping and few or no repeat locations. Again, such moderate clumping may not represent the kind of high-frequency noise that generates the observed differences in performance between LSCV and root-n estimators. However, we note that such noise may be characteristic of many animal location data sets where behavior promotes repeat locations. Thus, in many cases the original root-n will likely generate better UD estimates than LSCV, and particularly noisy data should be analyzed using the modified root-n.

Although LSCV never failed in our 12,000 simulated datasets, LSCV did not generate a bandwidth estimate in >20% of our red wolf samples; similar failures would have occurred in some cases with our repeat-point simulations and subsampled hawk data had we not imposed a random jitter (T. Roth, Indiana State University, unpublished data; T. Steury, unpublished data). Such problems often are not observed when data are analyzed with commercially available spatial analysis software because these packages typically force an arbitrary lower bound on LSCV estimates, often without informing software users. Users of such software should be aware that estimates at these boundaries are not true LSCV estimates. In such cases, few viable alternatives exist for accurate UD estimation, because other bandwidth estimators either perform poorly (e.g., reference bandwidth; Worton 1995, Seaman and Powell 1996) or suffer similar failures as LSCV (e.g., likelihood cross-validation, Silverman 1986; biased cross-validation, T. Steury, unpublished data). Although other, second-generation alternatives have recently been developed, they have yet to receive widespread use and in comparisons with LSCV have garnered equivocal support (Loader 1999, Gitzen et al. 2006). Thus, at the very least the root-n estimators we evaluated may serve as viable alternatives to LSCV in cases where the latter estimator fails.

Despite the general superiority of the root-n estimators over LSCV with noisy data, we note all 3 estimators performed poorly with very large data sets, such as the full hawk data sets. Although root-n estimates of h generally were double LSCV estimates with these data, root-n estimates often were more than an order of magnitude less than the value we would subjectively choose. Hawk home ranges generated using LSCV or root-n estimates were so under-smoothed that home-range estimates were polka-dots that barely covered the data. Analysis of a large data set collected from Global Positioning System transmitters deployed on wolves in Algonquin Provincial Park, Ontario, Canada (B. Patterson, Ontario Ministry of Natural Resources, unpublished data), confirmed that the problem was not exclusive to hawks. We speculate that the poor performance in such situations is not a limitation of the root-n estimator to handle the increased high-frequency noise of such large data sets, but a limitation of the fixed kernel to work adequately with both intensively used core areas (i.e., high peaks) and large home ranges (i.e., long distribution tails). Thus, further evaluation of the performance of adaptive kernels with the various bandwidth estimators, especially with large data sets, is warranted. In general, the increased ability to acquire such large samples of location data through new radiotelemetry technologies means that there is a substantial and growing need for bandwidth estimators, or other non–kernel-based methods (e.g., Getz and Wilmers 2004), that can provide unbiased and precise UD estimates from large samples of noisy data.

MANAGEMENT IMPLICATIONS

When using the kernel density method to estimate animal home ranges or UDs, choice of bandwidth estimator must be made carefully. The LSCV bandwidth estimator has been shown to perform well in many situations. However, as we have shown herein, root-n bandwidth estimators often perform better than LSCV, especially in situations where data sets have repeat locations or otherwise extreme clumping. In contrast, LSCV may be superior where the primary interest is in generating animal home ranges (rather than the UD) and data sets are small with limited clumping of locations. We note, however, that as in previous studies we found that no one bandwidth estimator was superior in all situations and that for any individual sample, considerable uncertainty exists as to which bandwidth estimator will return the best results (Gitzen et al. 2006). Thus, use of kernel density estimates for UD and home-range generation requires a careful evaluation of estimator performance on a case-by-case basis. Furthermore, selection of a bandwidth estimator should not preclude an investigator from testing other estimators to evaluate the effect of estimator choice on home range and UD estimates.

ACKNOWLEDGMENTS

We thank B. Kendall, R. Powell, G. White, and anonymous reviewers for comments that improved the manuscript. We are grateful to all individuals who collected red wolf telemetry data over the past 20 years, as well as those who assisted with tracking hawks. Funding for this study was provided by the United States Fish and Wildlife Service (D. L. Murray), the National Science Foundation (Grant IBN-0130758 to S. L. Lima and Grant DMS-0501079 to J. E. McCarthy), Indiana Academy of Sciences (T. C. Roth), Indiana State University School of Graduate Studies (T. C. Roth), and the Canada Research Chairs program (D. L. Murray). This work was made possible in part by a grant of high-performance computing resources and technical support from the Alabama Supercomputer Authority.

