Tick Deer Connectivity Methods

__*[1], __**[2], [3], [4], [3], [5], [6], [7], and [1]

[1]__
[2]__
[3]
[4]
[5]
[6]
[7]
*__
**__

Abstract
abstract text

Keywords

deer, rodent, connectivity

1 Methods

1.1 Estimating Preference and Probability of Occurrence

For the Roe Deer we made use of previously conducted analysis on their movements that aimed to determine their preference for various land use types, as well as how their movements are altered when within them . We took the coefficients of the population level poisson habitat preference model as a basis to estimate landscape level Roe Deer functional connectivity.

For Fallow Deer and Rodents, we lacked sufficient movement data to undertake the a similar procedure. Instead we followed the workflow outlined by Prima et al. (2024), using species distribution modelling to generate reasonable estimations of landscape resistance and then applying a circuit-theory approach <CITE - check wording> to determine how interconnected selected woodland patches were.

We collated occurrence data from GBIF and NBN Atlas for the Fallow Deer and Rodents, using searches from "*Dama dama*", "*Myodes glareolus*", "*Apodemus flavicollis*", "*Apodemus sylvaticus*", "*Sciurus carolinensis*", and "*Sciurus vulgaris*". We downloaded all records within the UK, for all species. We filtered the records as to only include those made via human observations, were recorded to sub-25m accuracy, and had verified species identities. Once filtered we used CoordinateCleaner to eliminate records that likely contained location errors on a species-by-species basis. Namely we removed records within 10km of a capital, within 1km of a country centroid, had identical decimal latitude and longitudes, had zero in either latitude or longitude, or were identified as outliers. We used the quantile method of identifying outliers, removing those records beyond 5 times the interquartile range of minimum distances to the next neighbour. Finally, we thinned the both the Fallow Deer and Rodent data so that no occurrence shared a raster cell with another occurrence. This process resulted in __ occurrence locations of Fallow Deer and ___ for rodents (__ *Myodes glareolus*; __ *Apodemus flavicollis*; __ *Apodemus sylvaticus*; __ *Sciurus carolinensis*; __ *Sciurus vulgaris*).

We generated a number pseudo-absences data points for both target taxa equal to three times the number of true-occurrences available; we repeated this process ___ times allowing for multiple model fits . In the case of Fallow Deer we limited the area pseudo-absences could occur to areas within 6.2km of a known Fallow Deer occurrence. As Fallow Deer have not uniformly spread across the UK, we excluded areas where they have not as of yet had the opportunity to colonise. For rodents we limited the pseudo-absences to Great Britain (minus Shetland and other outlying islands groups such as St. Kilda). We weighted the pseudo-absences using a bias layer to mitigate the impacts of the sampling biases associated with the true-occurrence data. For our bias layer, we used the human footprint index that combines a variety of human activities and land conversions into a single metric, and could serve as a proxy for the likelihood of observing our target species.

We used the biomod2 package to manage species distribution modelling . We fitted a variety of models: Artificial Neural Network, Boosted Regression Trees, Generalized Linear Model, and Random Forest. We intended to run Maximum Entropy and eXtreme Gradient Boosting Training, but the inclusion of categoric variables precluded their use. We used stratified cross validation to evaluate model performance, as we were not priorisitisng model transferability <CITE https://besjournals.onlinelibrary.wiley.com/doi/10.1111/2041-210X.12261>. We balanced stratification based on known occurrences both in x and y dimensions. We elected to use True Skill Statistic (TSS) as our primary metric of model performance . Once all models had been complete, we compiled all models into an ensemble model ready for projection to our target landscapes. We included all models that obtained TSS scores of over 0.25, and weighted models proportionately to their TSS score. Once the TSS weighted mean ensemble model had been created, we used it to project the probability of occurrence to the Aberdeenshire and Wessex landscapes.

1.2 Translating Preferences into Connectivity

For Roe Deer, 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.

