KERNELHR: a program for estimating animal home ranges

D. Erran Seaman, Brad Griffith, and Roger A. Powell

Abstract Kernel methods are state of the art for estimating animal home-range area and utilization distribution (UD). The KERNELHR program was developed to provide researchers and managers a tool to implement this extremely flexible set of methods with many variants. KERNELHR runs interactively or from the command line on any personal computer (PC) running DOS. KERNELHR provides output of fixed and adaptive kernel home-range estimates, as well as density values in a format suitable for in-depth statistical and spatial analyses. An additional package of programs creates contour files for plotting in geographic information systems (GIS) and estimates core areas of ranges.

Key words core area, home range, kernel method, nonparametric density estimation, utilization distribution

KERNEL estimators

The Kernel Home Range program KERNELHR is a tool for estimating the areal extent of home ranges and the intensity of use of different parts of the ranges. Although it was developed for animal home-range analysis, it is also appropriate for other applications, such as the range of a population or species, or for visualizing the distributional patterns of data other than the locations of organisms (e.g., distribution of data on 2 principal components axes).

KERNELHR uses nonparametric kernel smoothing methods for density estimation (Silverman 1986). These methods have a long history in statistical theory but have only recently gained popularity in applications as home-range estimators. Kernels are flexible estimators that can fit nonconvex, multimodal, irregularly shaped distributions (Fig. 1). The KERNELHR program gives great flexibility and control over the method and amount of smoothing, and over the content and form of the output.

Kernel methods produce a probability density estimate of a distribution based on a sample of points (for data of any number of d dimensions). In practice, kernels have been used primarily for data of only 1 or 2 dimensions, and we confine our discussion to 2-dimensional (X,Y) data. The probability density estimate that is produced by kernel methods may be directly interpreted as a utilization distribution (UD; Van Winkle 1975).

Worton (1989) introduced kernel methods as home-range estimators. Since then, several computer programs have been developed to implement kernel estimators, and they have been gaining popularity (Appendix A).

Kernel methods have several desirable qualities for home-range estimation: (1) they are nonparametric, and therefore have the potential to accurately estimate densities of any shape, provided that the level of smoothing is selected appropriately; (2) they produce a density estimate directly; and (3) they are not influenced by grid size or placement (Silverman 1986). Worton (1987) reviewed home-range estimators, including kernel methods. Worton (1995) and Seaman and Powell (1996) evaluated the accuracy of kernel methods for home-range estimation.

Kernel methods are statistical techniques for estimating the density of a distribution at any point. The density estimate is derived from the proximity of observations (sample points) to each evaluation point. Evaluation points may be the observations themselves,

Address for D. Erran Seaman: U.S. Geological Survey, Biological Resources Division, Forest and Rangeland Ecosystem Science Center, Olympic Field Station, 600 East Park Avenue, Port Angeles, WA 98362-6798, USA. Address for Brad Griffith: Alaska Cooperative Fish and Wildlife Research Unit, 209 Irving I Building, University of Alaska, Fairbanks, Fairbanks, AK 99775-7020, USA. Address for Roger A. Powell: Department of Zoology, North Carolina State University, Raleigh, NC 27695, USA.

