

# Roe Deer Movement

**Benjamin Michael Marshall<sup>\*1</sup>, —\*\*<sup>2</sup>, <sup>3</sup>, <sup>4</sup>, <sup>3</sup>, <sup>5</sup>, <sup>6</sup>, <sup>7</sup>, and <sup>1</sup>**

<sup>1</sup>—

<sup>2</sup>—

<sup>3</sup>

<sup>4</sup>

<sup>5</sup>

<sup>6</sup>

<sup>7</sup>

\*[benjamminmichaelmarshall@gmail.com](mailto:benjamminmichaelmarshall@gmail.com)

\*\*—

---

**Abstract**  
abstract text

---

## Keywords

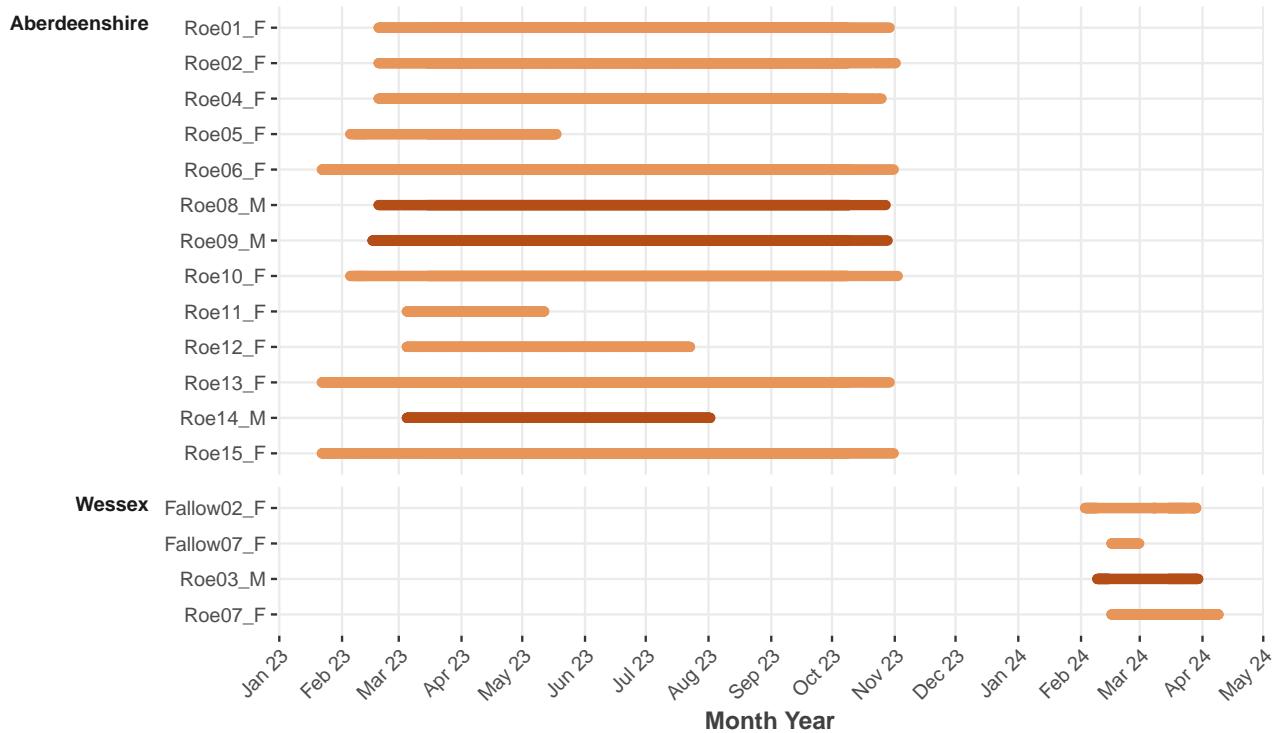
Movement ecology, step selection function, poisson, habitat preference, habitat selection, animal movement, roe deer

## 1 Introduction

## 2 Methods

```
## Warning: package 'dplyr' was built under R version 4.2.3  
  
## Warning: package 'stringr' was built under R version 4.2.3
```

