Minimum convex polygons

Minimum Convex Polygons (MCP)
 or Convex Hulls

In the past, it used to be a method to estimate home range. However, nowadays if used, it should be only employed as a method to describe the spatial extension of relocation data.

Convex shape based on the outermost positions from an organism' relocation data.

Relocation data- Crab 10

Y coordinate

X coordinate

Convex polygon is a not self intersecting flat shape consisting of straight lines, containing all set of points in a set, and where the internal angles do not exceed 180 degrees.
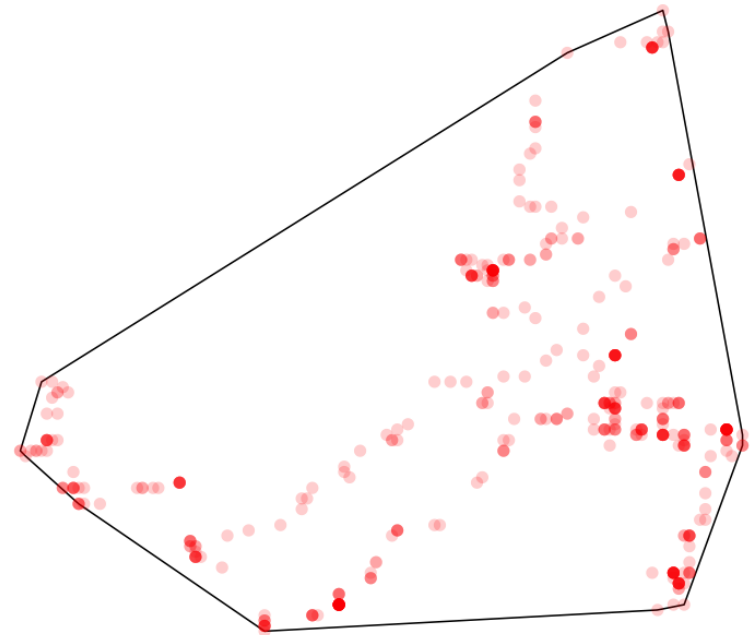
Minimum Convex Polygons (MCP)
or Convex Hulls

- One artifact is that the MCP area increase linearly as the number of sample relocations increase.

- Very sensitive to outliers.

- Bias to the spatial resolution of the data.

- Many ways to control these bias. However, these become cumbersome and are hard to justify in terms of ecological and biological basis.

- Other factors or bias affecting MCP are: underlying shape of species habitat, magnitude of errors in the relocation data (i.e. technique), frequency of sampling, and spatial distribution of samples.

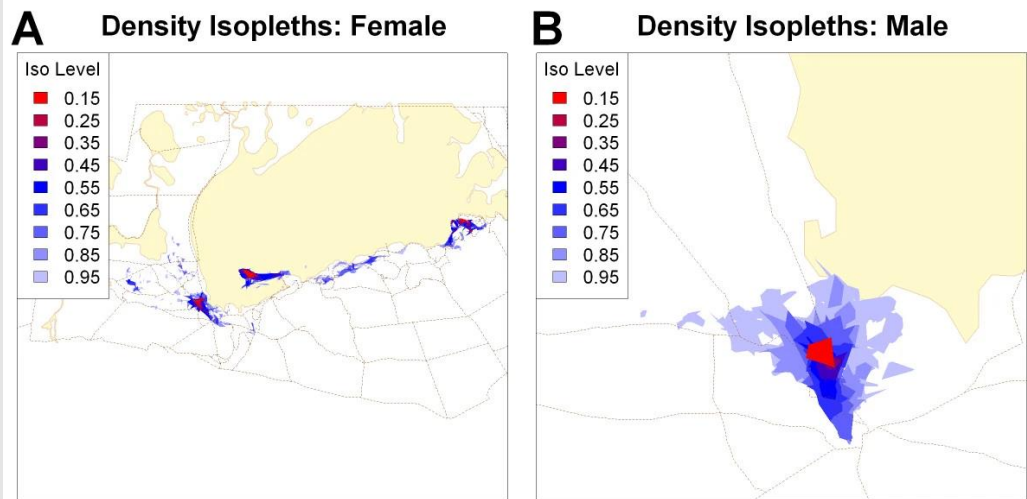- Borger et al 2006; Burgman & Fox 2003; Nilsen et al 2007; Burt 1943;

