



AniMove 2024, June 17th to 28th

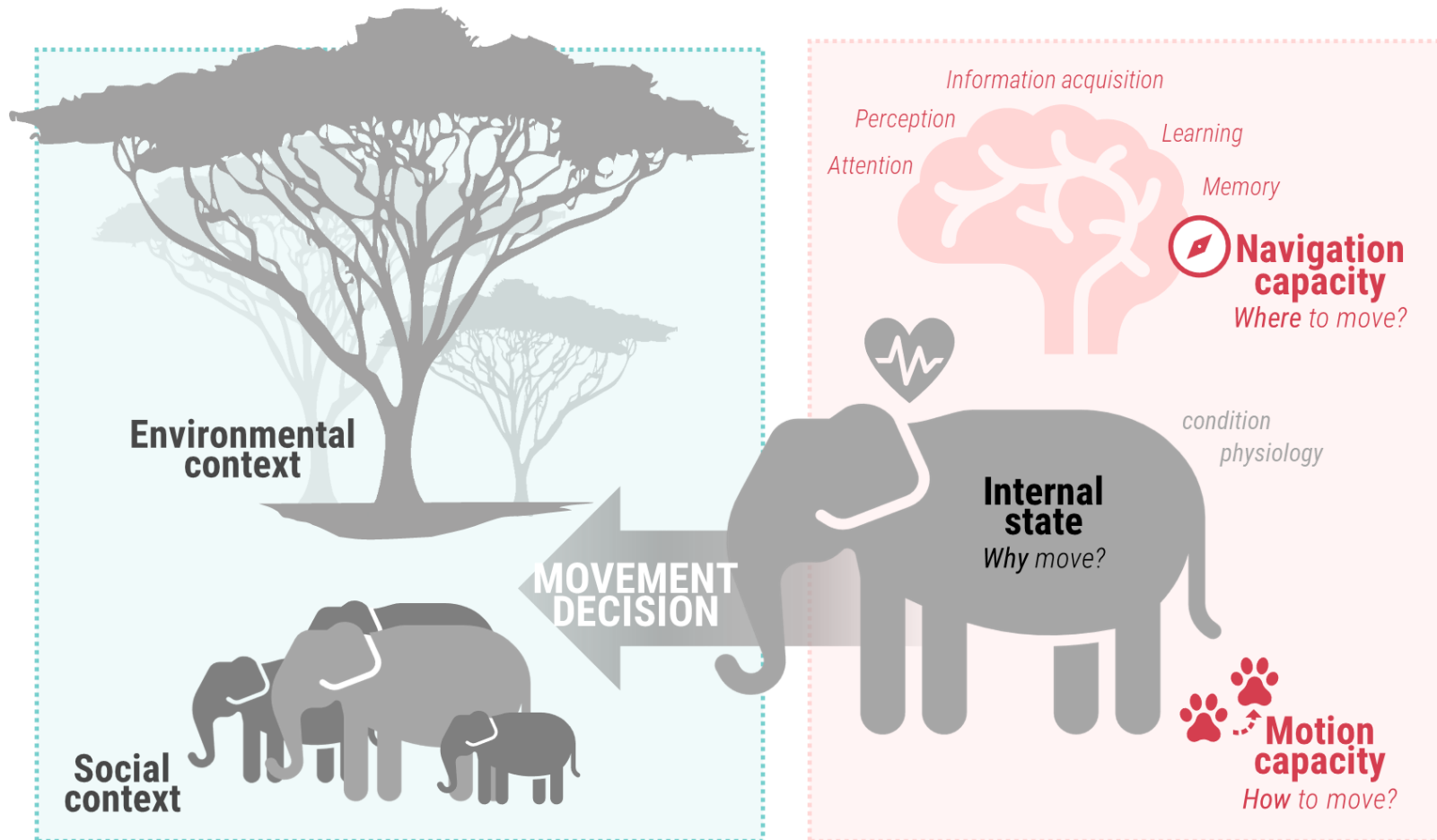
Home range estimation

Using the 'ctmm' R package

Inês Silva, Chris Fleming

✉ i.simoes-silva@hzdr.de





Adapted from **Lewis *et al.* (2021)**
DOI: 10.3389/fevo.2021.681704



What is a **home range**?

Home range behavior is a prevalent pattern in **space-use**.



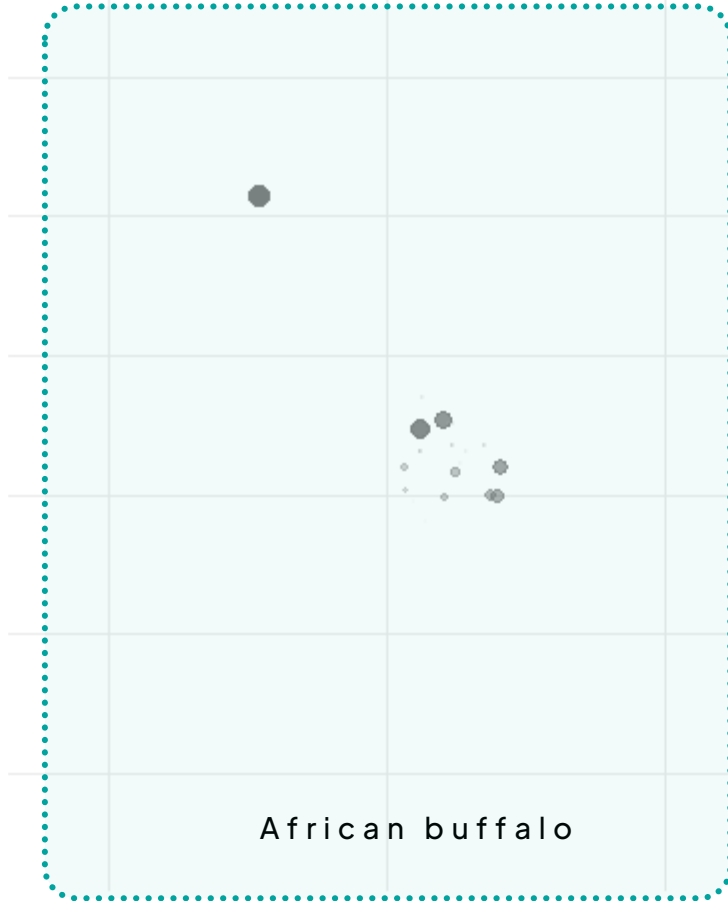
Rob Beechey

“

(...) it may be here remarked that most animals and plants keep to their proper homes, and **do not needlessly wander about**; we see this even with migratory birds, which almost always return to the same spot.

✍ Darwin (1861)

Area-restricted space-use behavior



vs.

Brownian motion
Unbounded space-use behavior

First defined as:

“

(...) the area traversed by the individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered as in part of the home range.

✍ Burt (1943)

Home range
not actively defended

≠

Territory
actively defended

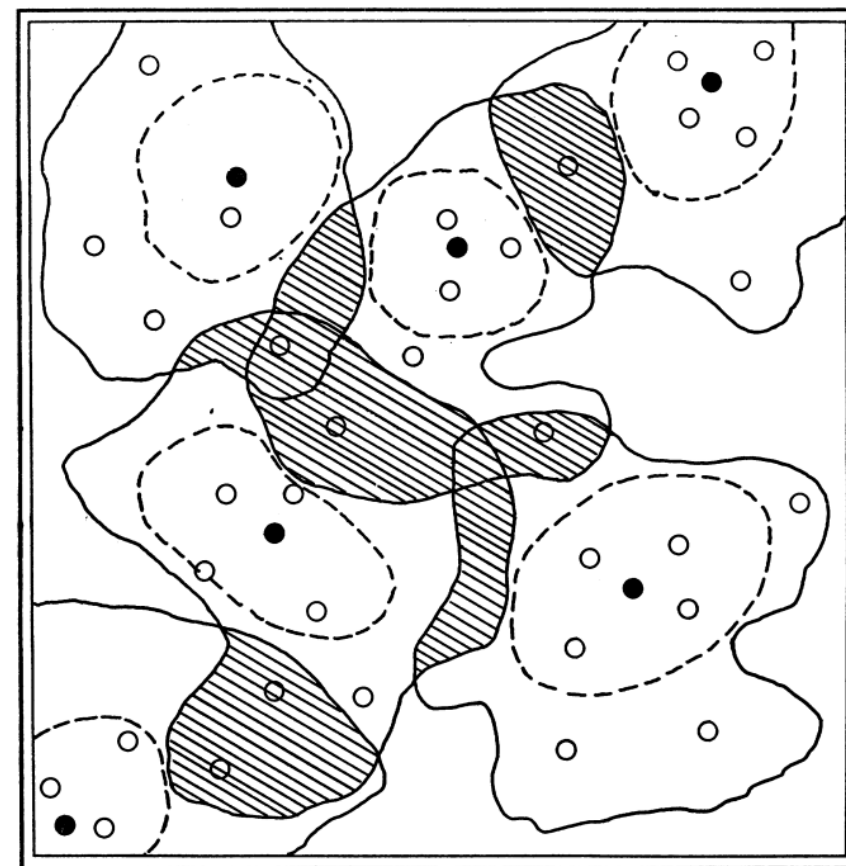


FIG. 1. Theoretical quadrat with six occupants of the same species and sex, showing territory and home range concepts as presented in text.

First defined as:

“

(...) the area traversed by the individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered as in part of the home range.

✍ Burt (1943)

- ▶ How to quantify home range area?
- ▶ What constitutes an exploratory movement?
- ▶ How to identify these exploratory movements?

In practice, it is hard to define when a move is purely **exploratory**.

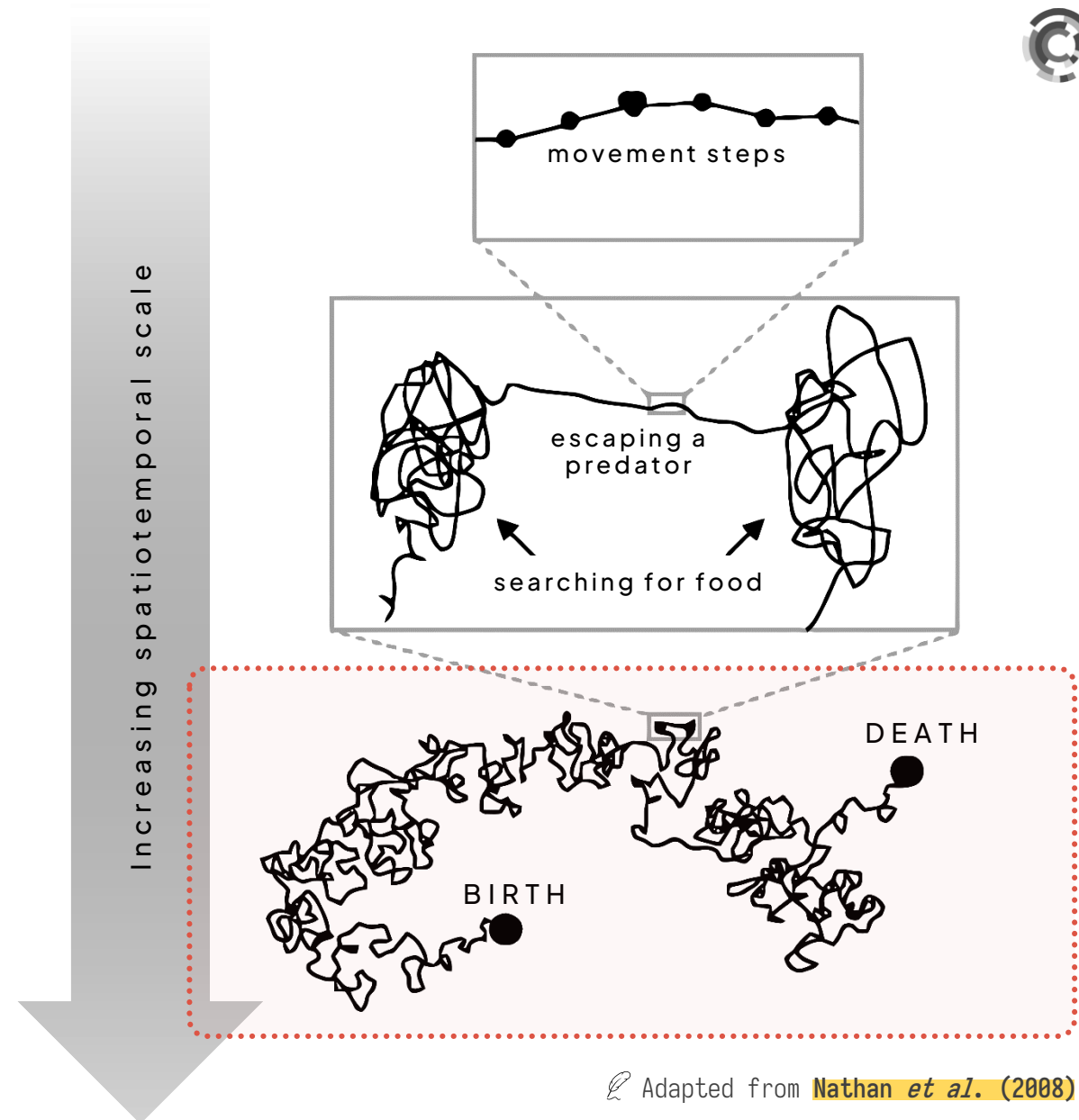
↪  Exclusion of infrequent
outlying relocation points.

What is a **home range**?

Here, we follow the definition of home range as the area repeatedly used throughout an animal's **lifetime** for all its **normal behaviors and activities**, excluding occasional **exploratory excursions**.

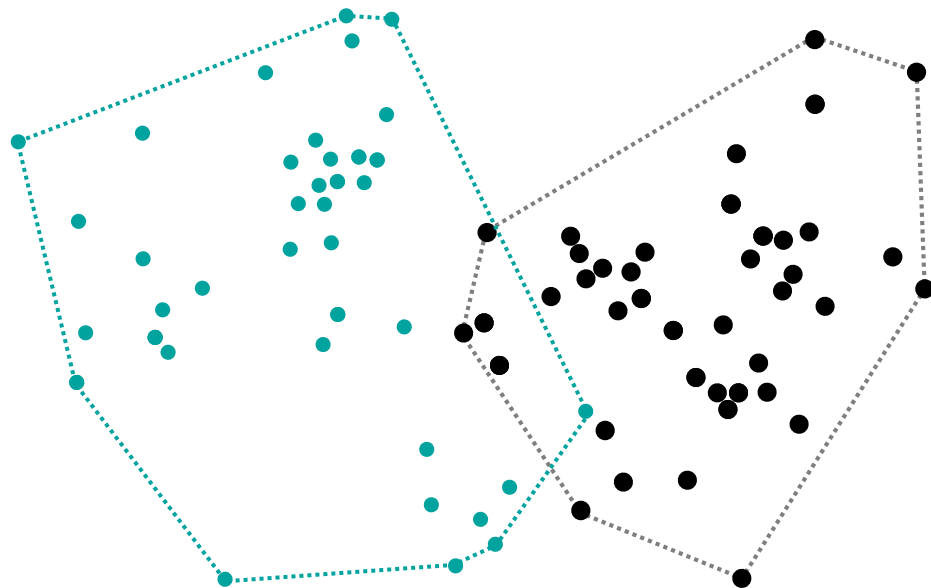


Home range area can be expected to include future locations.



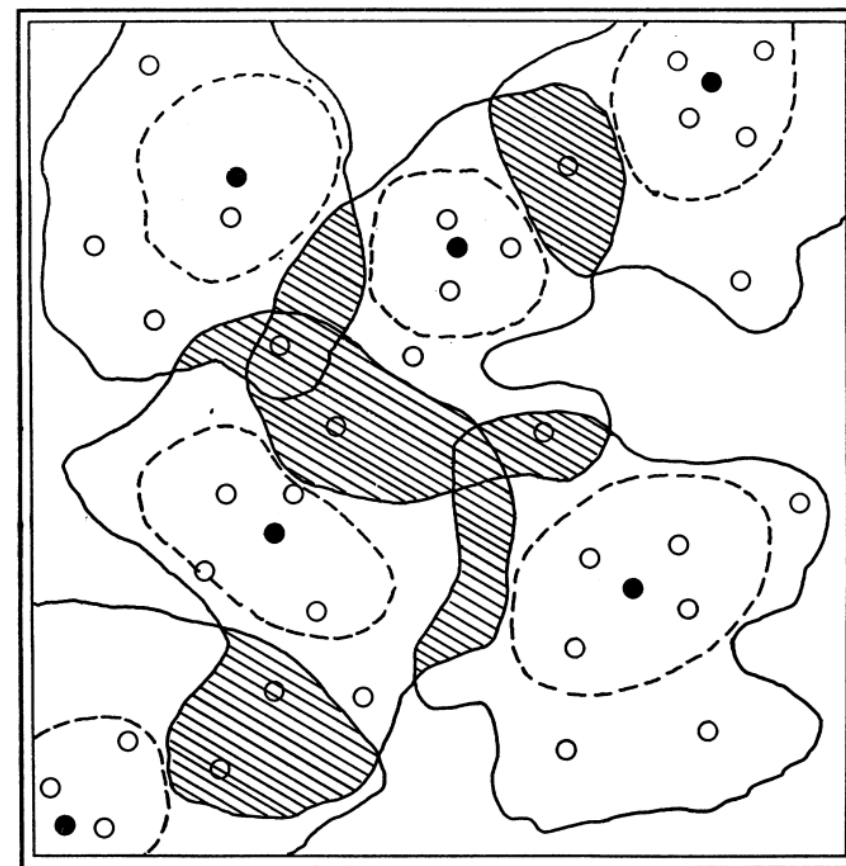
Adapted from **Nathan et al. (2008)**

 **Burt (1943)**



Minimum convex polygon (MCP)

The smallest polygon drawn around tracking locations with all interior angles less than 180 degrees.



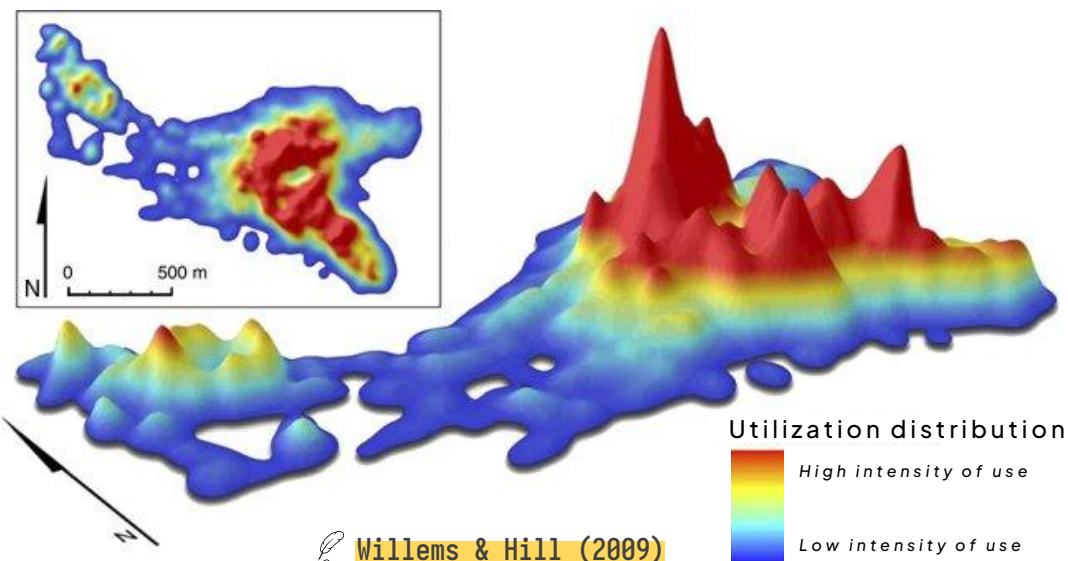
— HOME RANGE BOUNDARY  NEUTRAL AREA
- - - TERRITORIAL BOUNDARY ● NESTING SITE
BLANK--UNOCCUPIED SPACE ○ REFUGE SITE

FIG. 1. Theoretical quadrat with six occupants of the same species and sex, showing territory and home range concepts as presented in text.

“

It seems that an understanding of the biological significance of an animal's home range must include some knowledge of the **intensity of use**, by the animal, of various parts of the area.

Hayne (1949)



Worton (1989)

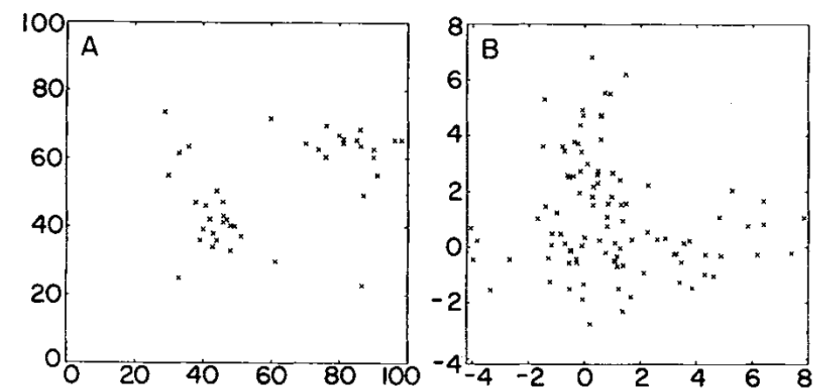


FIG. 1. Plots of (A) the DC data set and (B) the SIM data set.

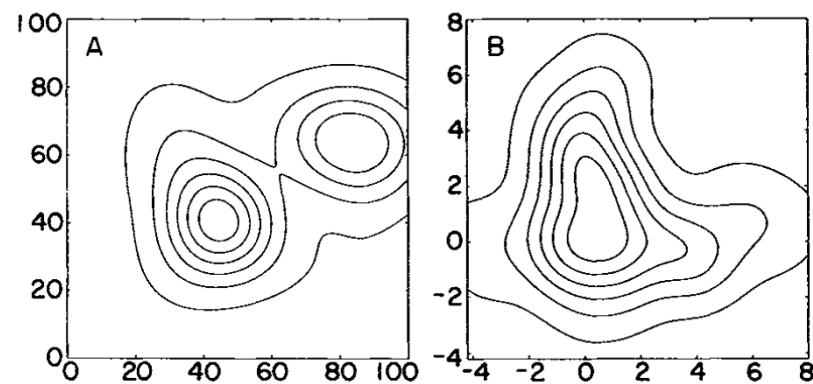


FIG. 2. **Fixed kernel density estimates** of the UD densities with the ad hoc choice of smoothing parameters for (A) the DC data set ($h = 10.$) and (B) the SIM data set ($h = 1.$).

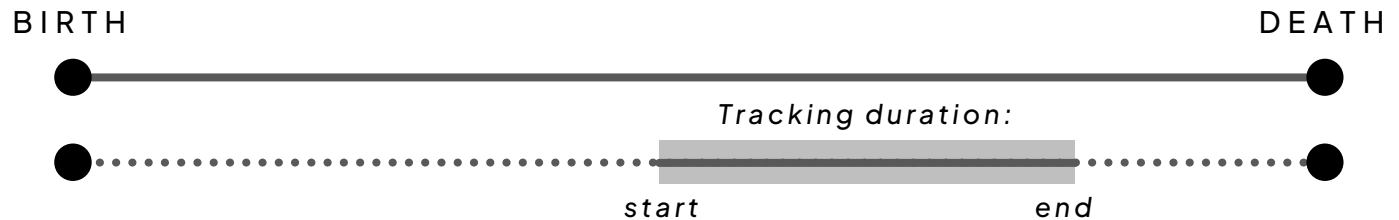


Home-range area estimates may inform:

- ▶ Protected area requirements,
- ▶ Land-use decisions,
- ▶ Conservation policy and initiatives,
(*e.g.*, related to human-wildlife conflict).



It is vital to accurately capture the area repeatedly used throughout an *animal's lifetime*.



How *representative* is this period?
Is the movement behavior *stationary*?

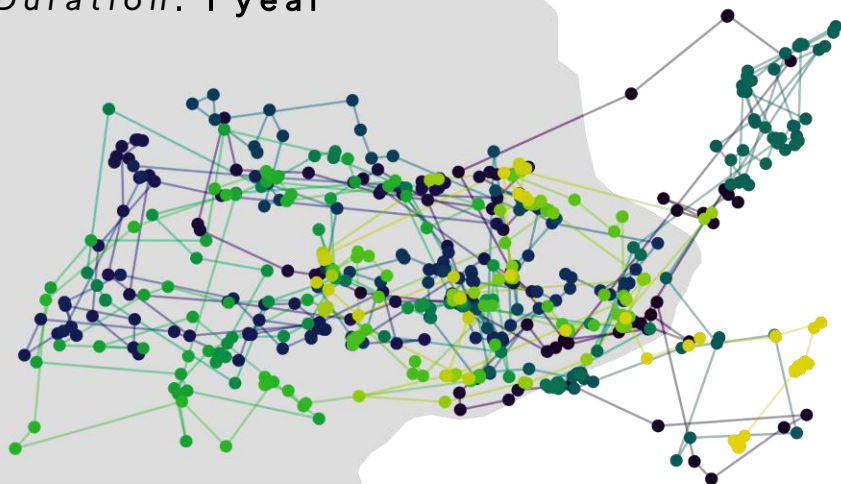
Home-range area estimates may inform:

- ▶ Protected area requirements,
- ▶ Land-use decisions,
- ▶ Conservation policy and initiatives,
(*e.g.*, related to human-wildlife conflict).



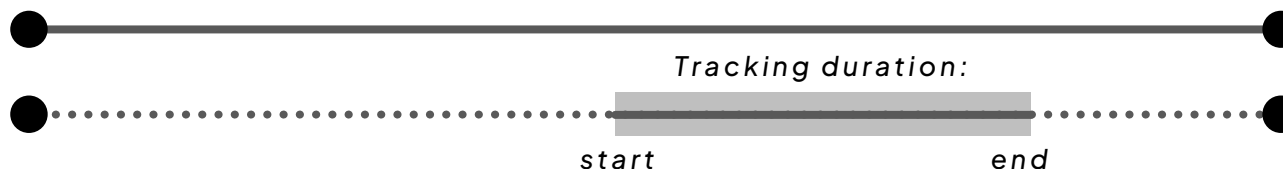
It is vital to accurately capture
the area repeatedly used throughout
an *animal's lifetime*.

Duration: 1 year



BIRTH

DEATH



How *representative* is this period?
Is the movement behavior *stationary*?

VU

SPECIES
KING COBRA
(*Ophiophagus hannah*)



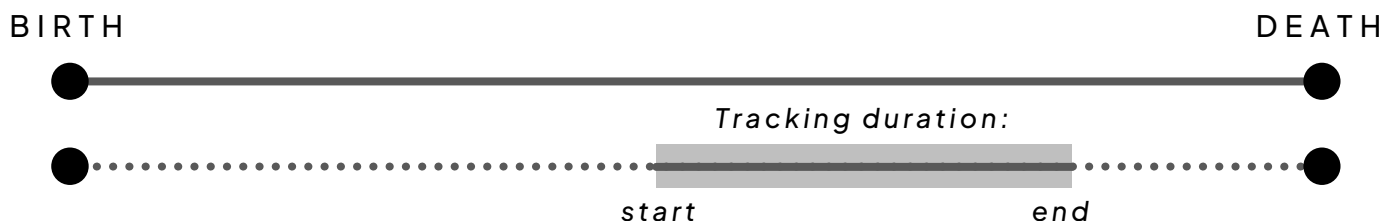
Home-range area estimates may inform:

- ▶ Protected area requirements,
- ▶ Land-use decisions,
- ▶ Conservation policy and initiatives,
(*e.g.*, related to human-wildlife conflict).

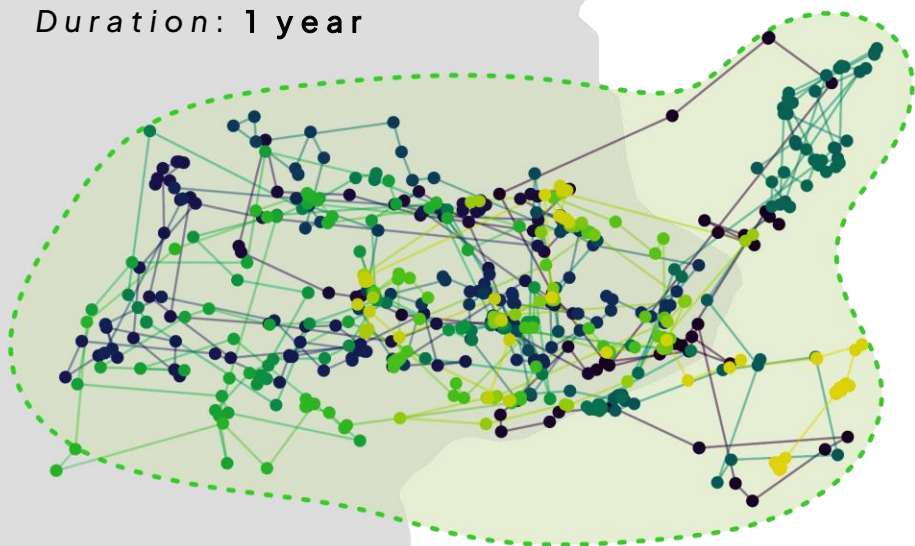


It is vital to accurately capture
the area repeatedly used throughout
an *animal's lifetime*.

Duration: 1 year



How *representative* is this period?
Is the movement behavior *stationary*?



