Appendix C

Cumulative influence of infrastructure on reindeer space use: fitting habitat selection models

Bernardo Niebuhr, Bram van Moorter, Manuela Panzacchi

21 March, 2022

Abstract

In this document we fit habitat selection models to wild reindeer GPS data to assess if and how the impacts of multiple infrastructure affect mountain reindeer (*Rangifer tarandus*) habitat selection during summer. We describe the modeling approach and present the results and predictions from the fitted models.

Contents

Introduction

In this document we describe the procedures to fit habitat selection models to wild reindeer GPS data and assess if and how the impacts of multiple infrastructure affect mountain reindeer (*Rangifer tarandus*) habitat selection during summer. We first briefly describe the study area, the GPS data handling, and the environmental variables used in the analysis. We then describe the calculation of the infrastructure-related covariates using both the measures of cumulative influence and the influence of the nearest feature. These measures quantify the zone of influence (ZoI) as well as how the influence varies with the distance to infrastructure. Then, we describe the structure of the statistical models and the fitting procedures and present the results in details. We explore qualitatively the interpretation of both influence measures in single-infrastructure models, and then estimate the magnitude of the impacts and the ZoI of each infrastructure in multi-infrastructure models, to finally assess the combined impacts of infrastructure on reindeer habitat selection.

Material and Methods

Study area

The study area was the Hardangervidda wild reindeer area in Southern Norway, where the largest remaining population of mountain reindeer is found (Fig. C1). During summer, the area is mainly used for tourism. Hardangervidda is a big plateau surrounded by large roads around its contour, which corresponds to the lower part of the area (Fig. C2). Towards the upper, central part, there are small access roads that link the large highways to tourist cabins and a multitude of private cottages, which are also connected by a network of trails (Fig. C2). The area has 26 large tourist cabins which are constantly visited by many tourists and 24 smaller public cabins. In contrast, 14154 private cottages are spread throughout Hardangervidda.

Due to their high density, most areas (90%) in Hardangervidda are closer than 3 km from any private cottage and 5 km from the closest trail (Table C1). In contrast, more than 50% of the areas are farther than 13 km from large tourist cabins and 10 km from small tourist cabins. There are also many areas far from roads towards the central part of the Hardangervidda (Table C1).

I kept the other infrastructure in the description of the area (Fig. C2 and Table C1), but in the next section here I deal with only the private and public cabins and do not mentioned the others anymore. I also do not mention them in the main text. Should I also remove them from the description or it is ok to keep it like that?