Relocation data- Crab 18

Question	Response
Give some background on how it works/what does it estimate	It calculates the minimum convex polygon enclosing all organism' relocations, or the minimum convex polygon enclosing a percentage of these relocations. Thus, outliers can be defined by the user as a percentage of the farthest away positions from the centre of the organism' relocation cloud. * The centre of the relocation cloud is calculated as the arithmetic mean from all organism' positions.
When/why would you use this and what data are appropriate?	MCP is highly criticized as a Home Range size estimator. Therefore, it should not be employed to estimate the area used by an animal. Not even for comparative purposes. It can be used as a way to describe the spatial extension of organisms' relocations (no ecological and biological meaning).
What are some of the strengths/weaknesses of this method?	Strength: Very simple to calculate, and seemingly to have a very intuitive interpretation. A lot of old/classic scientific literature had employed this method. It was and it is recommended by several organiziations (e.g. IUCN). Weaknesses: highly sensitive to a lot of sampling factors, and bias depending on data resolution. See previous slide.
What are the underlying assumptions	Assumes that a Home Range have a convex shape. Assumes that the space inside the polygon is used in an uniform fashion.
What packages/software can you use to create it?	AdehabitatHR, MOVE, AMT & RHR (Joo et al 2019)

Local Convex Hulls

Local Convex Hulls

Density Isopleths: Female / Density Isopleths: Male

Iso Level
- 0.15
- 0.25
- 0.35
- 0.45
- 0.55
- 0.65
- 0.75
- 0.85
- 0.95

- Non-parametric kernel method

- Generalises the MCP- applies MCPs to subsets of localized data in space -> local polygons or hulls

- How many do we join points together? Methods for nearest neighbours from reference point
  - K (a pre-set number)
  - R (all points within a fixed radius sphere of influence)
  - A (all points within adaptive sphere of influence so that distances of points to reference point sum to a value ≤ a)

  Convex hulls are joined to construct isopleths representing occupation distribution

More precise presentation of occupation as the sample of locations increases.
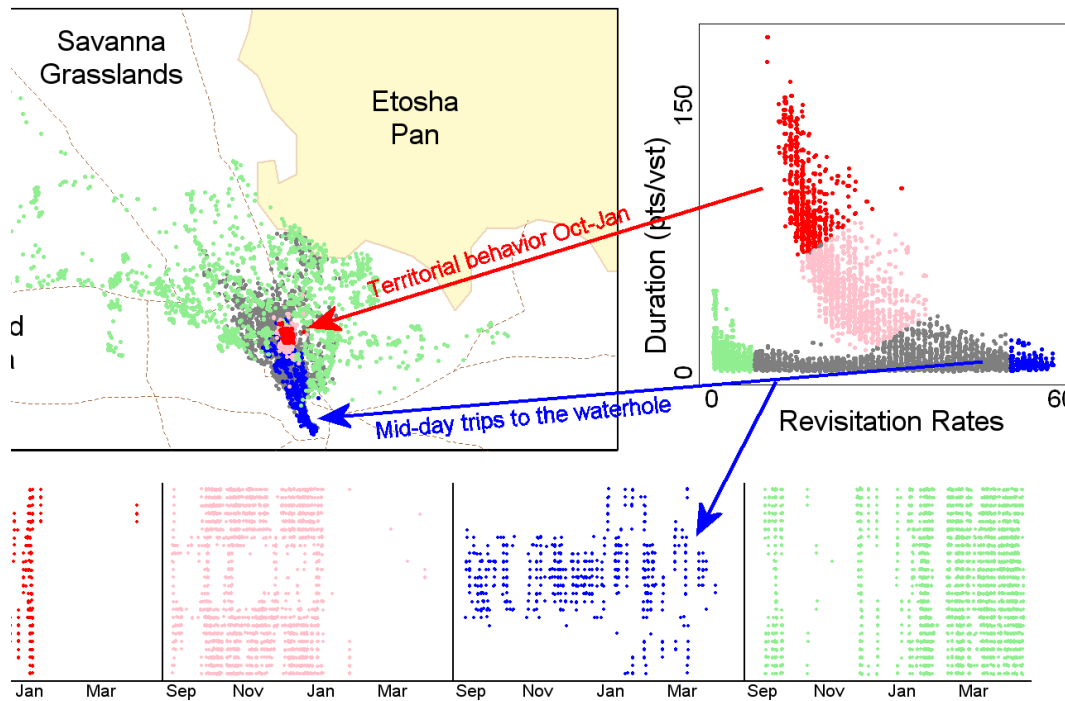
Can depict sharp boundaries (e.g. water bodies, fences)

Prone to type-II errors when sample size is small (depicts gaps where animals were not recorded simply by chance)