For the Fallow Deer and Rodents the underlying conductance layer was generated from the final ensemble species

distribution model. In the case of the Fallow Deer we have sufficient data to parametrise the road resistance based on the results of the Fallow Deer Poisson model; therefore, we alter the resistance of the cells that corresponded to the presence of a road in the same way as described for the Roe Deer. We did not have comparable road crossing data for the rodents, meaning that their conductance raster equalled one describing probability of occurrence.
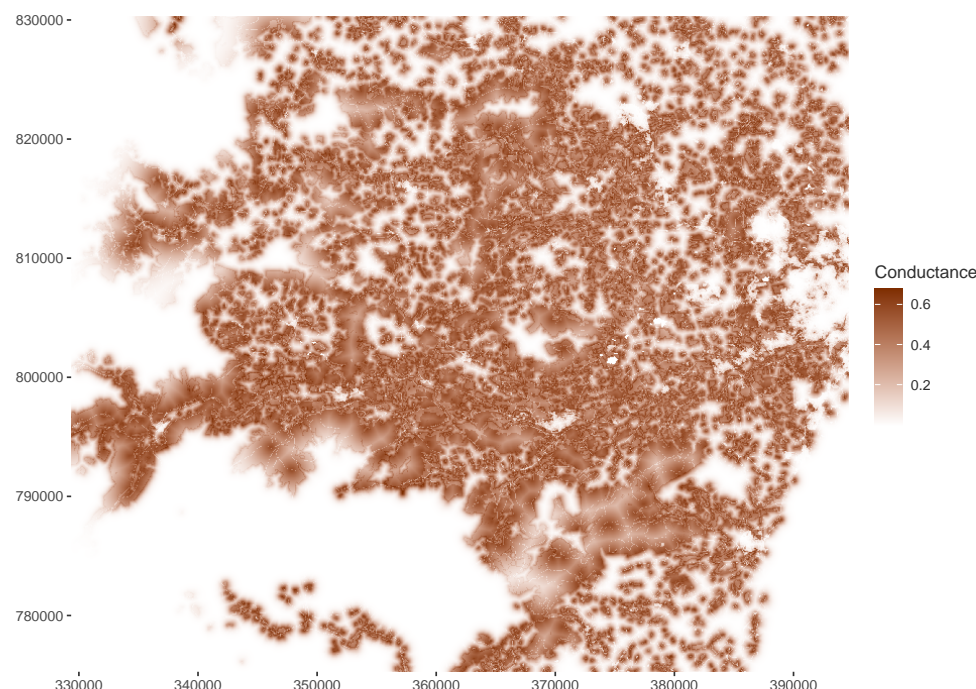

Figure 1. Conductance map of Aberdeen landscape developed from the coefficients of the Poisson habitat selection model. Higher values indicate easier movement for Roe Deer. Axis are m in British National Grid coordinate system.

For Roe Deer, we used random shortest paths, that consists of generating random walks between locations, to simulate potential connectivity across the landscapes based on the above described conductance/resistance raster. This method was used to show the connectivity of areas used by reindeer, and offers a mechanism for calibrating the connectivity maps to both the habitat selection and movement path characteristics (Panzacchi et al., 2016).

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 6 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 7564 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 complied 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. 2).
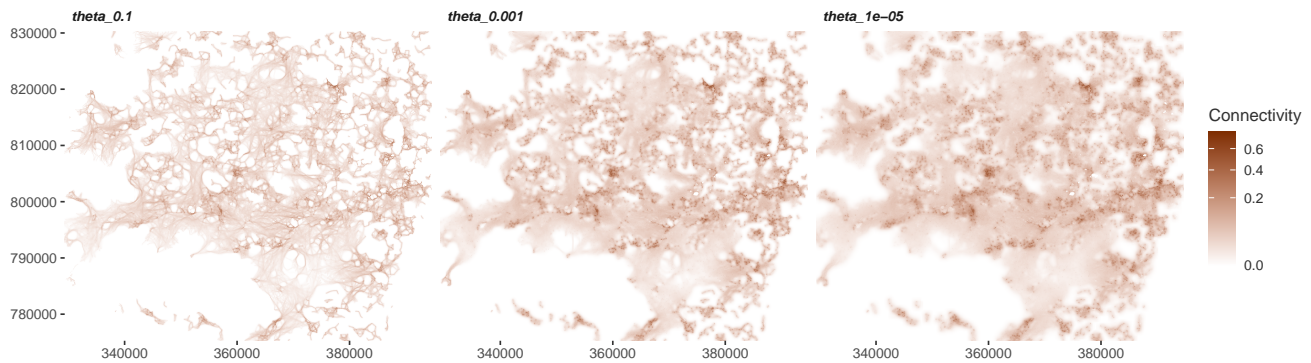
Figure 2. The connectivity maps under differing levels of theta (randomness) in the random shortest paths between patches. The colour scale is squarerooted to better differeniate low conenctivity areas. Axis are m in British National Grid coordinate system.