LITERATURE CITED

Amstrup, S. C., T. L. McDonald, and G. M. Durner. 2004. Using satellite radiotelemetry data to delineate and manage wildlife populations. Wildlife Society Bulletin 32:661–679.

Blundell, G. M., J. A. K. Maier, and E. M. Debevec. 2001. Linear home ranges: effects of smoothing, sample size, and autocorrelation on kernel estimates. Ecological Monographs 71:469–489.
Chiu, S. T. 1991. The effect of discretization error on bandwidth selection for kernel density estimation. Biometrika 78:436–441.
Chiu, S. T. 1992. An automatic bandwidth selector for kernel density estimation. Biometrika 79:771–782.
Getz, W. M., S. Fortmann-Roe, P. C. Cross, A. J. Lyons, S. J. Ryan, and C. C. Wilmers. 2007. LoCoH: nonparameteric kernel methods for constructing home ranges and utilization distributions. PLOS One 2:e207.
Getz, W. M., and C. C. Wilmers. 2004. A local nearest-neighbor convex-hull construction of home ranges and utilization distributions. Ecography 27:489–505.
Gitzen, R. A., and J. J. Millspaugh. 2003. Comparison of least-squares cross-validation bandwidth options for kernel home range estimation. Wildlife Society Bulletin 31:823–831.
Gitzen, R. A., J. J. Millspaugh, and B. J. Kernohan. 2006. Bandwidth selection for fixed-kernel analysis of animal utilization distributions. Journal of Wildlife Management 70:1334–1344.
Hemson, G., P. Johnson, A. South, R. Kenward, R. Ripley, and D. McDonald. 2005. Are kernels the mustard? Data from global positioning system (GPS) collars suggests problems with kernel home-range analyses with least-squares cross-validation. Journal of Animal Ecology 74:455–463.
Kernohan, B.J., R. A. Gitzen, and J. J. Millspaugh. 2001. Analysis of animal space use and movements. Pages 125–166 *in* J. J. Millspaugh and J. M. Marzluff, editors. Radio tracking and animal populations. Academic Press, San Diego, California, USA.
Loader, C. R. 1999. Bandwidth selection: classical or plug-in? Annals of Statistics 27:415–438.
Mack, Y. P., and M. Rosenblatt. 1979. Multivariate k-nearest neighbor density estimates. Journal of Multivariate Analysis 9:1–15.
Park, B. U., and J. S. Marron. 1990. Comparison of data-driven bandwidth selectors. Journal of the American Statistical Association 85:66–72.
Powell, R. A. 2000 Animal home ranges and territories and home range estimators. Pages 65–110 *in* L. Boitani and T. K. Fuller, editors. Research techniques in animal ecology: controversies and consequences. Columbia University Press, New York, New York, USA.
Roth, T. C., and S. L. Lima. 2007. Use of prey hotspots by an avian predator: purposeful unpredictability? American Naturalist 169:264–273.
Sain, S. R., K. A. Baggerly, and D. W. Scott. 1994. Cross-validation of multivariate densities. Journal of the American Statistical Association 89:807–817.
Seaman, D. E., J. J. Millspaugh, B. J. Kernohan, G. C. Brundige, K. J. Raedeke, and R. A. Gitzen. 1999. Effects of samples size on kernel home range estimates. Journal of Wildlife Management 63:739–747.
Seaman, D. E., and R. A. Powell. 1996. An evaluation of the accuracy of kernel density estimators for home range analysis. Ecology 77:2075–2085.
Silverman, B. W. 1986. Density estimation for statistics and data analysis. Chapman and Hall, London, United Kingdom.
Wand, M. P., and M. C. Jones. 1995. Kernel smoothing. Chapman and Hall, London, United Kingdom.
Worton, B. J. 1989. Kernel methods for estimating the utilization distribution in home-range studies. Ecology 70:164–168.
Worton, B. J. 1995. Using Monte-Carlo simulation to evaluate kernel-based home-range estimators. Journal of Wildlife Management 59:794–800.
Wu, T. J., and M. H. Tsai. 2004. Root n bandwidths selectors in multivariate kernel density estimation. Probability Theory and Related Fields 129:537–558.