Prone to issues in selecting value of k and a

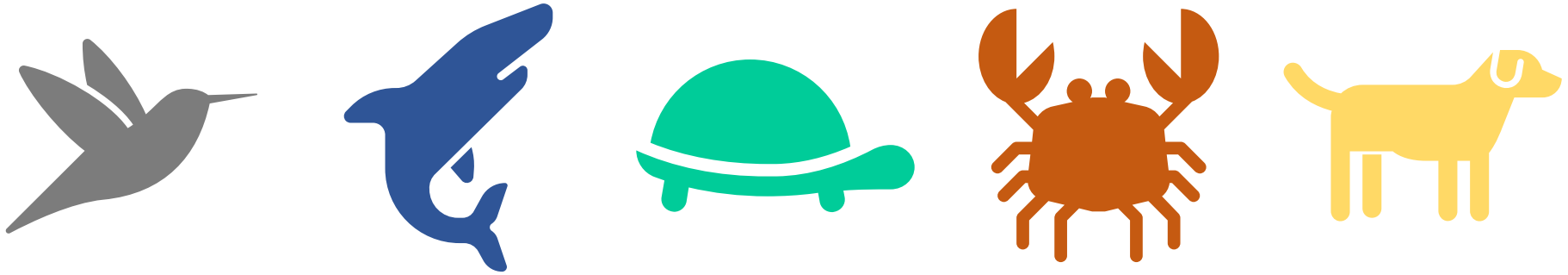k=√N for N<1000 and k= √N for N>1000 (recommended by Downs et al. 2012)

T-LoCoH

- Incorporates time into construction and aggregation of local kernels

- Nearest neighbor selection is based on time-scaled distance

- UDs capture temporal partitions of space and movement phase/time-use metrics

Interesting Resources

- Lyons, A.J., Turner, W.C. & Getz, W.M. (2013) Home range plus: a space-time characterization of movement over real landscapes. *Mov Ecol* **1,** 2. https://doi.org/10.1186/2051-3933-1-2

- Chirima, G.J. and Owen-Smith, N. (2017). Comparison of Kernel Density and Local Convex Hull Methods for Assessing Distribution Ranges of Large Mammalian Herbivores. *Trans. in GIS*, 21: 359-375. https://doi.org/10.1111/tgis.12193

Question	Response
Give some background on how it works/what does it estimate	Construct hulls (MCPs) based on local groups of points and joins these to create a utilization distribution (occupation)
When/why would you use this and what data are appropriate?	Any type of x-y data (and x-y-t for t-LoCoH) Useful for conspicuous species where absence of records indicates gaps in occurrence
What are some of the strengths/weaknesses of this method?	Can be used to produce UD that excludes hard boundaries, while KDE will extend ranges beyond observed records, obscuring gaps Poor accuracy when static Poor precision when sample sizes are small Need to define k or a parameters
What are the underlying assumptions	Classic LoCoH treats locations as independent
What packages/software can you use to create it?	adehabitatHR, t-LoCoH, AMT, RHR, ArcMap

Characteristic Hulls
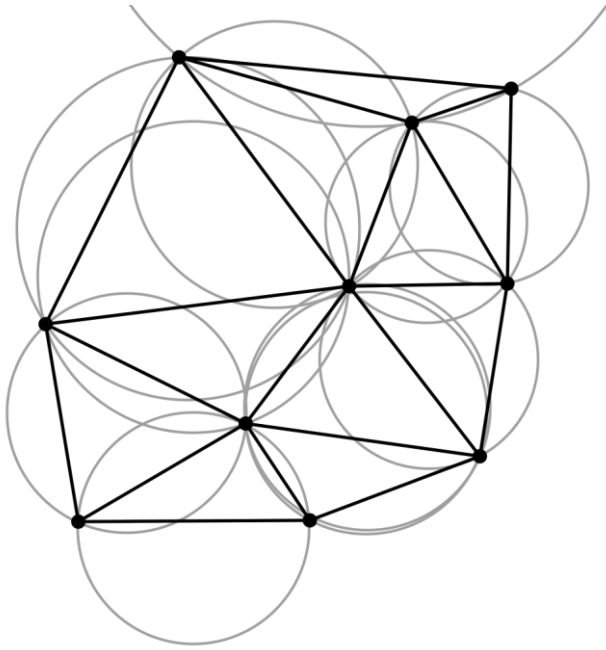
- Example of the Delaunay triangulation for a set of points with the circumcircles for each triangle shown