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 (Kranstauber et al., 2012; Kranstauber, Smolla & Scharf, 2024). 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 a 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.

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.

For the Fallow Deer and Rodents, we followed Prima et al. (2024) and used Omniscape.jl (; built on Circuitscape.jl To aid with computation costs, we selected a block size of __ for the Fallow Deer, and __ for the Rodents. We also varied the search radius between the groups, as each has different movement/dispersal abilities. In both instances we based our search radius on previously recorded estimates of home range size, based heavily on the data provide by the home range database . In both case we filtered the home range data to species/group and calculated a mean home range area. For the Fallow Deer their mean home range suggested a search radius of _; **whereas for Rodents a search radius of** was chosen. For both groups we used the previously denoted patches as sources/nodes (all with the same strength).
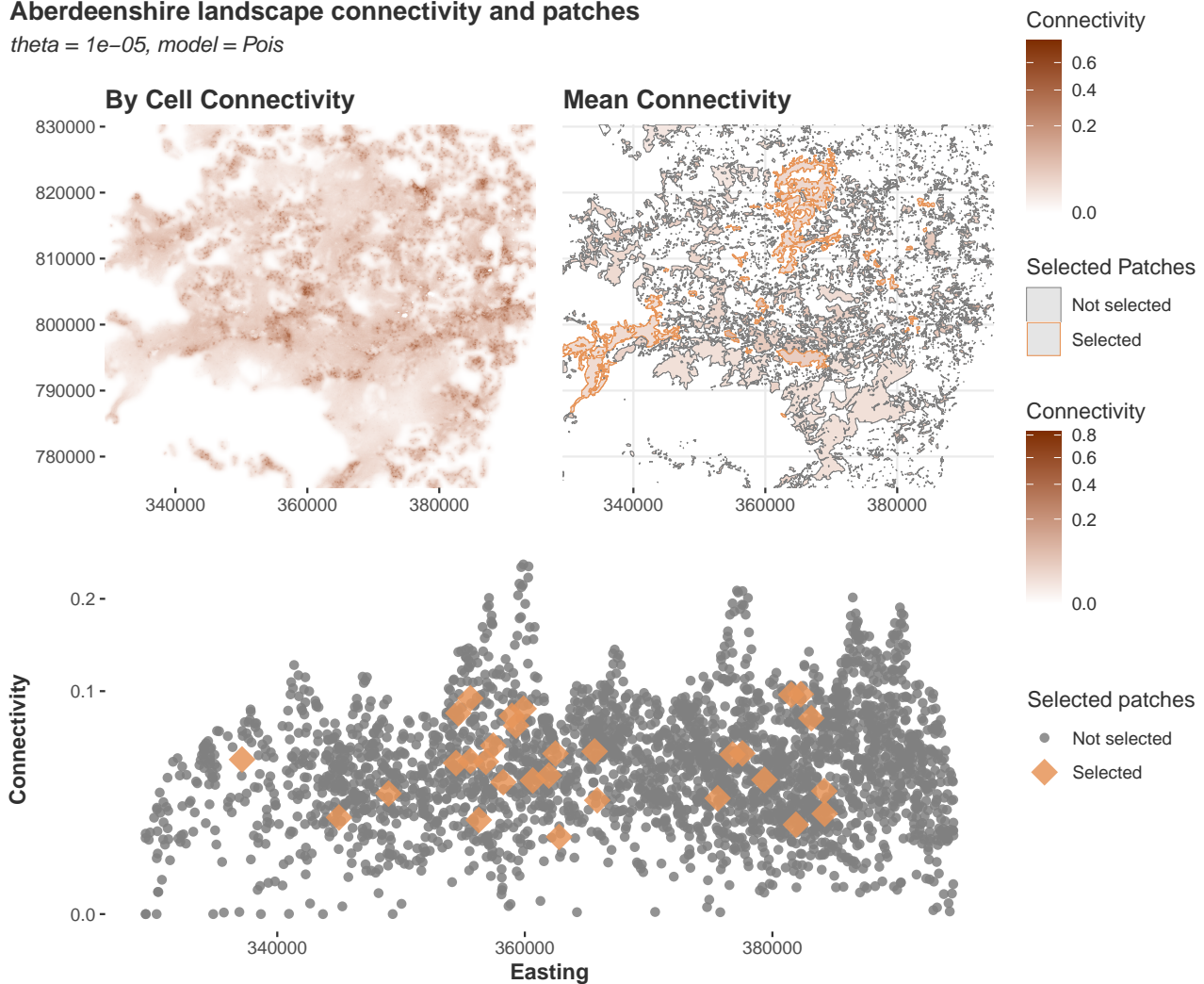
2   Results