### 2.1 Study Area



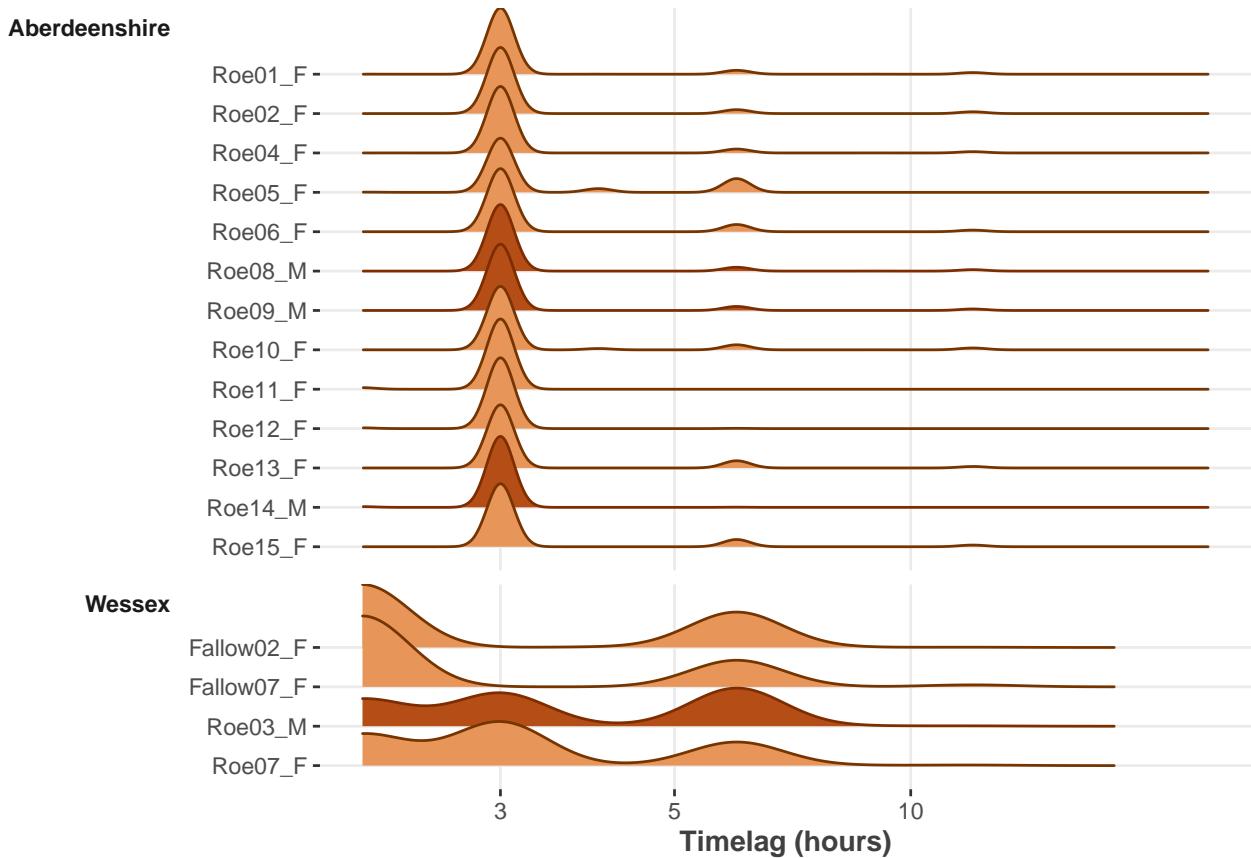
**Figure 1.** Duration

### 2.2 Home Range Esimation

We estimated roe deer home range using autocorrelated kernel density estimators (Fleming et al., 2015; Calabrese, Fleming & Gurarie, 2016; Fleming & Calabrese, 2017, 2023). This process consisted of fitting a number of continuous time movement models to an individuals movement data, selecting the best fitting movement model, and extracting a suitable range contour from the resulting utilisation distribution. We fit the following models (following the default process provided by the ctmm package): Ornstein-Uhlenbeck (OU), Ornstein–Uhlenbeck Foraging (OUF), and Independent Identically Distributed (IID), all in both isotropic and anisotropic forms. We elected to use perturbative hybrid residual maximum likelihood (pHREML) and (Fleming et al., 2019; Silva et al., 2022) AICc to determine the best fitting movement model on an individual basis, and used that single best fitting model for all further estimations.

Before committing to the estimations of home range size we examined whether the roe deer exhibited stable ranges through the visual inspection of variograms. A stable range should be reveal by a clear asymptote in the variogram. We paired these visual inspections with a judgement of effective sample size to help gauge our confident in the area estimates (effective sample size approximating the overall tracking duration divided by the mean time taken to cross the home range). (Silva et al., 2022). All our individuals showed effective sample sizes between 154.4 and 926.7, leading us to be confident in overall home range estimates.

Having determine the data suitability for home range estimations, we extracted the 95% and 99% contours from the weighted AKDE estimate, alongside 95% CI surrounding that contour. We selected 95% as a balance between a generous estimate of home range, while also avoiding the undue influence of the most extremely outlying movements. To generate an overall home range estimate for Aberdeen roe deer, we averaged the home ranges using the



**Figure 2.** Time lags.

weighted mean function provided by the ctmm package (Silva et al., 2022; Fleming & Calabrese, 2023). This way the home range mean is weighted by the confidence (i.e., ESS) surrounding each home range.

We retained 99% estimates for guiding the between patch connectivity to maximise the repeated between patch modelling, i.e., connecting as many patches as would be likely for an individual roe deer in the course of their life. We calculated the widest dimension of each home range polygon (or largest polygon if the home range area was non-contiguous), and halved that to serve as a proxy for the distance deer could travel between patches (756 m). A mean of this longest dimension was used to limit which patches could be considered likely to be travelled between by deer. To confirm this approach, we determined the distance from patch for every deer location that fell outside a patch. We examined the distribution of these distance values that revealed that 95% of all deer movements fell within the mean of half longest dimension of the 99% home range area (Fig. 3).

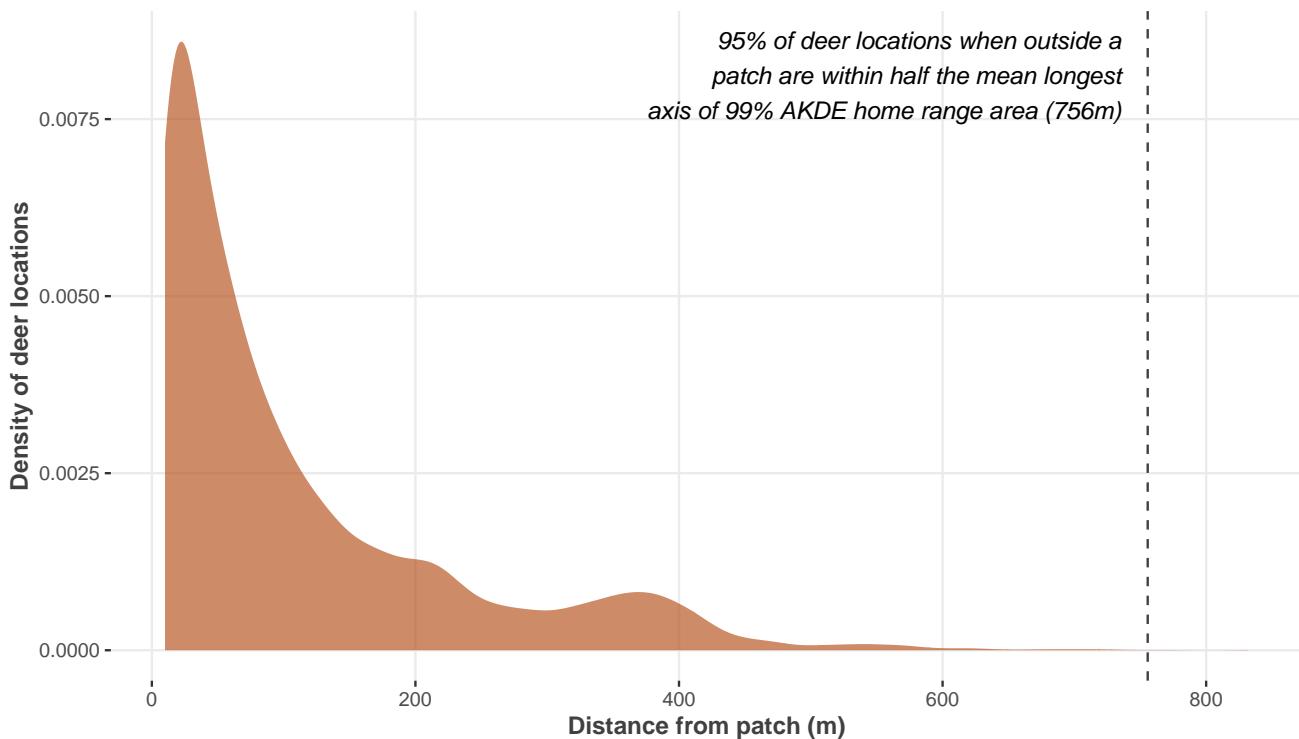
### 2.3 Habitat Selection

We used the reformulated Poisson model approach described by Muff et al to generate estimates of habitat selection as well as gauge deer's movement capabilities in relation to different aspects of the landscape.

The model required data pertaining to the used locations (i.e., GPS locations of the deer) and comparable data on randomly generated available points (i.e., randomly generated locations the deer could have travelled to). For each confirmed deer location we generated 10 random alternative locations they could have travelled to. The location of these random locations was governed by two distributions. A Gamma distribution that random step lengths were drawn from, and a Von Mises distribution that random turn directions were drawn from.

Once all random locations had been generated we extracted a suite of environmental conditions at all those locations. First was the land use type as described by the 2023 UKCEH land cover maps, which is a 25m resolution classified raster originally based on Sentinel-2 imagery. Validation of these data suggest 83% accuracy. The UKCEH land cover data comprises of 21 land cover classes, broadly following the Biodiversity Broad Habitats.

We recategorised these land use types categories into 10 more general categories that reduced instances of limited interaction with the deer movement data thereby aiding habitat selection model convergence and avoided extreme,



**Figure 3.** Distance.

unstable selection estimates.

In addition to land use types, we also acquired woody linear feature (i.e., hedgerows) data from UKCEH . This dataset maps the woody linear features across the UK (e.g., woodland) as polylines, based on Ordnance Survey maps and the 2007 UKCEH Land Cover Map .