- Circumcircle is a circle that touches all the vertices of a triangle
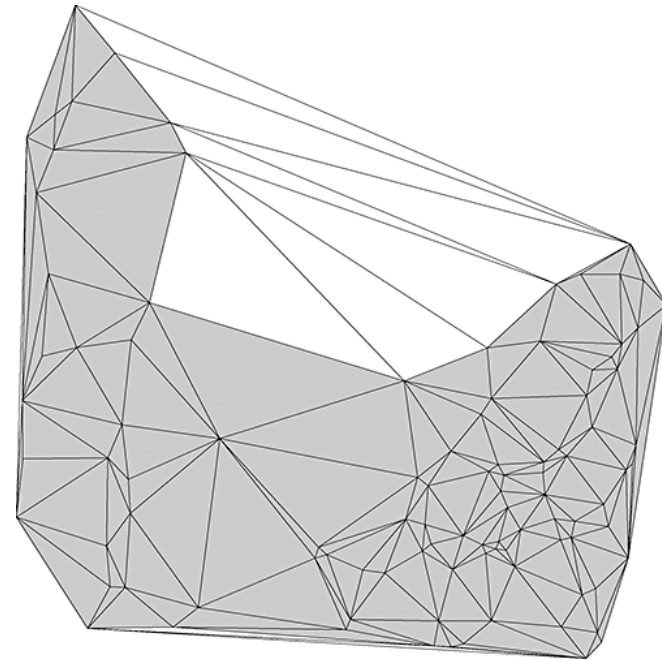
- Example of a characteristic hull polygon constructed using the Delaunay triangulation

- The characteristic hull contains 95% of the smallest triangles

- Removed the largest 5% of triangles by perimeter

For more information: **A Characteristic-Hull Based Method for Home Range Estimation** (Downs and Horner 2009)

https://doi.org/10.1111/j.1467-9671.2009.01177.x

Question	Response
Give some background on how it works/what does it estimate	Constructed using the Delaunay triangulation from a set of points – connecting neighbouring points to build triangles such that the circumcircle for each triangle contains no other points. Can be used to estimate home range from point relocations – usually by removing 5% of the largest triangles. Typically removes the largest triangles by perimeter which are generally large slender triangles along the edges of the point pattern.
When/why would you use this and what data are appropriate?	Can be used on any point relocation data.
What are some of the strengths/weaknesses of this method?	Can produce home range polygons with concave edges, disjoint regions, and areas of unoccupied space in the interior. Generally has smaller home-range estimates than other home range estimators (e.g. minimum convex polygon, kernel density estimation). Method not suitable for all point pattern shapes – best suited to irregular point pattern shapes (e.g. linear, disjoint, perforated). Performed well for seabirds that forage intensely in habitat patches separated by large distances.
What are the underlying assumptions	Have to specify the number of triangles to remove. Including 95% of the triangles is standard practice but may not be suitable for all point shapes/home range types.
What packages/software can you use to create it?	Can be applied in GIS software and in the adehabitatHR package.

Convex Polygons

| Question	Response
Give some background on how it works/what does it estimate	Concave Polygons are a type of characteristic Hull. It allows for inward bending shapes. They are generated from the Delaunay triangulation (DT) for a set of input points. The DT is a graph constructed by connecting neighbouring points to form triangles such that the minimum angle is the triangulation is maximised. The CHP is then extracted from the DT by removing some number of triangles from the graph. This number is usually 5% (95% confidence interval), meaning 5% of largest triangles are removed to retain the area where the animal most likely spends 95% of its time
When/why would you use this and what data are appropriate?	General home range estimations from animal location data.
What are some of the strengths/weaknesses of this method?	Are believed to be more accurate than KDE and MCP. However, still can underestimate by 5-20%. Positive is the inclusion of complex and flexible inward facing, disjointed, non-linear shapes
What are the underlying assumptions	?
What packages/software can you use to create it?	GIS, AdehabitatHR

Single-linkage clusters

Single-Linkage Clusters

First Generation Estimator

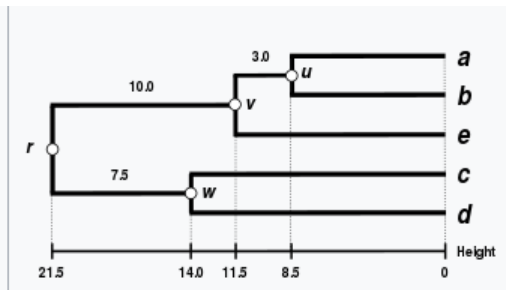Single link: $D(c_1, c_2) = \min_{x_1 \in c_1, x_2 \in c_2} D(x_1, x_2)$

- distance between closest elements in clusters
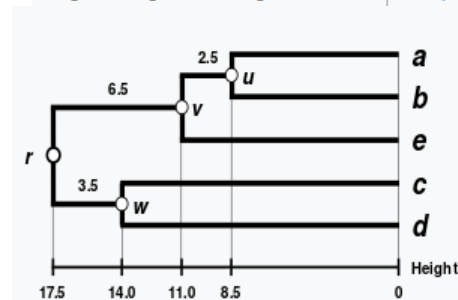- produces long chains a→b→c→…→z