2.1   Connectivity maps

(Fig. 3)

Aberdeenshire landscape connectivity and patches

theta = 1e−05, model = Pois

Figure 3. The resulting connectivity map for the Aberdeen landscape displayed on a cell by cell level and summarised by patch. Selected patches are highlighted in orange. Lower panel shows how the selected (tick sampled) patches relate to the overall distribution of patch connectivity values. N.b. connectivity scale is square-rooted. Axis are m in British National Grid coordinate system.
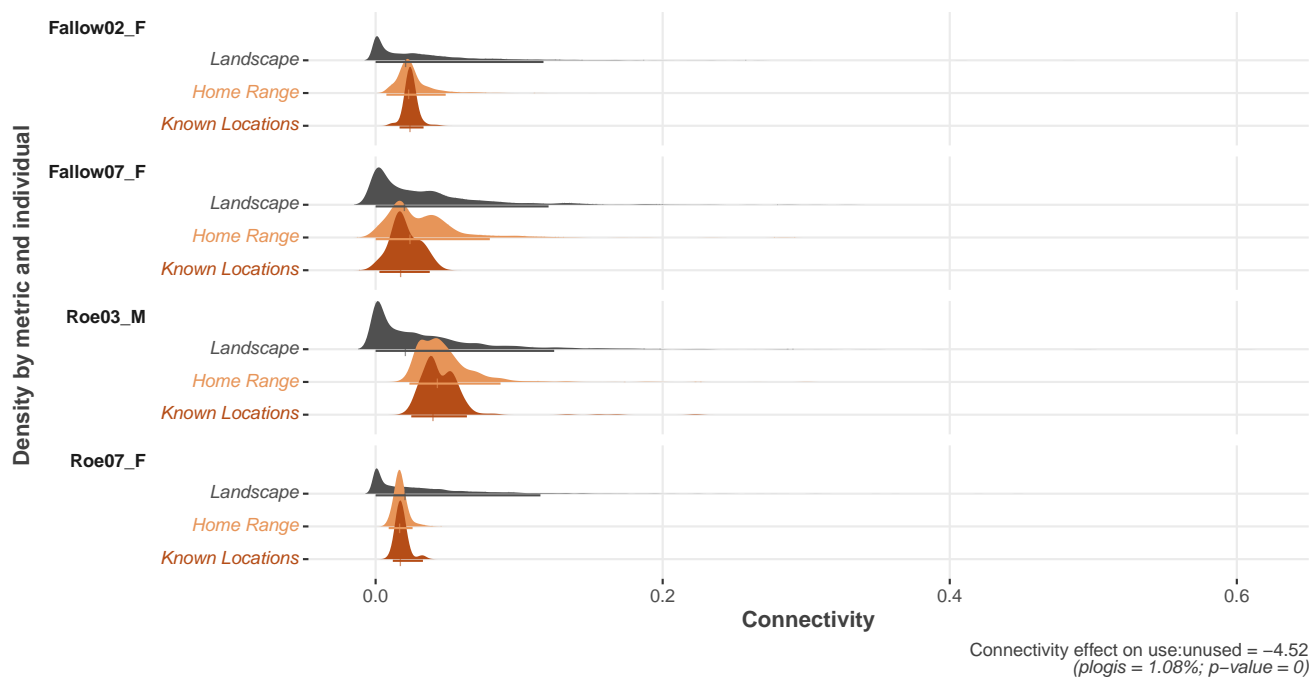
2.2   Validation