We converted the polygon spatial data into a raster, where 1 == hedgerow, and used that rasterisation to generate a distance to hedgerow raster for the entire study landscape. We conducted the same process to create a distance to woodland raster, where we calculated the distance from any area the UKCEH land use data classed as deciduous or coniferous woodland. These distance rasters allowed for easy extraction of the distance to the nearest hedgerow and woodland for all locations.

We acquired road data from OS map open GOV licensed . We created a binary variable describing crossing events for all steps, with all steps that crossed one or more of the roads being classed as 1. This binary variable allowed us to estimate the likelihood deer cross a road and therefore the level of barrier roads present.

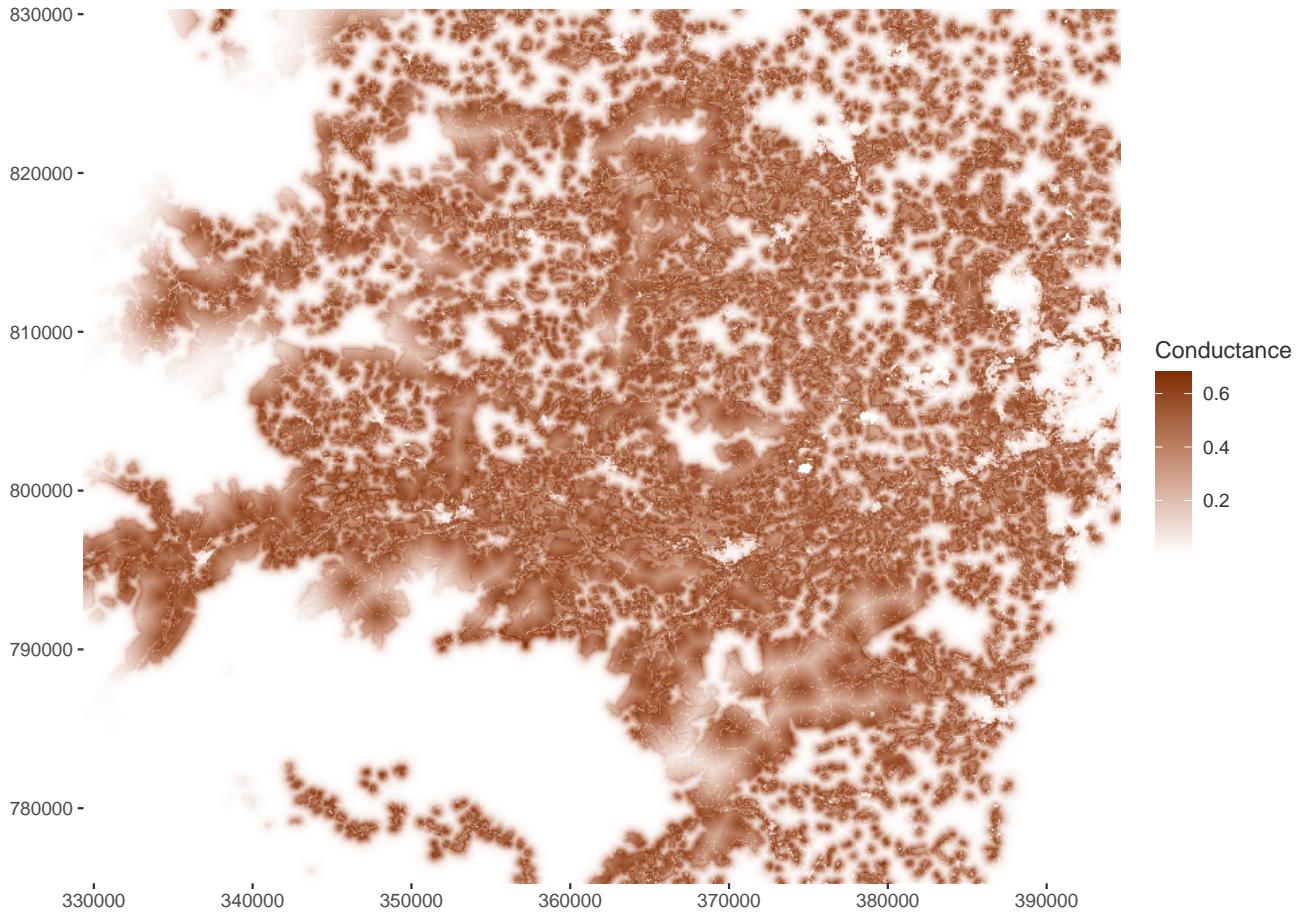
To ensure compatibility between all data sources, we ensured all data was projected into the British National Grid (BNG) coordinate reference system (OSGB36, epsg: 27700) before undertaking analysis.

The Poisson model formulae consisted of land use (a 8-term category variable formed into 7 dummy variables, with deciduous woodland placed as the reference category, barren and other excluded), distance to woodland (continuous in m), distance to hedgerow (continuous in m), road crossing (binary). In addition to these selection focused predictors, we several movement predictors including step length, log step length, and cos turn angle, as well as the interaction between step length and log of step length with all land use types. As the population Possion model contained all individuals the formula required in the inclusion of fixed Gaussian processes accounting for the time step as well as the individual. This formulation, namely the fixed Gaussian processes, as described by Muff et al allows for the efficient estimation of population level selection using integrated nested Laplace approximation (INLA) .

## 2.4 Translating Preferences into Connectivity

Once we had estimates of selection and effect for all the environmental aspects of interest, we spatially mapped those estimated covariates back onto the landscape, resulting in a map of conductance (Fig. @ref:(figconductancePlot)). This process took the form of generating a model matrix based on all environmental covariates at each cell in the

overall landscape raster and using the fitted Poisson model coefficients to calculate predicted deer use for each cell. For roads we rasterised the road polyline data and used the road crossing coefficient to create a conductance value for the road cells. For the covariates that involved movement interactions, we used the mean step and turn angle in calculations of predicted use. We replaced any areas with 0 conductance with 1e-12 to avoid any dead-ends or any absolute barriers. For we ignored the uncertainty surrounding the selection estimates, instead relying on just the point estimates for the resistance mapping.