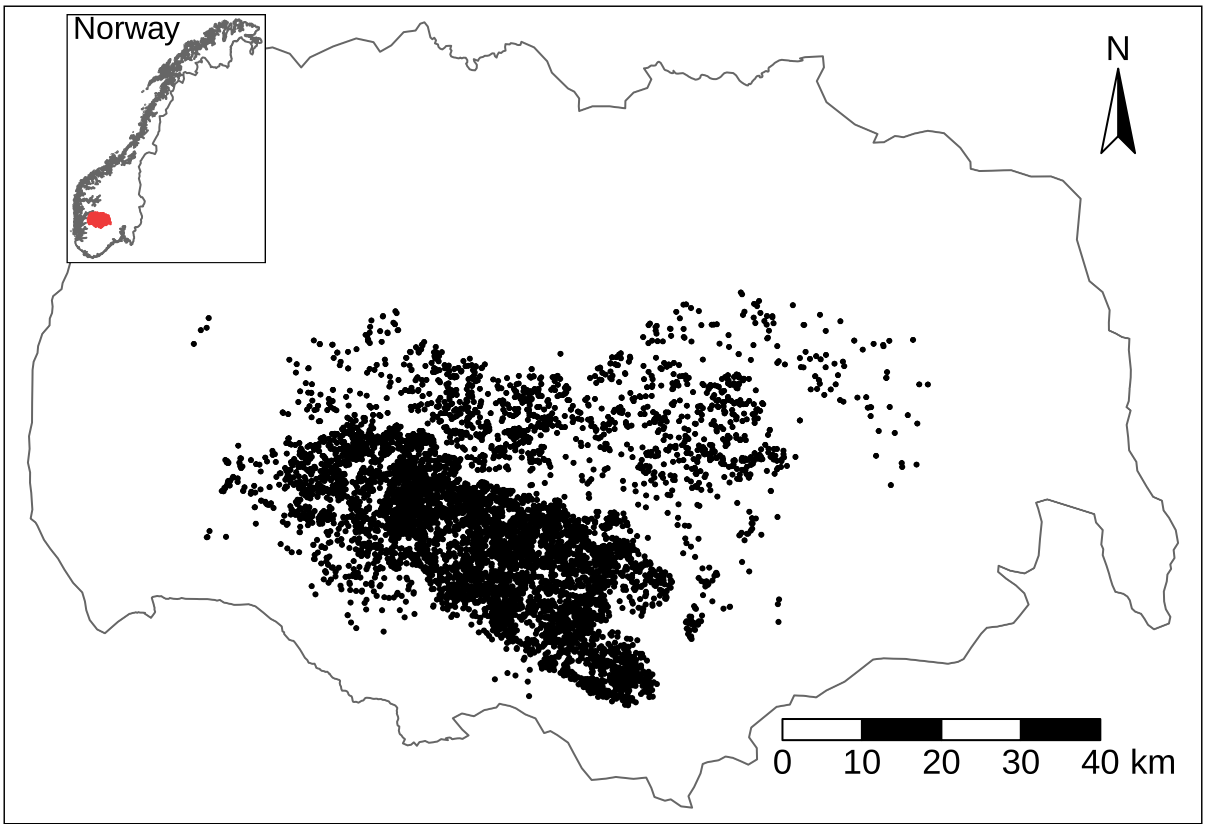

Figure C1: Hardangervidda reindeer area in Southern Norway and reindeer GPS locations used in this study.

Reindeer GPS data

Fourty-eight female reindeer were captured and monitored between 2001 and 2010. Reindeer were immobilized from helicopter (see details in Evans et al., 2013) and equipped with GPS collars with drop-off system. To regularize the fix rate among collars, we used 1 reindeer position every 6 hours. We analyzed only the data from July, selected here as a month representative of the summer, to avoid including reindeer positions during either the end of the calving season or during rut and autumn migration. For detailed data cleaning and preparation procedures, please see Panzacchi et al. (2015).

To perform habitat selection analyses, for each used GPS location we created a set of 9 locations available but not used by reindeer, spread uniformly within this wild reindeer area (Fig. C1). The combination of use and available locations was then annotated with environmental spatial data to assess the effects of the different infrastructure on reindeer space use.