- Hierarchical Clustering Method
- Agglomerative – bottom-up
- Nearest-neighbour

- Usually visualised as a dendogram
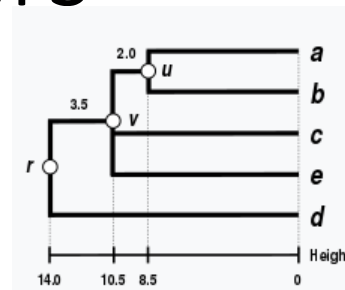
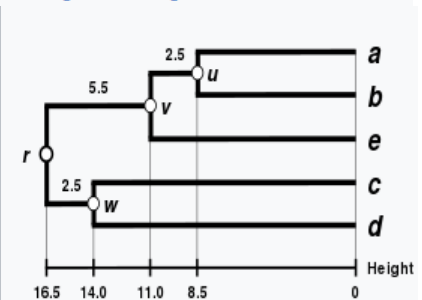Single-linkage clustering   Complete-linkage clustering

Average linkage clustering: WPGMA   Average linkage clustering: UPGMA

Minimum cluster size =3

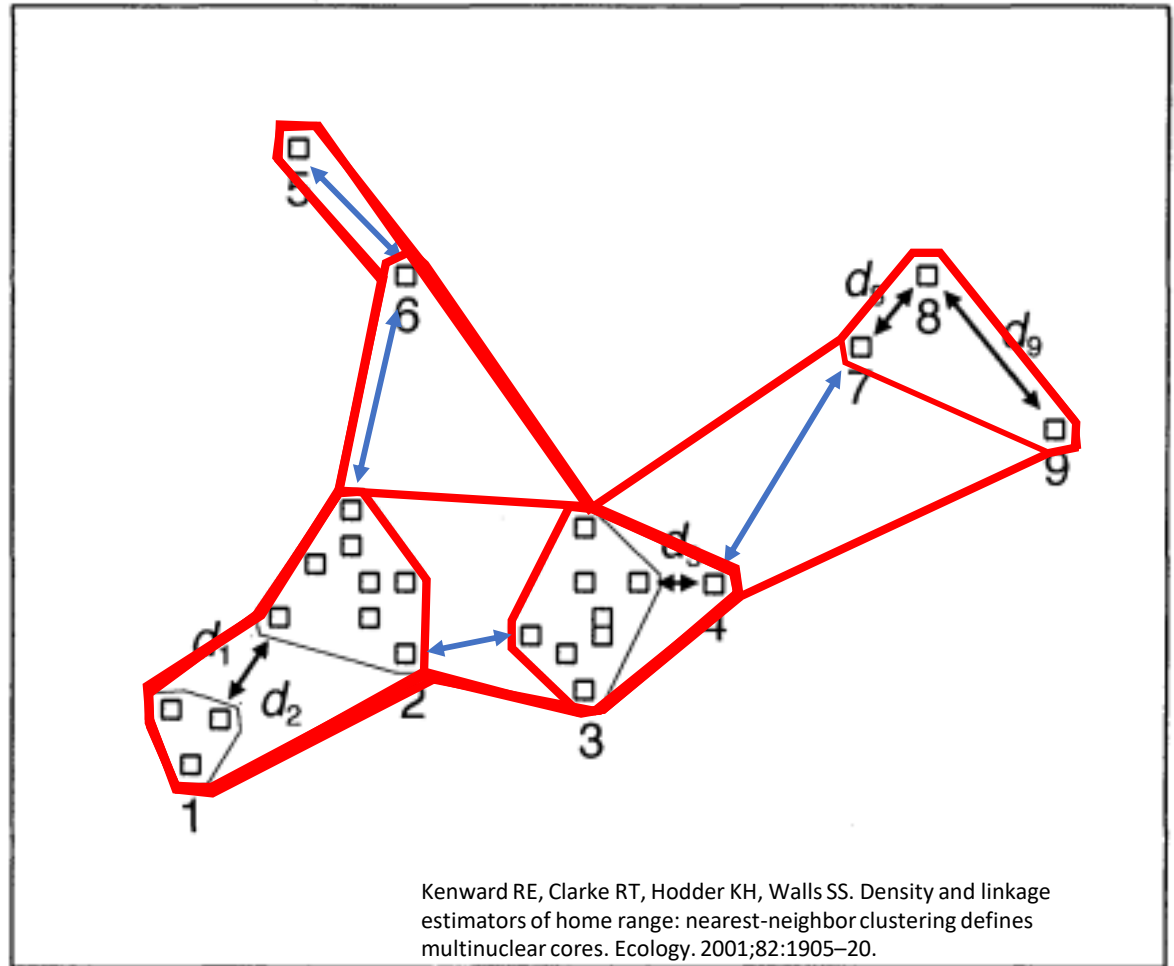Take 3 locations with minimum sum of nearest neighbour joining distances to for first cluster

Add 4th if distance to first clusters nearest outlier is < mean nearest neighbour in next potential cluster

Two clusters fuse if the outlier belonging to one is being assigned to another.