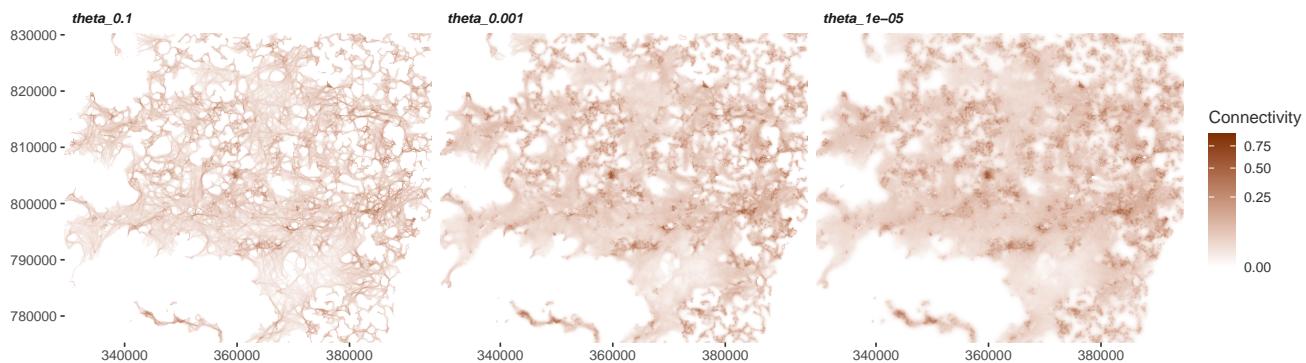
**Figure 4.** Conductance

We used random shortest paths, that consists of generating random walks between locations, to simulate potential connectivity across the landscapes. This method was used to show the connectivity of areas used by RIENDEER(?), and offers a mechanism for calibrating the connectivity maps to both the habitat selection and movement path characteristics. We used habitat patches as the sources of our random locations. We limited calculations of random shortest paths to pairs of patches that existed within 750 m of each other. We selected 750 m as that represented the mean longest axis of the 99% HR estimate of the deer (excluding outlying non-contiguous portions of the area polygons). For every pairing, we generated 3 start and end locations in each patch and ran random shortest paths between these locations. We repeated this for every pairing of patches. We elected not to generate start or end points in patches less than 0.5ha in areas, following . This was due to these patches being unlikely to host deer populations and the reduction in origin patches aided computational costs. The 0.5 ha threshold meant that 3415 Aberdeen patches and 154 Wessex patches were included in the generation of random shortest paths, thereby necessitating optimisations for the sake of compute time. Instead of considering the entire landscape when calculating the paths, the vast majority of which would not play a role between two neighbouring patches, we cropped the conductance raster to only include at area surrounding the randomly selected start and end points. This cropped area was 750m in all directions from the start and end points; where start and end points exceeded four times 750m apart from each other we extended the cropped area to half the distance between the two points in all directions. This

was required to avoid prohibitively computational intensive paths (i.e., those that encompassed nearly the entire landscape) caused by patches with large areas. A further optimisation was required to run the random shortest paths efficiently, we limited the distance between start and end location to four times the mean longest axis of the 99% HR estimate (4 times 756 m). This adaptive cropping allowed for much faster calculation, while retaining sufficient raster landscape for circuitous paths to be estimated.

Once every walk was complete, the resulting rasters describing the likelihood of a deer crossing a cell are compiled into a single landscape raster describing connectivity and standardised between 0 and 1 (where 0 is low connectivity and 1 is high connectivity).

A key consideration in these walks is how random the paths are. We elected to run walks at 3 different levels of randomness ( $\theta$ ; with 1 being close to a least cost path, and  $10^{-5}$  being the most random and diffuse walks; Fig. 5).



**Figure 5.** The connectivity maps under differing levels of  $\theta$  (randomness) in the random shortest paths between patches. The colour scale is squarerooted to better differentiate low connectivity areas.

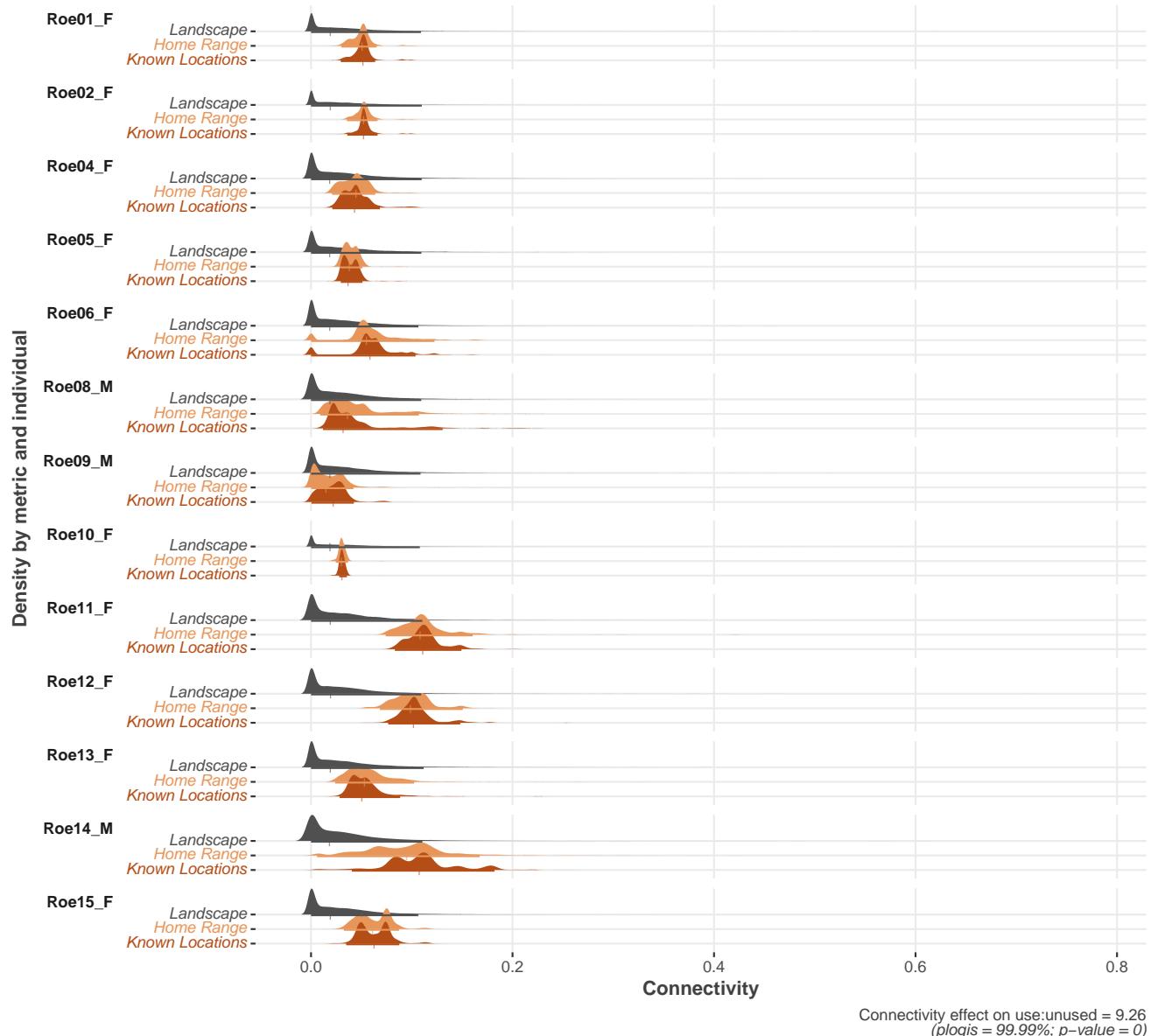
To determine what level of randomness best reflected the realised movements of the deer, we compared the resulting connectivity maps to dynamic Brownian Bridge Movement Models. Dynamic Brownian Bridge Movement Models run a series of random walks between defined start and end points, from the summation of these walks you can extract a rasterised occurrence distribution. Critically the dBBMM walks are calibrated to the movement capacity of the animal through rolling window (i.e., a number of data points) that summarises the movement rate during that time. Additionally within that window, a margin (a subset of data points) is used to detect any sudden changes in movement capacity that may be reflect of behavioural/movement mode changes. Therefore the dBBMMs provide an estimate of how diffuse the movements could be between known locations. We ran dBBMMs for all roe deer with a window size of 29 and a margin of 5, that provided estimates of motion variance on roughly a weekly sliding window with the allowance of sudden motion variance changes day to day (margin). We compared the dBBMMs to the connectivity maps constructed with varying levels of randomness, and used mean squared error to determine which level of randomness best fit the movement data. We used the connectivity map created using that  $\theta$  value for all subsequent analysis.