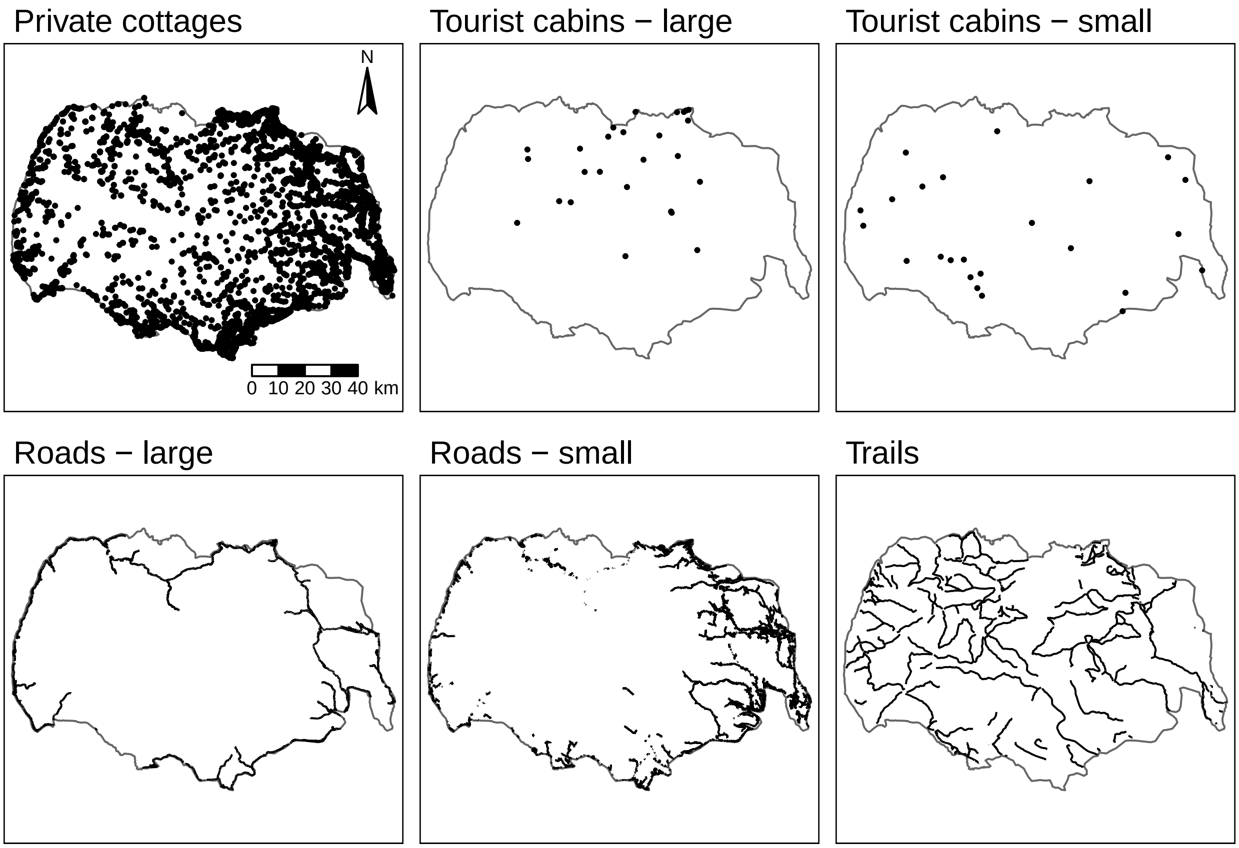

Figure C2: Main anthropogenic infrastructure in the Hardangervidda reindeer area, used to illustrate the area. Only private cottages and large tourist cabins were included in the analysis.

Table C1: Quantiles of the Euclidean distance from each 100 m-side cell in the Hardangervidda reindeer area to the nearest feature (in meters), for the main anthropogenic infrastructure present in the study area.

Infrastructure	0%	10%	25%	50%	75%	90%	100%
Private cottages	0	316	671	1265	2102	3178	7580
Large tourist cabins	0	4115	7117	13180	22496	29461	45393
Small tourist cabins	0	3833	6537	10308	14223	18065	31752
Large roads	0	906	2816	7235	14454	20873	29362
Small roads	0	316	1273	3970	9405	15180	26249
Trails	0	283	762	1709	3228	5124	11309

Environmental covariates

The locations of most types of infrastructure in Hardangervidda are correlated. Roads occur mostly in the lower parts of the area – and are correlated with the elevation and terrain variation – while other infrastructure occur closely together (e.g. small roads and cabins). For this reason, and for illustration purposes, for the analyses presented here we considered only the effects of private cottages and large tourist cabins.

The spatial data sources and details are described in Panzacchi et al. (2015). First, the vector representation for each kind of infrastructure was rasterized using a grid of 100 m resolution for an extent which included a buffer of 50 km around the study area; the buffer around the area was used to avoid edge effects in the influence measures' calculation. Then, both the influence of the nearest feature and the cumulative influence measures were calculated.

Since the infrastructure considered here (cottages and cabins) are represented as points, the input for influence calculation was the count of features within each grid cell.

Influence measures were calculated considering different influence functions (threshold, linear, Gaussian, and exponential decays; Fig. 1 in the main text), for a set of (irregularly distributed) zones of influence, from 100 m to 20 km. This allowed us to assess whether habitat selection is affected by either the cumulative influence or the influence of the nearest feature, while estimating the ZoI and accounting for the shape of the influence of the features of the different types of infrastructure within the ZoI. All influence functions were considered to have value 1 at the origin (where the infrastructure are located) and vary according to the different shapes (Fig. 1A in the main text). For the threshold and linear decay functions, the ZoI was defined as the distance at which the influence decreases to zero. For the Gaussian and exponential decay functions, which asymptotically approach zero, the ZoI was defined as the distance at which the functions reach 0.05.