VU

SPECIES
KING COBRA
(*Ophiophagus hannah*)



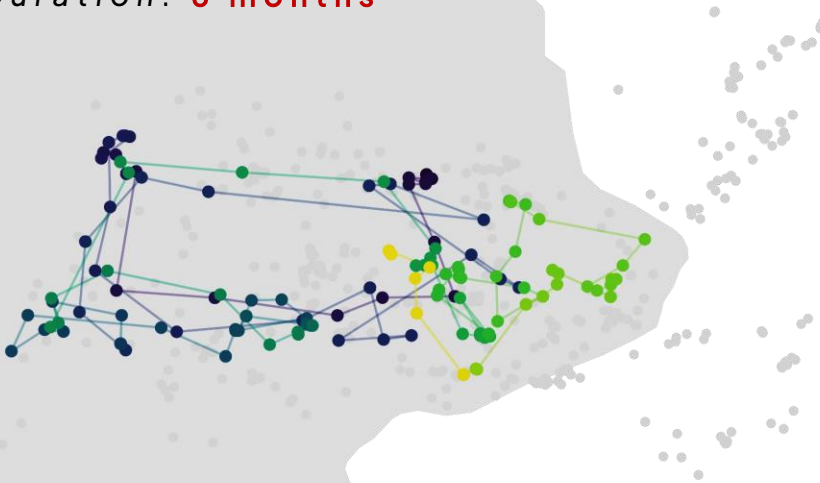
Home-range area estimates may inform:

- ▶ Protected area requirements,
- ▶ Land-use decisions,
- ▶ Conservation policy and initiatives,
(*e.g.*, related to human-wildlife conflict).



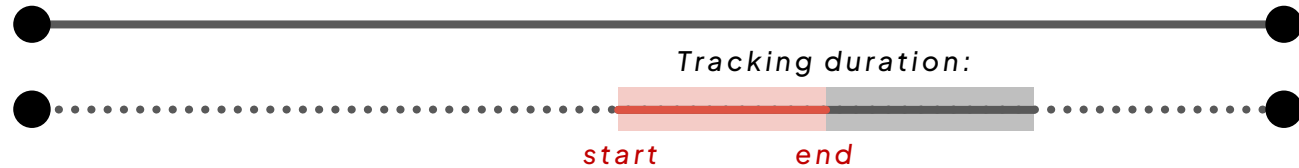
It is vital to accurately capture
the area repeatedly used throughout
an *animal's lifetime*.

Duration: 6 months



BIRTH

DEATH



How *representative* is this period?
Is the movement behavior *stationary*?

VU

SPECIES
KING COBRA
(*Ophiophagus hannah*)



Minimum Convex Polygon (MCP)

Kernel density estimator (KDE)

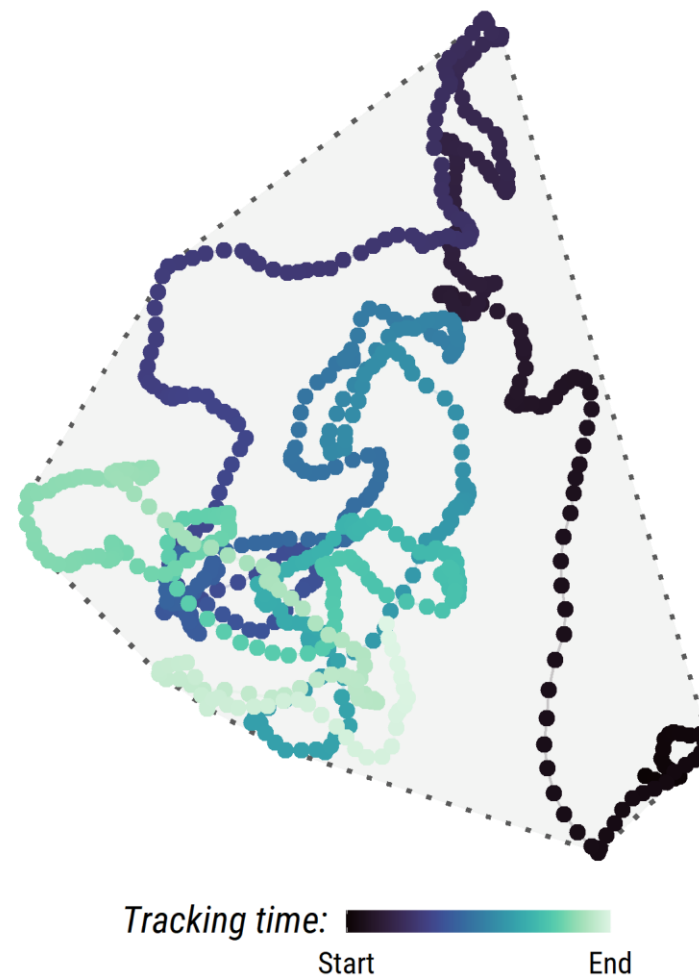


Tracking time: 
Start End

Minimum Convex Polygon (MCP)

Kernel density estimator (KDE)

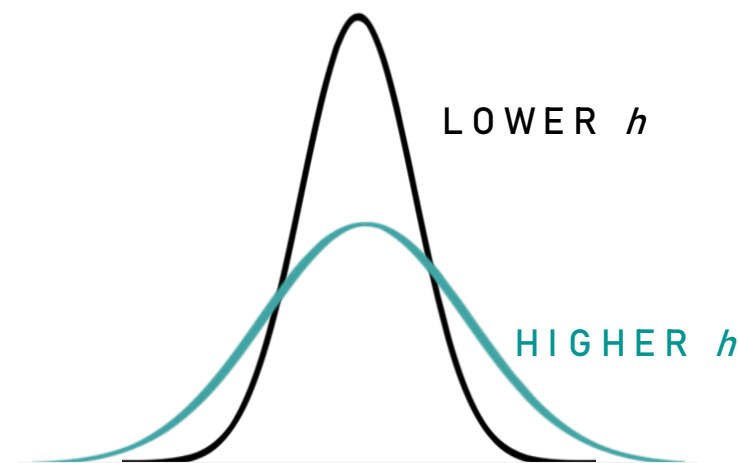
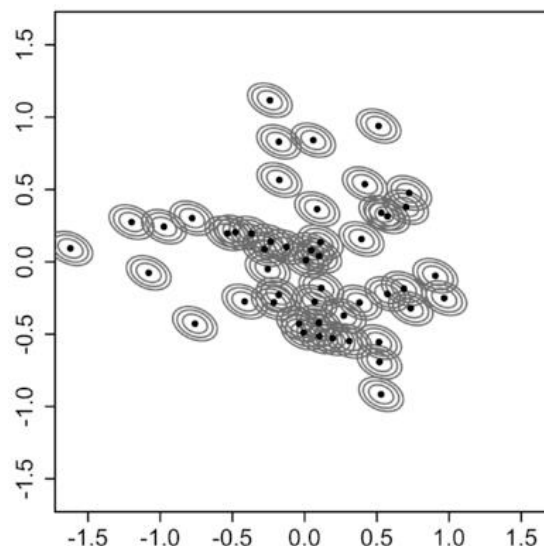
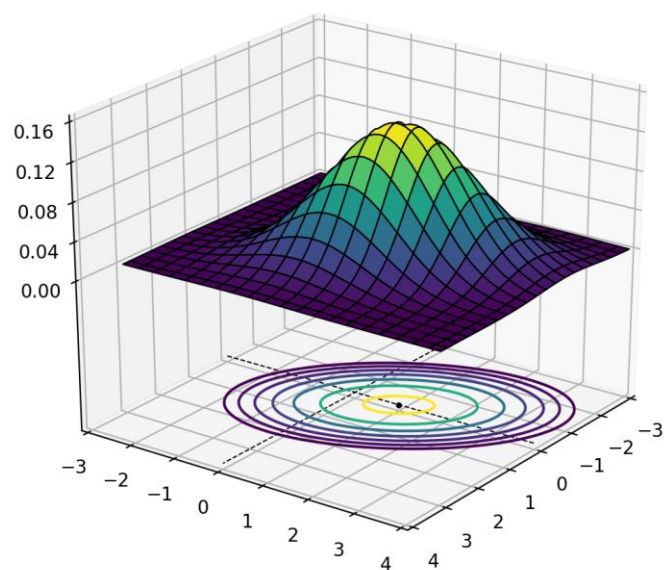
- ▶ Assumes **uniform use**,
- ▶ Assumes locations are **independent**;
- ▶ Sensitive to **outliers** and point geometry.



Minimum Convex Polygon (MCP)

Kernel density estimator (KDE)

Kernel density estimates describe not just the borders of the home range, but the probability of use.

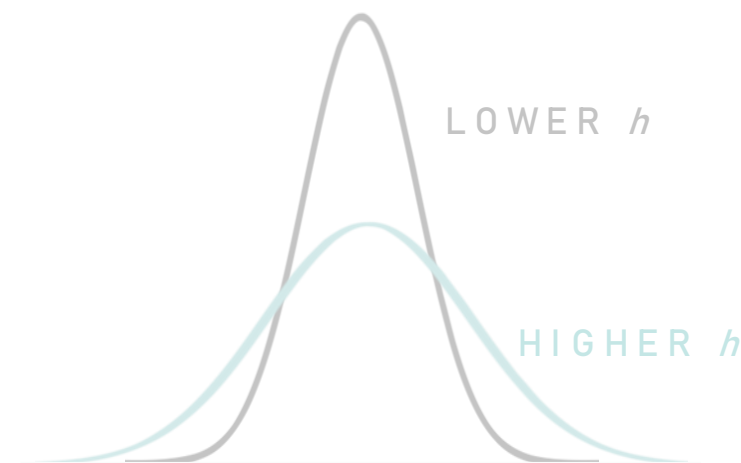
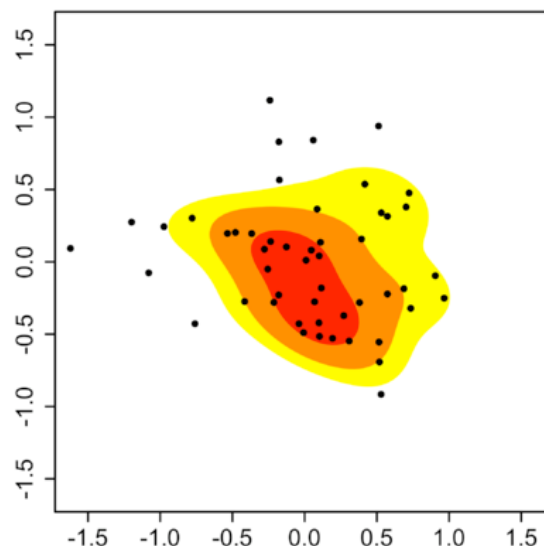
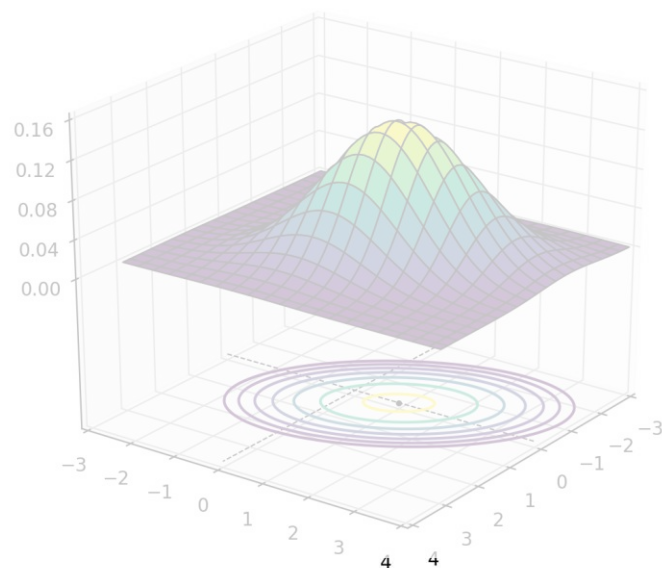


On an x-y plane, each location has a three-dimensional "hill", the **kernel**.
The shape and width of the kernel, called the **bandwidth** (h), can be selected using algorithms.

Minimum Convex Polygon (MCP)

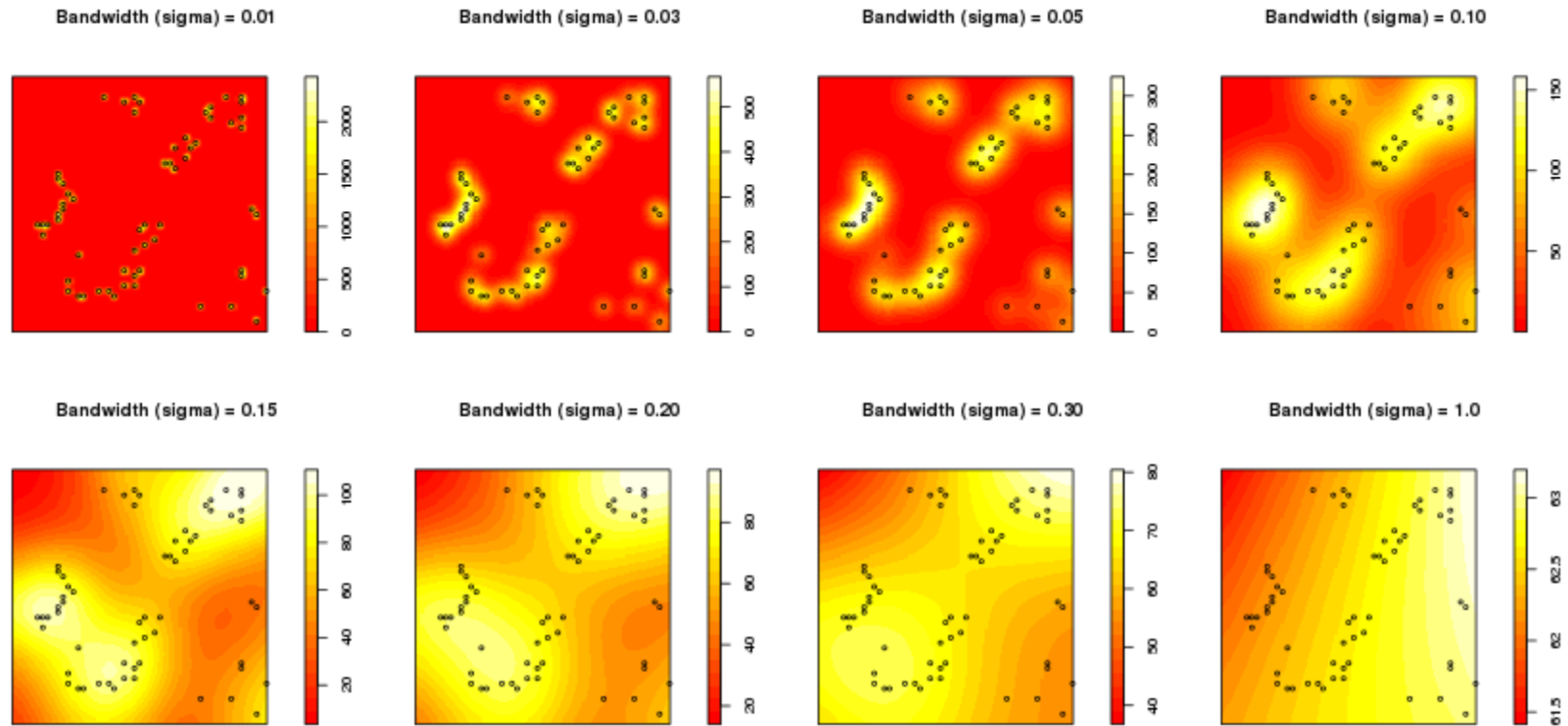
Kernel density estimator (KDE)

Kernel density estimates describe not just the borders of the home range, but the probability of use.



On an x-y plane, each location has a three-dimensional "hill", the **kernel**.
The shape and width of the kernel, called the **bandwidth** (h), can be selected using algorithms.

Minimum Convex Polygon (MCP)
Kernel density estimator (KDE)



bandwidth (h), or smoothing parameter

Minimum Convex Polygon (MCP)
Kernel density estimator (KDE)

- ▶ Assumes locations are **independent**;
- ▶ Sensitive to **bandwidth selection**.

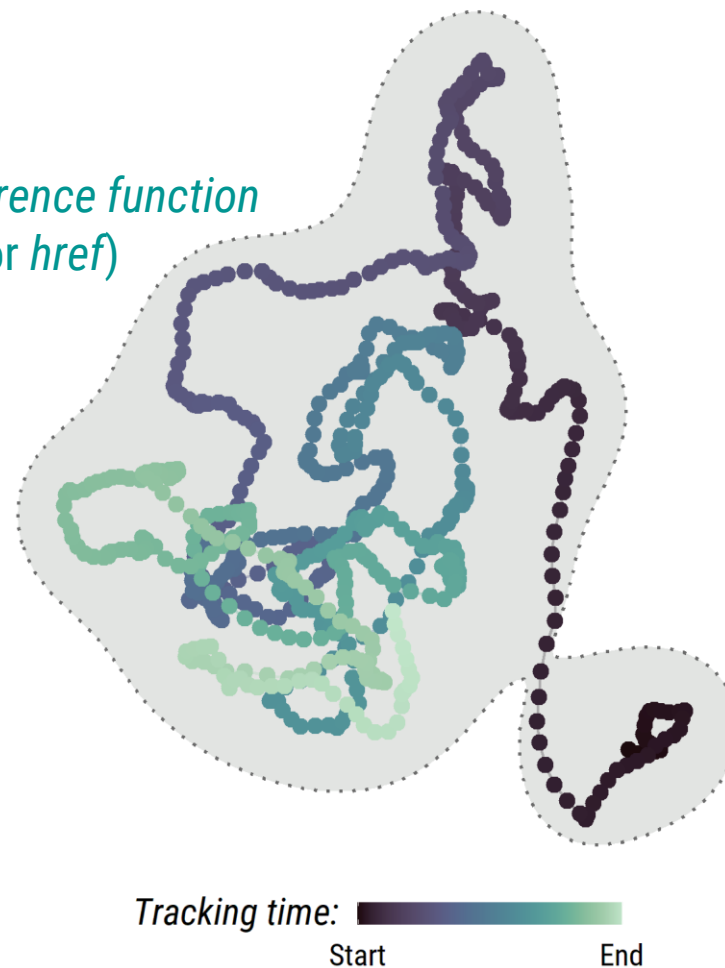


Tracking time: 
Start End

Minimum Convex Polygon (MCP)
Kernel density estimator (KDE)

*Gaussian reference function
(GFR or href)*

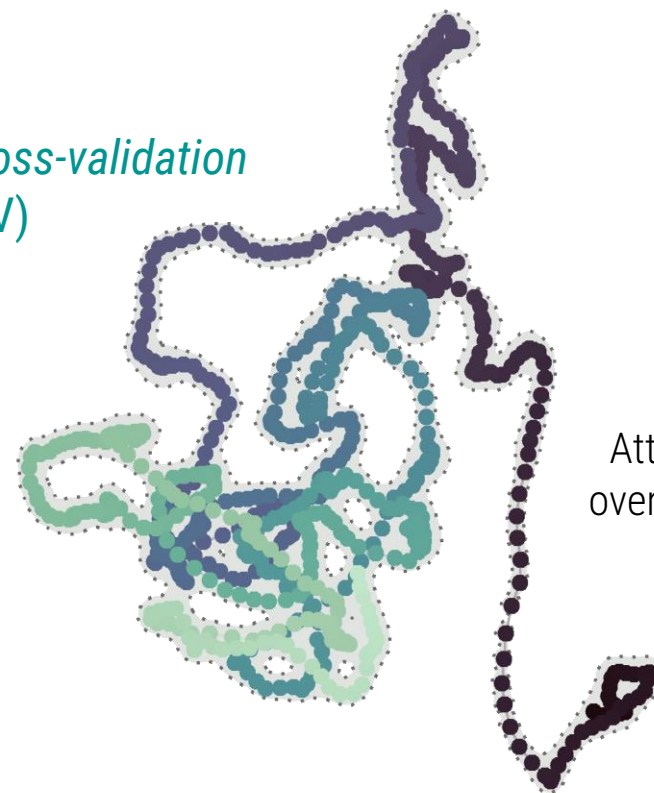
- ▶ Assumes locations are **independent**;
- ▶ Sensitive to **bandwidth** selection.



Minimum Convex Polygon (MCP)
Kernel density estimator (KDE)

*Least-squares cross-validation
(LSCV)*

- ▶ Assumes locations are **independent**;
- ▶ Sensitive to **bandwidth selection**.



ℙ **Rodgers & Kie (2010)**

Tracking time: 
Start End

! This algorithm performs poorly with **large sample sizes**, and still does not account for the locations' temporal structure.

Thinning the data...

Fig. Tracking data representing *hourly locations* over *one month*.

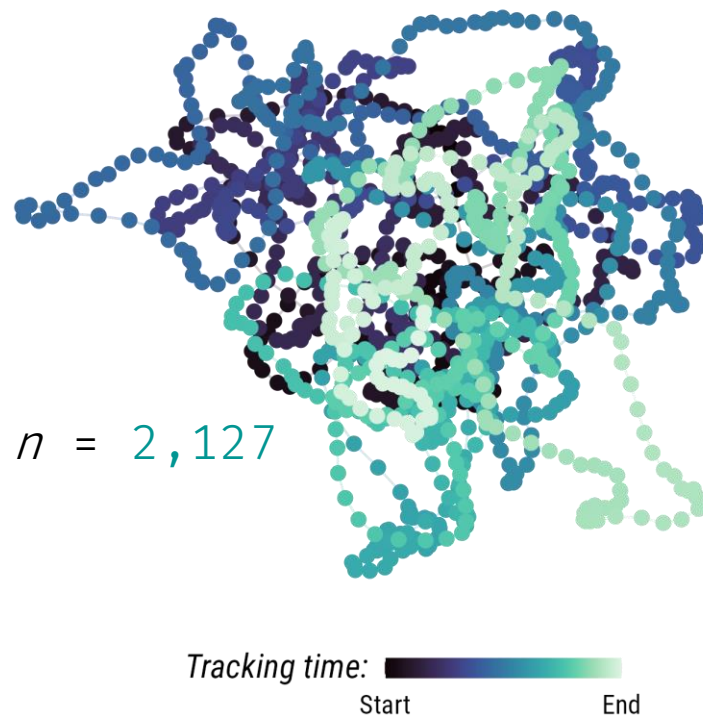
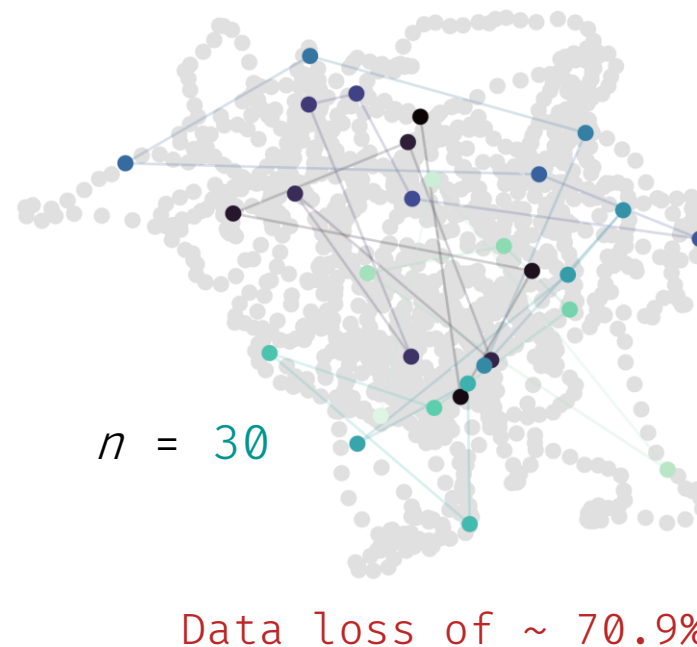


Fig. Tracking data subsampled so there is only *one location per day*.



Thinning the data...

Fig. Tracking data representing *hourly locations* over *one month*.

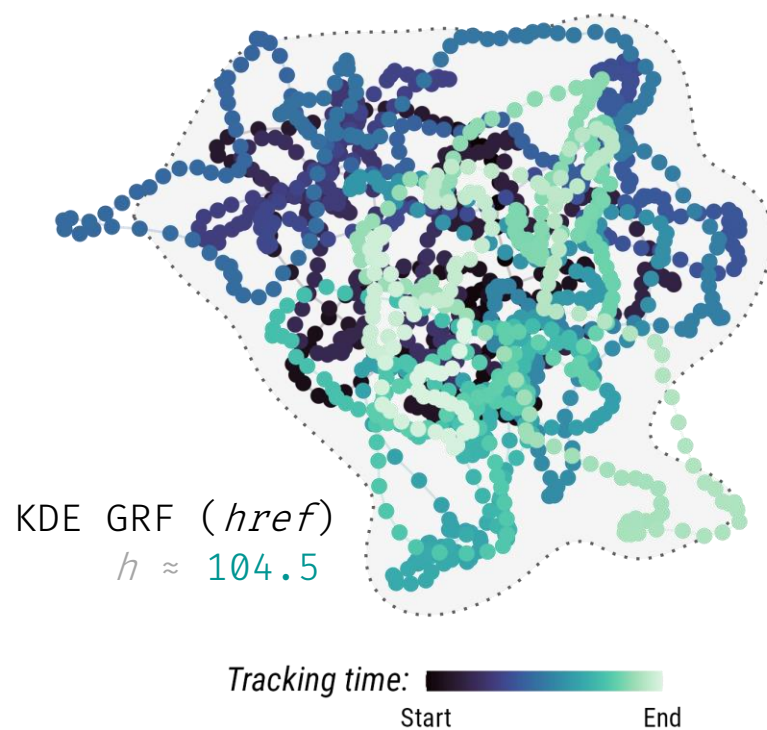
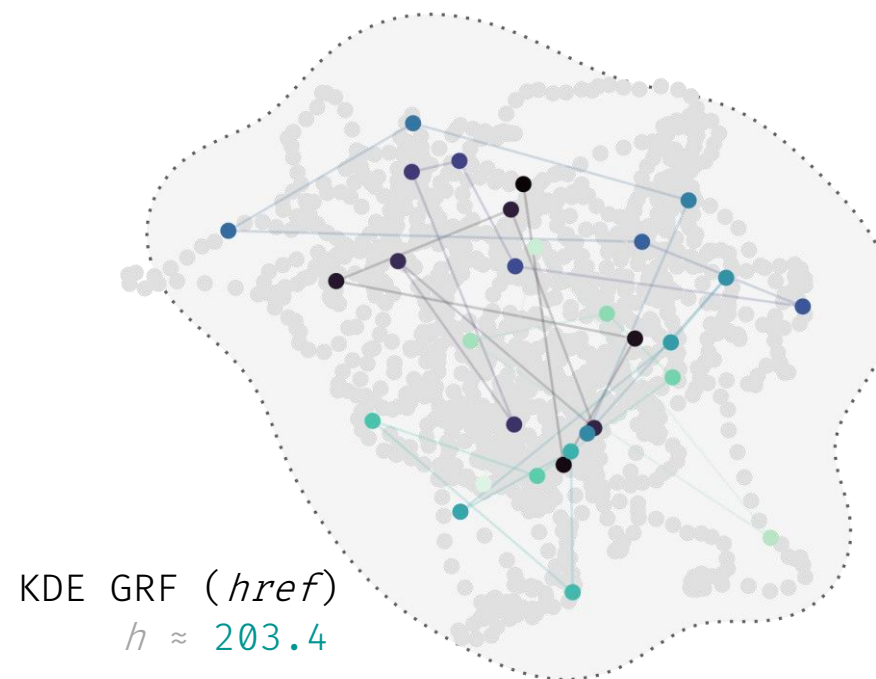


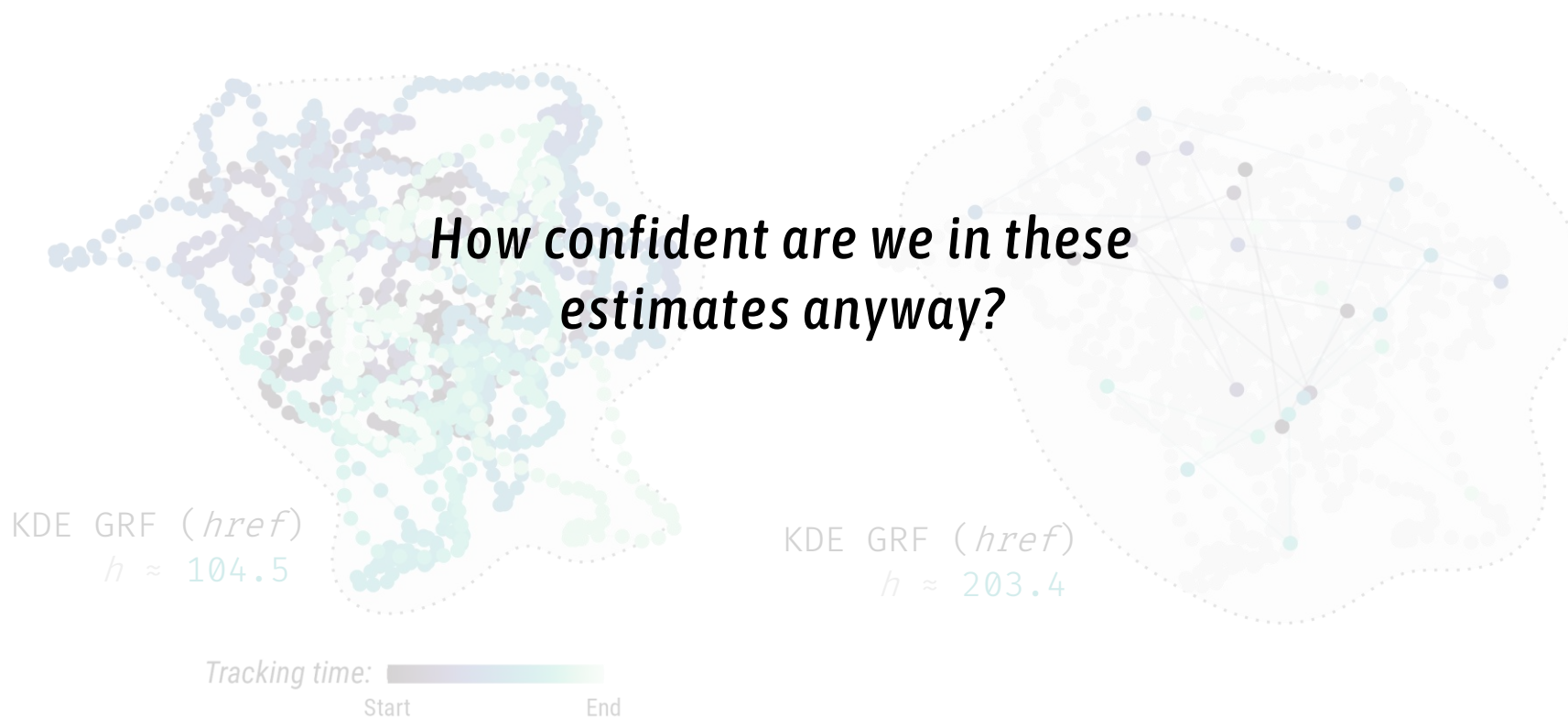
Fig. Tracking data subsampled so there is only *one location per day*.