## 2.5 Validation

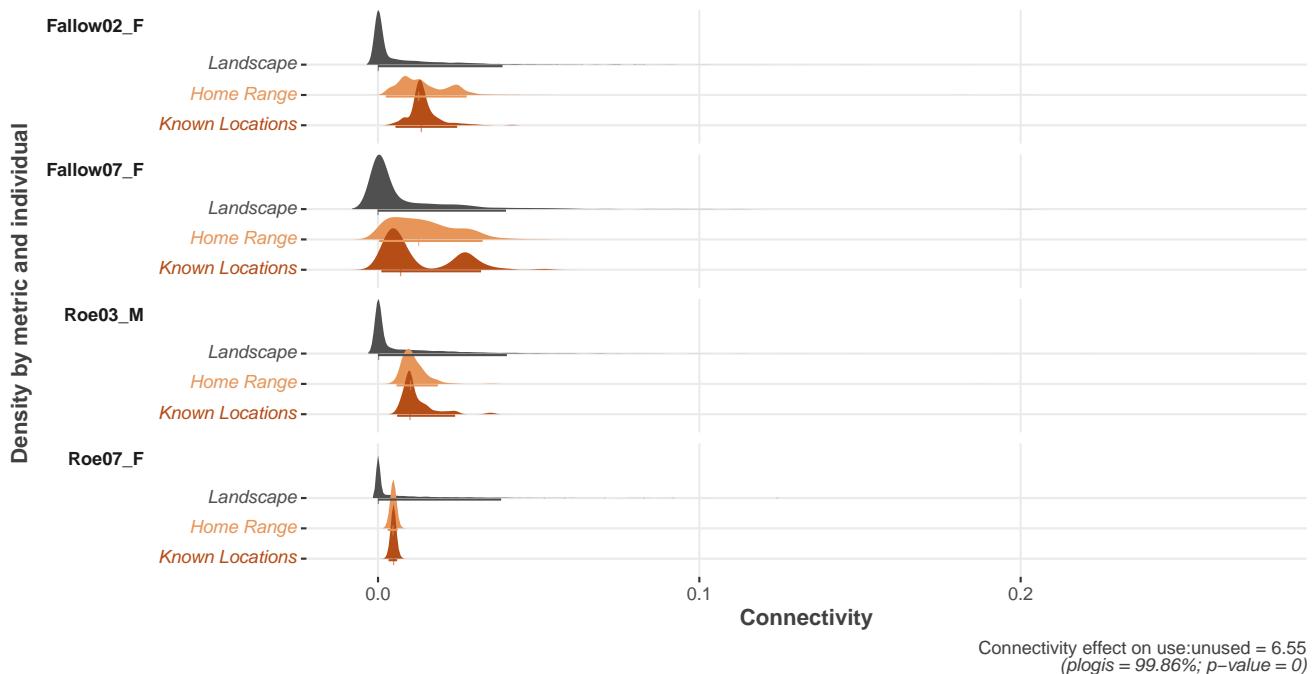
We examined whether the connectivity maps generated matched the observed movements of roe deer using a logistic regression. The model was supplied with the known locations of deer as well as 10 randomly generated points per known deer location across the landscape, all of which had associated connectivity values. We formulated the model to predict whether a point was used or random based on the connectivity values, and we included a random effect for deer ID. The expectation was that the model coefficients would indicate that deer locations were positively associated with higher connectivity values.

## 3 Results

For our Aberdeen Roe Deer home range sizes ranged from 32.2 to 122.5, with a weighted mean of 65.3 ha (95% CI 55.5-75.8). Largely the deer appeared range resident; however, a couple of individuals may show evidence of a range shift during the tracking period. All bar two individuals found OU models to fit best, with the remaining



**Figure 6.** Connectivity values distributions.



**Figure 7.** Connectivity values distributions.

two being better described by OUF models. All best fitting models were anisotropic, suggesting these Roe Deer are inhabiting non-uniform home ranges (i.e., not being as wide as they are tall).

The Poisson habitat selection model reveal a general tendency for Roe Deer to remain closer to the woodland patches (-0.0051; 95% CI -0.0082 to -0.0023), and with no significant care for hedgerows (-0.00091; 95% CI -0.0019 to 1e-04; Fig. 9)). The model also revealed that roads play a significant role in reducing connectivity across the landscape, with a significantly negative coefficient (-0.76; 95% CI -1.2 to -0.35).

Overall selection revealed significant selection for open shrubland (1.5; 95% CI 0.56 to 2.2) and tall grassland (0.57; 95% CI 0.17 to 1), with other relationships being less clear. Movement was most impacted by short grassland, tall grassland, and cropland, all showing the same pattern. These step lengths in these land uses to be lower, as seen in coefficients for log step length (short grassland 0.0044; 95% CI 0.0018 to 0.007; tall grassland 0.00049; 95% CI 5.4e-05 to 0.00092; cropland 0.0012; 95% CI 0.00064 to 0.0017) but with a larger tail to the gamma distribution (i.e. larger coefficient for step lengths; short grassland -0.76; 95% CI -1.2 to -0.32; tall grassland -0.18; 95% CI -0.25 to -0.12; cropland -0.16; 95% CI -0.24 to -0.076).

Estimates regarding human settlements were paired with very wide confidence intervals (-70; 95% CI -180 to 36), likely a result of minimal interaction with the tracked deer making estimation difficult. There are therefore treated as the point estimate suggests as areas of very low conductivity in further analysis.