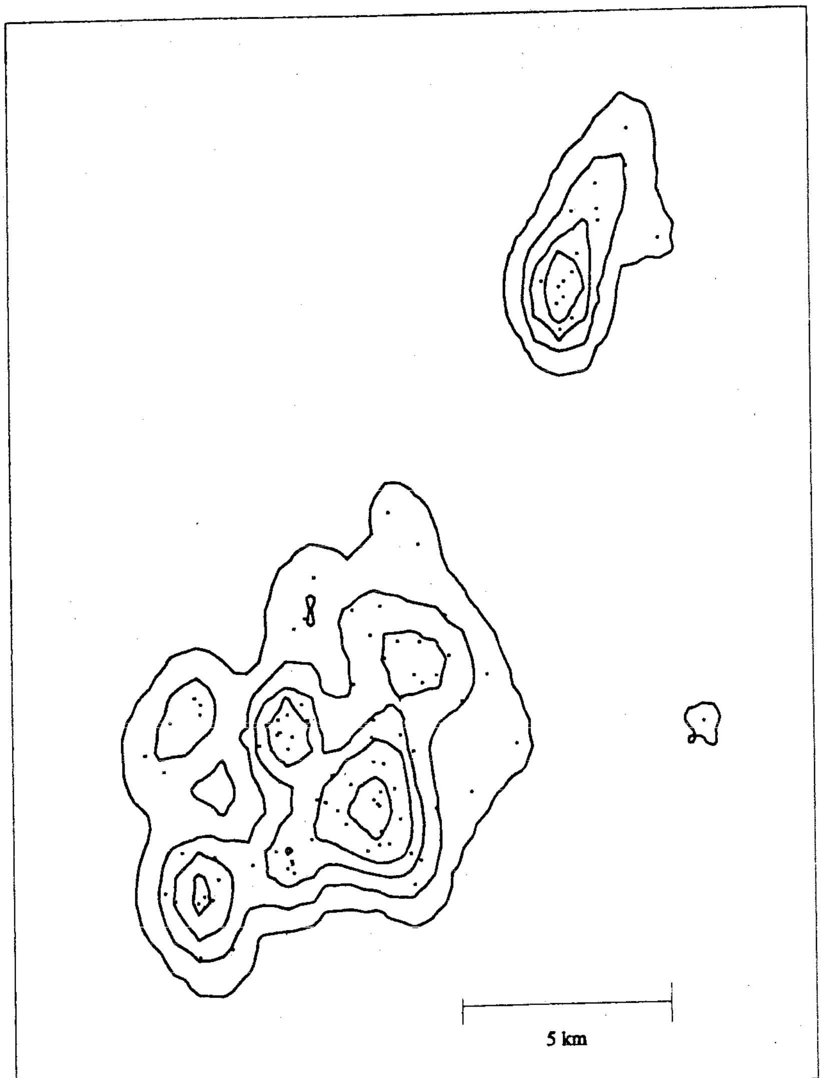


Fig. 1. Estimated home range of a black bear in North Carolina, illustrating lack of bivariate normality, as well as multiple peaks (centers of activity), disjunct sections, and nonconvexity. Contours represent 5, 27.5, 50, 72.5, and 95% of the utilization distribution (UD); dots indicate radiotelemetry locations. Data from Powell et al. (1997).

or a regular grid laid over the sample. The methods are described in more detail by Silverman (1986), Worton (1989), and Seaman and Powell (1996).

The methods impose a kernel over each observation. A familiar example of a kernel is the normal distribution, though kernels of other shapes exist and are often preferred over the normal for computational speed. The estimate of the density at any given evaluation point is essentially the average height of the kernels that overlap the evaluation point. The width (variance) of the kernels affects how much smoothing occurs and is referred to as the bandwidth, smoothing width, or smoothing parameter "h." Kernel estimates are highly sensitive to the bandwidth; thus, to get accurate results it is crucial to correctly select the value for this parameter.

A variety of methods are available for selecting the bandwidth. Two of these include: (1) the reference method (HREF; Worton 1995) which uses the theoretical optimum value, based on the assumption that the true distribution (from which the sample of observation points was selected) is normally distributed (Silverman 1986), and (2) least squares cross validation (LSCV), a jacknife method which uses an iterative approach to select the amount of smoothing that minimizes the estimated error for a given sample (Silverman 1986). LSCV makes no assumptions about the true distribution. Although HREF is frequently referred to as the "optimal" value, it is only optimal for bivariate normal distributions. Because home ranges are frequently far from normal (e.g., Fig. 1), the reference method tends to greatly overestimate home-range size, but LSCV produces a nearly unbiased estimate (Worton 1995, Seaman and Powell 1996). The amount of smoothing selected by LSCV will vary for different home-range shapes, but is frequently about 40% of the amount selected by HREF.

There are 2 major variants of kernel methods: the fixed kernel and the adaptive kernel. In the fixed kernel method, a single smoothing width is used on all the observations in the sample. In the adaptive kernel method, local adjustments are applied to the width of individual kernels. Observations in areas of high density get less smoothing (tighter fit), and observations in areas of low density get more smoothing (looser fit). Although the adaptive kernel was expected to produce better results, the fixed kernel generally produces estimates of home-range size and contours with lower bias than the adaptive kernel in simulation studies (Worton 1995, Seaman and Powell 1996).

A significant difficulty with kernel methods is their high sensitivity to the bandwidth. For many datasets, using LSCV to select the smoothing width produces good results. However, LSCV does not work acceptably for datasets with multiple observations at identical locations (e.g., trap grid data, animals in dens or nests; Silverman 1986, Tufto et al. 1996).

KERNELHR features and capabilities

KERNELHR was developed as a tool for detailed analysis of distributional patterns; it produces far more than just a summary home-range size estimate. It is designed to be used on 2-dimensional data by anyone with a basic knowledge of DOS. The user must create input files (ASCII) in the required format (3 columns: animal ID, *X*, *Y*), with *X*- and *Y*-coordinates in a decimal system (e.g., UTM). Although many features can be optionally controlled by the user, default selections have been chosen to generally provide the best results. The user is only required to specify the names of input and output files and the scale (no. of m/unit) of the data.