(Fig. 4) (Fig. 5) (Fig. 6)
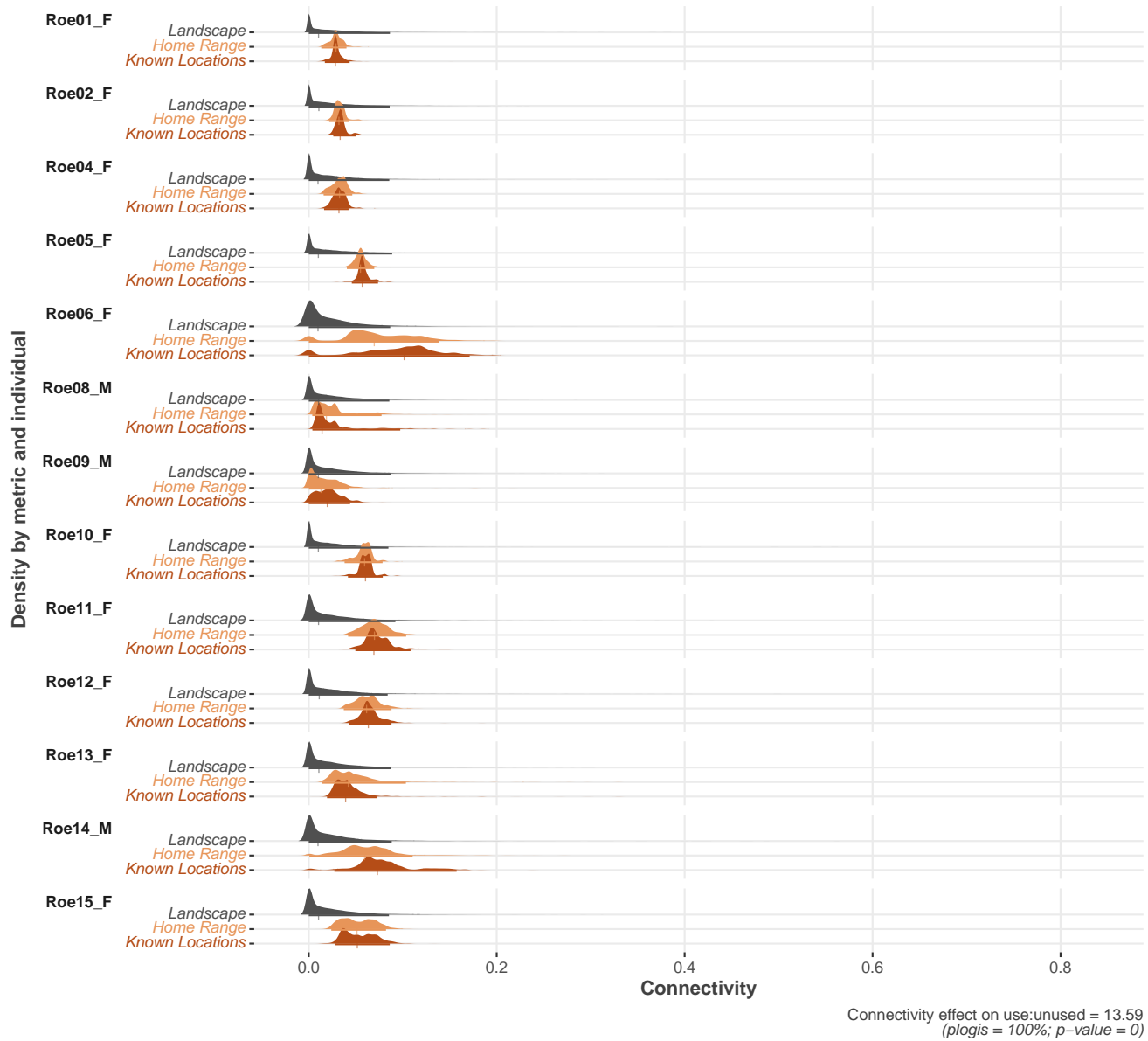
Figure 4. Distribution of connectivity values at known deer locations and randomly generated locations across the Aberdeen landscape, as well as random locations generated within the 95% home range. Lower caption describes the results of the logisitic regression testing whether connectivity can predict whether the location was random or a known deer location.