To account for bio-climatic variation in reindeer space use, we also included as covariates land cover and 4 principal components (PCA axes) from a large principal component analysis performed in Norway to understand patterns of bio-climatic-geographical variation across the country (Bakkestuen et al., 2008). We used the NORUT land cover map with 30 m resolution and 25 vegetation classes, which we further grouped for modeling purposes (see the final classes in Table C3). The four principal components represent gradients of (1) PC1 - continentality, (2) PC2 - altitude, (3) PC3 - terrain ruggedness, and (4) PC4 - solar radiation, and account for 75 - 85% of the bio-climatic variation in Norway, representing the major environmental gradients in the study area (Panzacchi et al., 2015). Prior to the analyses, the continuous variables (all but land cover) were standardized to mean 0 and standard deviation 1.

Habitat selection modeling

Reindeer habitat selection was modeled through habitat selection functions (HSF, eq. 1 in the main text) considering the additive effect of the covariates described above. We included a quadratic term for PC1 and PC2 to account for non-linear responses (Panzacchi et al., 2015). HSFs were fitted using the `coxph` function of the `survival` package (Therneau, 2020; Therneau & Grambsch, 2000).

The first step in the modeling approach was to fit HSFs considering one infrastructure type at a time in a procedure of variable selection (Burnham & Anderson, 2002), to infer which influence measures and zones of influence better explained habitat selection, while also checking for correlations among the predictors (an approach similar to Laforge et al., 2015, and Huais, 2018). These models included land cover and the bio-climatic PCAs, in addition to either the cumulative influence or the influence of the nearest feature of a single infrastructure type. Given that the influence measures could assume 2 representations (cumulative, nearest) and follow 4 different functions (threshold, linear, Gaussian, exponential decay) with 8 distinct ZoI values (100 m, 250 m, 500 m, 1 km, 2.5 km, 5 km, 10 km, 20 km), for each infrastructure type we fitted 64 HSFs. Additionally, we also fit HSFs considering the log-distance to the nearest feature, which is a predictor commonly used in statistical models to assess the impacts of anthropogenic infrastructure on biodiversity. Single-infrastructure HSFs were fit with the `multifit` approach (Huais, 2018). HSFs were compared though the Akaike information criterion (AIC), and for each infrastructure type the 15 influence measures that better explained habitat selection (lower AIC) were chosen to be included in the multi-infrastructure HSF (see below).

We considered variables to be correlated if the Pearson correlation coefficient between their values was higher than 0.6, and excluded models in which any of the infrastructure influence measures was correlated with the bio-climatic variables.

In the single-infrastructure HSFs, we also assessed the estimated coefficients related to the infrastructure influence variables. Even though the coefficient values were not a criterion for selecting the most parsimonious influence variables, they are important to indicate consistency in the influence measures across the scales. If the coefficient changes signs (from avoidance to selection, for instance) as the ZoI increases, this might be a warning be careful in the evaluation of the most parsimonious ZoI.

We fitted multi-infrastructure HSFs by combining the best influence measures for each infrastructure. Since not necessarily the best influence function and ZoI for single-infrastructure models will keep as the most likely in multi-infrastructure models, we selected the 15 best covariates for each infrastructure and fitted all possible combinations between them. For models in which the infrastructure covariates were correlated, and excluded those variables with higher Variance Inflation Factor (VIF, which measures how much the variance of an estimated regression coefficient is increased because of collinearity; Kutner, 2005). In total, we fitted $15^2 = 225$ multi-infrastructure HSFs, which were also compared through AIC.

To quantify the impacts of infrastructure, for the most likely model we used eq. 2 of the main text and multiplied the magnitude of the impacts – the coefficients of the fitted model – by the influence measures included the model. We then estimated habitat suitability by predicting the HSF (eq. 1 in the main text) over the space and rescaling the predicted values to the interval [0, 1].