Thinning the data...

Fig. Tracking data representing *hourly locations* over *one month*.

Fig. Tracking data subsampled so there is only *one location per day*.



Animal movement is *autocorrelated*. (violates independence assumptions)

Fig. Elevation
map of USA.



First Law of Geography

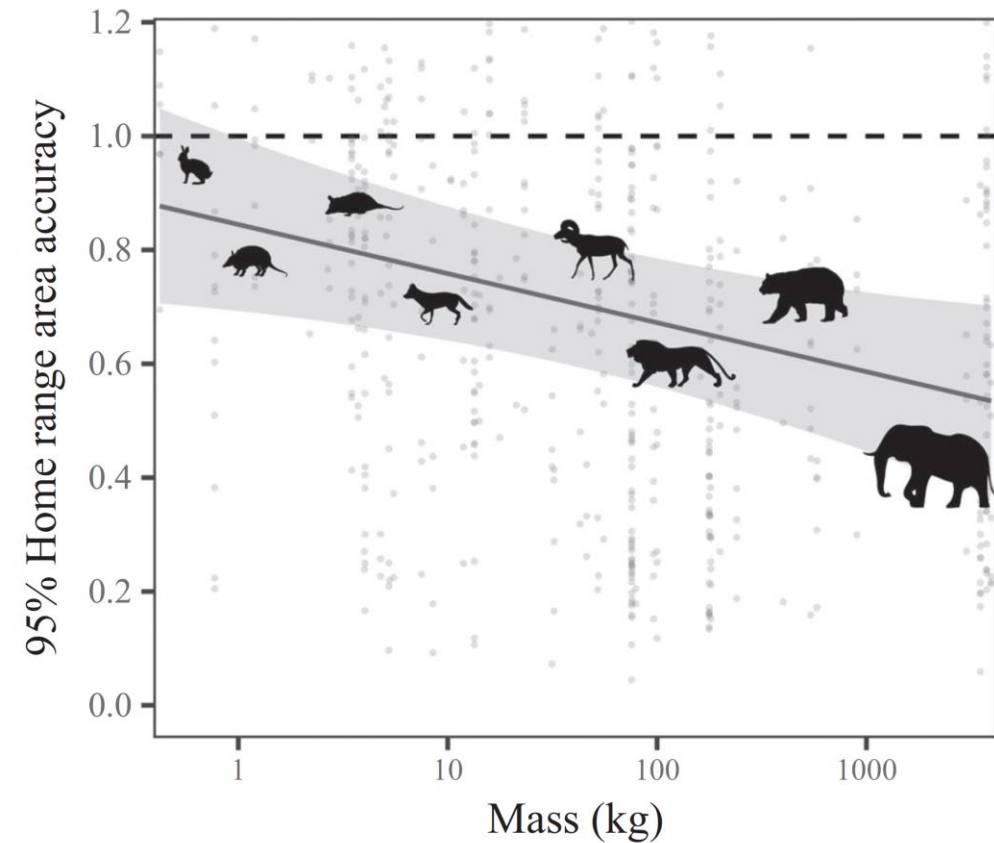
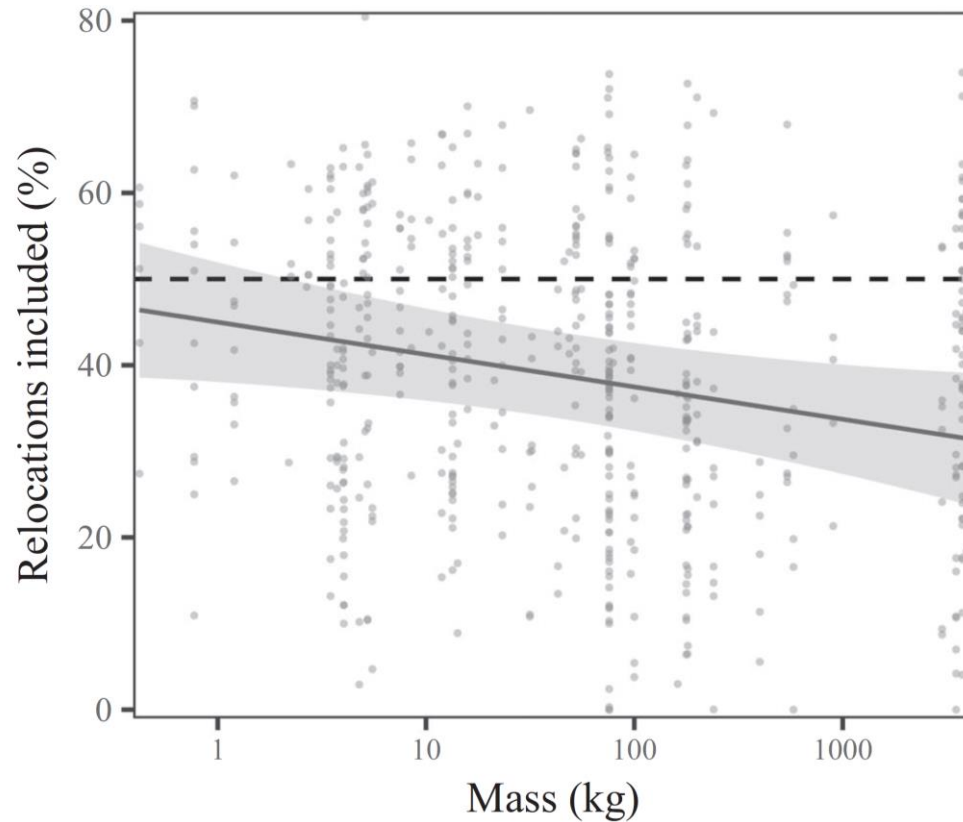
“

Everything is related to everything else, but
near things are more related than distant
things.

✍ Tobler (1970)

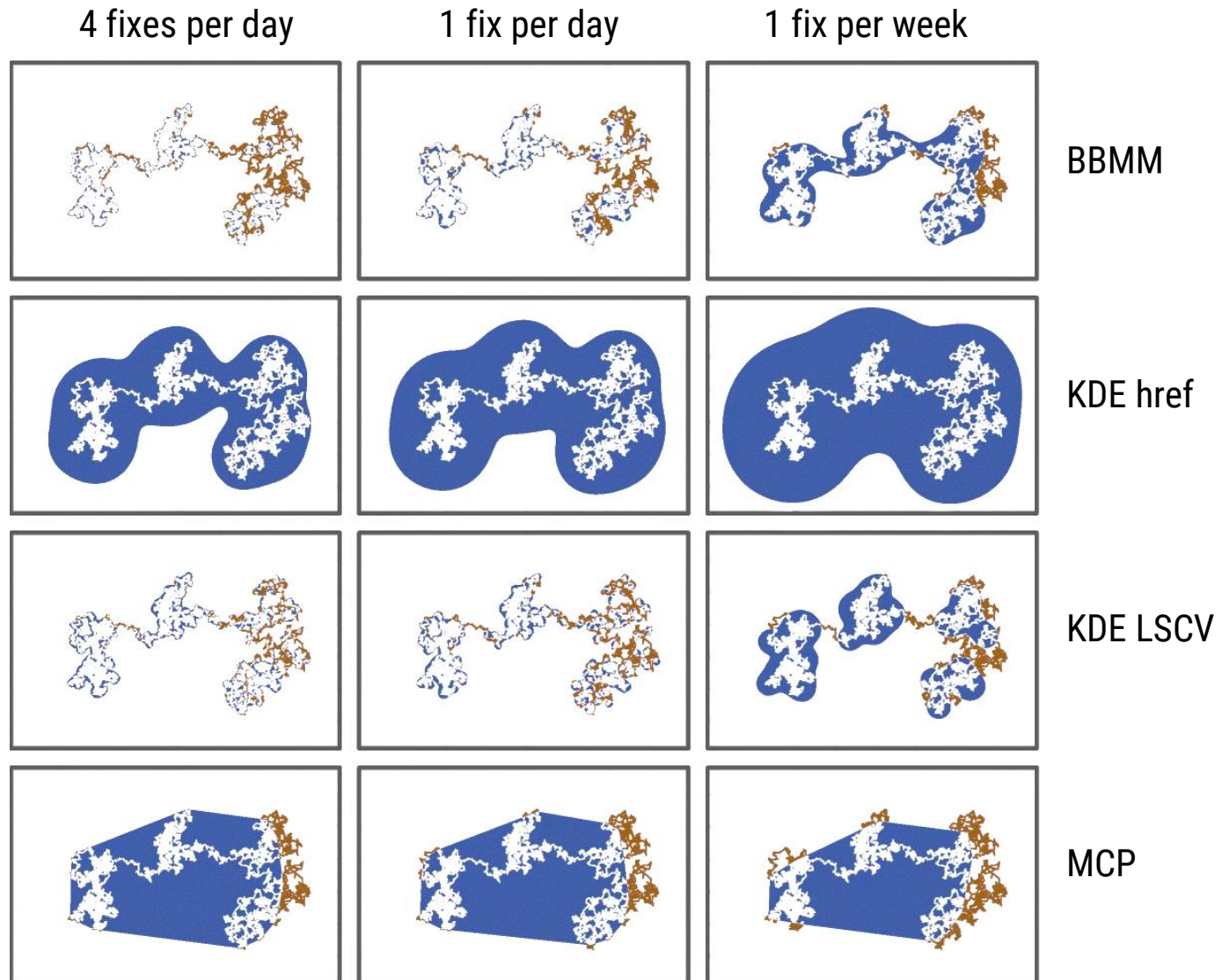
Kernel density estimator (KDE)

The magnitude of KDE's underestimation worsened as body mass increased.



What is each method
actually estimating?

ℓ Silva *et al.* (2021)



Utilization distribution

represents an animal's distribution and the probability of use throughout an area



Range distribution

extrapolate space use
into the future

“How much space does an animal
need over the long term?”

What is an animal's home range area?
What is their the population range area?
Are protected areas sufficiently large?

≠

Occurrence distribution

interpolate between data
points in the past

“Where did an animal go
during a period of observation?”

Where did an animal cross a linear feature?
How likely is it to visit a location of interest?
How much time did it spend in a specific habitat?

Utilization distribution

represents an animal's distribution and the probability of use throughout an area



Range distribution

extrapolate space use
into the future

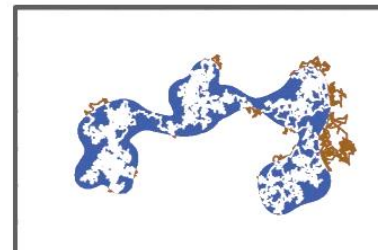
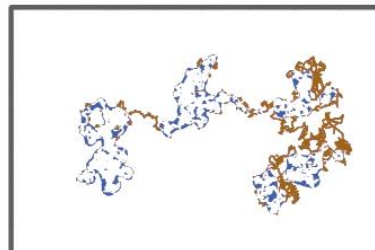
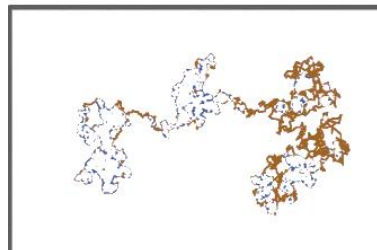
“How much space does an animal
need over the long term?”

What is an animal's home range area?
What is their the population range area?
Are protected areas sufficiently large?

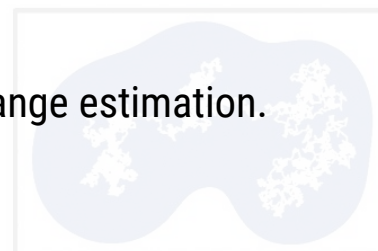
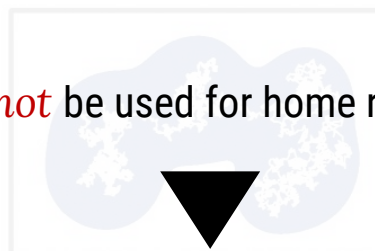
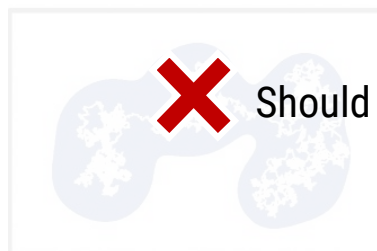
4 fixes per day

1 fix per day

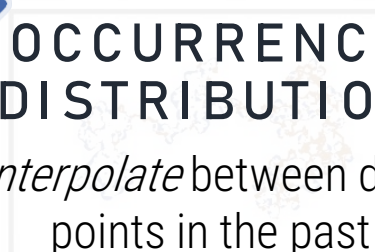
1 fix per week



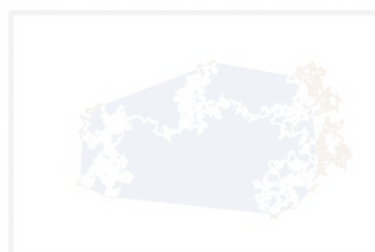
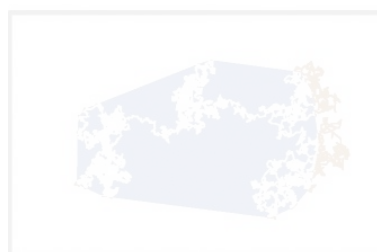
BBMM



KDE href



KDE LSCV



MCP

**However, not all
methods are
appropriate...**

Should **not** be used for home range estimation.

**OCCURRENCE
DISTRIBUTION**

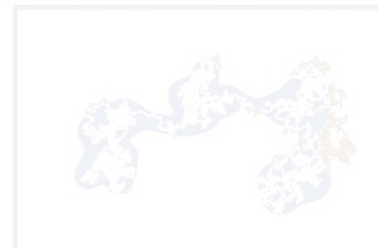
interpolate between data
points in the past

📧 Silva *et al.* (2021)

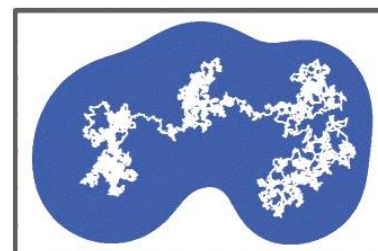
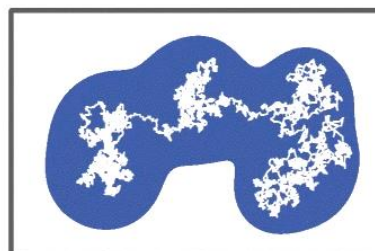
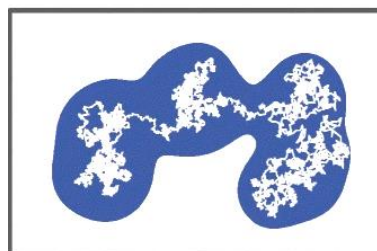
4 fixes per day

1 fix per day

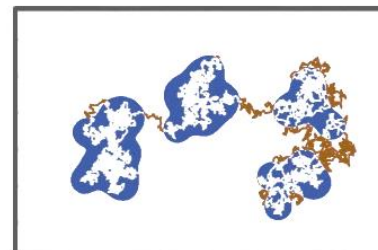
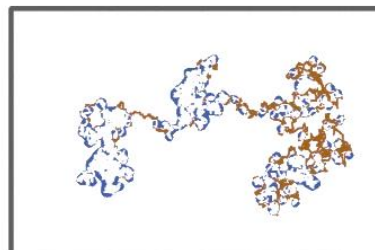
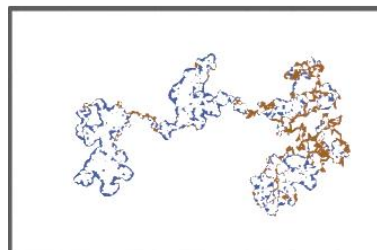
1 fix per week



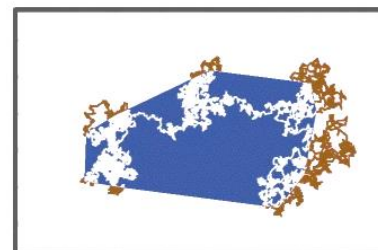
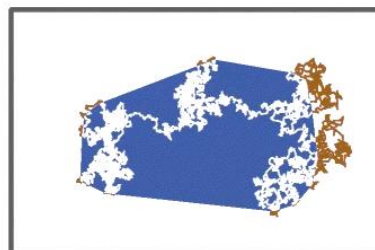
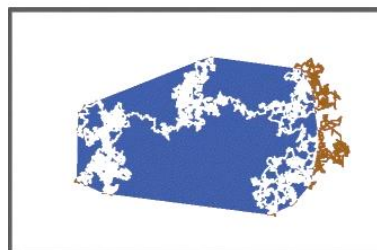
BBMM



KDE href
Kernel density estimator



KDE LSCV
Kernel density estimator



MCP
Minimum Convex Polygons

**However, not all
methods are
appropriate...**

 Silva *et al.* (2021)

Kernel density estimator (KDE)

KDE is the most statistically efficient **non-parametric distribution estimator**.

[non-parametric: not explicitly modeling the causes of space use]

The objective is typically to minimize the '**mean integrated square error**' (**MISE**):

$$\text{MISE}(H) = E \left[\iint (\hat{p}(x, y|H) - p(x, y))^2 dx dy \right]$$

with respect to the **bandwidth** or **smoothing**, ***h***



h is not a model parameter— there is no true value of *h* that best characterizes your animal!
You do not choose your bandwidth. ***You choose your bandwidth optimizer.***



01

Range residency **assumption**

*Checking if data is from a **range-resident** animal*

02

Movement models

*Selecting the best-fit movement model through **model selection***

03

Home range estimation

*Reconstructing **range distribution** from sampled locations*

04

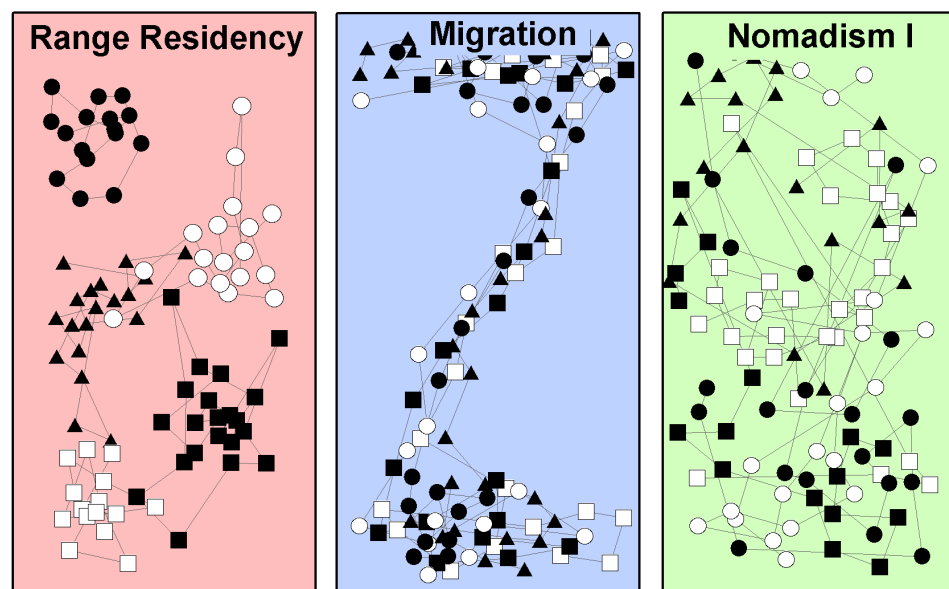
Mitigation measures

*Accounting for common **biases** in animal movement data*



There are three different behaviors that animals exhibit:

- ▶ **Resident** – individual occupies the same area throughout its lifetime.
- ▶ **Migratory** – regular movement to and from spatially disjoint ranges.
- ▶ **Nomadic** – does not follow regular temporal and spatial patterns.

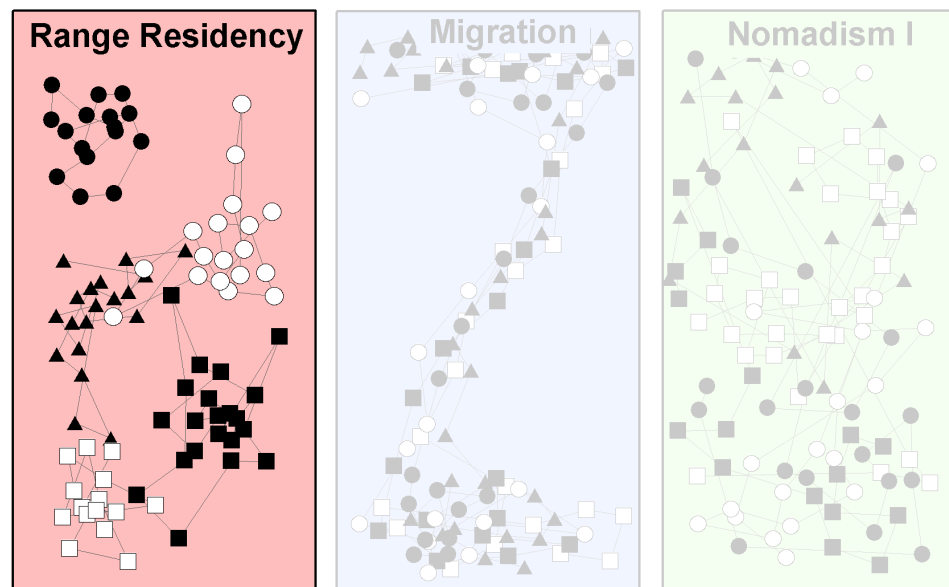


ℓ Mueller *et al.* (2008)



There are three different behaviors that animals exhibit:

- ▶ **Resident** – individual occupies the same area throughout its lifetime.
- ▶ **Migratory** – regular movement to and from spatially disjoint ranges.
- ▶ **Nomadic** – does not follow regular temporal and spatial patterns.

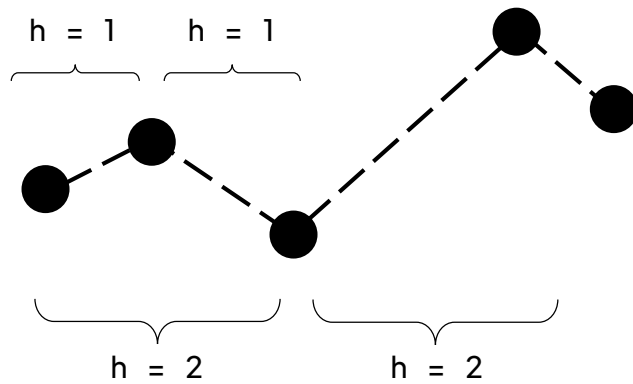


ℓ Mueller *et al.* (2008)

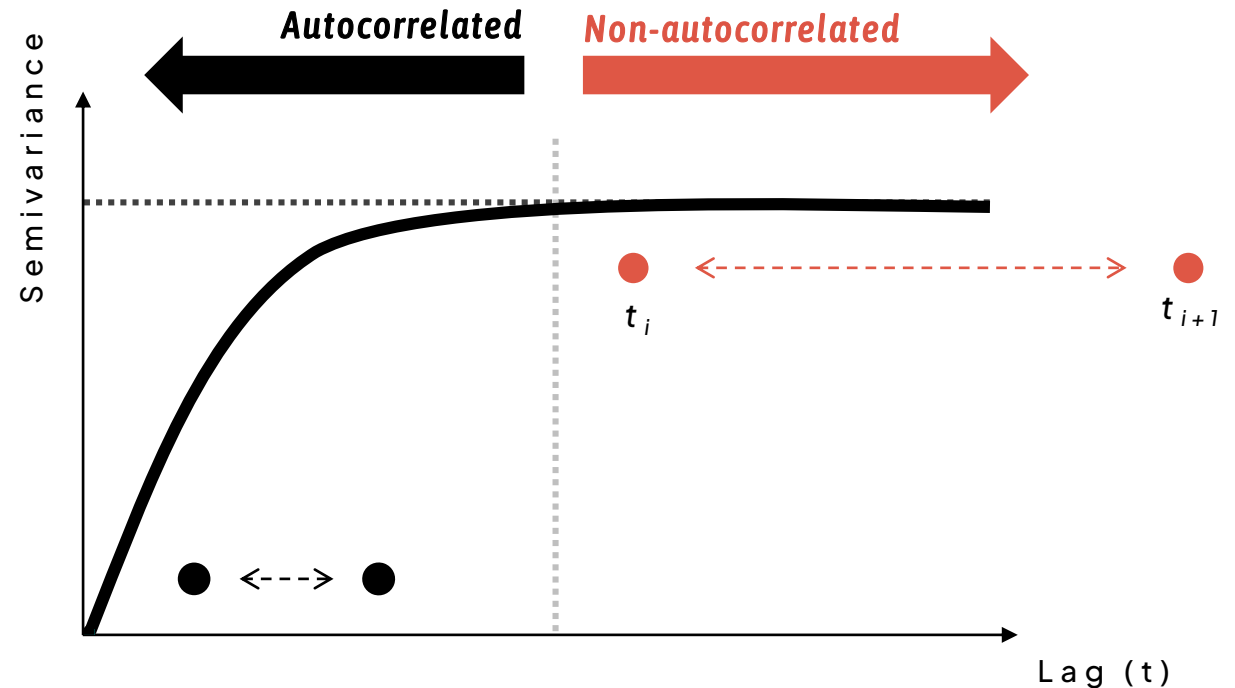


How can we measure and visualize autocorrelation?

Variogram, or **semivariogram**, plots time lags on the **x-axis** for all pairs of observations against their **semivariance** (average square distance between any two locations with a given lag) on the **y-axis**.



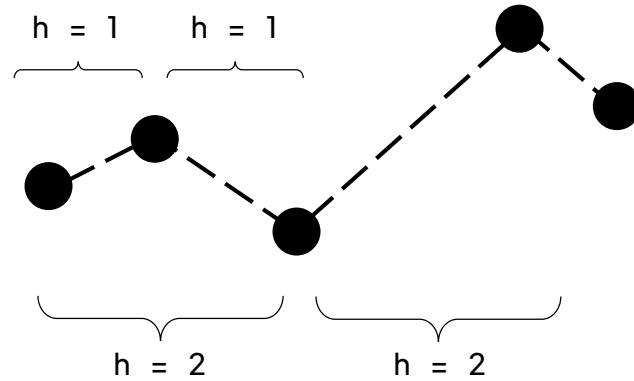
The more similar the pairs of locations are per lag, the lower the **semivariance** for that lag.



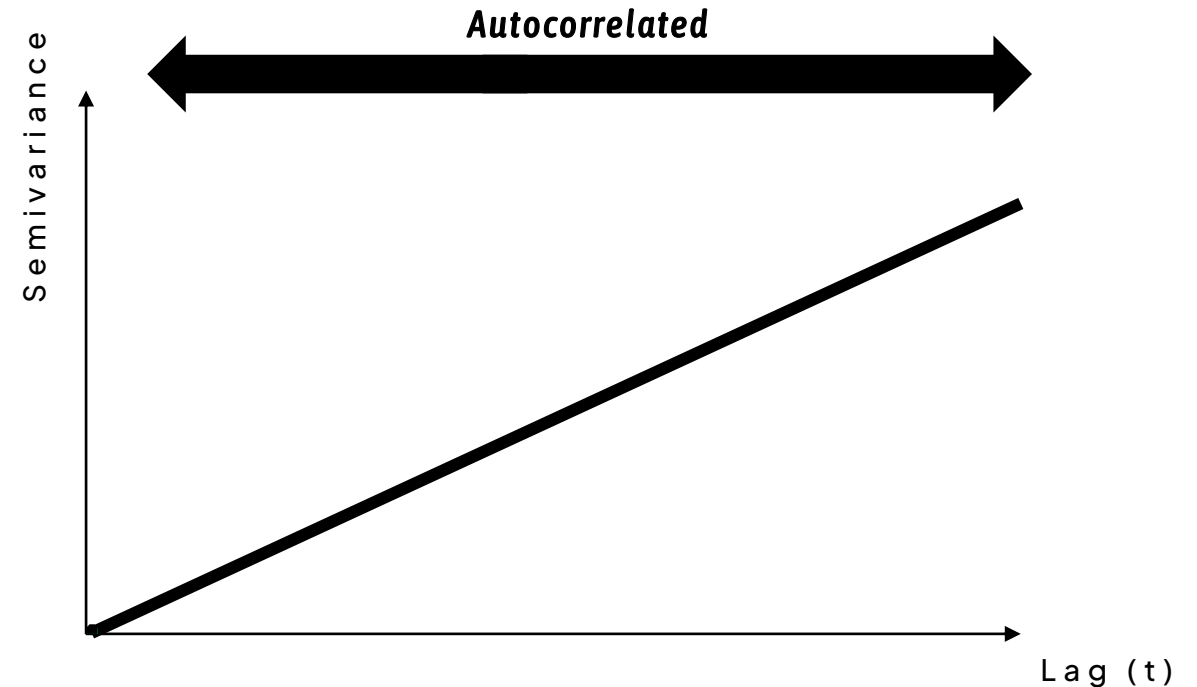
lag at which there is no evidence of dependence

How can we measure and visualize **autocorrelation**?