APPENDIX: ROOT-N FUNCTIONS FOR GAUSSIAN KERNELS

The modified score function proposed by Wu and Tsai (2004:542, eq 15), written here in the form for 2 dimensions, is as follows:

$$\hat{M}(h,T)=\frac{1}{(2\pi)^2}\int_{-T}^{T}\int_{-T}^{T}\left|\tilde{\phi}(t,s)\right|^2[1-\phi_{\mathbf{K}}(th,sh)]^2\,dtds - \frac{1}{(2\pi)^2 n}\int_{-T}^{T}\int_{-T}^{T}\left|\tilde{\phi}(t,s)\right|^2[\phi_{\mathbf{K}}(th,sh)]^2\,dtds + \frac{1}{(2\pi)^2 n}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[\phi_{\mathbf{K}}(th,sh)]^2\,dtds$$

where $\tilde{\phi}(t,s)$ is the Fourier transform of the data or the sample characteristic function, $\phi_{\mathbf{K}}(th,sh)$ is the Fourier transform of the 2-dimensional kernel $\mathbf{K}(x/h, y/h)$, and all other parameters are as defined in the text. The bandwidth estimate is the value of h that minimizes this score function. In this study, we used the Gaussian function for the kernel. Thus,

$$\tilde{\phi}(t,s)=\frac{1}{n}\sum_{j=1}^{n}e^{i\left(tX_j+sY_j\right)},$$

$$\mathbf{K}(x/h,y/h)=\frac{e^{\left(-\frac{x^2+y^2}{2h^2}\right)}}{2\pi h^2},$$

and

$$\phi_{\mathrm{K}}(th,sh)=e^{\left[-\frac{1}{2}\left(t^2h^2+s^2h^2\right)\right]}.$$

Thus, by substitution and simplification, the score function becomes

$$\hat{M}(h,T)=\frac{1}{(2\pi)^2}\int_{-T}^{T}\int_{-T}^{T}\frac{1}{n^2}\sum_{j,k=1}^{n}e^{i\left[t\left(X_j-X_k\right)+s\left(Y_j-Y_k\right)\right]}\left[1-e^{-\frac{1}{2}\left(t^2h^2+s^2h^2\right)}\right]^2dtds - \frac{1}{(2\pi)^2 n}\int_{-T}^{T}\int_{-T}^{T}\frac{1}{n^2}\sum_{j,k=1}^{n}e^{i\left[t\left(X_j-X_k\right)+s\left(Y_j-Y_k\right)\right]}\left[e^{-\frac{1}{2}\left(t^2h^2+s^2h^2\right)}\right]^2dtds+\frac{1}{4\pi nh^2}.$$

To simplify this function further, we first rearrange to

$$\hat{M}(h,T)=\frac{1}{(2\pi n)^2}\sum_{j,k=1}^{n}\int_{-T}^{T}\int_{-T}^{T}e^{i\left[t\left(X_j-X_k\right)+s\left(Y_j-Y_k\right)\right]}dtds - \frac{1}{(2\pi n)^2}\sum_{j,k=1}^{n}\int_{-T}^{T}\int_{-T}^{T}e^{i\left[t\left(X_j-X_k\right)+s\left(Y_j-Y_k\right)\right]}\left[2e^{-\frac{1}{2}\left(t^2h^2+s^2h^2\right)}\right]dtds + \frac{n-1}{n^3(2\pi)^2}\sum_{j,k=1}^{n}\int_{-T}^{T}\int_{-T}^{T}e^{i\left[t\left(X_j-X_k\right)+s\left(Y_j-Y_k\right)\right]}\left[e^{-\left(t^2h^2+s^2h^2\right)}\right]dtds+\frac{1}{4\pi nh^2}.$$