Connectivity maps and validation

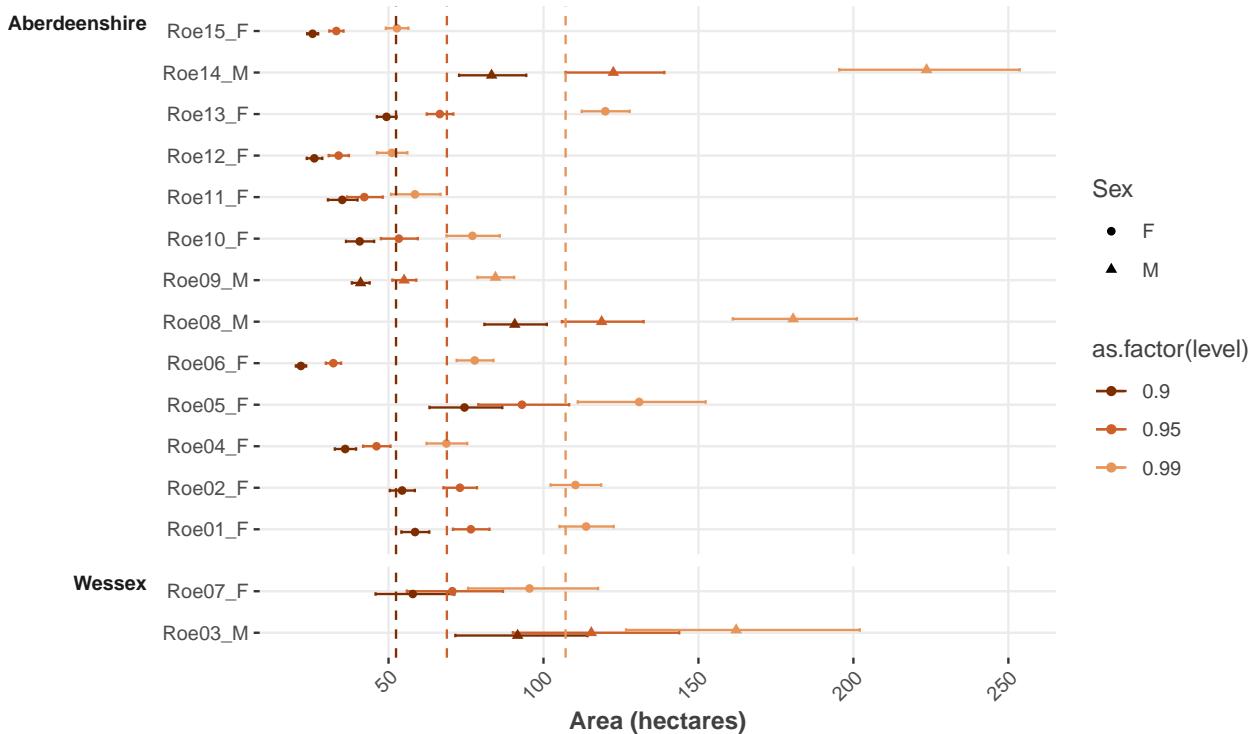
## 4 Discussion

### 4.1 Conlcusions

### 5 Acknowledgements

### 6 Software availability

For all analysis we used R (v.4.4.2) ([R Core Team, 2024](#)), and R Studio (v.2024.12.0+467) ([Posit team, 2024](#)). For analysis of animal movement data we used amt (v.0.2.2.0) ([Signer, Fieberg & Avgar, 2019](#)), ctmm (v.1.2.0) ([Fleming & Calabrese, 2023](#)), and move (v.4.2.6) ([Kranstauber, Smolla & Scharf, 2024](#)). For general data manipulation we used glue (v.1.8.0) ([Hester & Bryan, 2024](#)), sjmisc (v.2.8.10) ([Lüdecke, 2018](#)), tidyverse (v.2.0.0) ([Wickham et al., 2019](#)), and units (v.0.8.5) ([Pebesma, Mailund & Hiebert, 2016](#)). For project and code management we used here (v.1.0.1) ([Müller, 2020](#)), tarchetypes (v.0.11.0) ([Landau, 2021a](#)), and targets (v.1.9.0) ([Landau, 2021b](#)). For visu-



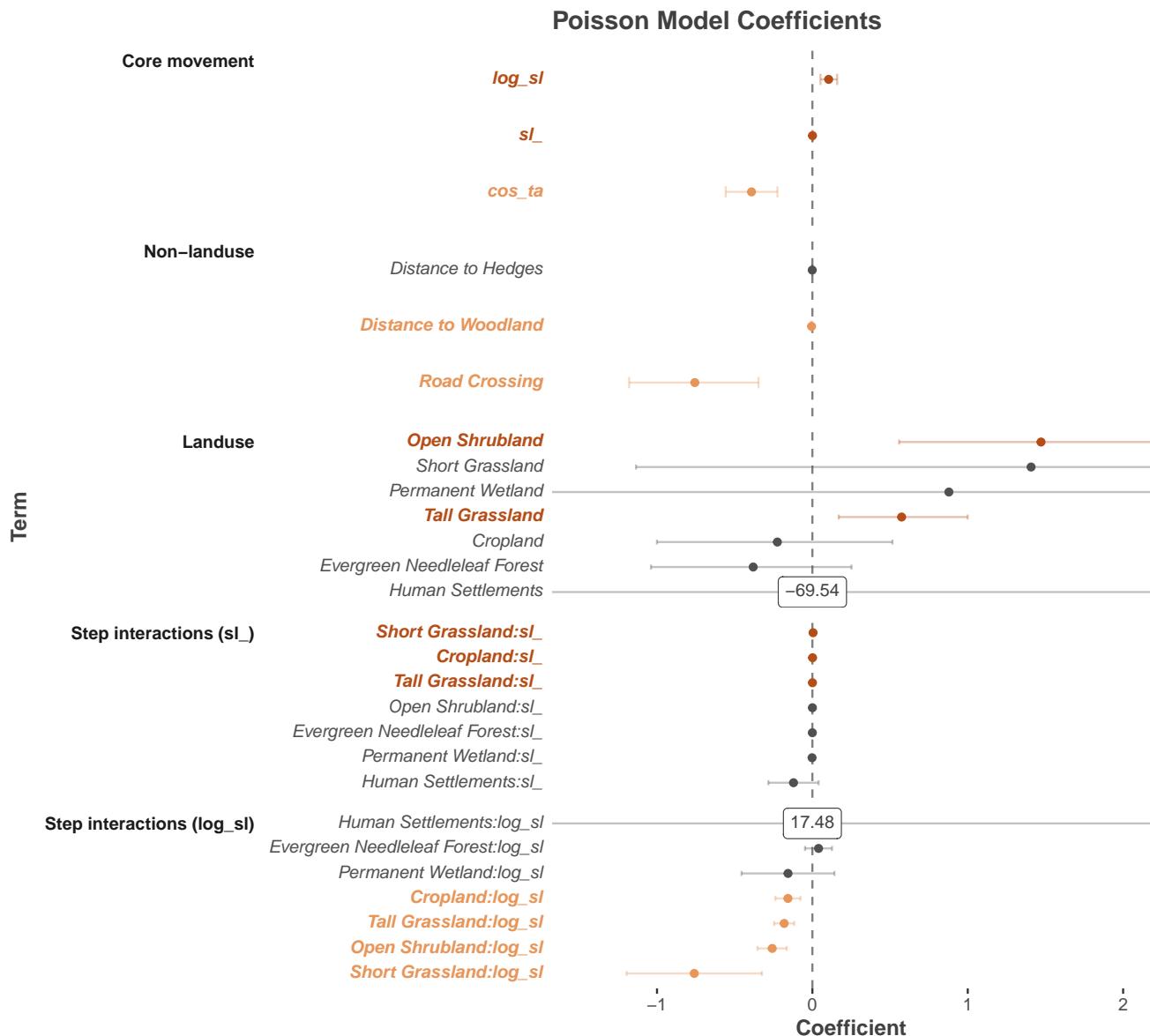
**Figure 8.** Home range size.