Variogram, or **semivariogram**, plots time lags on the **x-axis** for all pairs of observations against their **semivariance** (average square distance between any two locations with a given lag) on the **y-axis**.

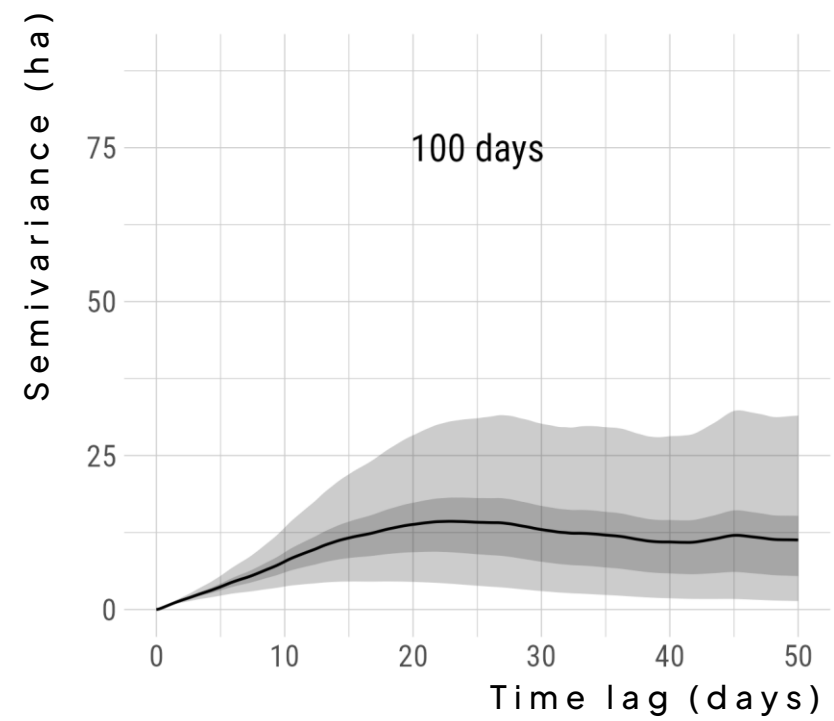
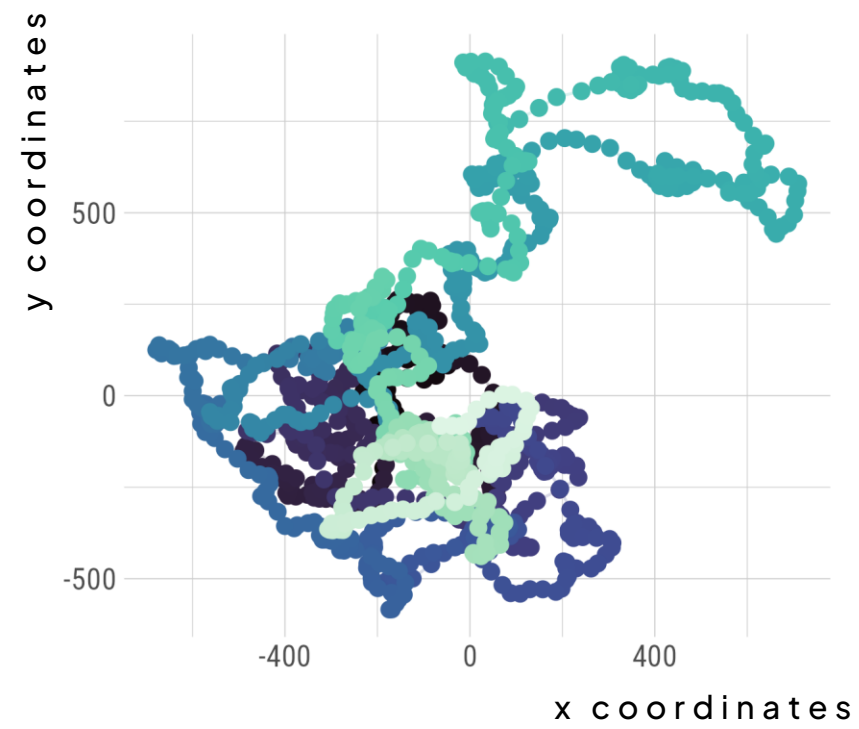


The more similar the pairs of locations are per lag, the lower the **semivariance** for that lag.



How can we measure and visualize autocorrelation?

Tracking time: 
Start End



01

Range residency assumption

Checking if data is from a **range-resident** animal

02

Movement **models**

Selecting the best-fit movement model through **model selection**

03

Home range estimation

Reconstructing **range distribution** from sampled locations

04

Mitigation measures

Accounting for common **biases** in animal movement data

CONVENTIONAL METHODS

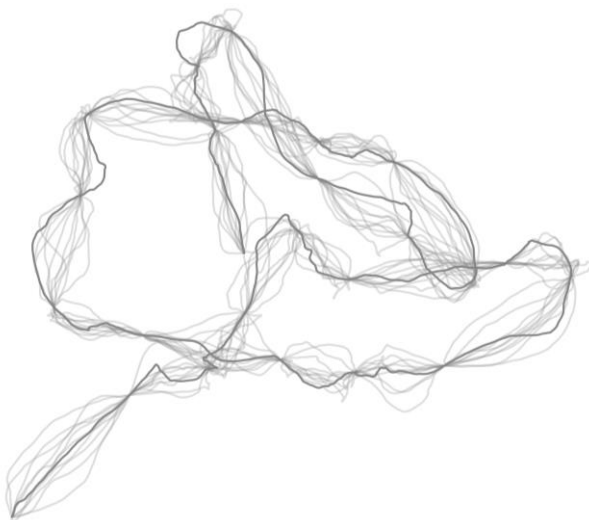


Assuming *independent locations*.

CONTINUOUS TIME METHODS



What *movement process* explains a particular animal movement dataset?



Current movement models available:

- 3.1. Independent and Identically Distributed (IID)
- 3.2. Brownian Motion (BM)
- 3.3. Ornstein-Uhlenbeck (OU)
- 3.4. Integrated Ornstein-Uhlenbeck (IOU)
- 3.5. Ornstein-Uhlenbeck with Foraging (OUF)

3.1. Independent and Identically Distributed (IID)

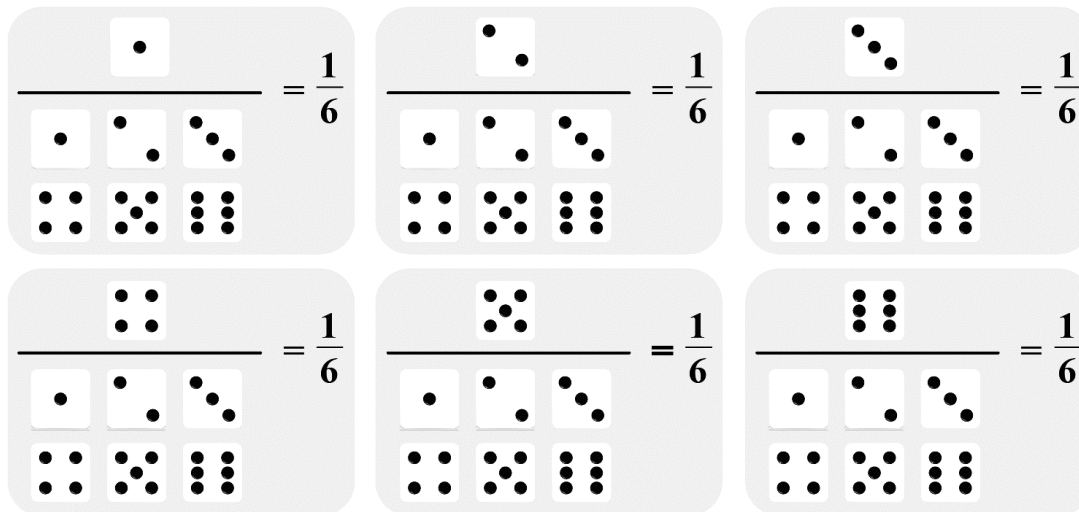


Stochastic process where each location has the same probability distribution as all others, and all are *mutually independent*.

SPATIAL DEPENDENCY

TEMPORAL DEPENDENCY

RESTRICTED



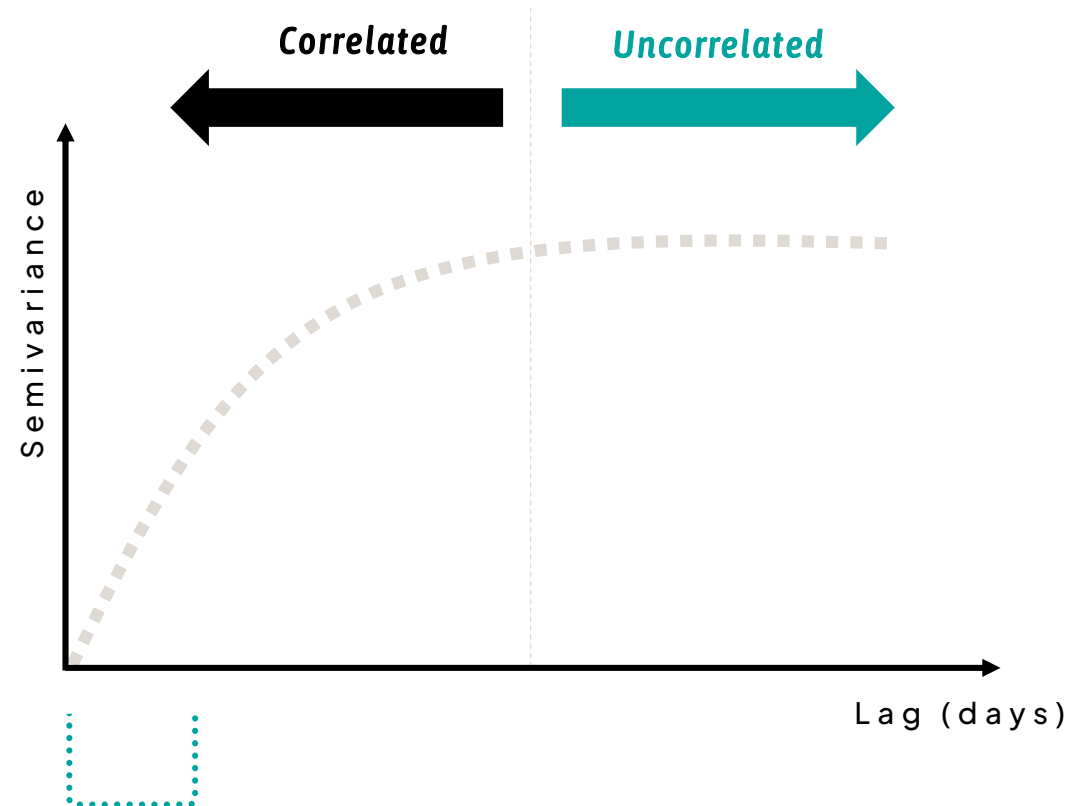
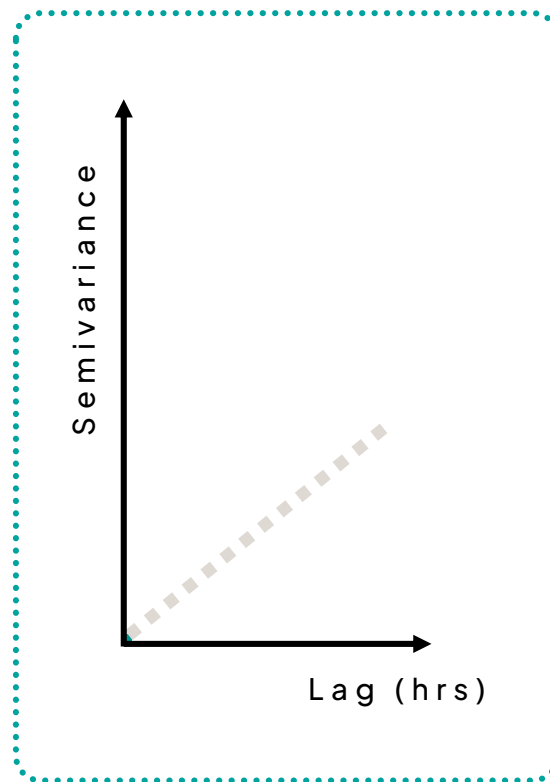
For example,



Dice rolls are *independent and identically distributed* (IID)

3.1. Independent and Identically Distributed (IID)

- How would the variogram of a **IID process** look like?



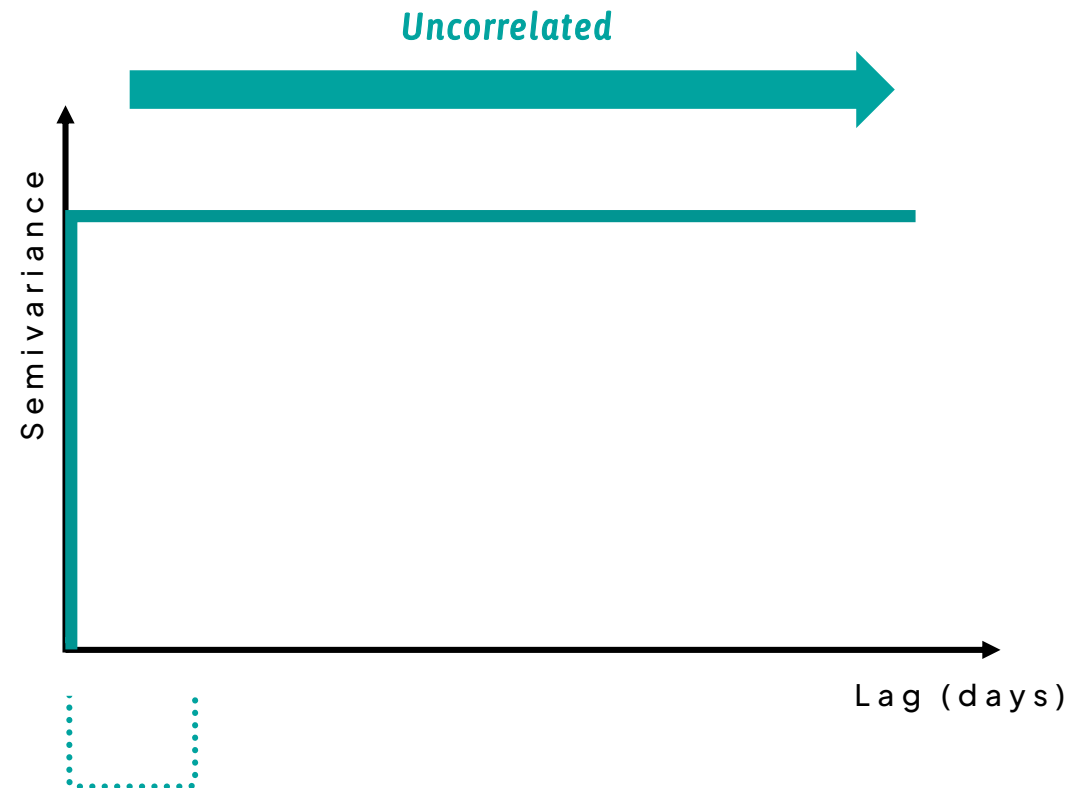
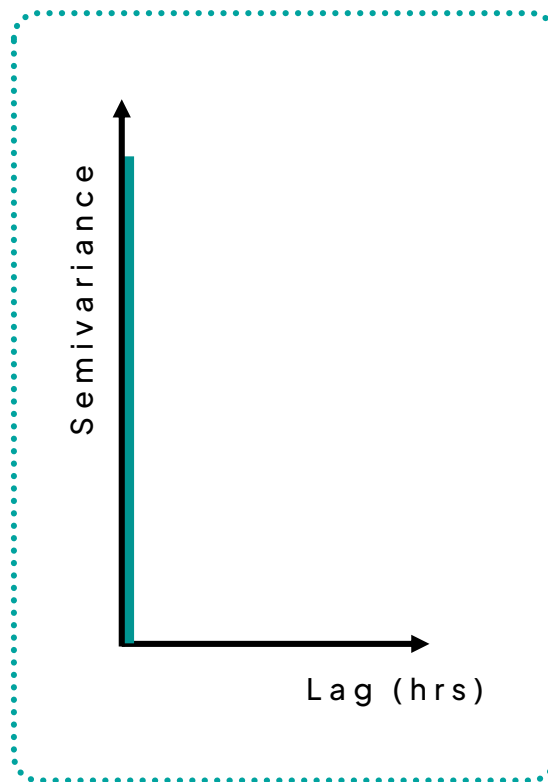
SPATIAL DEPENDENCY

TEMPORAL DEPENDENCY

RESTRICTED

3.1. Independent and Identically Distributed (IID)

- How would the variogram of a IID process look like?



SPATIAL DEPENDENCY

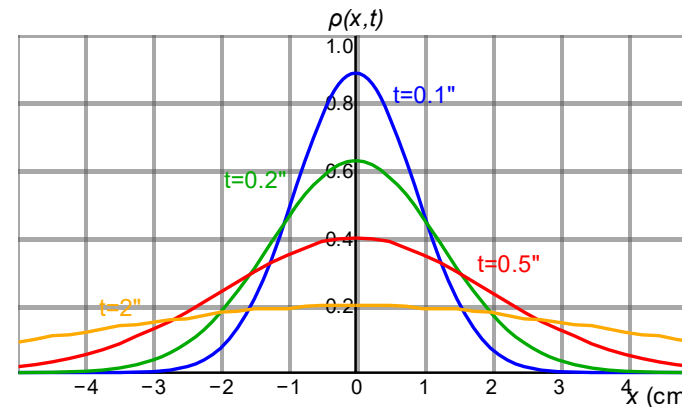
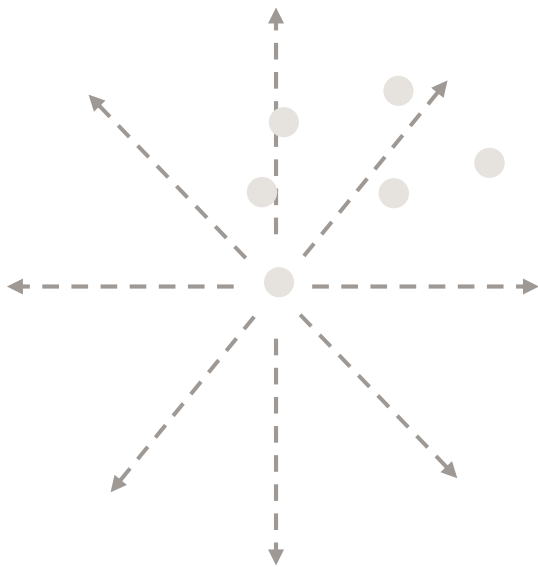
TEMPORAL DEPENDENCY

RESTRICTED

3.2. Brownian Motion (BM)



Stochastic process with stationary and **independent** increments, *i.e.*, no “memory” – the future behavior of a Brownian motion process **does not depend on its past**. Diffusion is **constant**.



As time goes on, the animal is more likely to be further away from its starting location.

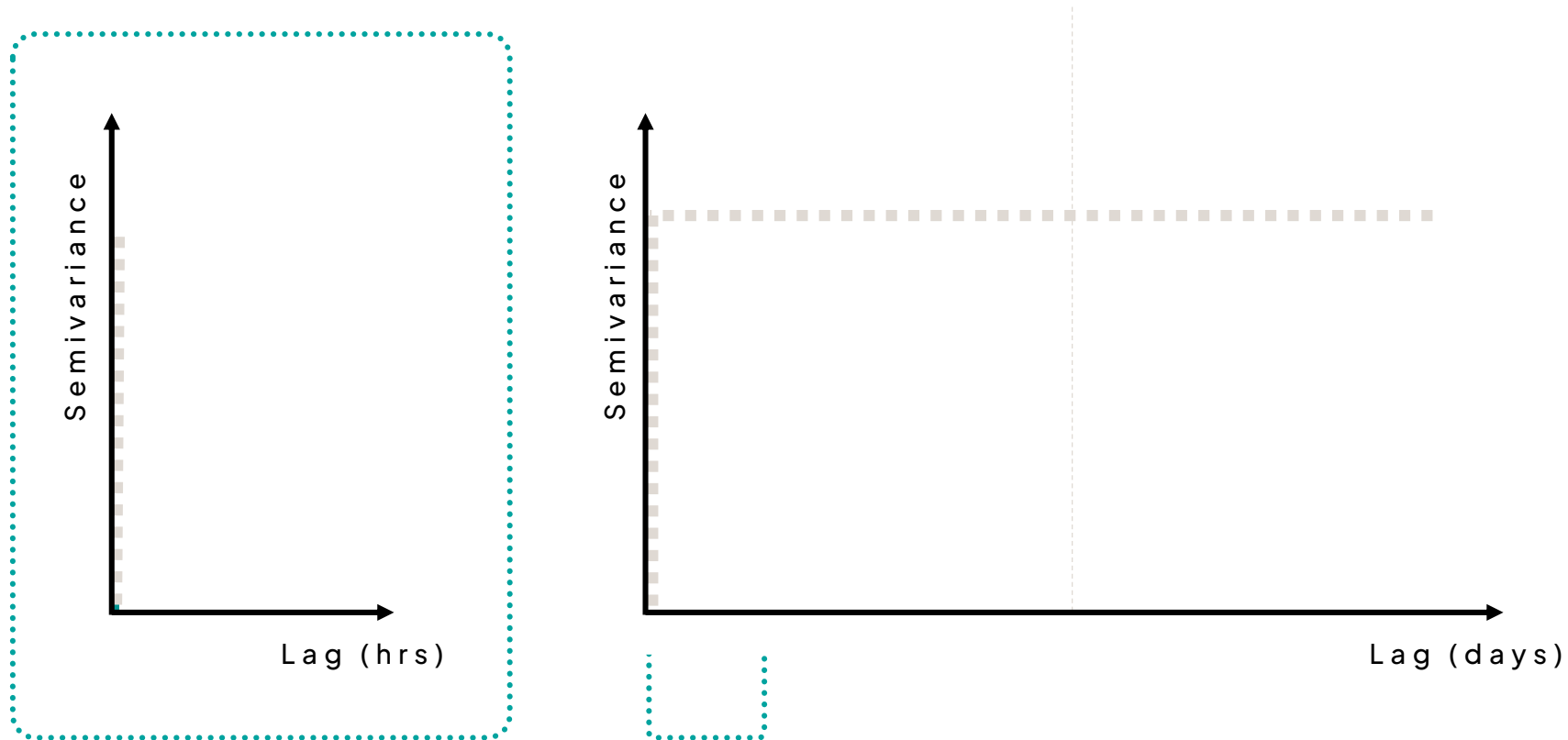
SPATIAL DEPENDENCY

TEMPORAL DEPENDENCY

RESTRICTED

3.2. Brownian Motion (BM)

- How would the variogram of a BM process look like?



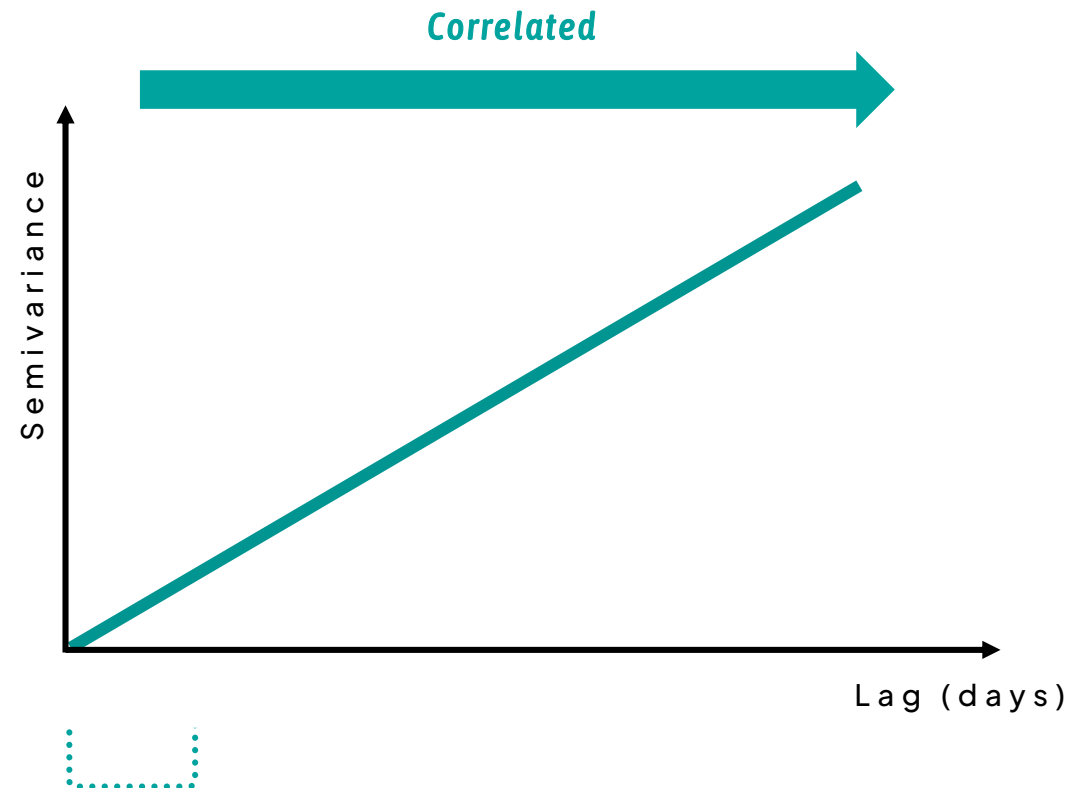
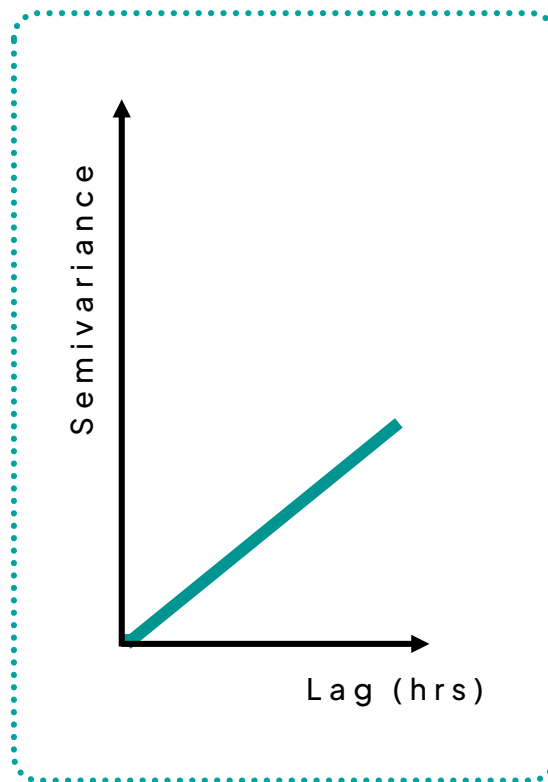
SPATIAL DEPENDENCY

TEMPORAL DEPENDENCY

RESTRICTED

3.2. Brownian Motion (BM)

- How would the variogram of a BM process look like?



SPATIAL DEPENDENCY

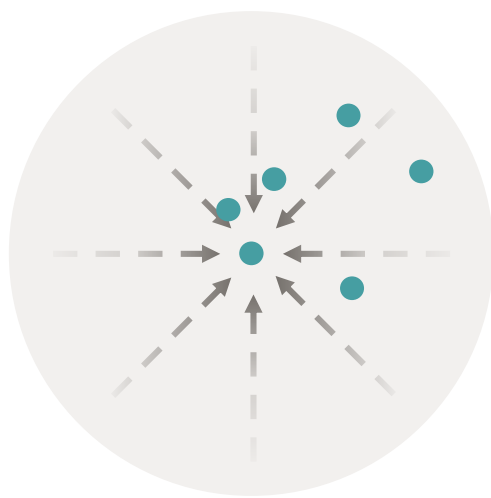
TEMPORAL DEPENDENCY

RESTRICTED

These processes are all modifications of Brownian motion:

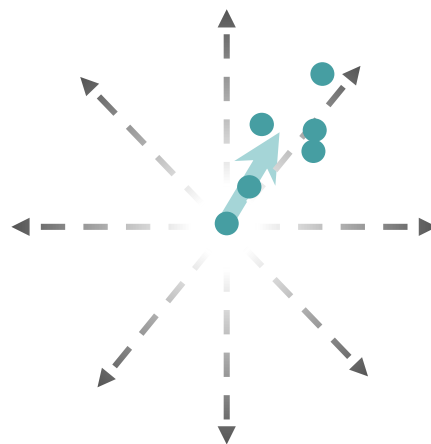
3.3. Ornstein-Uhlenbeck (OU)

Unlike Brownian motion, OU tends towards a central location, with greater attraction the further away from the center (bounded).