We summarized and compared specific features of the KERNELHR program to other kernel programs (Table 1). Sources for the kernel programs are listed in Appendix A. Other reviews have compared home-range estimation methods (Worton 1987, 1995; Seaman and Powell 1996) or home-range programs but without detailed review of kernel capabilities (Larkin et al. 1994, Lawson and Rodgers 1997). A detailed list and discussion of KERNELHR's features and their significance follows.

Output of results

Output is available for virtually everything calculated by the program, for both the fixed and adaptive kernel estimators.

Fixed and adaptive range-size estimate. KERNELHR gives the area of any requested percentage (e.g., 95%) of the estimated UD. This is a more accurate method than "point percentages" (PP) for gauging the home-range size (J. Baldwin, USFS PSW Res. Lab., Albany, Calif., pers. commun.).

Fixed and adaptive utilization distribution values at grid intersection points. This is perhaps the most useful output from the program, allowing detailed analyses of intensity of use of different areas, different habitats, and 3-dimensional overlap (i.e., space and time overlap) between individuals (or overlap between species, where location data are observations of individuals of different species). These data provide a basis for home-range estimation that is not sensitive to grid size, within the range of grid sizes that gives valid estimates.

Fixed- and adaptive-density estimates at the observation points. An alternative to grid cell output, densities and values of lambda (the local smoothing factor for the adaptive kernel) are available at the observation locations.

Fixed and adaptive utilization distribution values for ≤11 requested contour intervals. The contour values include the stated percentage of the volume of the UD; they are not point percentage values. These values are available for use when plotting contours with other software.

Table 1. Comparison of selected features of kernel home-range programs. These programs also perform other functions; this is not a complete list of their capabilities. See Appendix A for program sources.

	KERNELHR	CALHOME	RANGESV	GRID	TRACKER	KERNEL
Methodology						
Implements automatic LSCV[a]	+[b]	–[c]	+	–	–	+
Maximum no. of locations	≈ 4,000	500	>6,000	5,000[d]	3,000	3,000
Maximum grid size, cells	≈ 235 x 235	50 x 50	50 x 50	360 x 360	NA[e]	70 x 70
UD or point % (PP) estimated	UD	PP	UD, PP	UD	PP	UD
Output						
LSCV score	+	+	+	–	–	–
Vol. of UD[f] estimate	+	–	+	+	–	+
Area (fixed, adaptive)	F, A	A	F, A	F	A	F, A
Density[g]	O, G, F, A	–	G, F, A	O, G, F	–	G, F, A
Graphical	–	+	+	+	+	–

[a] Least squares cross validation.
[b] Feature is present.
[c] Feature is absent.
[d] With optional module, 11,000 observations can be used.
[e] Information not available.
[f] Utilization distribution.
[g] Density available at observations (O), at grid points (G), from fixed (F) or adaptive kernels (A).

Control of bandwidth

Several methods are available for choosing the bandwidth in KERNELHR:

Automatic LSCV. LSCV is performed automatically, using a "smart" minimization routine. This is performed on normalized data, making it suitable even for nonlocation data with different units in the 2 dimensions.

Reference method. This is often called the "optimal bandwidth," a name that is misleading because it is only optimal for bivariate normal data. It is far from optimal for most actual location data, and is not recommended for home-range analysis.

Manual selection. The user can input the bandwidth, either as absolute units, or as a percentage of the reference method.

Program capabilities

Large number of observations. KERNELHR does not limit the number of observations per home range. It uses all conventional memory available and can handle about 4,000 observations with 640K of RAM (depending on the system configuration and other software running in a multitasking environment).

Multiple ranges in a single input file. KERNELHR can accept multiple animals per file. You do not need to create separate files for each individual.

Independent bandwidths in 2 dimensions. If a dataset has less spread in 1 dimension than in the other, that dimension will get less smoothing and, thus, will not be distorted by oversmoothing (Fig. 2).