Results

Single-infrastructure HSF

We start by describing how much support the different influence measures presented in explaining reindeer habitat selection in the single-infrastructure models. By doing so, we aim at showing qualitatively what the different influence measures represent and how one would interpret them within an ecological context.

Private cottages

For private cottages, the most parsimonious HSF included the cumulative influence with Gaussian decay and ZoI = 10 km, but the support for the cumulative influence with the same ZoI of 10 km but other functions was also relatively high (low relative difference in AIC, Fig. C3A). Overall, the models including cumulative influence measures (regardless of the influence function and in great part of the ZoI) performed much better than the ones including the influence of the nearest private cottage, what shows a strong evidence that the impacts of multiple private cottages on reindeer habitat selection accumulate. The coefficients were consistently negative across ZoI values (Fig. C3B), which indicates the ZoI values with minimum AIC presented in the x axis of the Fig. C3A are also consistent.

We also go beyond the simple statistical variable selection and interpret the most parsimonious models considering the influence of the nearest feature. In this case, regardless of the influence function, the ZoI varies from 500 m to 1000 m (Fig. C3A). Combining the results, we can say that, if we consider the closest private cottage only, reindeer generally avoid being closer than 1 km from any cottage, but since many areas have a high density of cottages (Fig. C2, Table C1), they respond to the combined impact of individual cottages at a larger extent - a zone of influence of 10 km. This might also be related to how the cottages are used. Tourists who stay in a cottage hardly walk farther than a few kilometers from it, since they must return to the cottage in the end of the day. Then, the ZoI of a single cottage is shorter. However, in areas where many private cottages are clustered, there is a much wider area used by tourists and the ZoI of this combined cluster of cottages is higher.