Figure 5. Distribution of connectivity values at known deer locations and randomly generated locations across the Wessex landscape, as well as random locations generated within the 95% home range. Lower caption describes the results of the logisitic regression testing whether connectivity can predict whether the location was random or a known deer location.
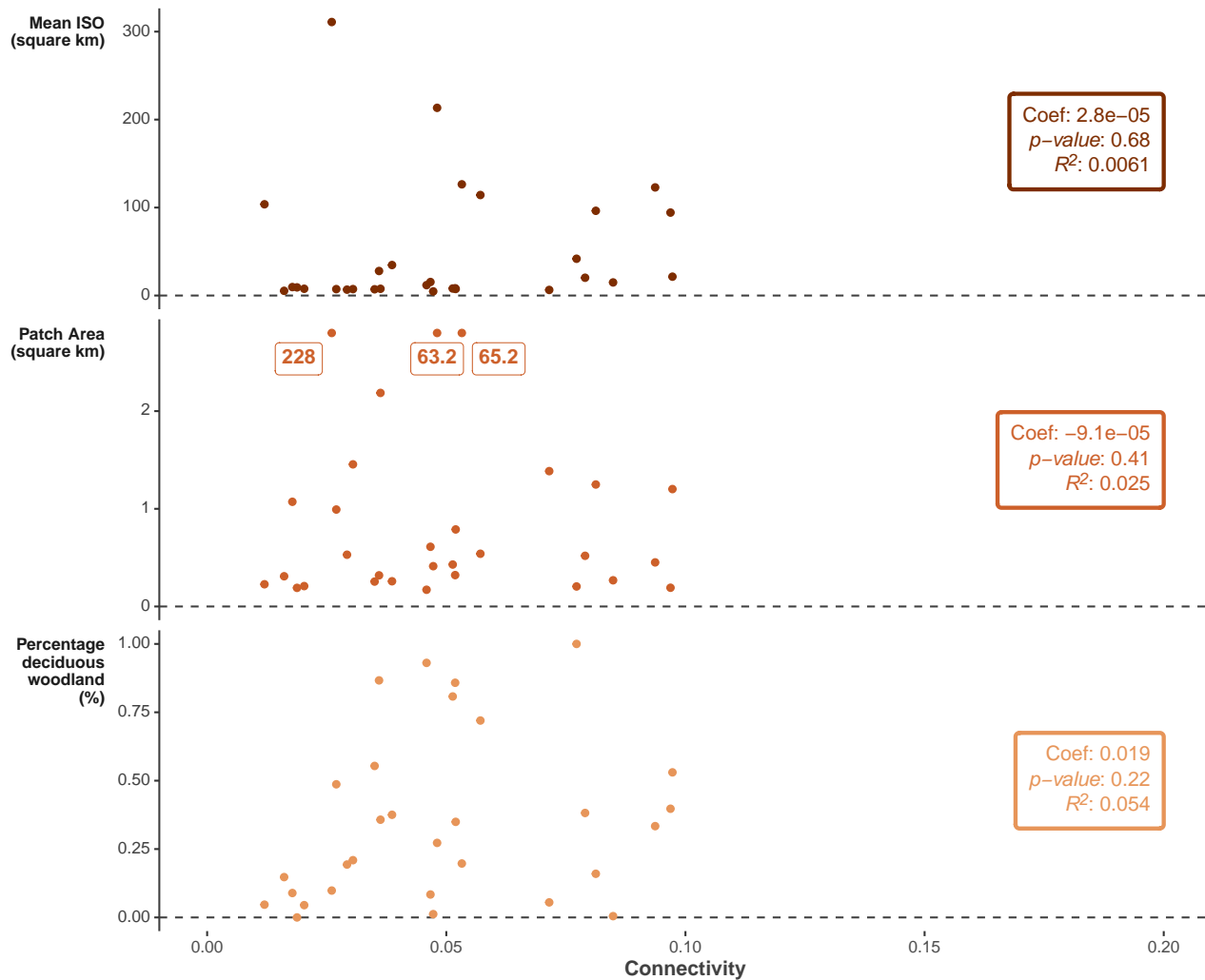
Figure 6. A comparison between the mean connectivity values derived from deer movement data and the structural metrics of connectivity on a patch by patch basis. Outlying patch areas are replaced with labels indicating their size in square kilometres. Boxes on the right describe the linear relationship between the two variables.