alisation we used the following as expansions from the tidyverse suite of packages: `ggdist` (v.3.3.2) (Kay, 2024a,b), `ggridges` (v.0.5.6) (Wilke, 2024), `ggeffect` (v.0.1.2) (Wilke & Wiernik, 2022), `patchwork` (v.1.3.0) (Pedersen, 2024), and `scales` (v.1.3.0) (Wickham, Pedersen & Seidel, 2023). Other packages we used were `boot` (v.1.3.31) (A. C. Davison & D. V. Hinkley, 1997; Angelo Canty & B. D. Ripley, 2024), `doParallel` (v.1.0.17) (Corporation & Weston, 2022), `foreach` (v.1.5.2) (Microsoft & Weston, 2022), and `usethis` (v.3.0.0) (Wickham et al., 2024). To generate typeset outputs we used `bookdown` (v.0.42) (Xie, 2016, 2025), and `rmarkdown` (v.2.29) (Xie, Allaire & Grolemund, 2018; Xie, Dervieux & Riederer, 2020; Allaire et al., 2024). To manipulate and manage spatial data we used `gdistance` (v.1.6.4) (van Etten, 2017), `raster` (v.3.6.30) (Hijmans, 2024a), `sf` (v.1.0.19) (Pebesma, 2018; Pebesma & Bivand, 2023), `sp` (v.2.1.4) (Pebesma & Bivand, 2005; Bivand, Pebesma & Gomez-Rubio, 2013), `terra` (v.1.7.83) (Hijmans, 2024b), and `tidyterra` (v.0.6.1) (Hernangómez, 2023). To run models and explore model outputs we used `effects` (v.4.2.2) (Fox, 2003; Fox & Hong, 2009; Fox & Weisberg, 2018, 2019), `INLA` (v.24.6.27) (@ Martins et al., 2013; @ Lindgren & Rue, 2015; De Coninck et al., 2016; Rue et al., 2017; Verbosio et al., 2017; Bakka et al., 2018; Kourounis, Fuchs & Schenk, 2018), `lme4` (v.1.1.35.5) (Bates et al., 2015), and `performance` (v.0.12.4) (Lüdecke et al., 2021).