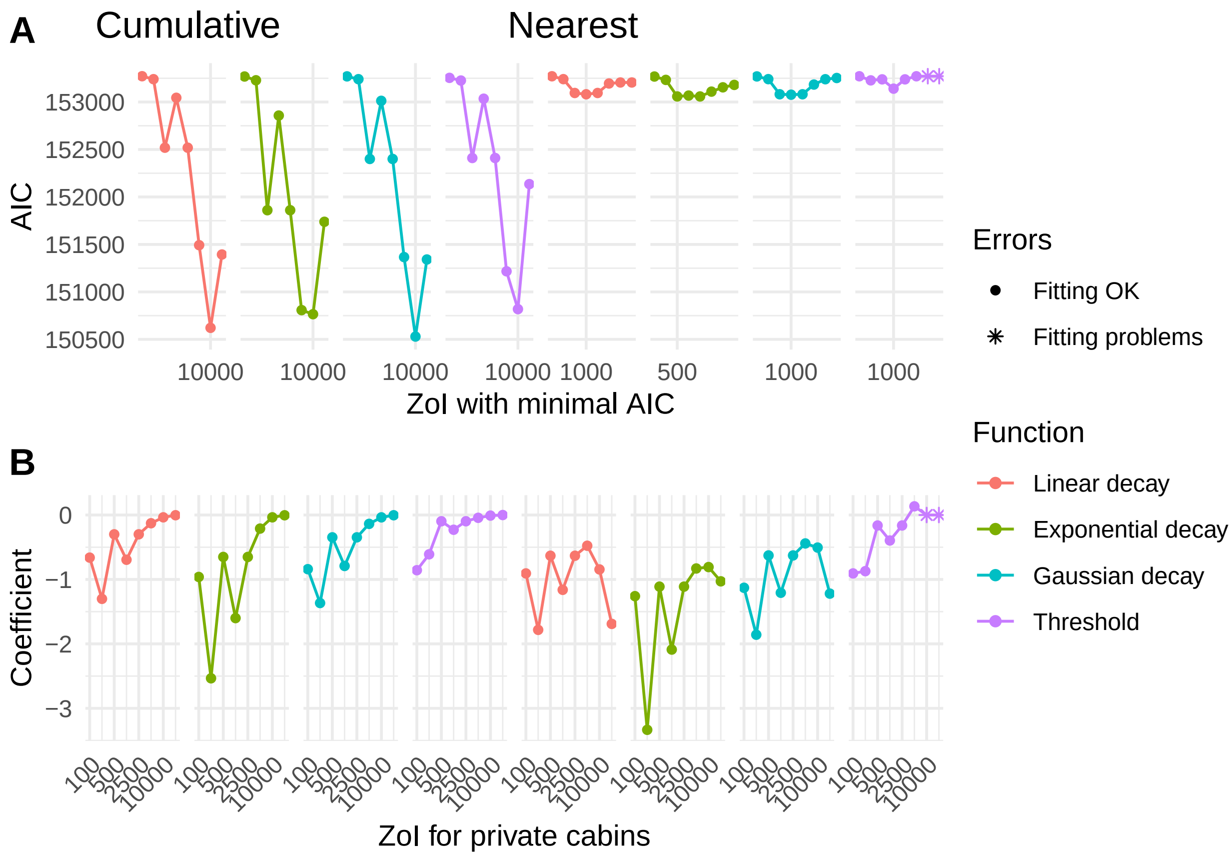

Figure C3: AIC (A) and coefficients (B) estimated for the influence of private cottages in models including only this type of infrastructure. The plots show the AIC and the coefficients (scaled back to the original range of the covariates) of the cumulative influence and influence of the nearest feature for different influence shapes and ZoI values (see the x axis in B for some of the candidate ZoI values, in meters, which varied from 100 m to 20 km). The x axis in A show the ZoI at which the AIC was minimal for each type and function of the influence measure. Scales marked with '*' represent errors in the model fitting (e.g. threshold influnece of the nearest feature for ZoI = 10 km or 20 km, when the variable is constant = 1 over the whole study area).

Large tourist cabins

In the single-infrastructure models, there was also evidence of cumulative impacts of large tourist cabins with a 20 km zone of influence (Fig. C4). The most supported influence variable was the cumulative influence with exponential decay shape and ZoI of 20 km. This ZoI value was selected regardless of the influence function, and even for the influence of the nearest feature the most common ZoI was 20 km (see the x axis in Fig. C4A).

The closer correspondence between the selected ZoI of cumulative influence and nearest influence measures, in comparison to the private cottages, might be due to several factors. First, tourist cabins are present at a much lower density in Hardangervidda, with high median Euclidean distance to the closest cabin (Fig. C2 and Table C1). As a consequence, there is not so much difference between what the cumulative and nearest influences represent (as for the private cottages), since the influence function of each feature only start to accumulate for larger ZoIs. Therefore, the ZoI with smaller AIC is closer between cumulative and nearest influence measures. Second, large tourist cabins are used very differently by tourists than private cottages. Each cabin is visited by a large load of tourists at a time, which consequently use a much higher area around the cabins, compared to single private cottage. As a consequence, the ZoI of large tourist cabins tend to be higher.

As for private cottages, in the models with tourist cabins the beta was also consistently negative, representing reindeer avoidance to public tourist cabins.