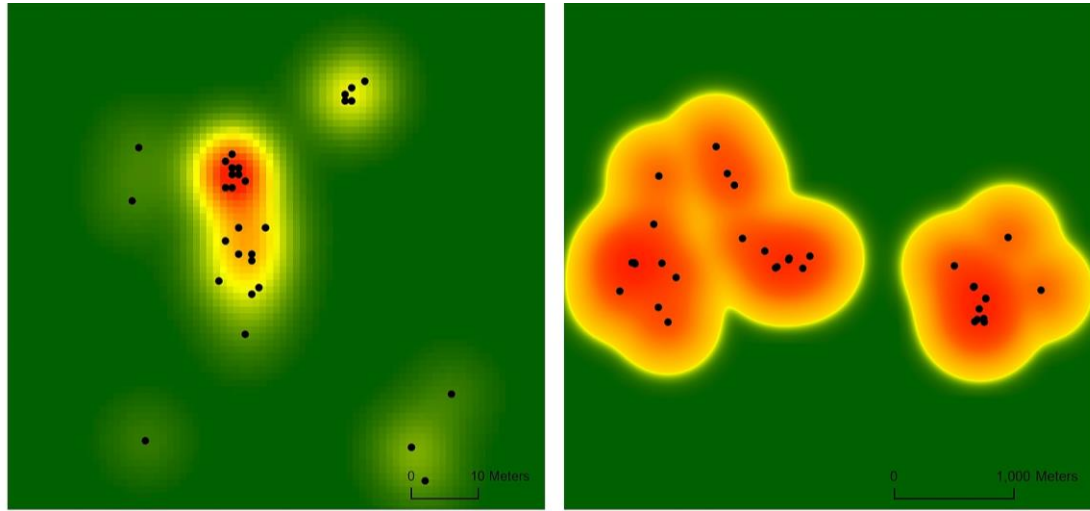Outlines are either convex polygons or concave polygons.

Rules can be tweaked e.g. centroid distancing,

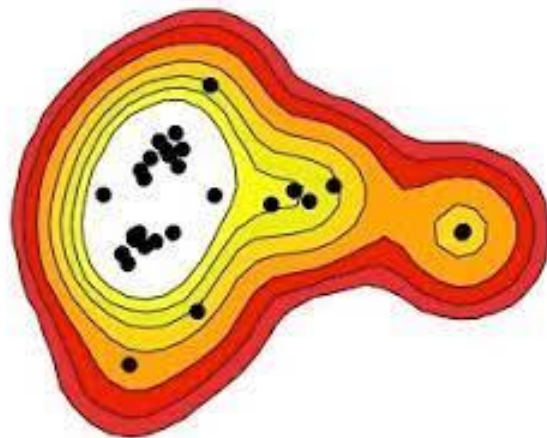Might be possible to include other metrics using scaling (e.g. temporal?)

Kenward RE, Clarke RT, Hodder KH, Walls SS. Density and linkage estimators of home range: nearest-neighbor clustering defines multinuclear cores. Ecology. 2001;82:1905–20.

Question	Response
Give some background on how it works/what does it estimate	- Clusters nearest neighbours into groups using an agglomerative approach, this can be used to identify spatial patterns or behavioural modes. More common as a tool in machine learning. - "entities within a cluster should be as similar as possible and entities in one cluster should be as dissimilar as possible from entities in another." – Datacamp
When/why would you use this and what data are appropriate?	- x/y data - Generally not recommended, although other clustering methods are better (Rencher et al. 2002). - Usually used to identify similarities between data in machine learning, rather than for spatial analysis - Can be used to identify similar behaviours. - Can work well in areas with barriers as there can be less "bleeding" into inaccessible areas
What are some of the strengths/weaknesses of this method?	- Simple interpretation, simple assumptions - Can show internal structure of range compared to MCP - Walter et al. 2015 – considered the least preferred estimator due to length of time to provide estimate and >1,000 locations often failed to produce home ranges. - Prone to chaining, where new clusters move closer to individual observations, so observations end up joining other clusters rather than making a distinct cluster - Sensitive to errors in distances between observations - Polygons "fill in gaps"
What are the underlying assumptions	- no correlation, need to account for temporal autocorrelation - Points that are closer together are more similar - Animals probably route their initial visits to a new location from one nearby familiar location
What packages/software can you use to create it?	- stats(), dextend(), and with adehabitatHR

Kernel density estimates (KDE)

Example KDE plot

Example KDE plot https://jamesepaterson.github.io/jamespatersonblog/04_trackingworkshop_kernels

Relies heavily on bandwidth assumptions (the amount of area that each point influences)
- Too high will over-smooth the data and result in unrealistically high estimates
- Too low under-smooths the data and results in unrealistically low estimates
- No universally agreed upon bandwidth or method of selecting a bandwidth
- Affected by data distribution, assumptions about movement and use, amount of data available.