3.4. Integrated OU (IOU)

Like Brownian motion, the integrated OU process exhibits unbounded diffusion, but with persistence of motion.



3.5. OU with Foraging (OUF)

Unlike Brownian motion, OUF is bounded, but with persistence of motion.

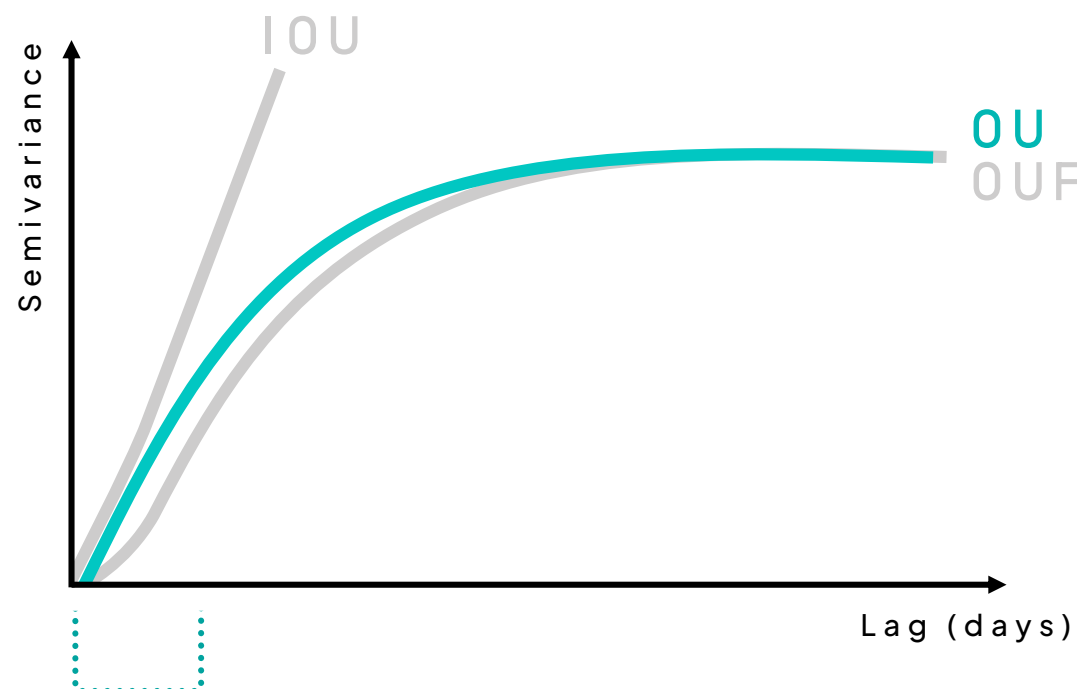
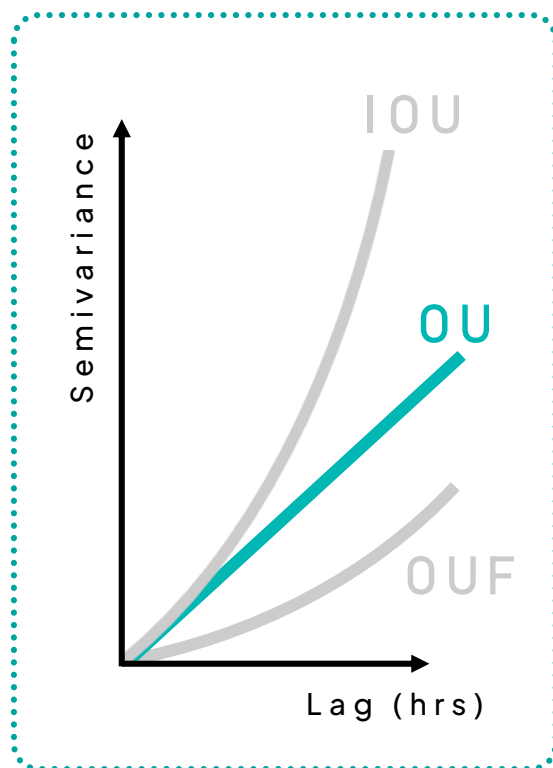


3.3. Ornstein-Uhlenbeck (OU)

3.4. Integrated OU (IOU)

3.5. OU with Foraging (OUF)

- How would the variograms of **OU processes** look like?



SPATIAL DEPENDENCY

TEMPORAL DEPENDENCY

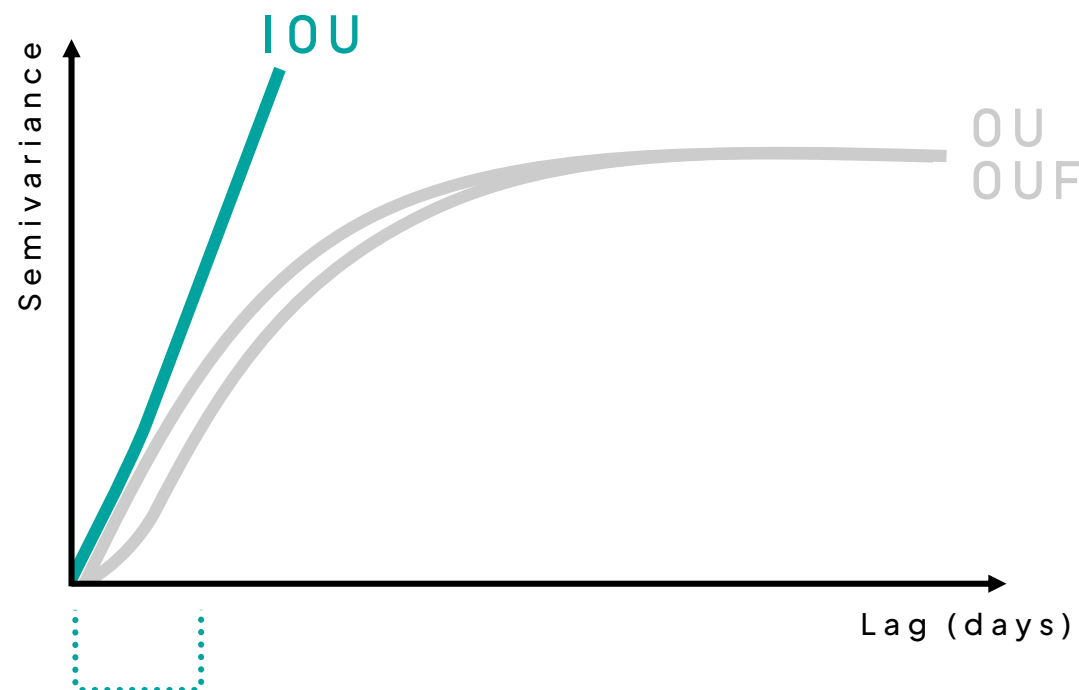
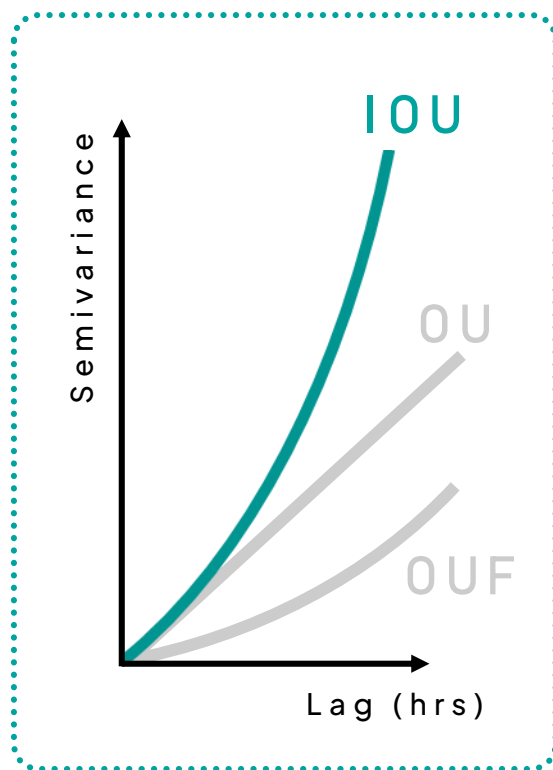
RESTRICTED

3.3. Ornstein-Uhlenbeck (OU)

3.4. Integrated OU (IOU)

3.5. OU with Foraging (OUF)

- How would the variograms of **OU processes** look like?



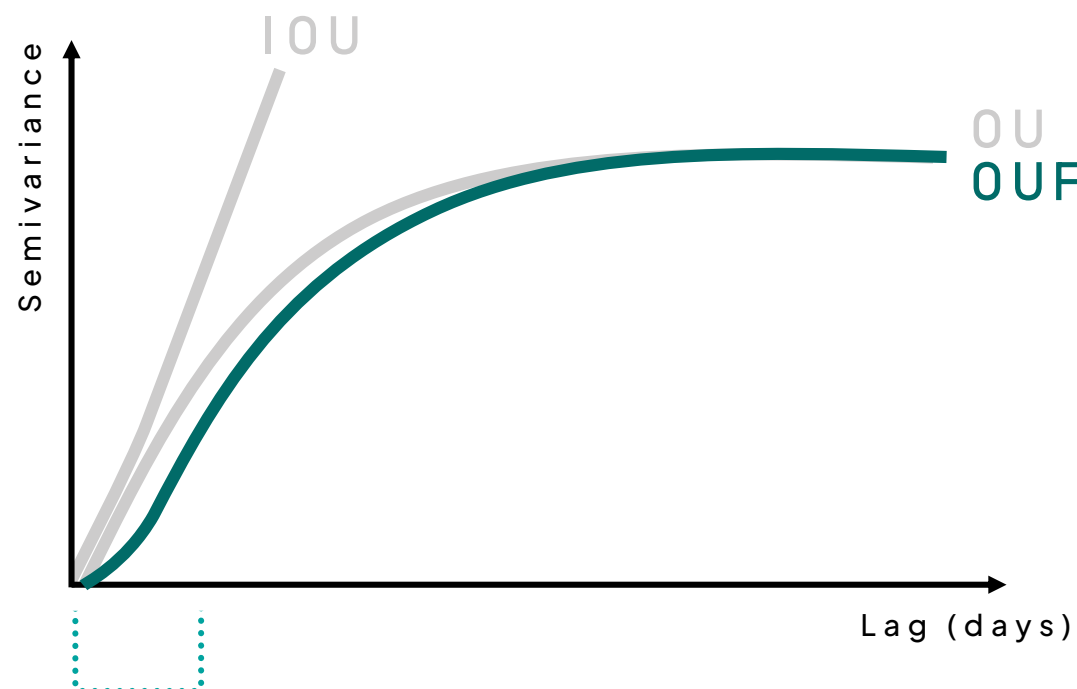
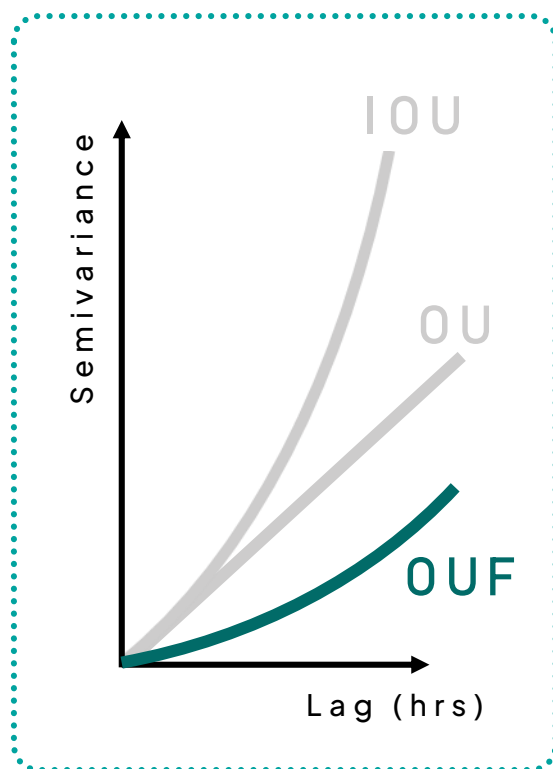
SPATIAL DEPENDENCY
TEMPORAL DEPENDENCY
RESTRICTED

3.3. Ornstein-Uhlenbeck (OU)

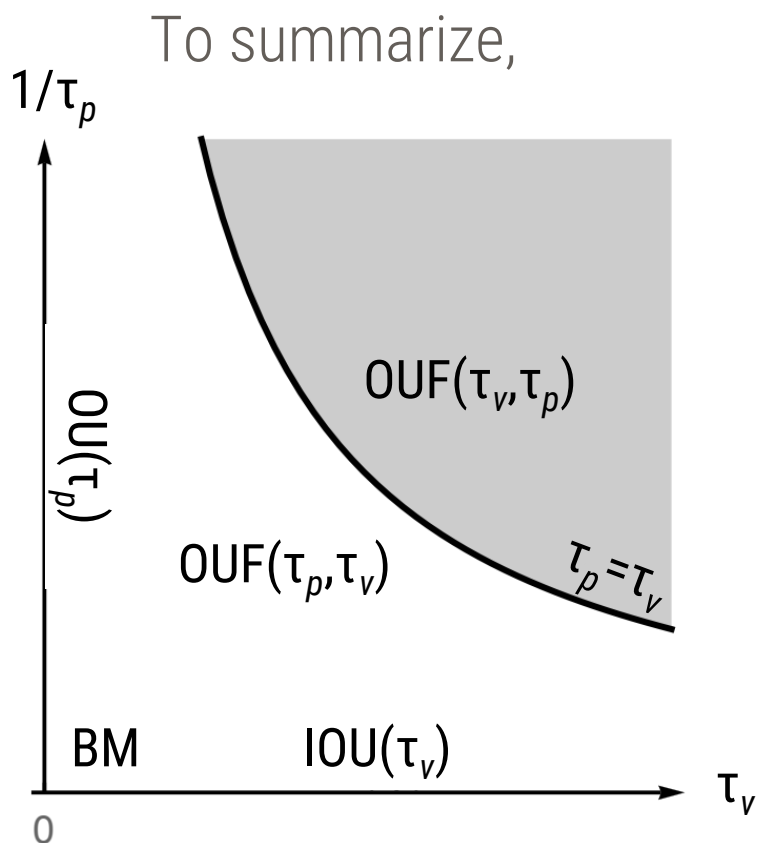
3.4. Integrated OU (IOU)

3.5. OU with Foraging (OUF)

- How would the variograms of **OU processes** look like?



SPATIAL DEPENDENCY
TEMPORAL DEPENDENCY
RESTRICTED

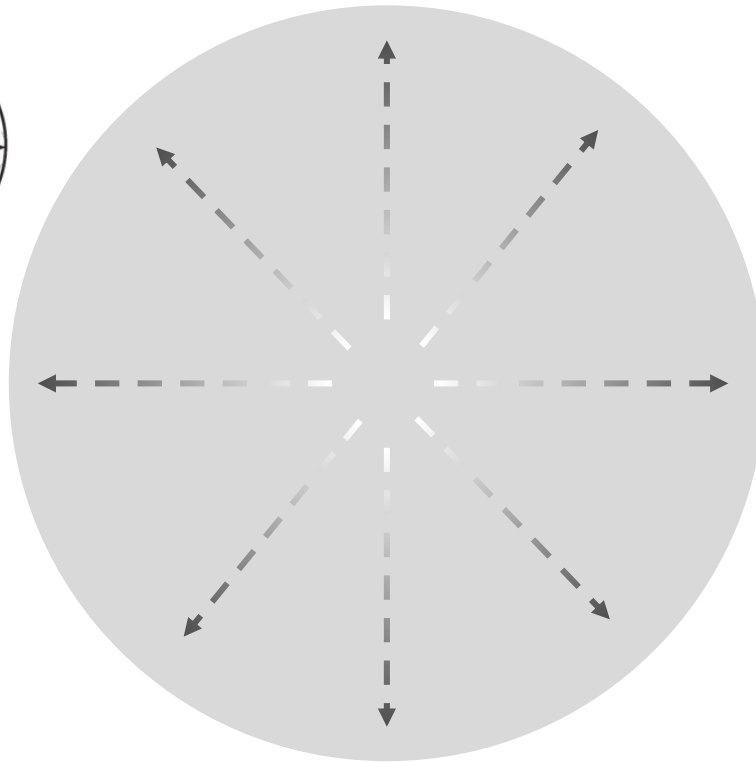
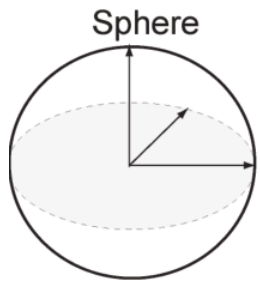


Model	Autocorrelation			Parameters:
	Position	Velocity	Restricted	
IID	No	No	Yes	$\tau = \text{NULL}$
BM	Yes	No	No	$\tau = \infty$
OU	Yes	No	Yes	$\tau = \tau_p$
IOU	Yes	Yes	No	$\tau = \{\infty, \tau_v\}$
OUF	Yes	Yes	Yes	$\tau = \{\tau_p, \tau_v\}$

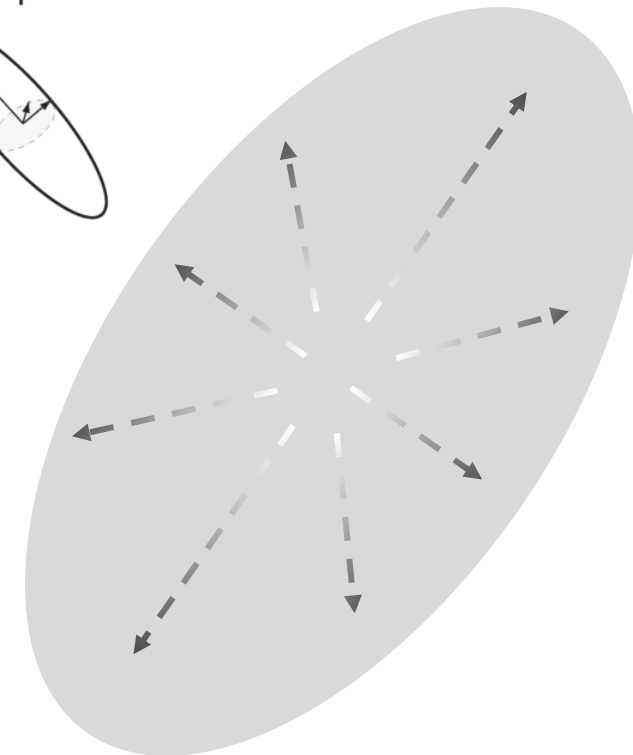
“ All models are wrong, but some are useful.

Box *et al.* (1987)

Isotropic refers to the properties of a material which is independent of the direction; whereas **anisotropic** is direction-dependent.



Isotropic diffusion



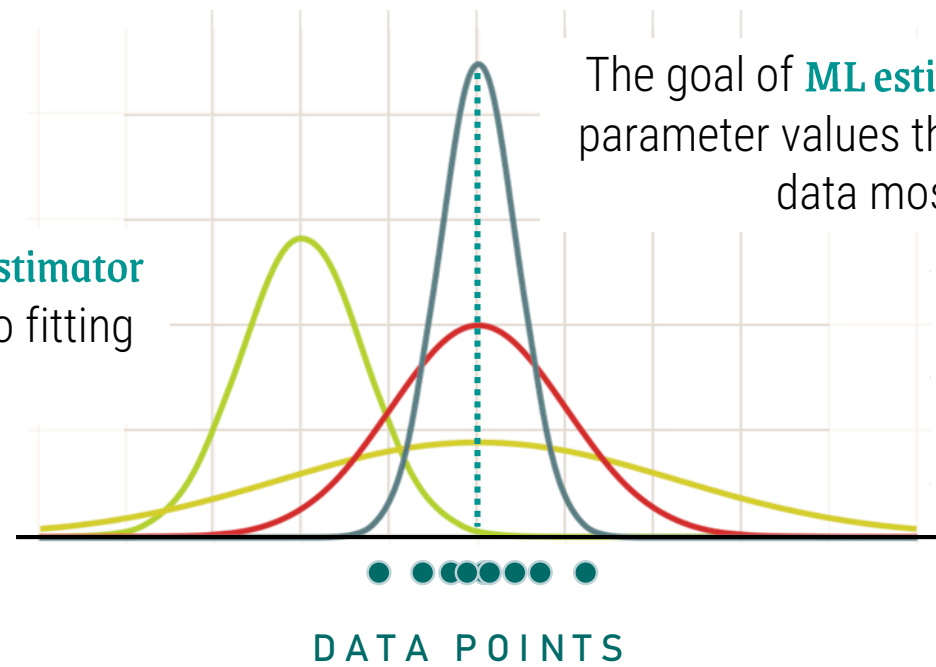
Anisotropic diffusion



What **movement model parameters** are most likely to characterize a given tracking dataset?

MAXIMUM LIKELIHOOD:

Maximum Likelihood (ML) estimator is the standard approach to fitting movement models.



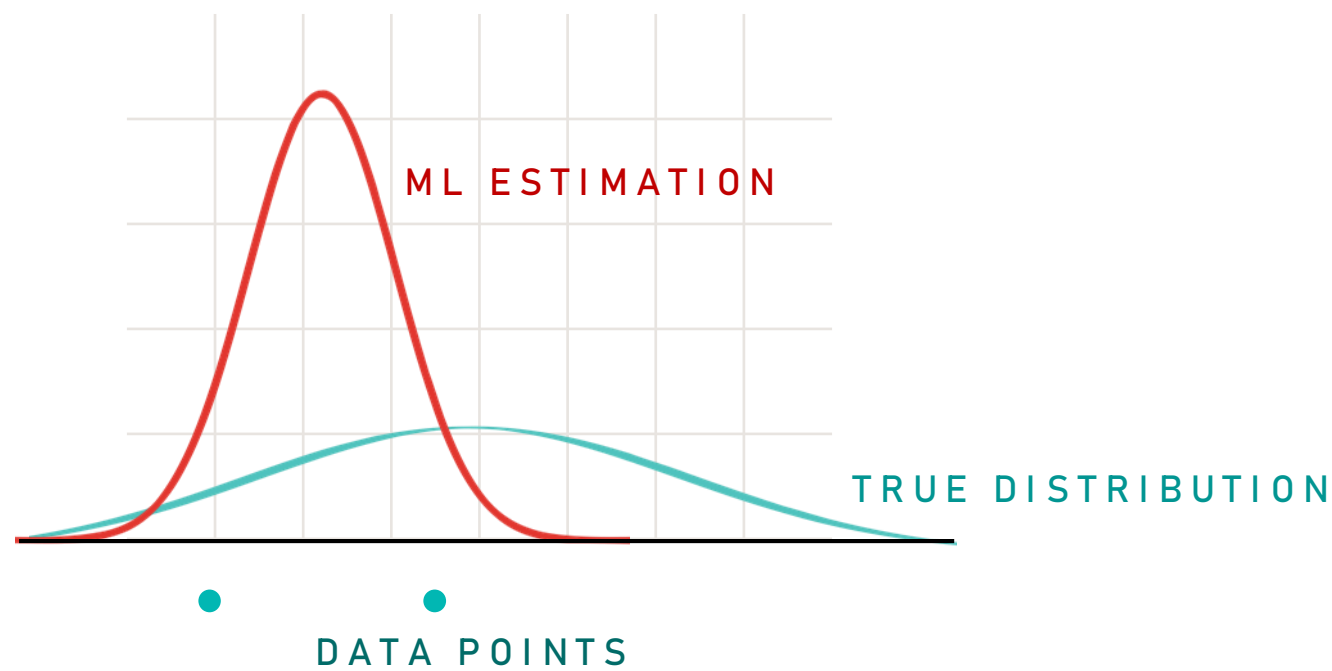


What movement model parameters are most likely to characterize a given tracking dataset?

MAXIMUM LIKELIHOOD:

Unfortunately, ML performs poorly at small sample sizes

(Cressie, 1993)



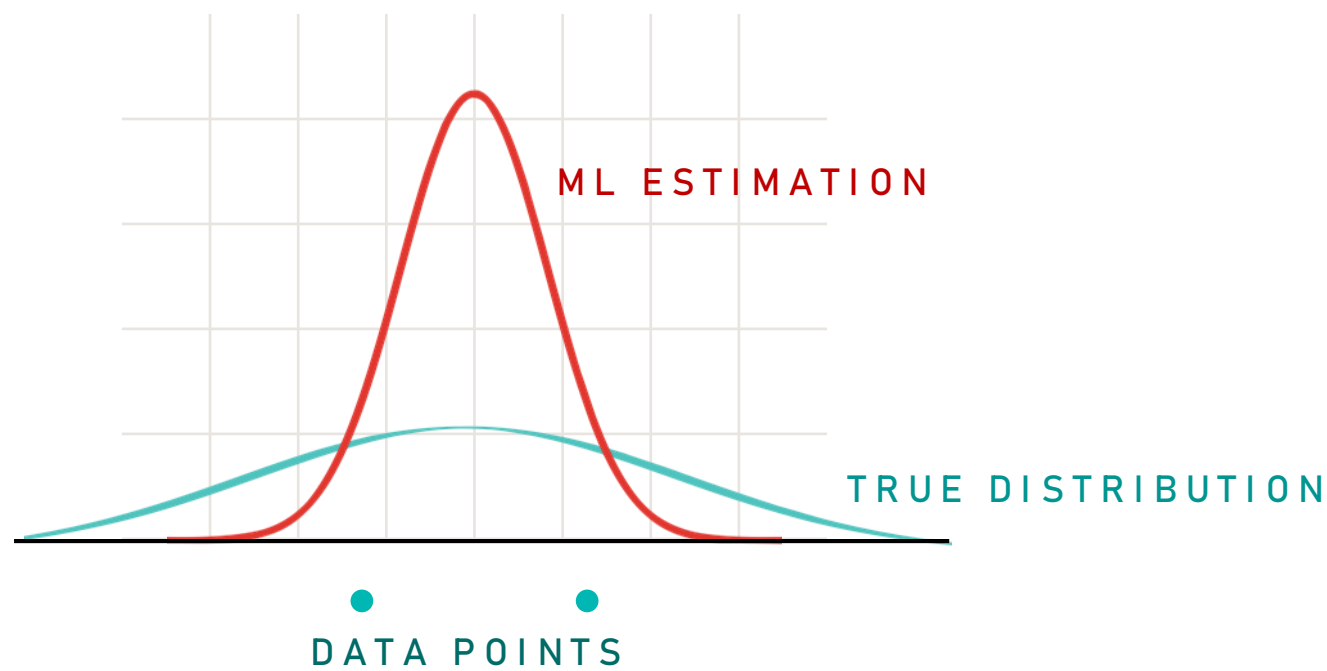


What movement model parameters are most likely to characterize a given tracking dataset?

MAXIMUM LIKELIHOOD:

Unfortunately, ML performs poorly at small sample sizes

(Cressie, 1993)



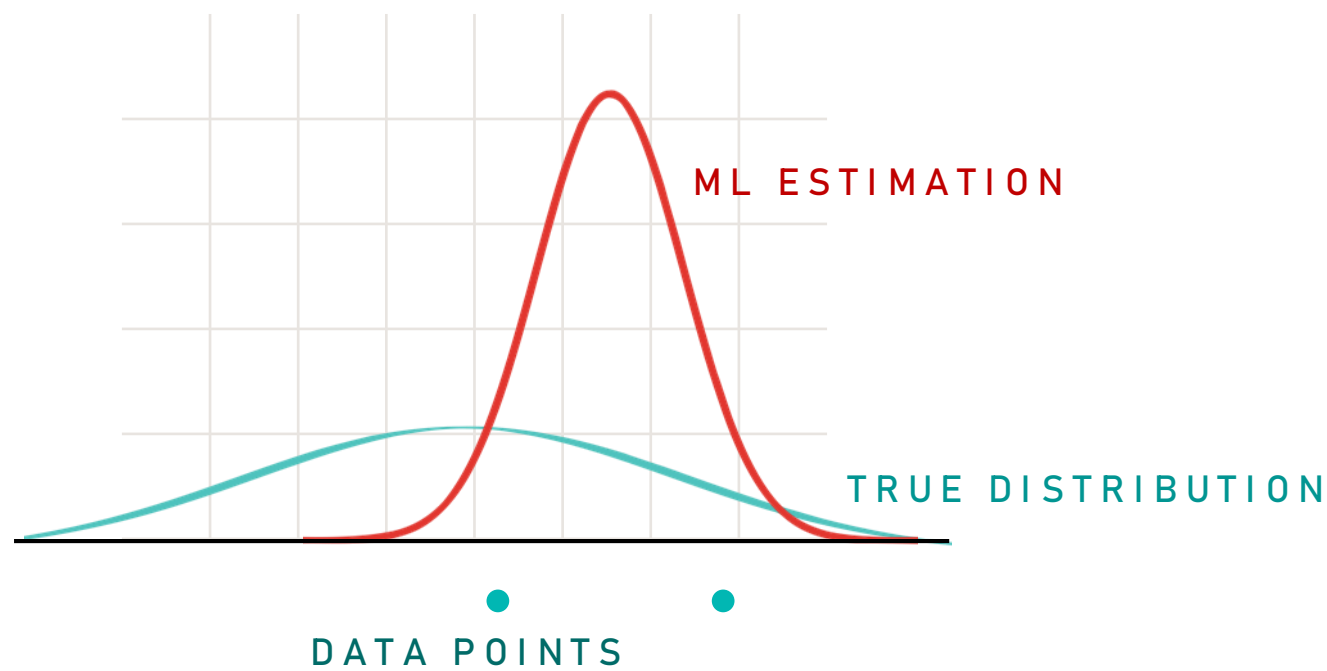


What movement model parameters are most likely to characterize a given tracking dataset?