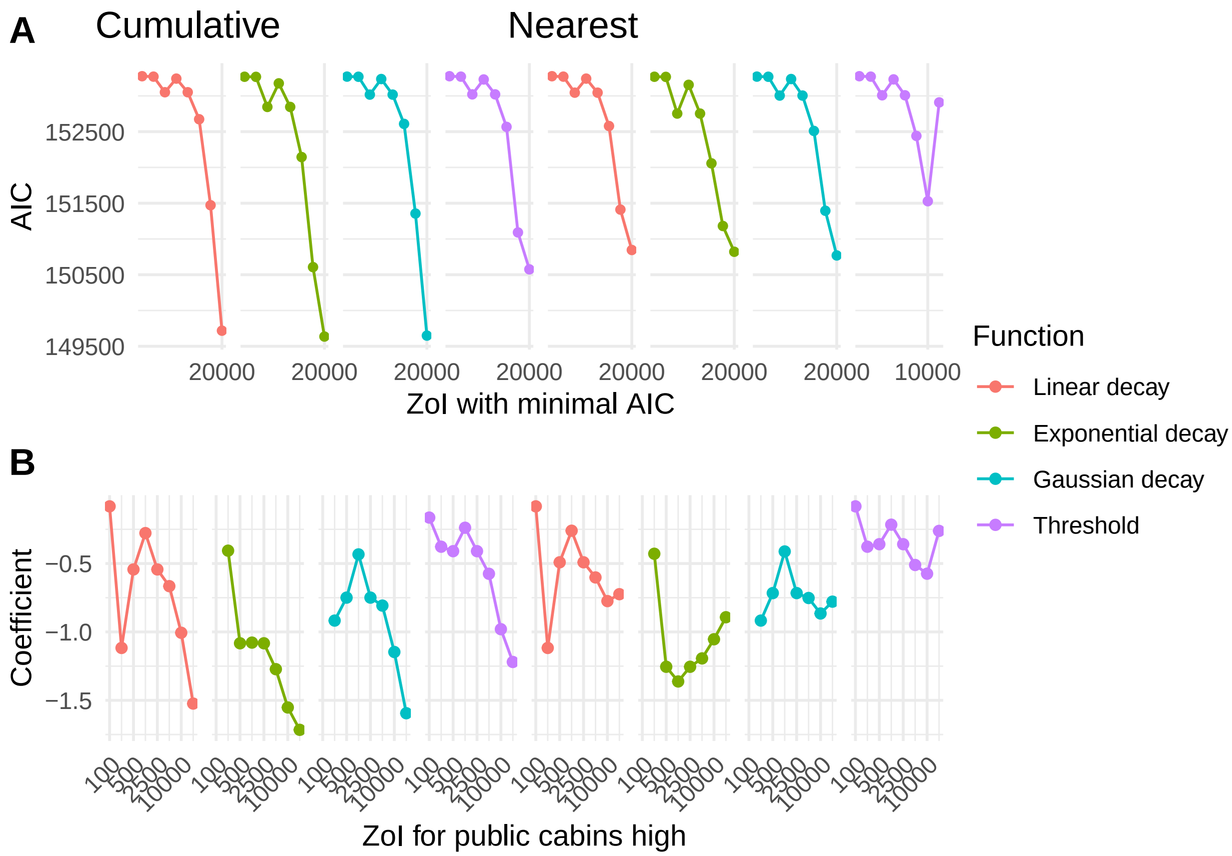

Figure C4: AIC (A) and coefficients (B) estimated for the influence of large tourist cabins in models including only this type of infrastructure. The plots show the AIC and the coefficients (scaled back to the original range of the covariates) of the cumulative influence and influence of the nearest feature for different influence shapes and ZoI values (see the x axis in B for some of the candidate ZoI values, in meters, which varied from 100 m to 20 km). The x axis in A show the ZoI at which the AIC was minimal for each type and function of the influence measure.

Multi-infrastructure HSF

The most parsimonious multi-infrastructure model included the cumulative influence of private cottages with threshold decay and ZoI = 10 km and the cumulative influence of multiple tourist cabins with exponential decay and ZoI = 20 km ($\Delta AIC = 31.9$, $wAIC = 1$; Table C2). Notice that, as parameterized here, for the tourist cabins an exponential decay with ZoI of 20 km means that the influence of cabins decrease to half of its maximum value when one walks 5 km away from the infrastructure (exponential half life is $\sim ZoI/4$ here). The best ranked model with a covariate for the influence of the nearest feature was ranked 25$^{\text{th}}$ in the model selection, and the best ranked model including the log-distance to the nearest feature was ranked 34$^{\text{th}}$ (Table C2). This presents a strong support for the cumulative impacts of both private cottages and tourist cabins on the reindeer habitat selection in Hardangervidda.

Table C2: Infrastructure variables included in the most parsimonious models. For each model we show the type of influence ("cumulative", "nearest"), the influence function ("exponential decay", "gaussian decay", "threshold", "Bartlett or linear decay"), and the ZoI (in km) for that covariate included in the model. For each model we also present the AIC, the difference in AIC to the most parsimonious model (dAIC), and the AIC weight. The last lines show the most parsimonious model which included any variable with the influence of the nearest feature and the log-distance to the nearest feature (in this case, for tourist cabins).