Question	Response
Give some background on how it works/what does it estimate	In a nutshell, it estimates use as a probability around points (i.e., a probability that an animal uses a given pixel in a grid). This results in a probability cloud, from which you can calculate home range as a percentage of the cloud.
When/why would you use this and what data are appropriate?	Probably the most popular home range estimator; very widely applicable.
What are some of the strengths/weaknesses of this method?	Pro: Fairly easy to use, applicable to a wide range of data sets (VHF, GPS, sonic tags, etc.), often gives results that make more biological sense than a method like MCP. Con: Reliant on bandwidth assumptions, often excludes areas that an animal had to move through at some point
What are the underlying assumptions	Non-parametric (for some bandwiths). Multiple types, but they **usually** don't make much difference. Bandwidth (smoothing) types have a large effect (bandwith = the amount that each point influences the usage distribution). Choice of bandwidth depends on distribution and amount of data (e.g., GPS vs VHF may differ).
What packages/software can you use to create it?	adehabitatHR (tutorials: https://jamesepaterson.github.io/jamespatersonblog/04_trackingworkshop_kernels)

Brownian bridges/Biased random bridges

Movement-based methods to calculate space use

Brownian bridges

- "A Brownian bridge is a **continuous-time stochastic** model of movement in which the probability of being in an area is conditioned on starting and ending locations, the **elapsed time** between those points, and the **mobility or speed** of movement." (Horne et al. 2007)

- Conditional random walk

- Brownian motion: conditioned on the beginning and ending locations of each pair

- Three common uses:

    - home range estimation

    - estimating migration routes

    - analysing fine-scale resource selection

Biased random bridges

- Based on the biased random walk model

- Similar to Brownian bridge approach, but adds a drift/advection component (i.e. general tendency to move in direction of next relocation), rather than just a diffusive component (i.e. random)

- angular distribution is not uniform, i.e. there is a preferred direction of movement

- Bottom line: BRB is more directed, whereas BB is more random

Question	Response
Give some background on how it works/what does it estimate	Based on continuous time-based movement • BB: just based on diffusive component • BRB: advective + diffusive component
When/why would you use this and what data are appropriate?	Estimate home range, estimate migration routes, analyse resource selection BB can also be used in mathematical finance
What are some of the strengths/weaknesses of this method?	BB: • can be used to analyse data collected at short time intervals, can account for location error, can describe long-distance movements (e.g. migration routes) • Can be too simple and unrealistic movement models (especially as time interval increases) • Relevant only for wandering animals in a uniform environment BRB: • Acknowledges and incorporates directed movement (drift) in location/movement
What are the underlying assumptions	BB: random movement between locations BRB: space use intensity is proportional to activity time
What packages/software can you use to create it?	BB: adehabitatHR, move, Sim.DiffProc, sde BRB: adehabitatHR

Useful sources

- Horne, J. S., E. O. Garton, S. M. Krone, and J. S. Lewis. (2007) Analyzing animal movements using Brownian bridges. Ecology 88:2354–2363.

- Benhamou, S. (2011) Dynamic approach to space and habitat use based on biased random bridges. PLOS One, 6, 1–8.

- Benhamou, S. and Cornelis, D. (2010) Incorporating movement behavior and barriers to improve biological relevance of kernel home range space use estimates. Journal of Wildlife Management, 74, 1353–1360.

- Calenge C (2006). The package adehabitat for the R software: tool for the analysis of space and habitat use by animals. Ecological Modelling, 197, 1035.