MAXIMUM LIKELIHOOD:

Unfortunately, ML performs poorly at small sample sizes

(Cressie, 1993)



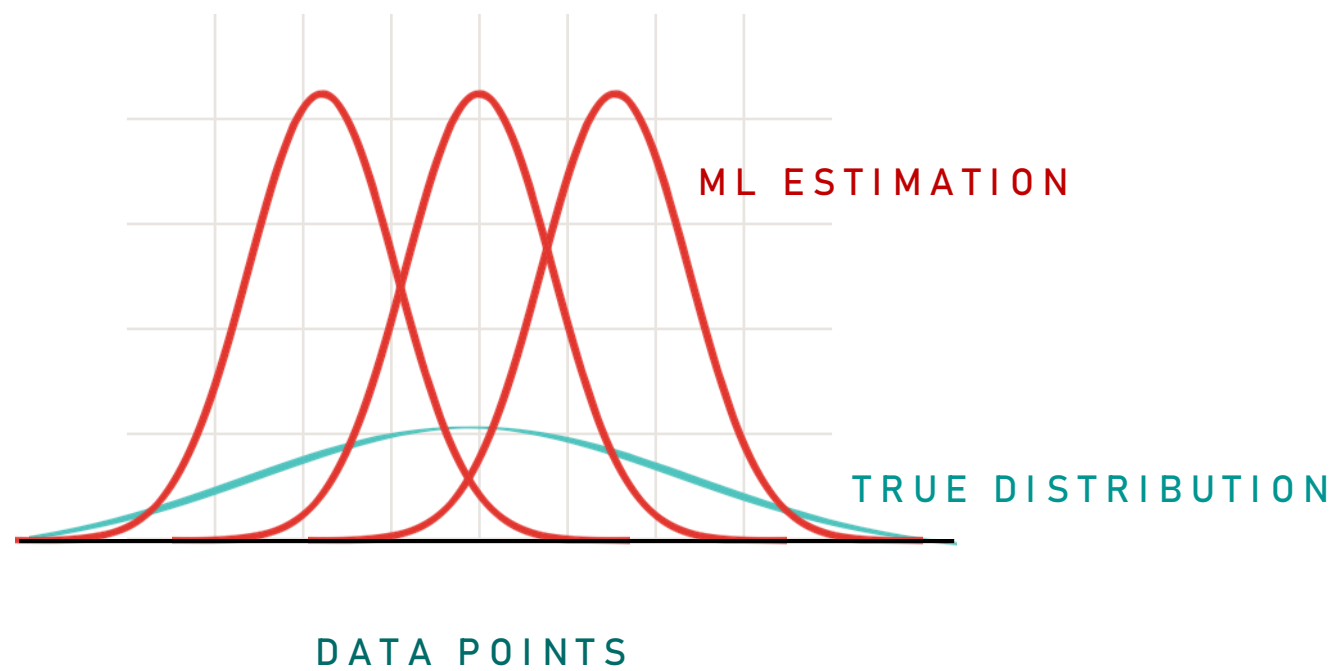


What movement model parameters are most likely to characterize a given tracking dataset?

MAXIMUM LIKELIHOOD:

Unfortunately, ML performs poorly at small sample sizes

(Cressie, 1993)



We can tackle the ML bias in several ways.
For example:

Residual ML (or REML)

Widely used method for reducing bias in ML variance estimation, by maximizing the likelihood of residuals rather than the data.

Essentially, it trades *reduced bias* for *increase variability in parameter estimates*.

(Bartlett, 1937)



It is important to distinguish two **sample size** concepts:

SAMPLE SIZES

Absolute sample size

\neq

Effective sample size

n

T

+

Δt

Sampling duration

Sampling frequency

How long is an animal tracked for?

How frequently are locations collected?



Total number of locations

It is important to distinguish two sample size concepts:

SAMPLE SIZES

Absolute sample size

n

T

+

Δt

Sampling duration

How long is an animal tracked for?

Sampling frequency

How frequently are locations collected?

Total number of locations

\neq

Effective sample size

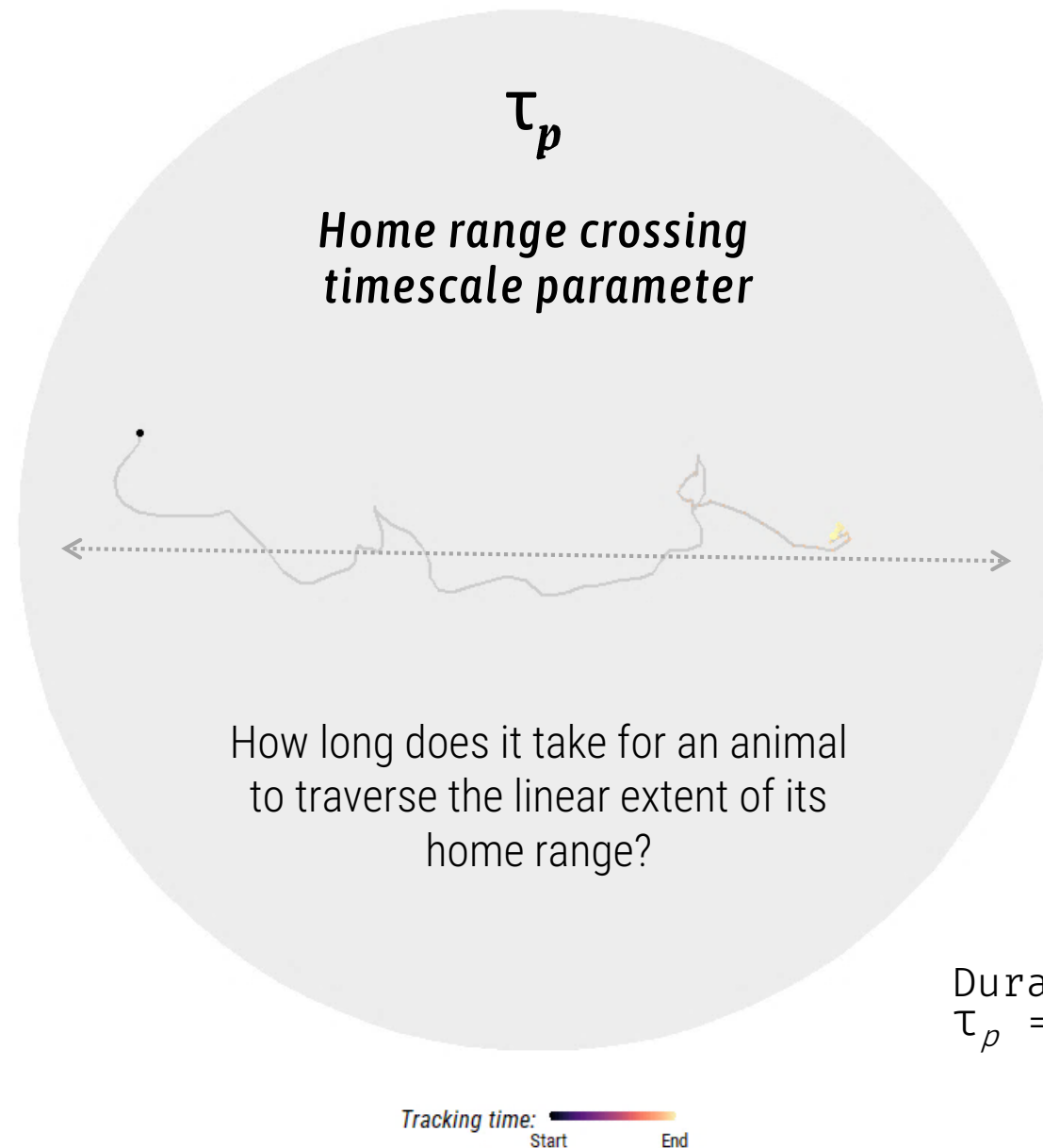
N_{area}

N_{speed}

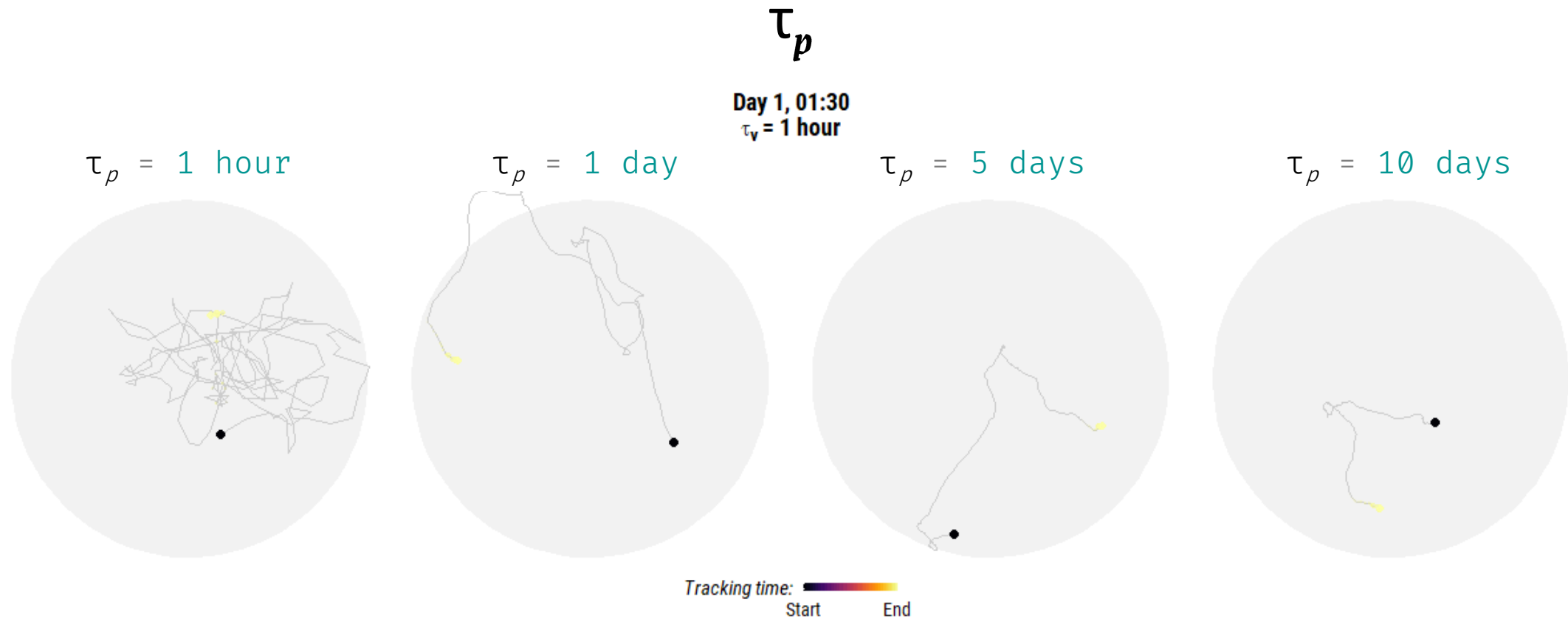
roughly estimated as T/τ_p

T is the sampling duration

τ_p is the average home range crossing time



Duration = 1 day
 $\tau_p = 1$ day



Effective sample size (N) decreases as the *home range crossing time parameter* (τ_p) increases.



LC

Sampling duration = tracked for 389 days

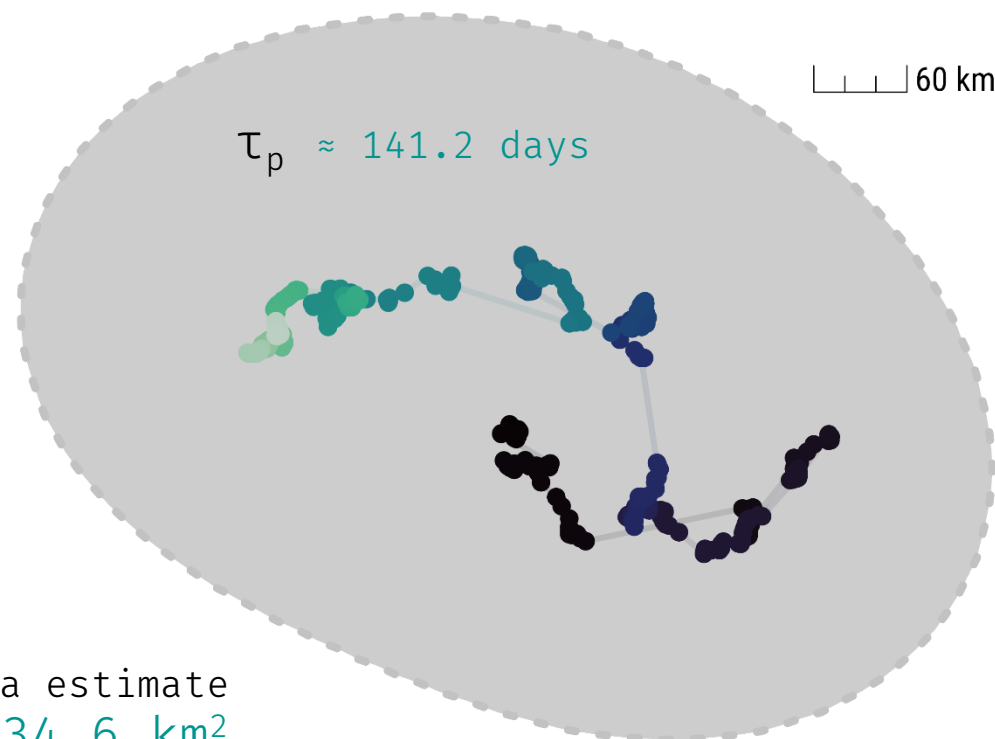
Sampling frequency \approx 1 fix every 5 hours

MONGOLIAN GAZELLE
PROCOPRA GUTTUROSA

Absolute sample size $n = 710$ locations

Effective sample size $N \approx 2.7$ locations

Home range area estimate
 $303,434.6 \text{ km}^2$



For **independent** data,

$$n = N$$

For **autocorrelated** data,

$$n \gg N$$

n = absolute sample size

N = effective sample size

Many biases, including most that affect home range estimation, are exacerbated by ***small sample sizes***.



01

Range residency assumption

Checking if data is from a **range-resident** animal

02

Movement models

Selecting the best-fit movement model through **model selection**

03

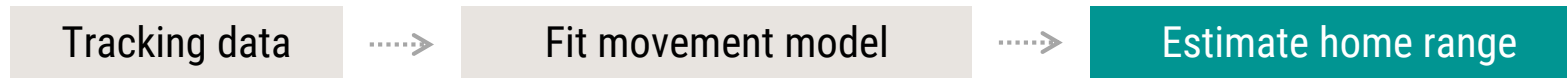
Home range **estimation**

Reconstructing **range distribution** from sampled locations

04

Mitigation measures

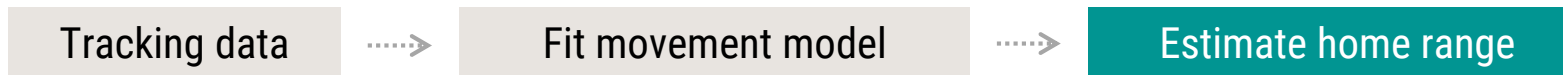
Accounting for common **biases** in animal movement data



Autocorrelated Kernel Density Estimator (AKDE)

Adapted from Fleming *et al.* (2015)
DOI: 10.1890/14-2010.1

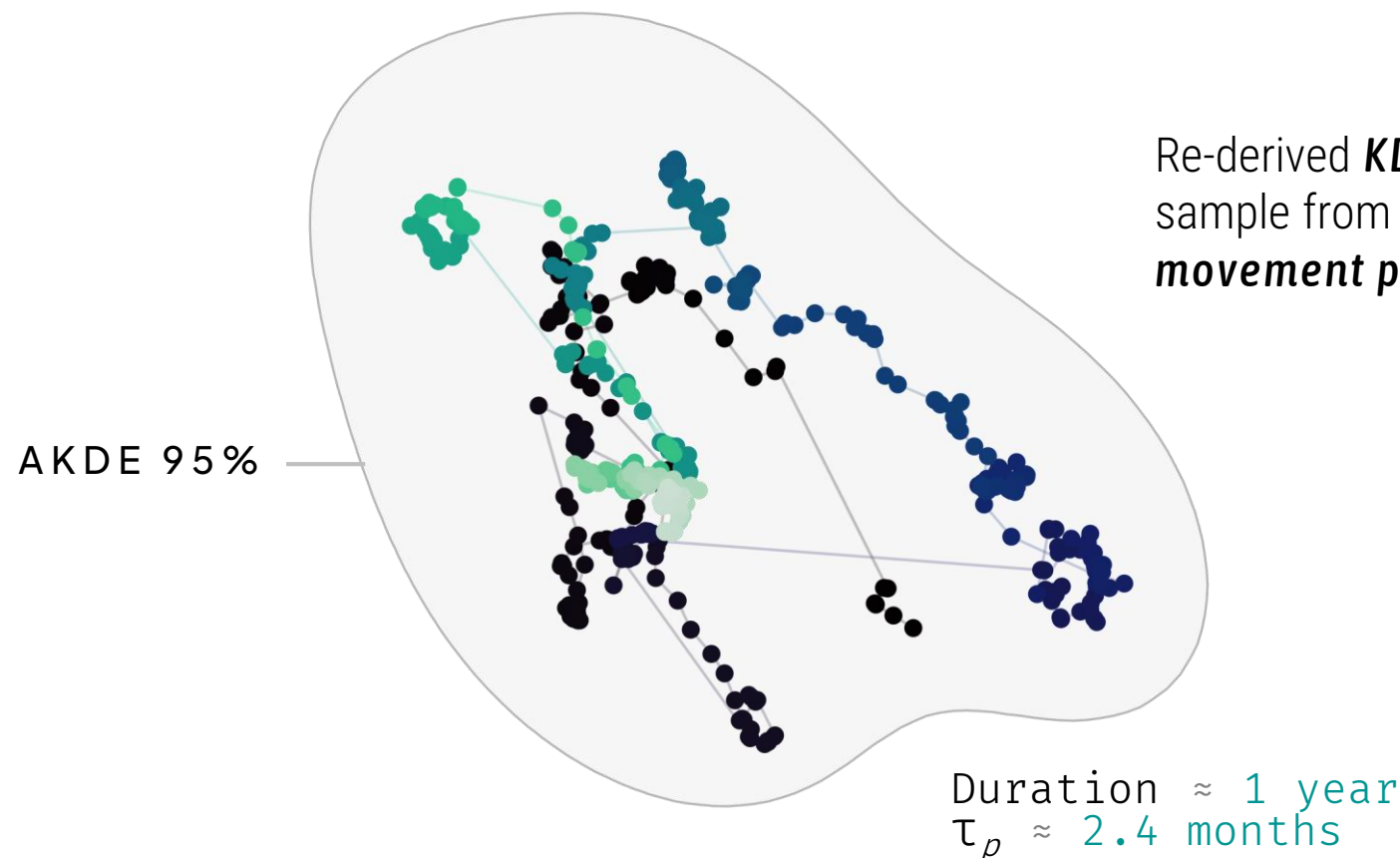
Re-derived **KDE** that explicitly assumes the data represents a sample from a **nonstationary, autocorrelated, continuous movement process**.

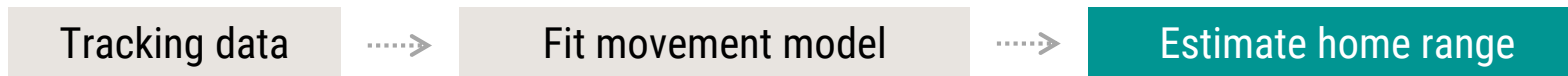


Autocorrelated Kernel Density Estimator (AKDE)

 Adapted from **Fleming *et al.* (2015)**
DOI: 10.1890/14-2010.1

Re-derived **KDE** that explicitly assumes the data represents a sample from a **nonstationary, autocorrelated, continuous movement process**.

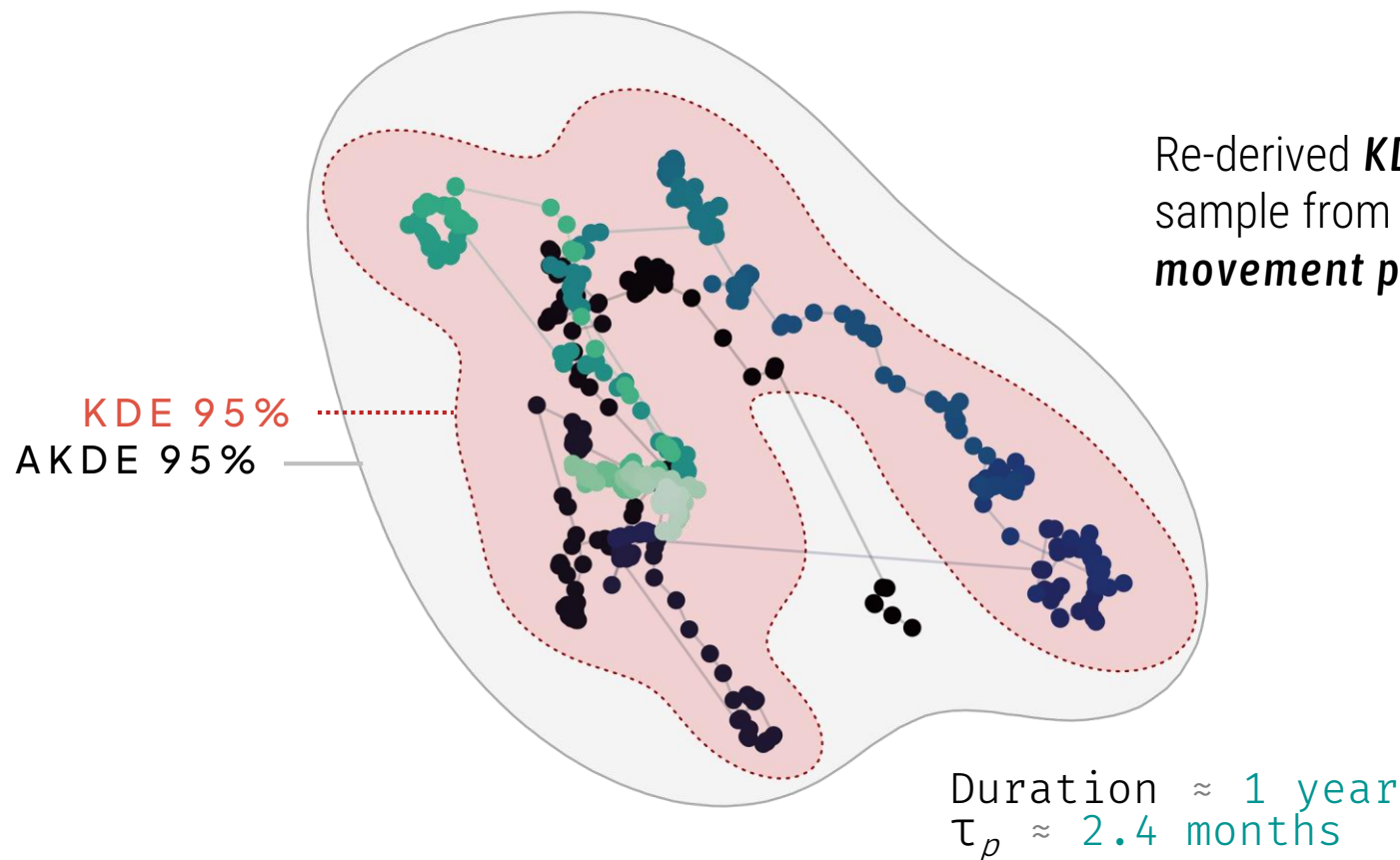


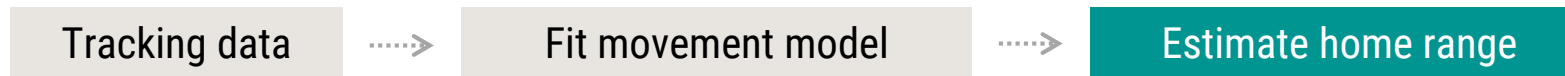


Autocorrelated Kernel Density Estimator (AKDE)

Adapted from **Fleming *et al.* (2015)**
DOI: 10.1890/14-2010.1

Re-derived **KDE** that explicitly assumes the data represents a sample from a **nonstationary, autocorrelated, continuous movement process**.

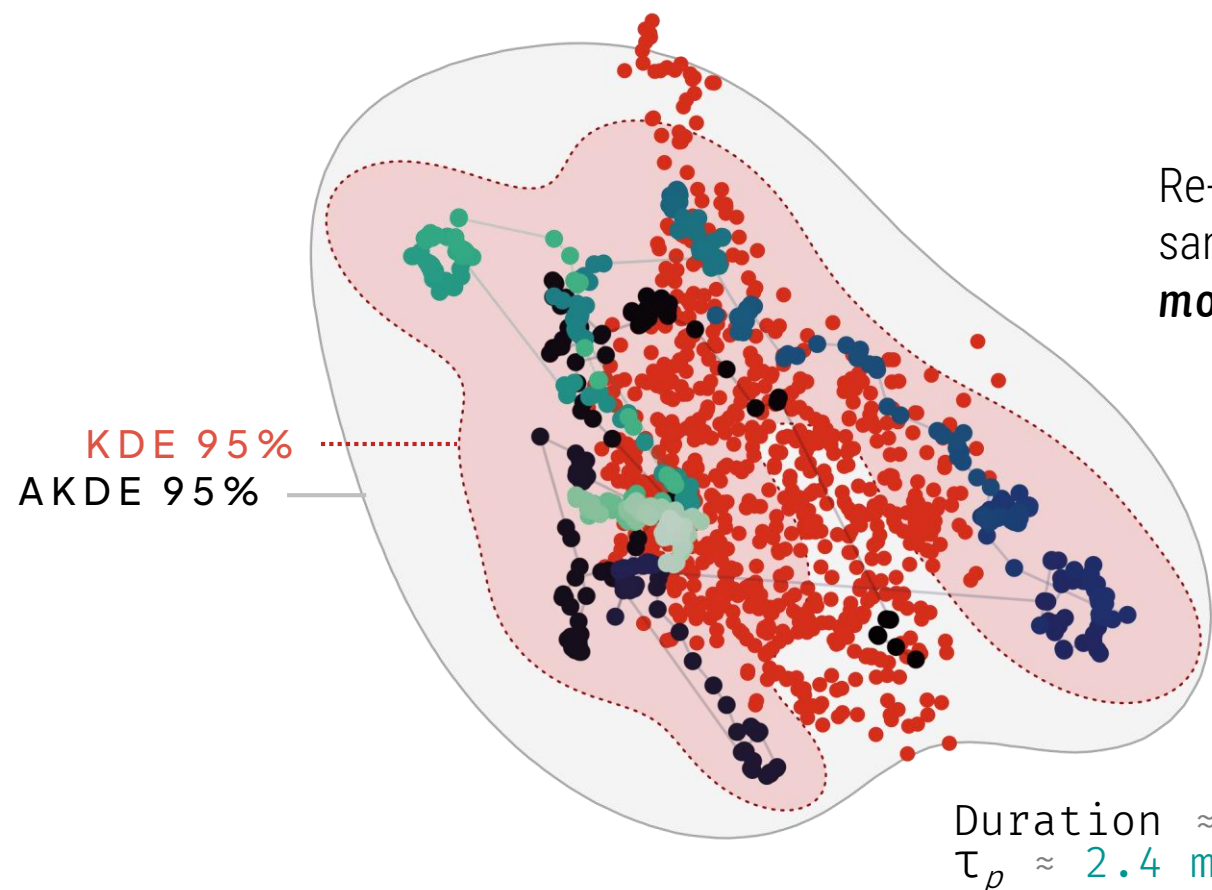




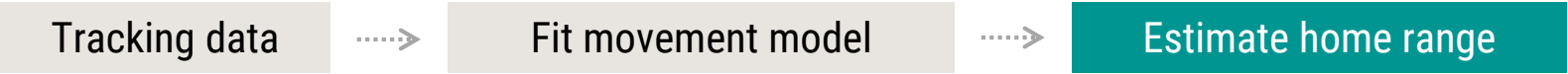
Autocorrelated Kernel Density Estimator (AKDE)

Adapted from **Fleming *et al.* (2015)**
DOI: 10.1890/14-2010.1

Re-derived **KDE** that explicitly assumes the data represents a sample from a **nonstationary, autocorrelated, continuous movement process**.