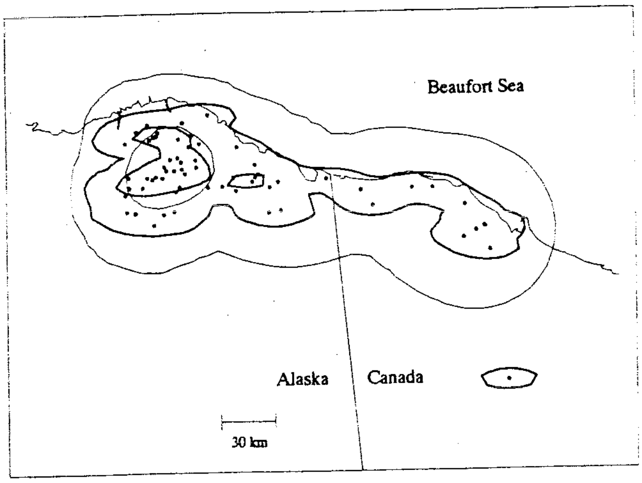


Fig. 2. Distribution of caribou calving locations for multiple animals on the north slope of Alaska. This illustrates the use of population data and greater location variance along the east-west axis than the north-south axis. Plotted contours encompass core areas and 95% of the utilization distribution (UD). Heavy lines illustrate the fixed kernel with least squares cross validation (LSCV) smoothing, and different amounts of smoothing in the 2 dimensions. This choice yielded a tight fit to the points for the core contour and included little nonhabitat (the Beaufort Sea) in the 95% contour estimate. Thin lines illustrate the adaptive kernel with reference method (HREF) smoothing and equal smoothing in both dimensions. This choice produced distortion (oversmoothing) in the north-south dimension and included much nonhabitat in the 95% contour estimates. (B. Griffith, unpubl. data).

Output of additional information

- Bandwidth.
- Method of selecting the bandwidth.
- Ratio of LSCV:HREF, when LSCV is chosen.
- Minimum score from LSCV, when LSCV is chosen.
- Grid spacing and number of cells in each dimension.
- Volume of the estimate. This is a diagnostic tool for the estimate. Since the estimated UD should, by definition, integrate to 1.00, a value far from 1.00 indicates a poor estimate. This most frequently results from using an overly coarse grid, but other factors can also cause occasional inaccurate estimates. Without this diagnostic, it can be difficult to determine whether estimates are reliable.

Optional control of output

Units. The home-range size estimate may be reported in the appropriate scale (but only metric units are available: km^2, ha, m^2).

Grid spacing. It can be controlled automatically or manually. KERNELHR does not limit the grid size, therefore a fine grid and high accuracy can be obtained even for animals or populations with very large ranges. The program uses all conventional memory available and can handle a grid of about 235 x 235 cells with 640K of RAM (depending on the system configuration and other software running under a multitasking operating system).

- Automatic grid size: the program chooses a fine grid (minimum of 45 x 45 cells), sized to the dimensions of each individual home range. Coarser grids frequently result in poor estimates (i.e., the resulting vol. of the estimate deviates from 1.00). The grid is automatically dimensioned to fit the data and will extend to a rectangle (e.g., 45 x 82 cells) for an elongated range.
- Manual grid size: the user specifies any grid spacing, and all ranges will be on the same grid spacing with the same origin. This makes the ranges comparable on a point-by-point basis (e.g., for 3-dimensional overlap analysis).

An independently developed contouring package is available (from the second author) which processes KERNELHR grid output into vector polygons for plotting contours in a geographic information system (GIS), identifies "core" areas based on greater than expected density (compared to a bivariate uniform distribution, c.f. Samuel et al. 1985), calculates areas of selected contour levels, and assigns observations to contour levels. An additional utility converts files of latitude-longitude locations to UTM coordinates for the western hemisphere.

Using KERNELHR

KERNELHR is a simple DOS program; it does not have menus nor does it use a mouse. It can be run either from DOS or in a DOS window within a multitasking environment (e.g., Windows 3.x, Windows 95, WindowsNT 3.x, 4.x). KERNELHR can be run interactively or from the command line. When used interactively, it presents the user with a series of prompts. When used from the command line, the user must select options in the form of command line switches. The command line option facilitates processing a large number of files by allowing the program to be operated from a batch file.

All input and output is in the form of ASCII text files; no graphical output is available. No other software is needed if the only interest is in home-range size estimates. However, further analysis of the density patterns must be performed by the user, with other software such as statistical, graphics, or GIS packages. KERNELHR is distributed with a documentation file, a sample data file, and a sample batch file.

KERNELHR will run on any DOS computer; a 386 processor or higher is recommended, as is a math coprocessor. The amount of hard-drive space required can be quite minimal, but will depend on how much output is to be stored. The program itself requires only 46K of hard-disk space, but grid-output files can require 0.2-1.0 MB or more per home range (depending on grid size). Since most of the internal data storage takes place in dynamic memory, very large datasets (thousands of observations per home range) will require 640K of conventional memory, but most home-range analyses should be able to run on computers with ≤512K RAM.

KERNELHR is available free of charge by sending a blank, DOS-formatted, 3.5" diskette to the first author, or from the internet at http://www.im.nbs.gov/tws/cse.html.