3   Acknowledgements

3.1   Software availablity

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 visualisation 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), ggtext (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), circular (v.0.5.1) (Agostinelli & Lund, 2024), doParallel (v.1.0.17) (Corporation & Weston, 2022), foreach (v.1.5.2) (Microsoft & Weston, 2022), knitr (v.1.49) (Xie, 2014, 2015, 2024), 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), lme4 (v.1.1.35.5) (Bates et al., 2015), and performance (v.0.12.4) (Lüdecke et al., 2021).

3.2   Data availabilty

3.3   Author Contributions

4   Supplementary Material

References

A. C. Davison, D. V. Hinkley. 1997. *Bootstrap methods and their applications*. Cambridge: Cambridge University Press.

Agostinelli C, Lund U. 2024. *R package `circular`: Circular statistics (version 0.5-1)*.

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*.

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.

Bivand RS, Pebesma E, Gomez-Rubio V. 2013. *Applied spatial data analysis with R, second edition*. Springer, NY.

Corporation M, Weston S. 2022. *doParallel: Foreach parallel adaptor for the "parallel" package*.

Fleming CH, Calabrese JM. 2023. *ctmm: Continuous-time movement modeling*.

Fox J. 2003. Effect displays in R for generalised linear models. *Journal of Statistical Software* 8:1–27. DOI: 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.

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.

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.

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.

Kay M. 2024b. *ggdist: Visualizations of distributions and uncertainty*. DOI: 10.5281/zenodo.3879620.

Kranstauber B, Kays R, Lapoint SD, Wikelski M, Safi K. 2012. A dynamic Brownian bridge movement model to estimate utilization distributions for heterogeneous animal movement. *Journal of Animal Ecology* 81:738–746. DOI: 10.1111/j.1365-2656.2012.01955.x.

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.

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.

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*.

Panzacchi M, Van Moorter B, Strand O, Saerens M, Kivimäki I, St. Clair CC, Herfindal I, Boitani L. 2016. Predicting the *continuum* between corridors and barriers to animal movements using Step Selection Functions and Randomized Shortest Paths. *Journal of Animal Ecology* 85:32–42. DOI: 10.1111/1365-2656.12386.

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.

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: 10.1201/9780429459016

Pebesma E, Mailund T, Hiebert J. 2016. Measurement units in R. *R Journal* 8:486–494. DOI: 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.

Prima M, Renaud J, Witté I, Suarez L, Rouveyrol P, Fernando M, Sacchi A, Cosentino F, Santini L, Maiorano L, Moreira F, Dertien J, Fernández N, Thuiller W. 2024. A comprehensive framework to assess multi-species landscape connectivity. *Methods in Ecology and Evolution* 15:2385–2399. DOI: 10.1111/2041-210X.14444.

R Core Team. 2024. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

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.

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.

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.

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. 2014. knitr: A comprehensive tool for reproducible research in R. In: Stodden V, Leisch F, Peng RD eds. *Implementing reproducible computational research*. Chapman; Hall/CRC,.

Xie Y. 2015. *Dynamic documents with R and knitr*. Boca Raton, Florida: Chapman; Hall/CRC.

Xie Y. 2016. *bookdown: Authoring books and technical documents with R markdown*. Boca Raton, Florida: Chapman; Hall/CRC.

Xie Y. 2024. *knitr: A general-purpose package for dynamic report generation in r*.

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.