Because we seek to minimize the equation with respect to h, we can ignore the first term in the function. As for the second and third terms, we note that

$$e^{fi} = \cos(f) + i\sin(f),$$

and that

$$\int_{-T}^{T} e^{-t^2 g}\left[i\sin(tf)\right]dt = 0,$$

where f and g are any functions not involving t. Thus, we can rewrite the score function in the following form:

$$\begin{aligned}\hat{M}(h,T) = &-\frac{1}{2(\pi n)^2}\sum_{j,k=1}^{n}\int_{-T}^{T}\int_{-T}^{T} e^{-\frac{1}{2}\left(t^2h^2+s^2h^2\right)} \\ &\left\{\cos\left[t\left(X_j-X_k\right)\right]\right\}\left\{\cos\left[s\left(Y_j-Y_k\right)\right]\right\}dtds \\ &+\frac{n-1}{n^3(2\pi)^2}\sum_{j,k=1}^{n}\int_{-T}^{T}\int_{-T}^{T} e^{-\left(t^2h^2+s^2h^2\right)} \\ &\left\{\cos\left[t\left(X_j-X_k\right)\right]\right\}\left\{\cos\left[s\left(Y_j-Y_k\right)\right]\right\}dtds \\ &+\frac{1}{4\pi nh^2}.\end{aligned}$$

The above equation is relatively simple, in that it does not include any imaginary terms. However, the double integral and summation terms make it computationally expensive. An alternative is to use the error function (erf), a built-in function in many programming languages, where

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x} e^{-t^2}dt.$$

We can thus write the score function as follows:

$$\begin{aligned}\hat{M}(h,T) = &\frac{1}{4\pi nh^2} \\ &-\frac{1}{\pi(2nh)^2}\sum_{j,k=1}^{n}\left\{\begin{array}{l} e^{-\frac{\left(X_j-X_k\right)^2+\left(Y_j-Y_k\right)^2}{2h^2}} \\ \times\left[\mathrm{erf}\left(\frac{hT}{\sqrt{2}}-\frac{i\left(X_j-X_k\right)}{h\sqrt{2}}\right)-\mathrm{erf}\left(\frac{-hT}{\sqrt{2}}-\frac{i\left(X_j-X_k\right)}{h\sqrt{2}}\right)\right] \\ \times\left[\mathrm{erf}\left(\frac{hT}{\sqrt{2}}-\frac{i\left(Y_j-Y_k\right)}{h\sqrt{2}}\right)-\mathrm{erf}\left(\frac{-hT}{\sqrt{2}}-\frac{i\left(Y_j-Y_k\right)}{h\sqrt{2}}\right)\right]\end{array}\right\} \\ &+\frac{n-1}{\pi n^3(4h)^2}\sum_{j,k=1}^{n}\left\{\begin{array}{l} e^{-\frac{\left(X_j-X_k\right)^2+\left(Y_j-Y_k\right)^2}{4h^2}} \\ \times\left[\mathrm{erf}\left(hT-\frac{i\left(X_j-X_k\right)}{2h}\right)-\mathrm{erf}\left(-hT-\frac{i\left(X_j-X_k\right)}{2h}\right)\right] \\ \times\left[\mathrm{erf}\left(hT-\frac{i\left(Y_j-Y_k\right)}{2h}\right)-\mathrm{erf}\left(-hT-\frac{i\left(Y_j-Y_k\right)}{2h}\right)\right]\end{array}\right\}.\end{aligned}$$

However, we note that this particular function can perform poorly when the h that minimizes the score function is small (T. Steury, unpublished data). As a third alternative, one could use the fast-Fourier transform (see Chiu [1992] and Wu and Tsai [2004] for such methods).

As defined by Wu and Tsai (2004:543, eq 16), we estimate the variable T in the above equations by minimizing the following function, again written here in the form for 2 dimensions, with respect to T:

$$CV^{\infty}(T) = \frac{8T^2}{n+1} - \int_{-T}^{T}\int_{-T}^{T}\left|\tilde{\phi}(t,s)\right|^2 dtds.$$

Standardizing the dispersion in X and Y values not only allows for us to calculate a single bandwidth (see text), but also allows for one value of T (Wu and Tsai 2004). In the specific case we described herein, the function becomes

$$CV^{\infty}(T) = \frac{8T^2}{n+1} - \frac{4}{n^2}\sum_{j,k=1}^{n}\frac{\sin\left[T\left(X_j-X_k\right)\right]}{\left(X_j-X_k\right)}\frac{\sin\left[T\left(Y_j-Y_k\right)\right]}{\left(Y_j-Y_k\right)}.$$

Notice that if $X_j = X_k$ or $Y_j = Y_k$, ≥ 1 of the denominators in the above equation will equal zero. For these cases, note that

$$\lim_{f\to 0}\frac{\sin\left[T\times f\right]}{f} = T,$$

where f is any function not involving T.

Associate Editor: White.