- 'adehabitatHR' package manual: https://cran.r-project.org/web/packages/adehabitatHR/adehabitatHR.pdf

Accounting for autocorr. makes a big difference
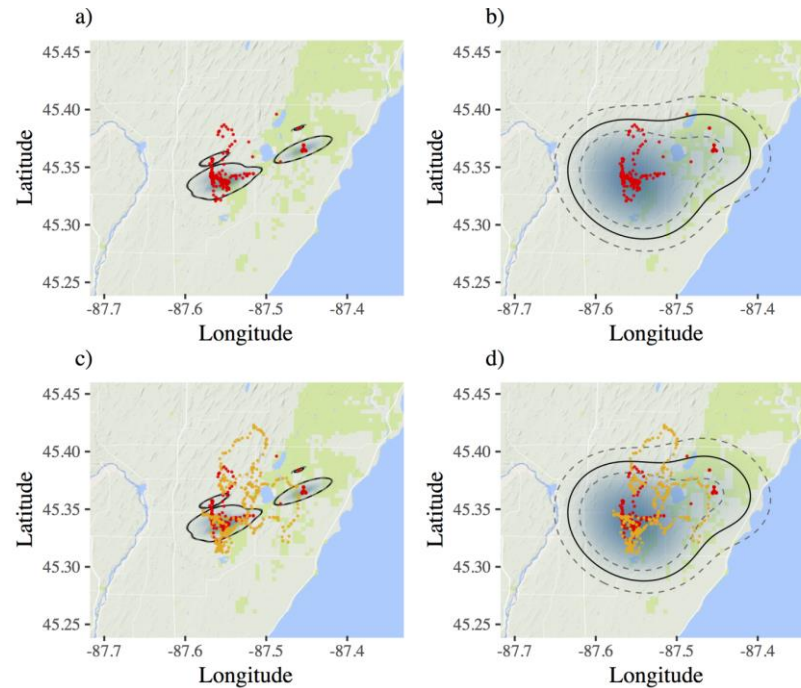
Movement Model	Pos. AC	Vel. AC	H. Range
Ind. Ident. Distr. (IID)	No	No	Yes
Brownian Motion (BM)	Yes	No	No
Ornstein-Uhlenbeck (OU)	Yes	No	Yes
Integrated OU (IOU)	Yes	Yes	No
Ornstein-Uhlenbeck F (OUF)	Yes	Yes	Yes

Autocorrelated KDE

Johnny
Tracked for
5 months with an
hourly fix
Ne= 7.767 (N= 3540)
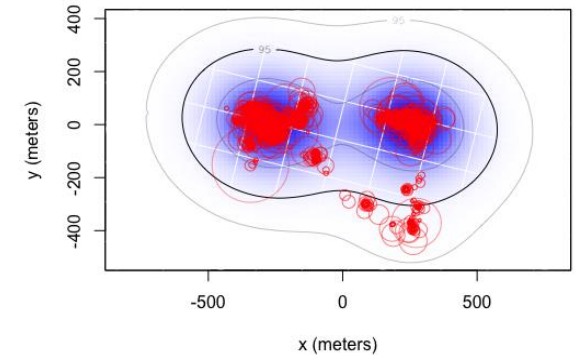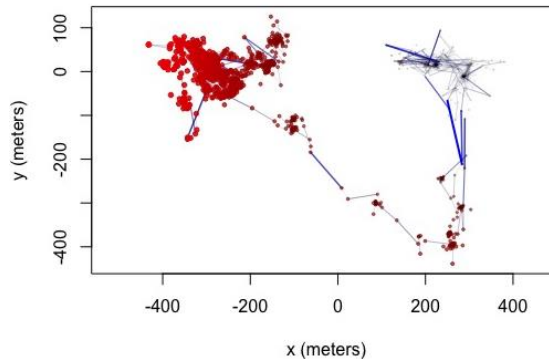
Parameters	low	ML	high
area (hectares)	26.786	62.993	114.414
τ[position] (days)	6.561	22.194	75.070
τ[velocity] (minutes)	29.323	29.897	36.908
speed (meters/day)	351.697	365.073	378.442

Question	Response
Give some background on how it works/what does it estimate	This is the same as a conventional kernel density estimator but this includes movement models beyond the standard IID which allows you to account for autocorrelation if it is present in your data.
When/why would you use this and what data are appropriate?	This will be particularly needed when you have very fine temporal data or the animal movement patterns lead to strong autocorrelation in your data. Any movement data with x,y and t is appropriate, you would use this when you want a more accurate estimate of space use.
What are some of the strengths/weaknesses of this method?	Strengths: You get estimates which are more realistic and unbiased, you can fit different movement models, you get confidence intervals, you can also estimate HR overlap and other metrics (speed, HR time etc.) Weaknesses: Models can be more complicated to fit than previous methods, you need more computing power, can take a longer time to get model estimates (few hours to days depending on the size of the track and how many individuals).
What are the underlying assumptions	That you have a range resident - The the assumptions per movement model can be a bit different
What packages/software can you use to create it?	'ctmm' in R, they also have a web interface run through R 'ctmmweb' *Calabrese et. al, (2016) Ctmm: an R Package for Analyzing Animal Relocation Data as a Continuous-Time Stochastic Process, Methods in Ecol & Evol*