Duration \approx 2 years → Showing an extra year of data
 $\tau_p \approx$ 2.4 months



Autocorrelation

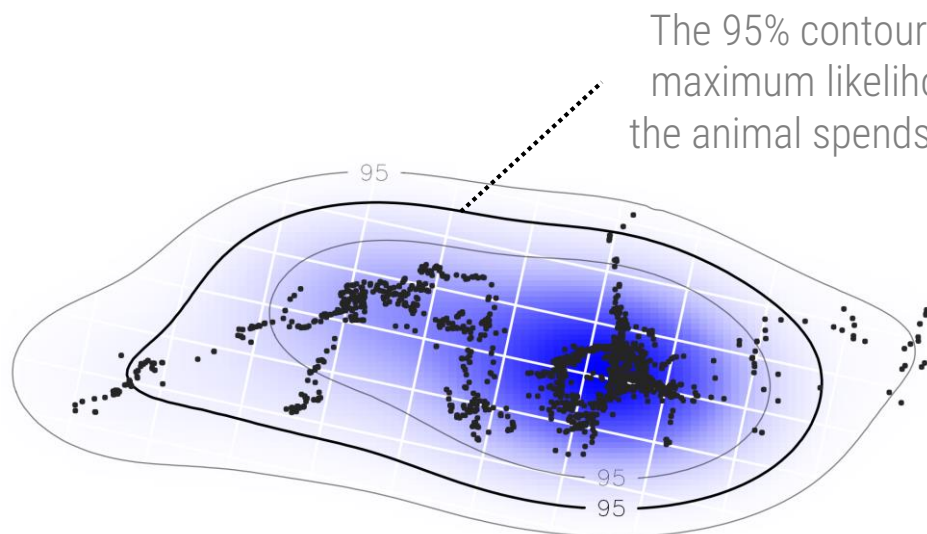
Model	Position	Velocity	Restricted	Parameters:
IID	No	No	Yes	$\tau = \text{NULL}$
BM	Yes	No	No	$\tau = \infty$
OU	Yes	No	Yes	$\tau = \tau_p$
IOU	Yes	Yes	No	$\tau = \{\infty, \tau_v\}$
OUF	Yes	Yes	Yes	$\tau = \{\tau_p, \tau_v\}$

AKDEs explicitly requires a movement model that accounts *autocorrelated* data.

AKDE reduces to the **conventional KDE** in the limit where autocorrelation vanishes, and locations are truly *independent*.



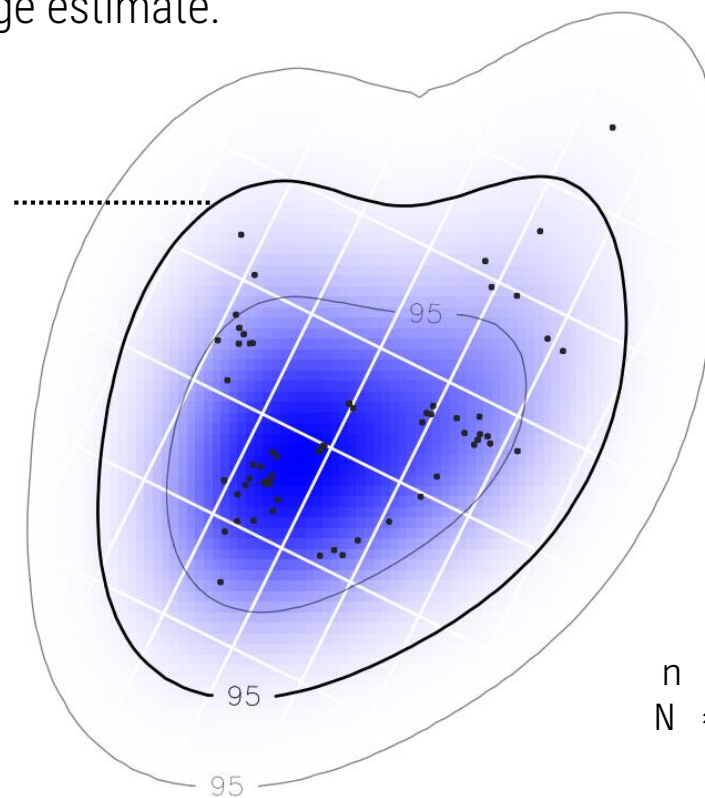
AKDEs also provide accurate **confidence intervals** that can diagnose situations where the data are insufficient to provide a reasonable home range estimate.



$n = 1,725$
 $N \approx 15.7$

Area estimate
757.5 km² (430.0 – 1,176.2)

The 95% contour represents the maximum likelihood area where the animal spends 95% of its time.



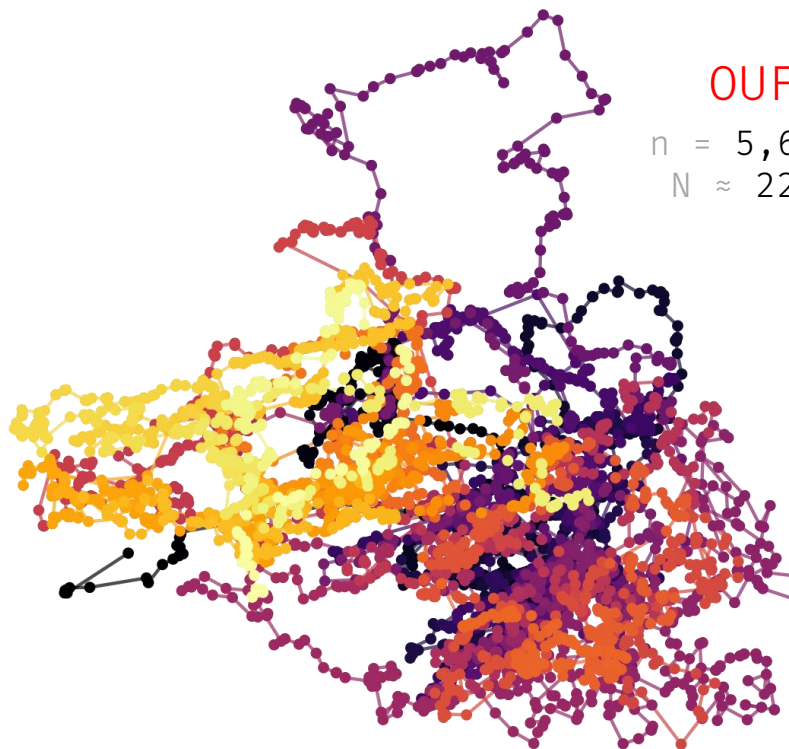
$n = 66$
 $N \approx 6.9$

Area estimate
4,232.3 km² (1,681.7 – 7,939.5)



NT

AFRICAN BUFFALO
(SYNCERUS CAFFER)



OUF process:

$n = 5,677$ locations

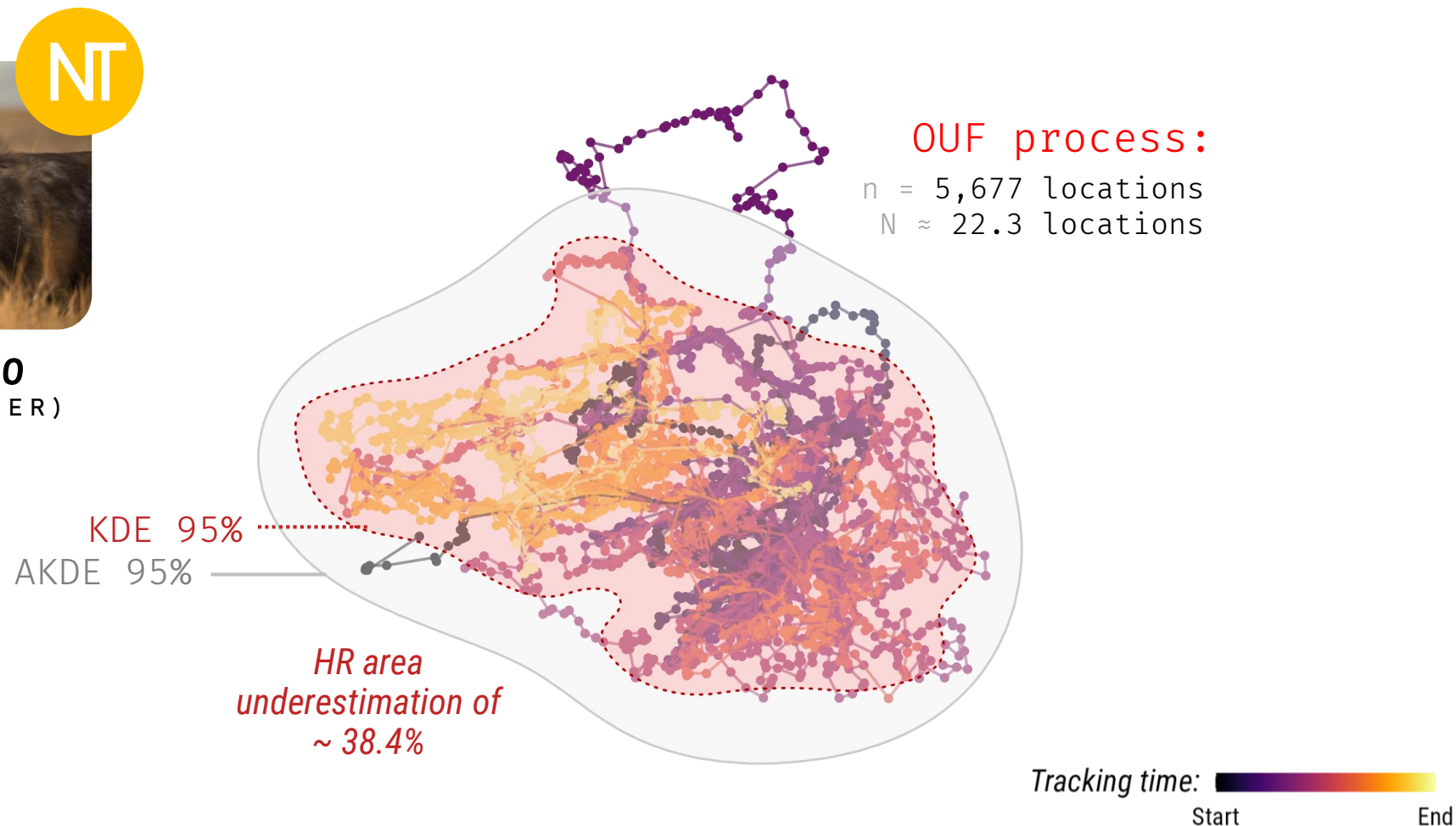
$N \approx 22.3$ locations

Tracking time: 
Start End



NT

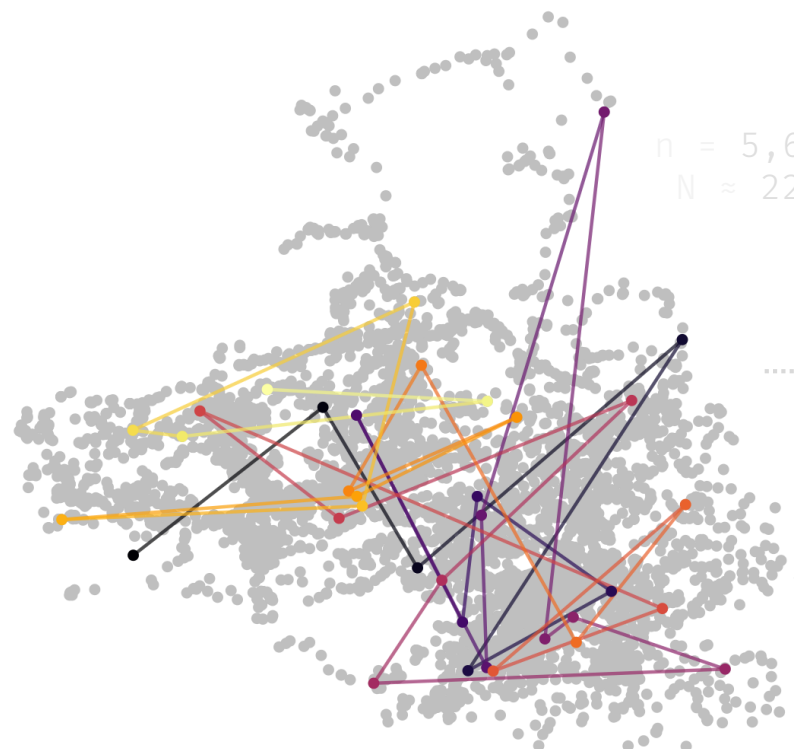
AFRICAN BUFFALO
(SYNCERUS CAFFER)





NT

AFRICAN BUFFALO
(SYNCERUS CAFFER)



$n = 5,677$ locations
 $N \approx 22.3$ locations

IID process:

$n = 35$ locations
 $N \approx 35$ locations

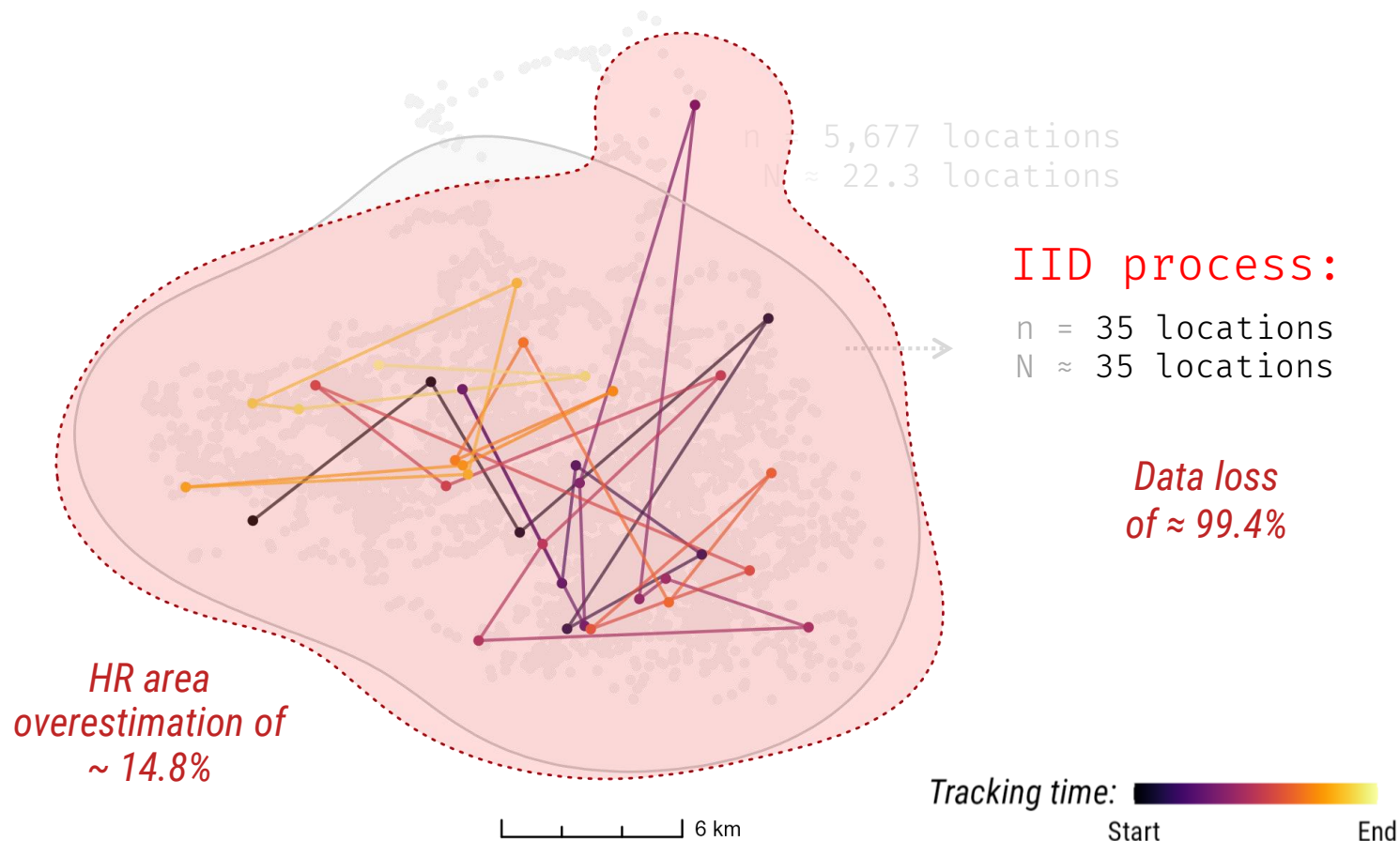
*Data loss
of $\approx 99.4\%$*

Tracking time: 
Start End



NT

AFRICAN BUFFALO
(SYNCERUS CAFFER)





01

Range residency assumption

Checking if data is from a **range-resident** animal

02

Movement models

Selecting the best-fit movement model through **model selection**

03

Home range estimation

Reconstructing **range distribution** from sampled locations

04

Mitigation **measures**

Accounting for common **biases** in animal movement data



Many biases, including most that affect home range estimates, are exacerbated by **small sample sizes**. Conversely, **large sample sizes** in modern tracking datasets exacerbate autocorrelation.

Bias sources (in order of their general importance): ***Mitigation measures:***

Unmodelled autocorrelation ▶ AKDE,

Oversmoothing ▶ AKDE_c (default)

Autocorrelation estimation bias ▶ pHREML (default)

Parametric bootstrapping

Unrepresentative sampling in time ▶ Weighted AKDE, or wAKDE



Many biases, including most that affect home range estimates, are exacerbated by **small sample sizes**. Conversely, **large sample sizes** in modern tracking datasets exacerbate autocorrelation.

Bias sources (in order of their general importance): *Mitigation measures:*

Unmodelled autocorrelation ▶ AKDE,

Oversmoothing ▶ AKDE_c (default)

Autocorrelation estimation bias ▶ pHREML (default)

Parametric bootstrapping

Unrepresentative sampling in time ▶ Weighted AKDE, or wAKDE



Area-corrected AKDE or $AKDE_c$

Deals with: **oversmoothing**



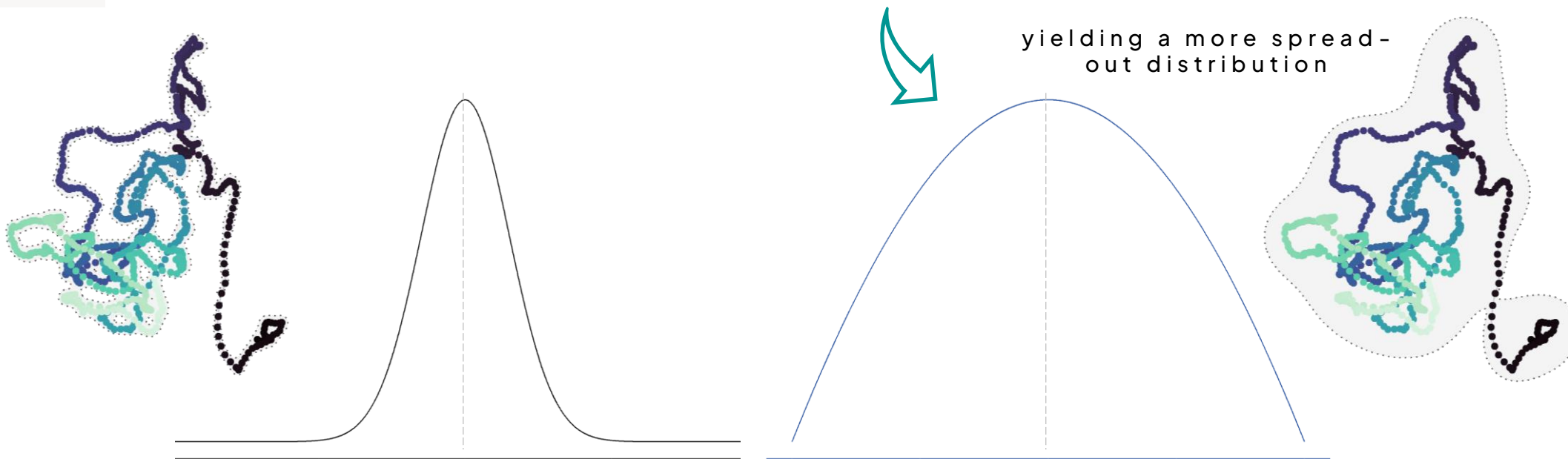
Even when we account for autocorrelation, GRF-KDEs remain biased due to the natural tendency of the GRF approximation to **oversmooth**.

AKDE_c

PHREML

WAKDE

BOOTSTRAP





Area-corrected AKDE or $AKDE_c$

Deals with: **oversmoothing**



Even when we account for autocorrelation, GRF-KDEs remain biased due to the natural tendency of the GRF approximation to **oversmooth**.

$AKDE_c$

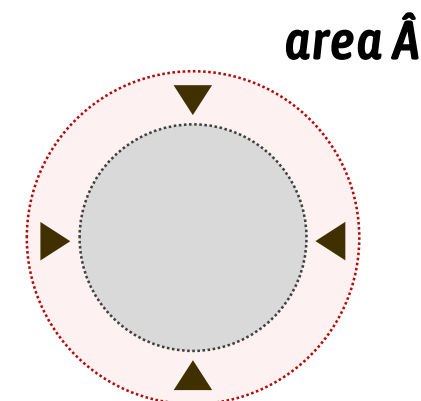
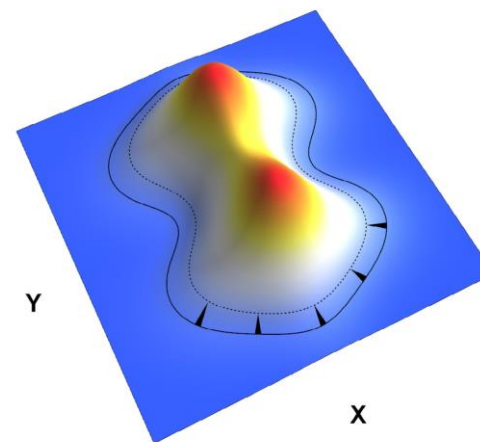
PHREML

WAKDE

BOOTSTRAP

Derived an **improved (A)KDE** that pulls the contours of the location distribution estimate inward towards the data without distorting its shape.

 Fleming & Calabrese (2017)

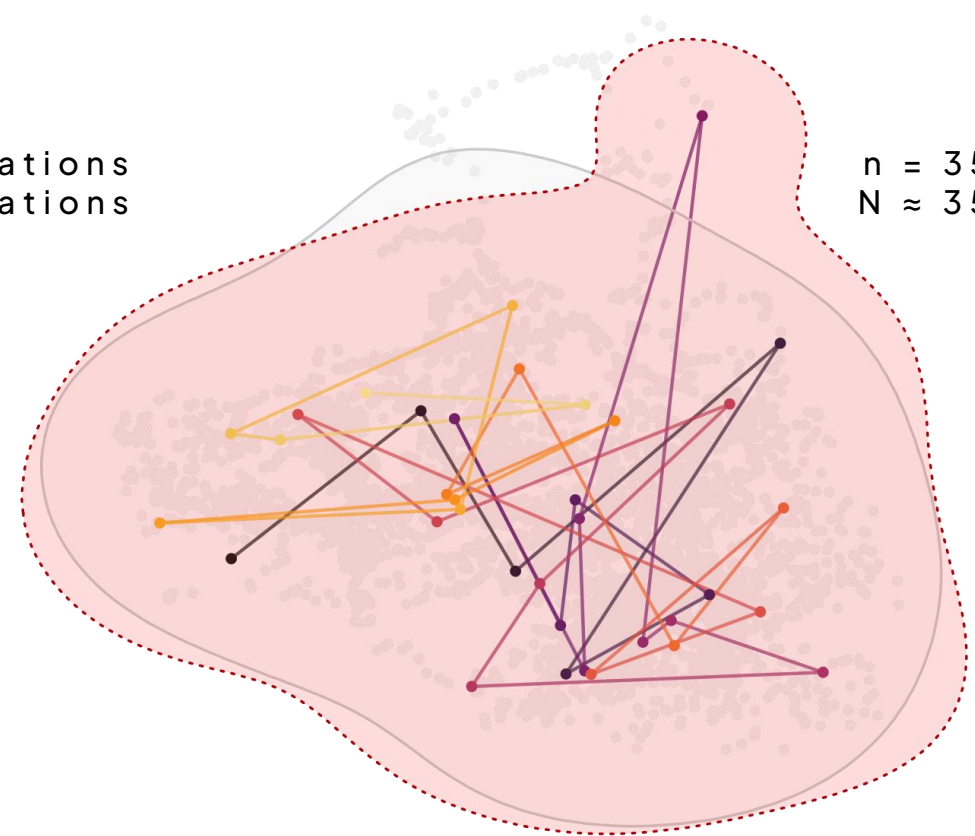
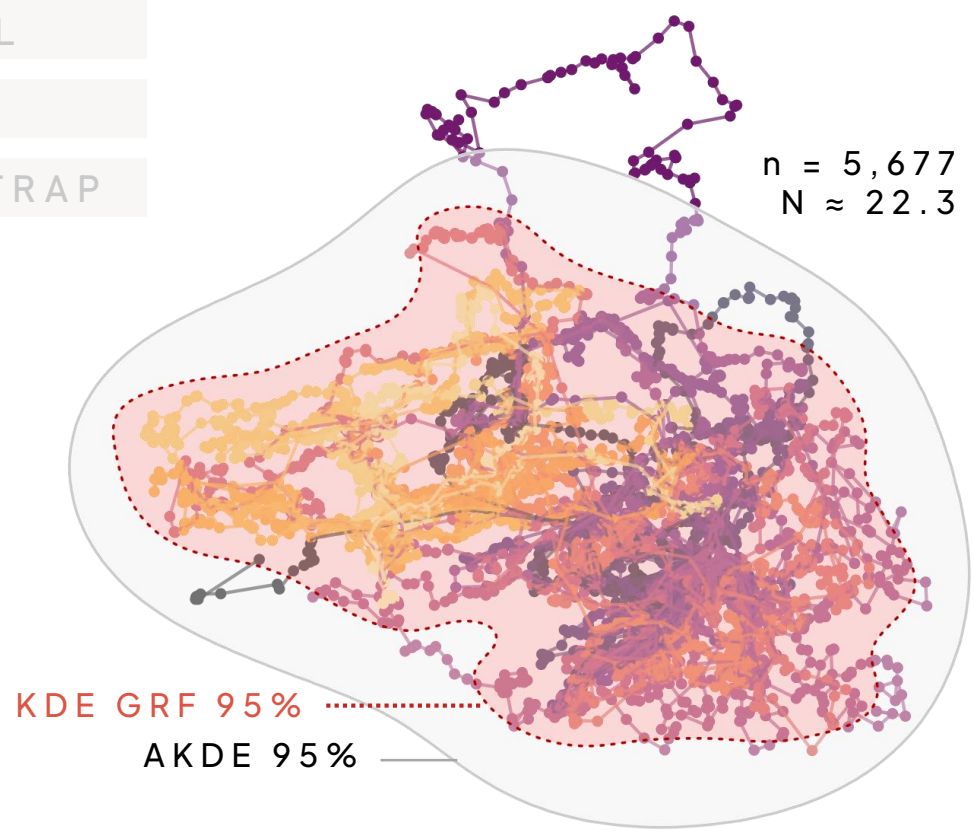




The oversmoothing (positive) bias can be **masked** by the often-stronger negative bias caused by unmodeled autocorrelation.

AKDE_c

- PHREML
- WAKDE
- BOOTSTRAP



6 km



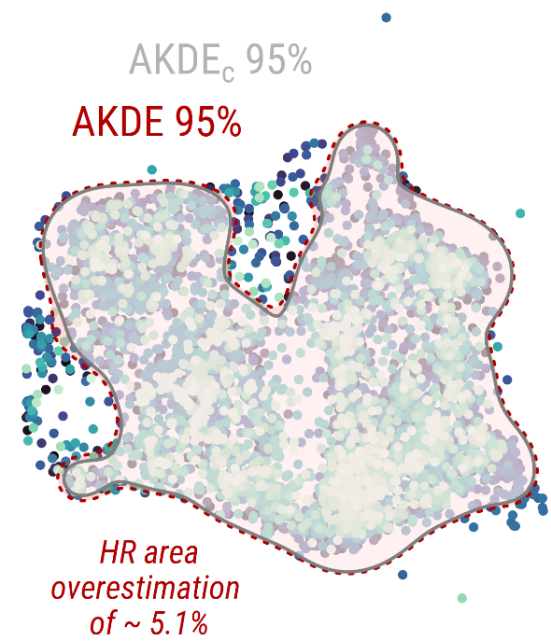
AKDE_c

PHREML

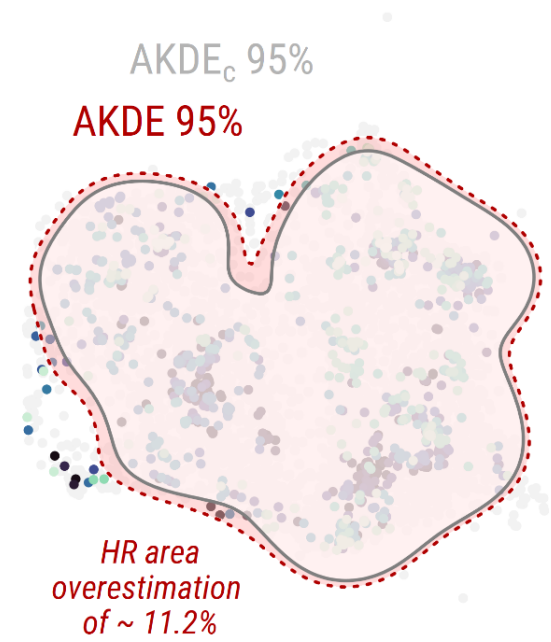
WAKDE

BOOTSTRAP

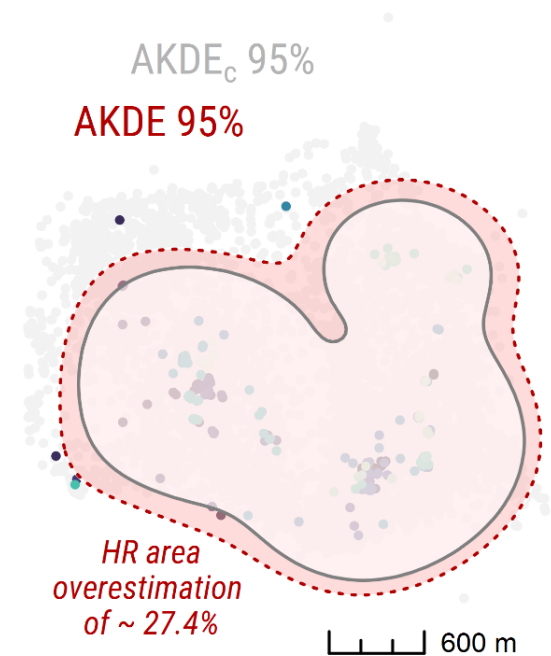
Large effective sample size
Sampling duration \approx 1 year



Medium effective sample size
Sampling duration \approx 3 months



Small effective sample size
Sampling duration \approx 15 days





pHREML AKDE

Deals with: autocorrelation estimation bias



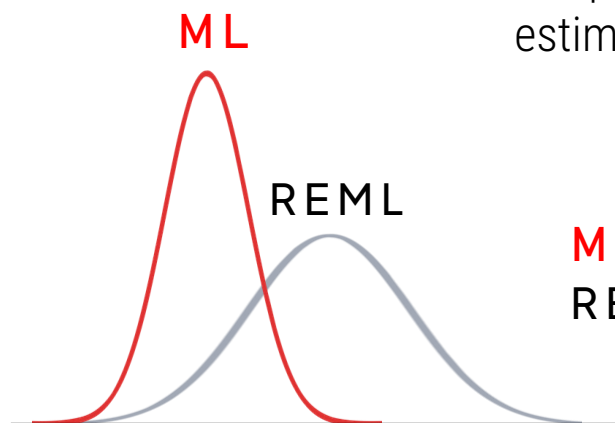
For optimal performance, we need to estimate autocorrelation correctly.