Acknowledgments. We are grateful to D. Nychka for encouraging the use of the kernel method as a home-range estimator, and to B. Silverman for invaluable assistance in understanding the method. J. Greslin programmed the contouring package. We thank J. Kie, B. Ackerman, and K. Jenkins for reviewing earlier drafts of this manuscript.

Literature cited

Larkin, R. P., and D. Halkin. 1994. A review of software packages for estimating animal home ranges. Wildlife Society Bulletin 22:274-287.

Lawson, E. J. G., and A. R. Rodgers. 1997. Differences in home-range size in commonly used software programs. Wildlife Society Bulletin 25:721-729.

Powell, R. A., J. W. Zimmerman, and D. E. Seaman. 1997. Ecology and behaviour of North American black bears: home ranges, habitat, and social organization. Chapman & Hall, Limited, London, England.

Samuel, M. J., D. J. Pierce, and E. O. Garton. 1985. Identifying areas of concentrated use within the homerange. Journal of Animal Ecology 54:711-719.

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, Limited, London, England.

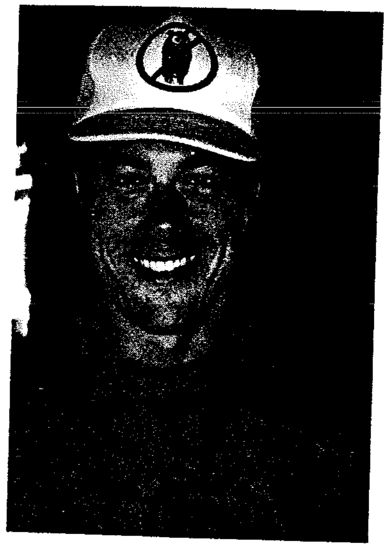
Tufto, J., R. Andersen, and J. Linnell. 1996. Habitat use and ecological correlates of home range size in a small cervid: the roe deer. Journal of Animal Ecology 65:715-724.

Van Winkle, W. 1975. Comparison of several probabilistic home-range models. Journal of Wildlife Management 39:118-123.

Worton, B. J. 1987. A review of models of home range for animal movement. Ecological Modelling 38:277-298.

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.



D. Erran Seaman (photo) is a research ecologist with the U.S. Geological Survey (USGS), Biological Resources Division, Olympic Field Station. He received a B.A. in linguistics from Cornell University and a Ph.D. in zoology from North Carolina State University. His research interests are in home-range estimation, population estimation, and monitoring population size and demography, particularly in pristine wilderness systems. ***Brad Griffith*** is the Assistant Leader of the Alaska Cooperative Fish and Wildlife Research Unit, Biological Resources Division, USGS. He received his B.A. in zoology from the University of Missouri, his M.S. in wildlife science from Oregon State University, and his Ph.D. in forestry, range, and wildlife science from the University of Idaho. He previously held positions with the U.S. Fish and Wildlfe Service (Coop.

Fish and Wildl. Res. Units and Alaska Sci. Cent.) and the U.S. Forest Service. His research interests include habitat and nutritional ecology, landscape processes, remote sensing, and stochastic population processes. ***Roger A. Powell*** is an associate professor of zoology and wildlife at North Carolina State University. He received his B.A. from Carleton College and his Ph.D. from the University of Chicago. His research is focused on the effects of limiting resources on social organization, mating patterns, sexual dimorphism in body size, choice of habitat, and choice of home range.

APPENDIX A

KERNELHR. Available from D. Erran Seaman, U.S. Geological Survey, Biological Resources Division, Olympic Field Station, 600 E. Park Ave., Port Angeles, WA 98362, Erran_Seaman@NPS.GOV, or from The Wildlife Society home page http://www.im.nbs.gov/tws/cse.html.

CALHOME. Available from John Kie, Starkey Project, Forestry and Range Sciences Lab, 1401 Gekeler Lane, La Grande, OR 97850.

RANGES V. Available from: http://www.nmw.ac.uk/ite/ranges.html.

GRID. Available from: Beat Naef-Daenzer, Swiss Ornithological Institute, CH-6204 SEMPACH, Switzerland, naefb@orninst.ch.

TRACKER. Available from: Radio Location Systems AB, P.O. Box 2131, S-141 02 Huddinge, SWEDEN.

KERNEL. Available from http://www.nina.no/jtu/kernel/.

These programs are also available on the Wildlife Telemetry Clearinghouse web site (http://www.uni-sb.de/philfak/fb6/fr66/tpw/telem/telem.htm), but the most recent versions generally will be available from the author of the individual program.

Associate Editor: Smith