## 7 Data availability

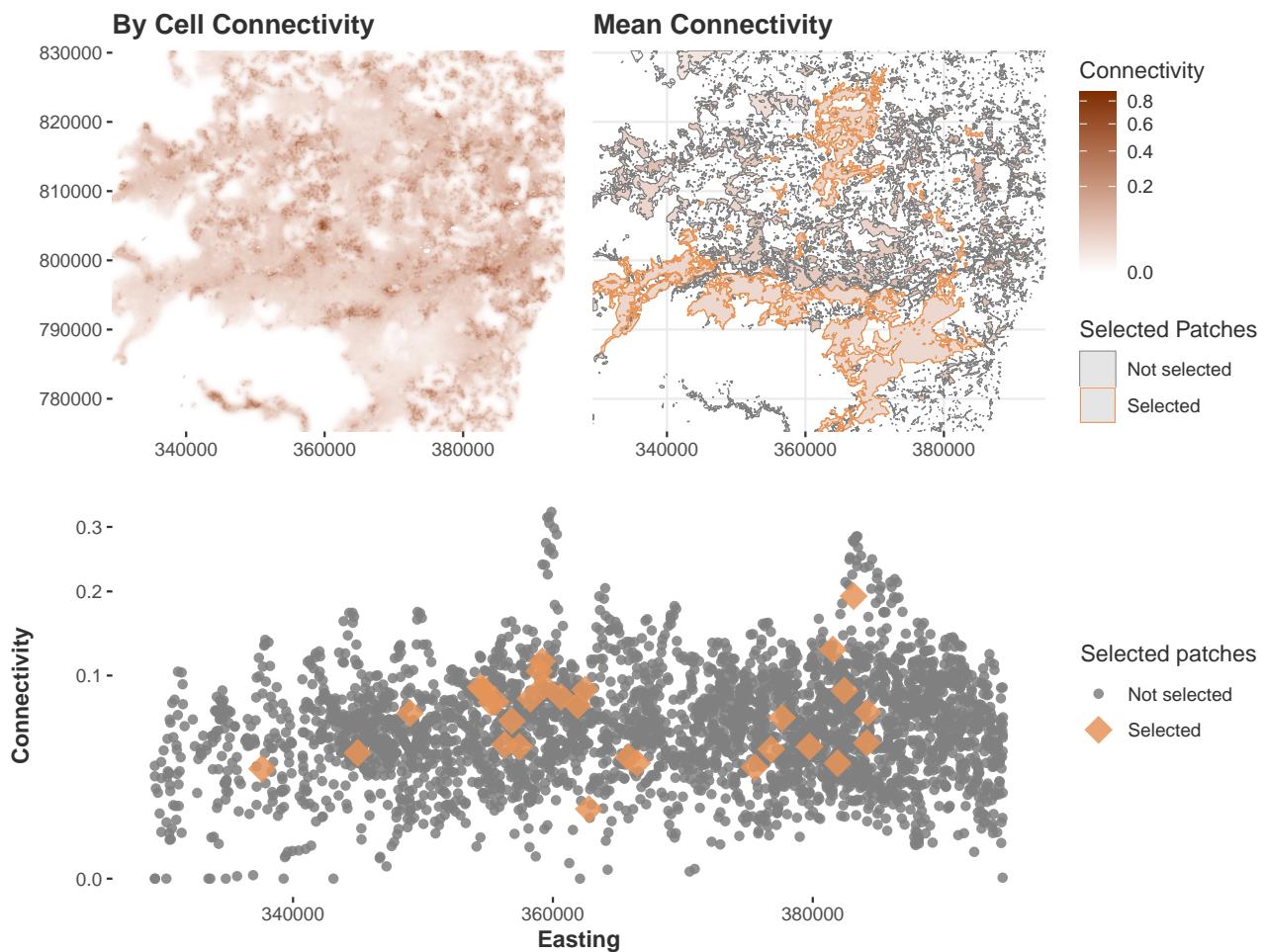
## 8 Author Contributions

## 9 Supplementary Material

**Figure 9.** Poisson coefficients.

### Aberdeenshire landscape connectivity and patches

$\theta = 1e-05$ , model = Pois



**Figure 10.** Poisson coefficients.

## References

- A. C. Davison, D. V. Hinkley. 1997. *Bootstrap methods and their applications*. Cambridge: Cambridge University Press.
- Allaire J, Xie Y, Dervieux C, McPherson J, Luraschi J, Ushey K, Atkins A, Wickham H, Cheng J, Chang W, Iannone R. 2024. *rmarkdown: Dynamic documents for r*.
- Angelo Canty, B. D. Ripley. 2024. *boot: Bootstrap r (s-plus) functions*.
- Bakka H, Rue H, Fuglstad G-A, Riebler AI, Bolin D, Illian J, Krainski E, Simpson DP, Lindgren FK. 2018. *Spatial modelling with INLA: A review*. *WIREs (Invited extended review)* xx:xx–xx.
- Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67:1–48. DOI: [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- Bivand RS, Pebesma E, Gomez-Rubio V. 2013. *Applied spatial data analysis with R, second edition*. Springer, NY.
- Calabrese JM, Fleming CH, Gurarie E. 2016. Ctmm: An R Package for Analyzing Animal Relocation Data As a Continuous-Time Stochastic Process. *Methods in Ecology and Evolution* 7:1124–1132. DOI: [10.1111/2041-210X.12559](https://doi.org/10.1111/2041-210X.12559).
- Corporation M, Weston S. 2022. *doParallel: Foreach parallel adaptor for the “parallel” package*.
- De Coninck A, De Baets B, Kourounis D, Verbosio F, Schenk O, Maenhout S, Fostier J. 2016. Needles: Toward large-scale genomic prediction with marker-by-environment interaction. *Genetics* 203:543–555. DOI: [10.1534/genetics.115.179887](https://doi.org/10.1534/genetics.115.179887).
- Fleming CH, Calabrese JM. 2017. A new kernel density estimator for accurate home-range and species-range area estimation. *Methods in Ecology and Evolution* 8:571–579. DOI: [10.1111/2041-210X.12673](https://doi.org/10.1111/2041-210X.12673).
- Fleming CH, Calabrese JM. 2023. *ctmm: Continuous-time movement modeling*.
- Fleming CH, Fagan WF, Mueller T, Olson KA, Leimgruber P, Calabrese JM. 2015. Rigorous home range estimation with movement data: A new autocorrelated kernel density estimator. *Ecology* 96:1182–1188.
- Fleming CH, Noonan MJ, Medici EP, Calabrese JM. 2019. Overcoming the challenge of small effective sample sizes in home-range estimation. *Methods in Ecology and Evolution* 10:1679–1689. DOI: [10.1111/2041-210X.13270](https://doi.org/10.1111/2041-210X.13270).
- Fox J. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8:1–27. DOI: [10.18637/jss.v008.i15](https://doi.org/10.18637/jss.v008.i15).
- Fox J, Hong J. 2009. Effect displays in R for multinomial and proportional-odds logit models: Extensions to the effects package. *Journal of Statistical Software* 32:1–24. DOI: [10.18637/jss.v032.i01](https://doi.org/10.18637/jss.v032.i01).
- Fox J, Weisberg S. 2018. Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software* 87:1–27. DOI: [10.18637/jss.v087.i09](https://doi.org/10.18637/jss.v087.i09).
- Fox J, Weisberg S. 2019. *An r companion to applied regression*. Thousand Oaks CA: Sage.
- Hernangómez D. 2023. Using the tidyverse with terra objects: The tidyterra package. *Journal of Open Source Software* 8:5751. DOI: [10.21105/joss.05751](https://doi.org/10.21105/joss.05751).
- Hester J, Bryan J. 2024. *glue: Interpreted string literals*.
- Hijmans RJ. 2024a. *raster: Geographic data analysis and modeling*.
- Hijmans RJ. 2024b. *terra: Spatial data analysis*.
- Kay M. 2024a. ggdist: Visualizations of distributions and uncertainty in the grammar of graphics. *IEEE Transactions on Visualization and Computer Graphics* 30:414–424. DOI: [10.1109/TVCG.2023.3327195](https://doi.org/10.1109/TVCG.2023.3327195).
- Kay M. 2024b. *ggdist: Visualizations of distributions and uncertainty*. DOI: [10.5281/zenodo.3879620](https://doi.org/10.5281/zenodo.3879620).
- Kourounis D, Fuchs A, Schenk O. 2018. Towards the next generation of multiperiod optimal power flow solvers. *IEEE Transactions on Power Systems* PP:1–10.
- Kranstauber B, Smolla M, Scharf AK. 2024. *move: Visualizing and analyzing animal track data*.
- Landau WM. 2021a. *tarchetypes: Archetypes for targets*.
- Landau WM. 2021b. *The targets r package: A dynamic make-like function-oriented pipeline toolkit for reproducibility and high-performance computing*. *Journal of Open Source Software* 6:2959.
- Lindgren F, Rue H. 2015. Bayesian spatial modelling with R-INLA. *Journal of Statistical Software* 63:1–25.
- Lüdecke D. 2018. sjmisc: Data and variable transformation functions. *Journal of Open Source Software* 3:754. DOI: [10.21105/joss.00754](https://doi.org/10.21105/joss.00754).
- Lüdecke D, Ben-Shachar MS, Patil I, Waggoner P, Makowski D. 2021. performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software* 6:3139. DOI: [10.21105/joss.03139](https://doi.org/10.21105/joss.03139).
- Martins TG, Simpson D, Lindgren F, Rue H. 2013. Bayesian computing with INLA: New features. *Computational Statistics and Data Analysis* 67:68–83.
- Microsoft, Weston S. 2022. *foreach: Provides foreach looping construct*.
- Müller K. 2020. *here: A simpler way to find your files*.
- Pebesma E. 2018. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* 10:439–446. DOI: [10.32614/RJ-2018-009](https://doi.org/10.32614/RJ-2018-009).
- Pebesma EJ, Bivand R. 2005. *Classes and methods for spatial data in R*. *R News* 5:9–13.
- Pebesma E, Bivand R. 2023. *Spatial Data Science: With applications in R*. Chapman and Hall/CRC. DOI: [10.1201/9780429459010](https://doi.org/10.1201/9780429459010).

- Pebesma E, Mailund T, Hiebert J. 2016. Measurement units in R. *R Journal* 8:486–494. DOI: [10.32614/RJ-2016-061](https://doi.org/10.32614/RJ-2016-061).
- Pedersen TL. 2024. *patchwork: The composer of plots*.
- Posit team. 2024. *RStudio: Integrated development environment for r*. Boston, MA: Posit Software, PBC.
- R Core Team. 2024. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rue H, Riebler AI, Sørbye SH, Illian JB, Simpson DP, Lindgren FK. 2017. Bayesian computing with INLA: A review. *Annual Reviews of Statistics and Its Applications* 4:395–421.
- Signer J, Fieberg J, Avgar T. 2019. Animal movement tools (amt): R package for managing tracking data and conducting habitat selection analyses. *Ecology and Evolution* 9:880–890.
- Silva I, Fleming CH, Noonan MJ, Alston J, Folta C, Fagan WF, Calabrese JM. 2022. Autocorrelation-informed home range estimation: A review and practical guide. *Methods in Ecology and Evolution* 13:534–544. DOI: [10.1111/2041-210X.13786](https://doi.org/10.1111/2041-210X.13786).
- van Etten J. 2017. R package gdistance: Distances and routes on geographical grids. *Journal of Statistical Software* 76:1–21. DOI: [10.18637/jss.v076.i13](https://doi.org/10.18637/jss.v076.i13).
- Verbosio F, De Coninck A, Kourounis D, Schenk O. 2017. Enhancing the scalability of selected inversion factorization algorithms in genomic prediction. *Journal of Computational Science* 22:99–108.
- Wickham H, Averick M, Bryan J, Chang W, McGowan LD, François R, Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM, Müller K, Ooms J, Robinson D, Seidel DP, Spinu V, Takahashi K, Vaughan D, Wilke C, Woo K, Yutani H. 2019. Welcome to the tidyverse. *Journal of Open Source Software* 4:1686. DOI: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).
- Wickham H, Bryan J, Barrett M, Teucher A. 2024. *usethis: Automate package and project setup*.
- Wickham H, Pedersen TL, Seidel D. 2023. *scales: Scale functions for visualization*.
- Wilke CO. 2024. *ggridges: Ridgeline plots in “ggplot2”*.
- Wilke CO, Wiernik BM. 2022. *ggtext: Improved text rendering support for “ggplot2”*.
- Xie Y. 2016. *bookdown: Authoring books and technical documents with R markdown*. Boca Raton, Florida: Chapman; Hall/CRC.
- Xie Y. 2025. *bookdown: Authoring books and technical documents with r markdown*.
- Xie Y, Allaire JJ, Grolemund G. 2018. *R markdown: The definitive guide*. Boca Raton, Florida: Chapman; Hall/CRC.
- Xie Y, Dervieux C, Riederer E. 2020. *R markdown cookbook*. Boca Raton, Florida: Chapman; Hall/CRC.