AKDEC

PHREML

WAKDE

BOOTSTRAP



ML — performs poorly at small sample sizes.

REML — performs poorly at small *effective* sample sizes.



pHREML AKDE

Deals with: **autocorrelation estimation bias**



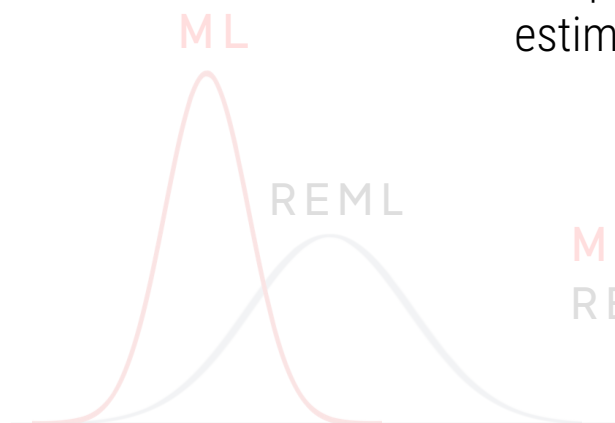
For optimal performance, we need to estimate autocorrelation correctly.

AKDEC

PHREML

WAKDE

BOOTSTRAP



ML — performs poorly at small sample sizes.

REML — performs poorly at small *effective* sample sizes.

Focus on:

small **effective** sample sizes

small **absolute** sample sizes

small **absolute** and **effective** sample sizes

As such, we consider other parameter estimation methods:

perturbative REML (pREML)

Hybrid REML (HREML)

perturbative Hybrid REML (pHREML)

ℓ Fleming et al. (2019)



pHREML AKDE

Deals with: **autocorrelation estimation bias**



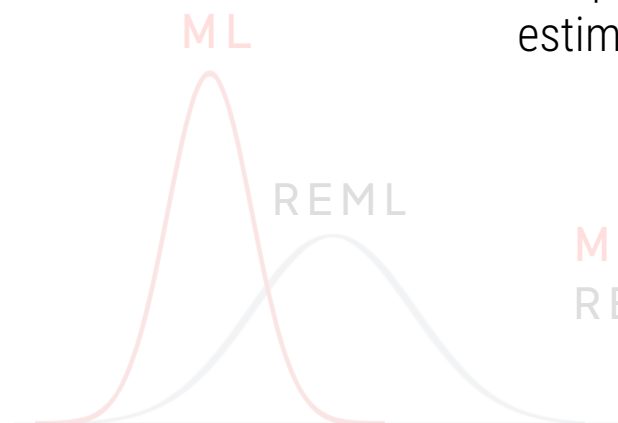
For optimal performance, we need to estimate autocorrelation correctly.

AKDEC

PHREML

WAKDE

BOOTSTRAP



ML — performs poorly at small sample sizes.

REML — performs poorly at small **effective** sample sizes.

Focus on:

small **effective** sample sizes

small **absolute** sample sizes

small **absolute** and **effective** sample sizes

As such, we consider other parameter estimation methods:

----- **perturbative REML (pREML)**

----- **Hybrid REML (HREML)**

----- **perturbative Hybrid REML (pHREML)**

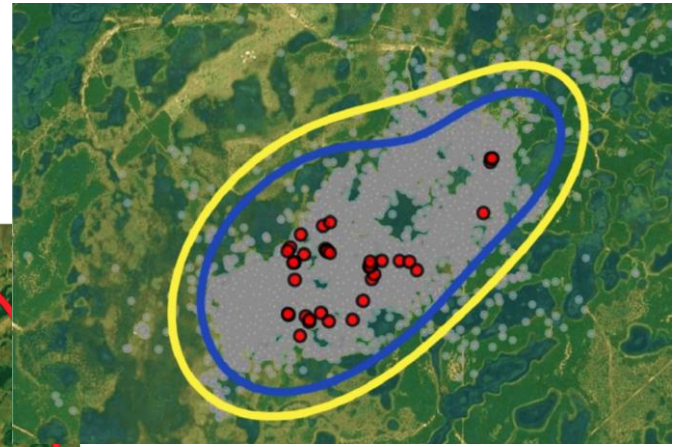
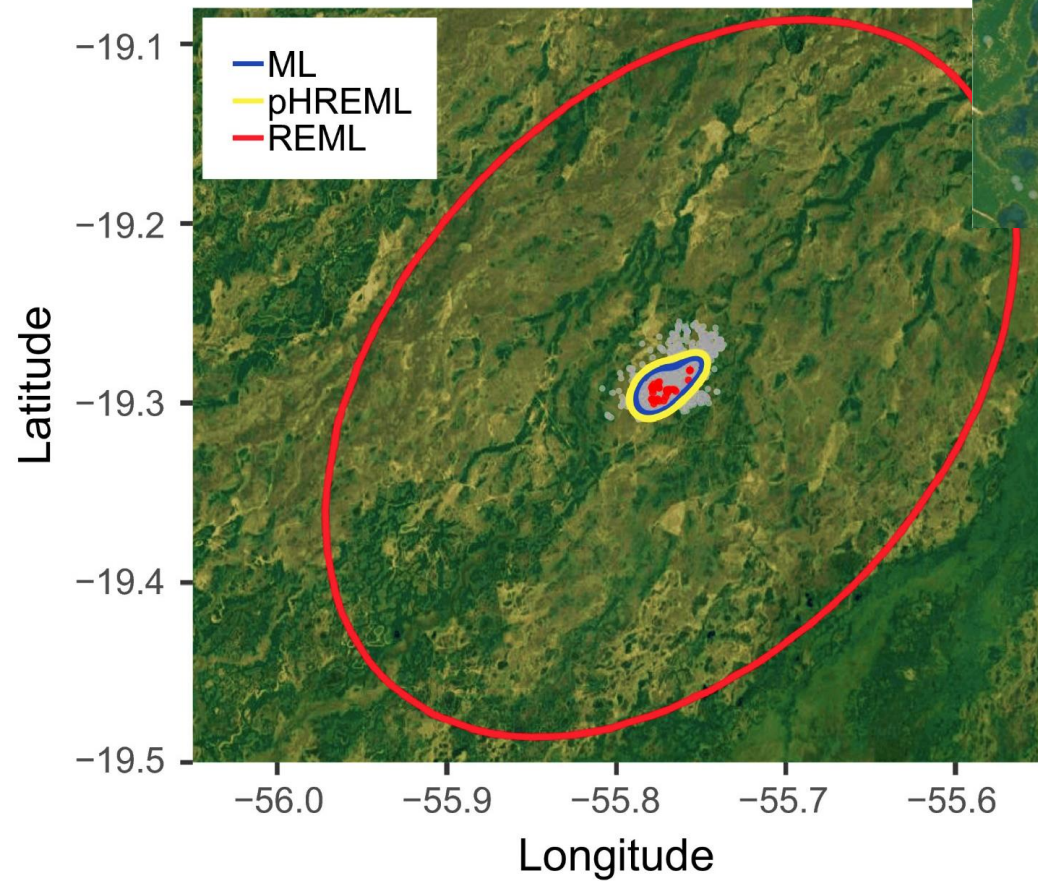
ℓ Fleming et al. (2019)



Tracking data (1-hr intervals for 19 months), reduced to 2 days.



LOWLANDTAPIR
TAPIRUS TERRESTRIS



- ML
- pHREML
- REML



AKDEC

PHREML

wAKDE

BOOTSTRAP

Weighted AKDE or wAKDE

Deals with: **unrepresentative sampling in time**

- ▶ Many real-world issues can lead to **irregular sampling**:

duty-cycling tags to avoid wasting battery,
acceleration-informed sampling,
device malfunction,
habitat-related signal loss,
and other causes.

- ▶ Shifting **sampling schedules** (based on behavioral or seasonal patterns) is also a common strategy.

wAKDE optimally **upweights** observations that occur during **under-sampled times**, while optimally **downweighting** observations occurring during **over-sampled times**.



Weighted AKDE or wAKDE

Deals with: *unrepresentative sampling in time*

AKDEC

PHREML

wAKDE

BOOTSTRAP

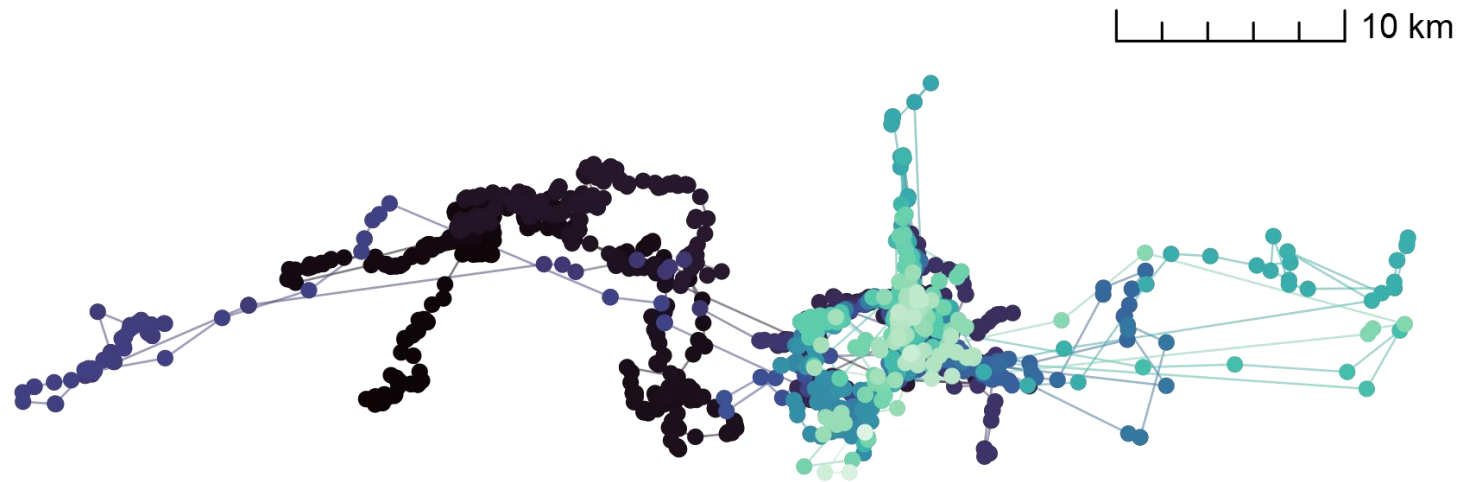


Fig.

African buffalo dataset (nicknamed “Pepper”) with an irregular sampling schedule.



Weighted AKDE or wAKDE

Deals with: *unrepresentative sampling in time*

AKDEC

PHREML

wAKDE

BOOTSTRAP

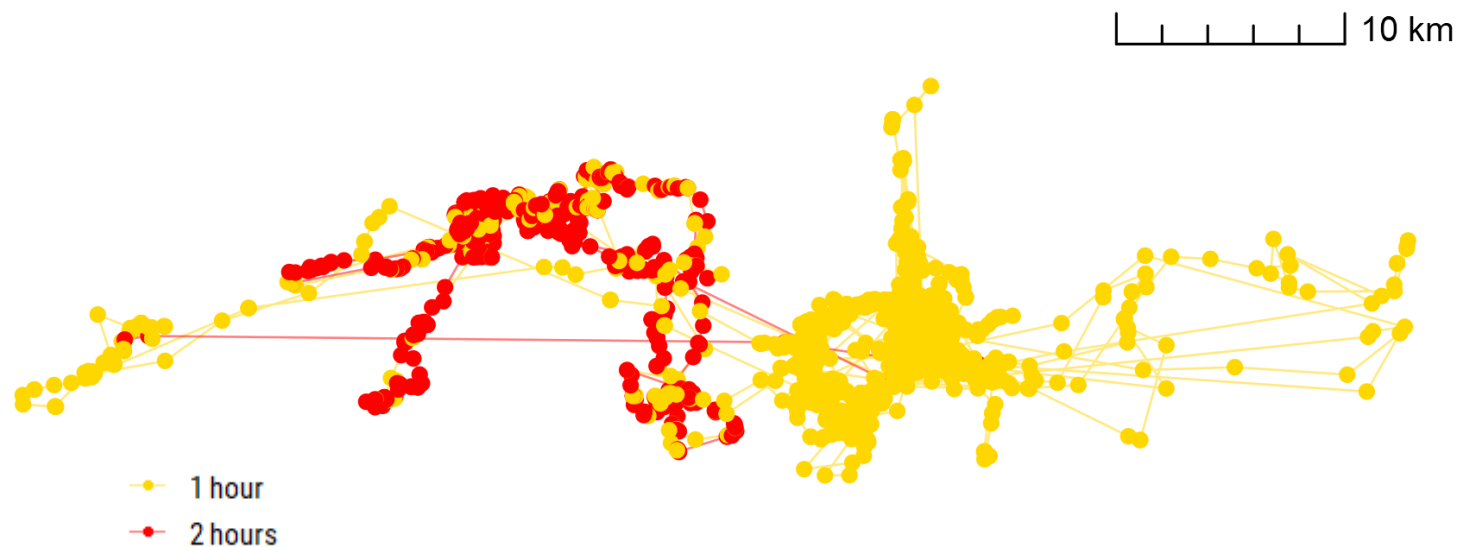


Fig.

African buffalo dataset (nicknamed "Pepper") with an irregular sampling schedule. Sampling rate shifted from 1 fix every hour to 1 fix every 2 hours.



Weighted AKDE or wAKDE

Deals with: *unrepresentative sampling in time*

AKDEC

PHREML

wAKDE

BOOTSTRAP

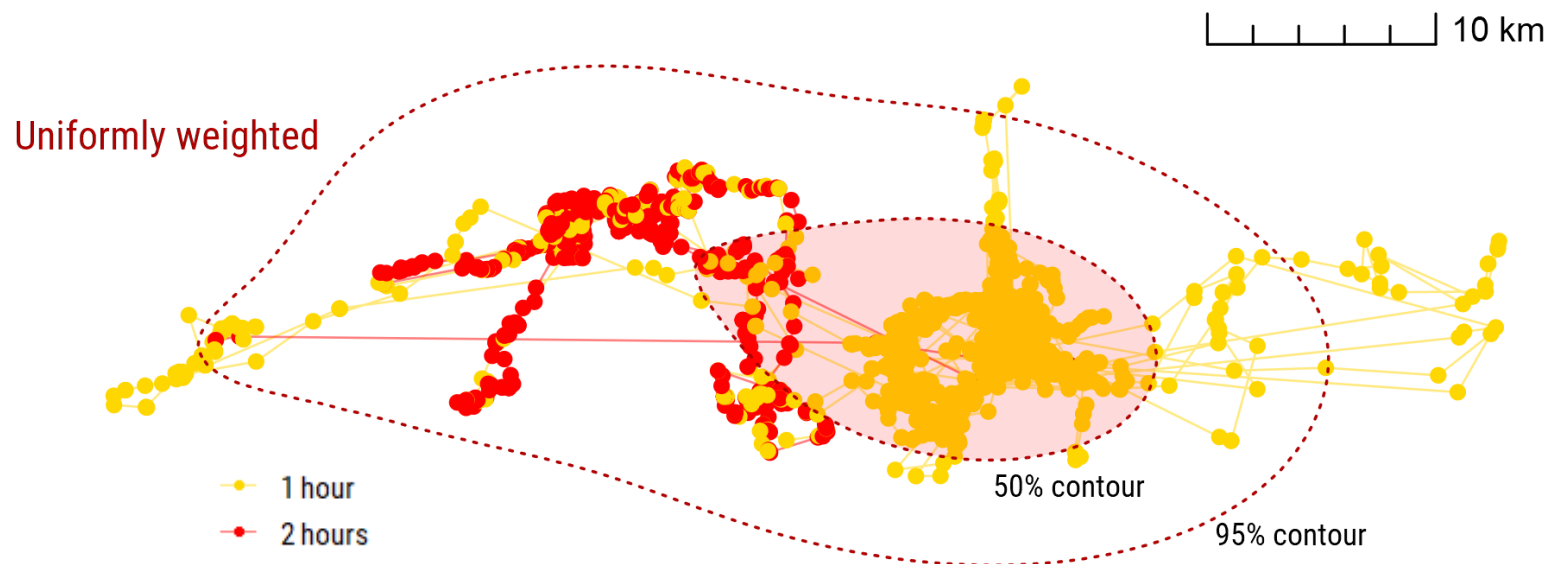


Fig.

African buffalo dataset (nicknamed "Pepper") with an irregular sampling schedule.
Sampling rate shifted from 1 fix every hour to 1 fix every 2 hours.



Weighted AKDE or wAKDE

Deals with: *unrepresentative sampling in time*

AKDEC

PHREML

wAKDE

BOOTSTRAP

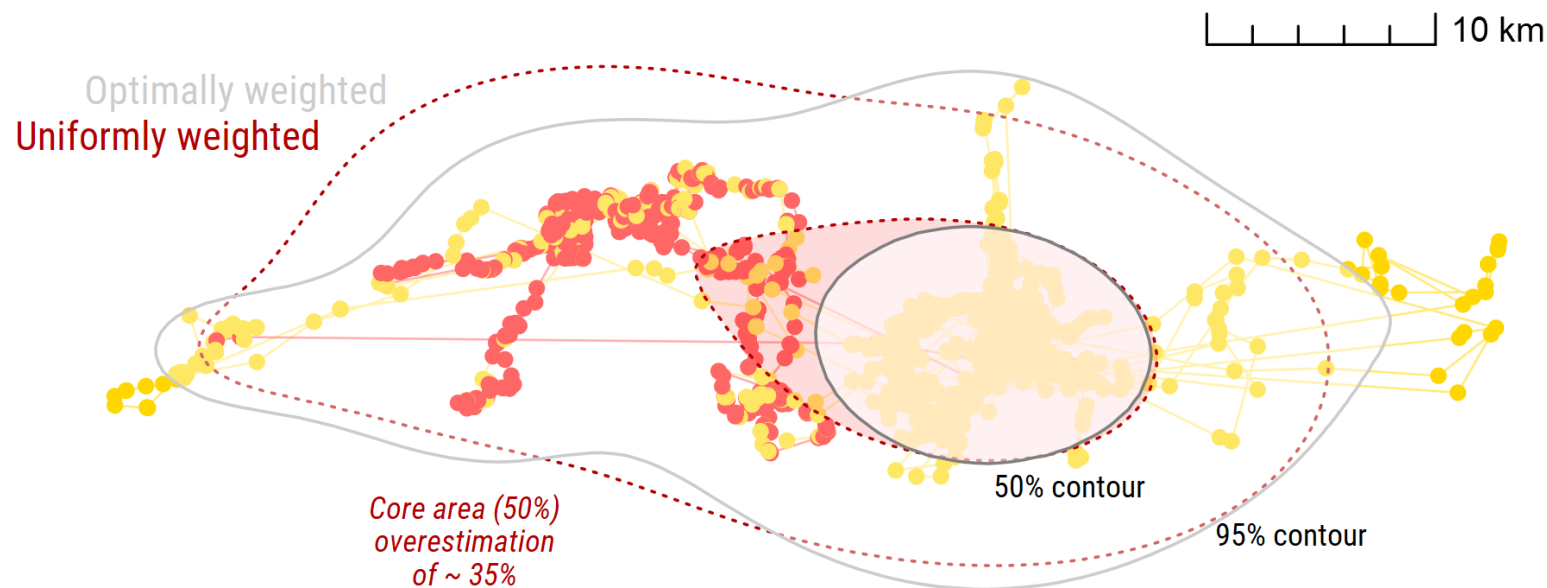


Fig.

African buffalo dataset (nicknamed "Pepper") with an irregular sampling schedule.
Sampling rate shifted from 1 fix every hour to 1 fix every 2 hours.



Parametric bootstrapping AKDE

Deals with: very low effective sample size

AKDEC

PHREML

WAKDE

BOOTSTRAP

Bartlett (1937)

Residual ML (or REML)

Fleming *et al.* (2019)

perturbative Hybrid REML (pHREML)



Efron & Efron (1982)

Parametric bootstrapping

The parametric bootstrap estimates the bias and variance of an estimator by approximating the sampling distribution of the true movement model with that of the best-fit model.

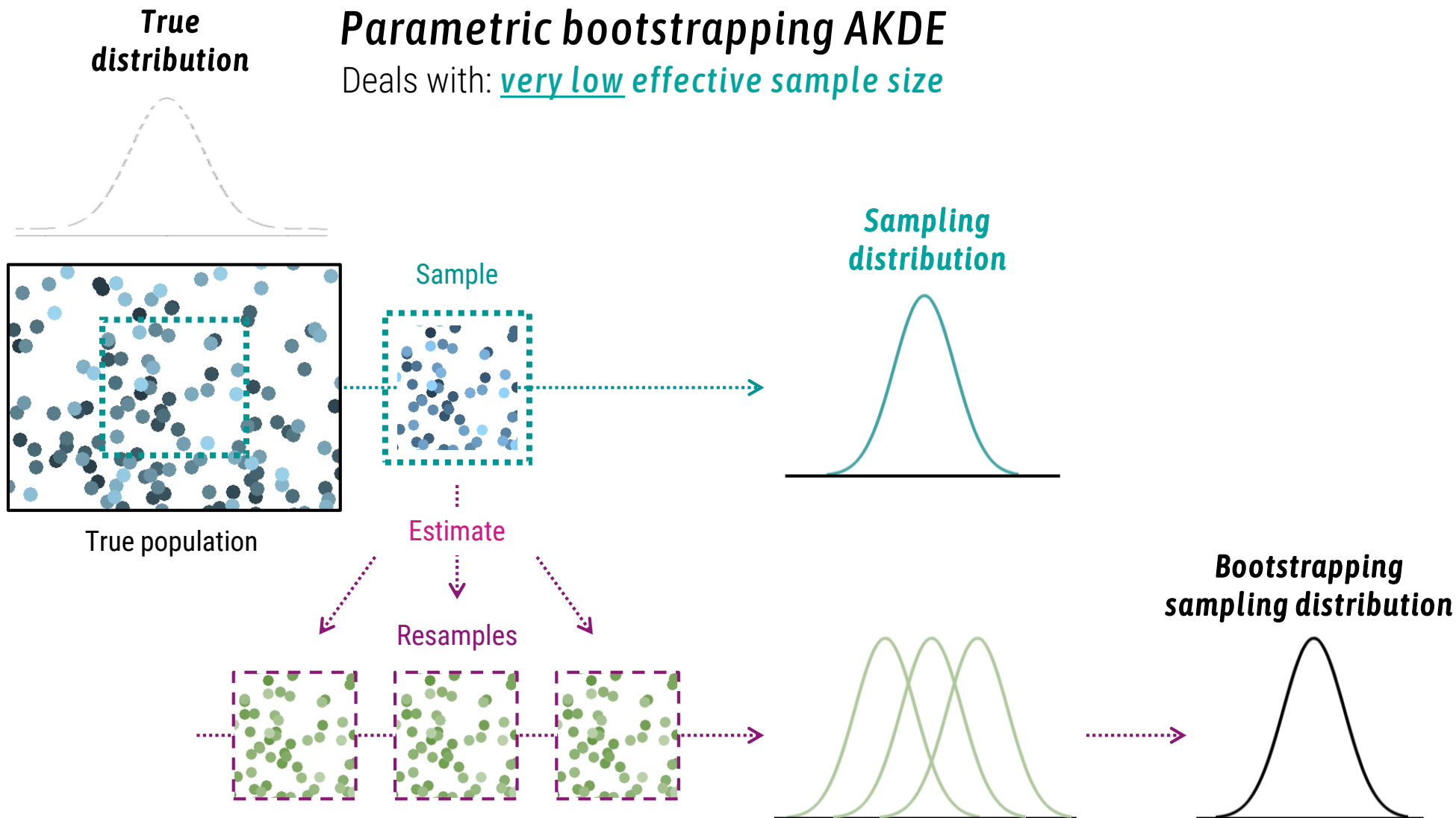


AKDEC

PHREML

WAKDE

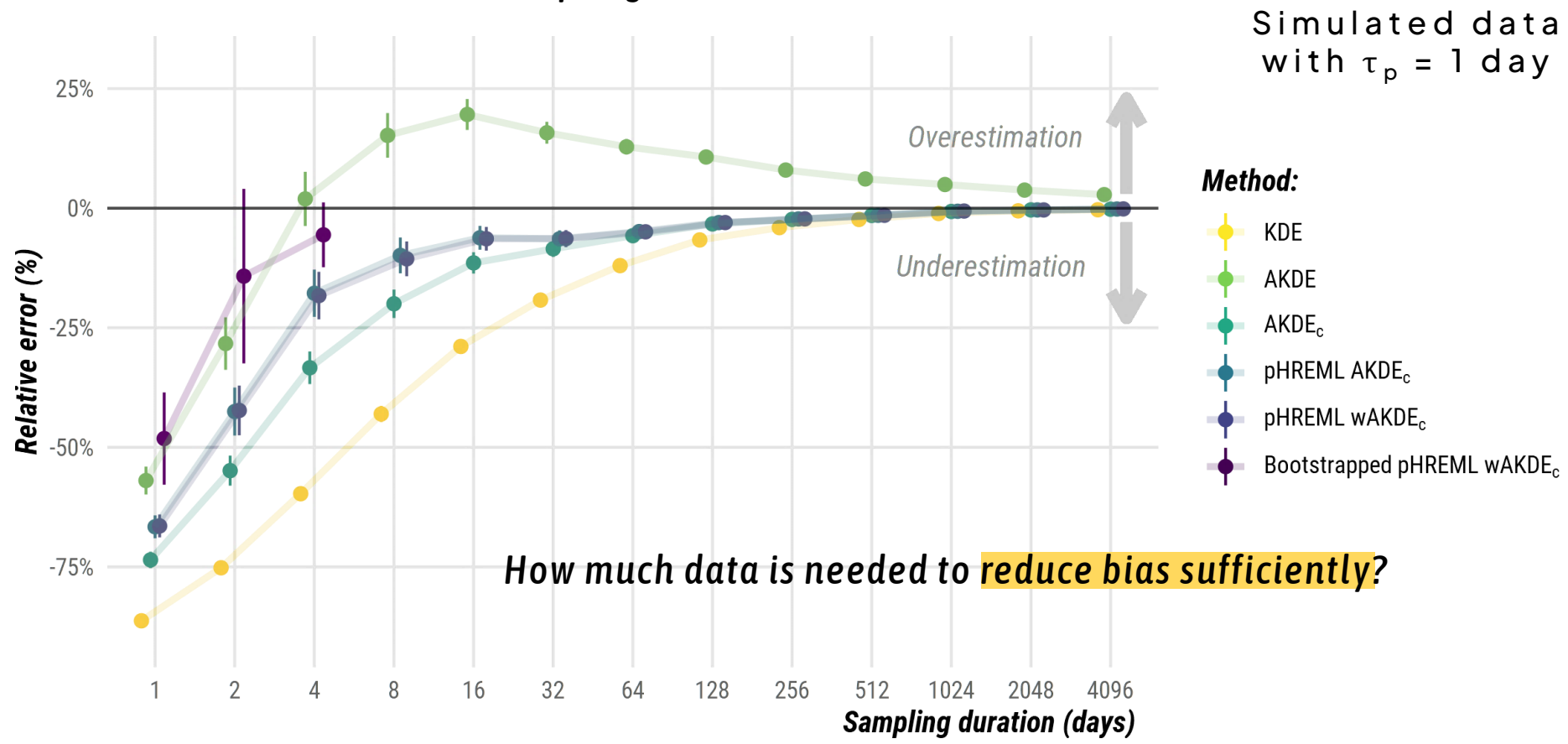
BOOTSTRAP





AKDE family

Relative error vs. sampling duration





AKDE family

Computational cost vs. sampling duration