Rank	Private cottages	Large tourist cabins	AIC	dAIC	wAIC
1	cumulative, threshold, 10	cumulative, exp decay, 20	147578	0	1
2	cumulative, Gauss, 10	cumulative, exp decay, 20	147610	31.9	<0.001
3	cumulative, exp decay, 10	cumulative, exp decay, 20	147634	55.2	<0.001
4	cumulative, bartlett, 10	cumulative, exp decay, 20	147667	88.3	<0.001
5	cumulative, threshold, 10	cumulative, Gauss, 20	147702	123.1	<0.001
25	cumulative, exp decay, 10	nearest, bartlett, 20	148363	784.7	<0.001
34	cumulative, exp decay, 10	nearest, log nearest, NA	148493	914.5	<0.001

Looking closely to the most parsimonious multi-infrastructure HSF (after unscaling the coefficients back to the original variation of the infrastructure influence predictors), we see both infrastructure are avoided by reindeer. Their influence vary differently across space since their influence functions and ZoI differ – a threshold function with 10 km ZoI for private cottages and an exponential decay with 20 km ZoI for tourist cabins –, but also because the estimated magnitude of the impact of a single private cottage ($\beta_{\text{private cottage}} = -0.00746$) was much smaller than that of a single tourist cabin ($\beta_{\text{tourist cabin}} = -2.233$; Table C3, Fig. C6A). However, since private cottages occur at much higher densities, in some areas their overall impact might be higher than that of tourist cabins (Fig. C6, Fig. 4 in the main text). Comparing an area with only 1 cottage in a 10 km radius with an area with only 1 tourist cabins in a 20 km radius, and assuming all other conditions are similar, the impact – measured here as the product between the magnitude of the impact and the influence covariate (eq. 4 from the main text) – is much smaller for private cottages (Fig. C6A). In contrast, if we take the areas with higher influence in Hardangervidda – where the number of private cottages sum to 2664 and the (exponentially weighted) number of tourist cabins sum to 5 – the impact of private cottages agglomerates can be several times higher than that of tourist cabins (Fig. C6C). Following the HSF coefficient interpretation from Fieberg et al. (2021), and considering that all other conditions are kept similar but each of these influence variables is changed by 1 standard deviation (1 SD = 491 for private cottages and 1 SD = 0.79 for tourist cabins), reindeer are $exp(491 \cdot 0.00746) = 38.97$ times more likely to select an area with less private cottages and $exp(0.79 \cdot 2.233) = 5.83$ times more likely to select an area with less tourist cabins.

Table C3: Magnitude of the impact (model coefficients) of the most parsimonious model of space use for reindeer, including private cottages and tourist cabins. The table show the coefficient estimates (scaled back to the scale of variation of the original data), their standard error (SE), and the significance (p).

Covariate	Estimate	SE	p
private cottages (cumulative, threshold, 10km)	-0.00746	0.13	< 0.0001
tourist cabins (cumulative, exponential, 20km)	-2.23282	0.04	< 0.0001
exposed ridges	0.29668	0.14	0.0346
grass ridges	0.89062	0.14	< 0.0001
heather ridges	0.89877	0.13	< 0.0001
lichen	1.11598	0.17	< 0.0001
heather	0.92703	0.13	< 0.0001
heathland	0.85826	0.13	< 0.0001
meadows	0.90285	0.15	< 0.0001
early snowbed	0.61045	0.13	< 0.0001
late snowbed	0.45201	0.14	0.0008
bog	0.87809	0.15	< 0.0001
glacier	-0.32682	0.32	0.3087
other	-5.54769	164.61	0.9731
water	-1.47964	0.2	< 0.0001
poly(pc1, 2)1	282.07133	10.05	< 0.0001
poly(pc1, 2)2	-233.30191	10.57	< 0.0001
poly(pc2, 2)1	-512.34001	37.4	< 0.0001
poly(pc2, 2)2	-173.26455	22.47	< 0.0001
pc3	-30.6969	23.94	0.1998
pc4	19.1936	24.08	0.4254

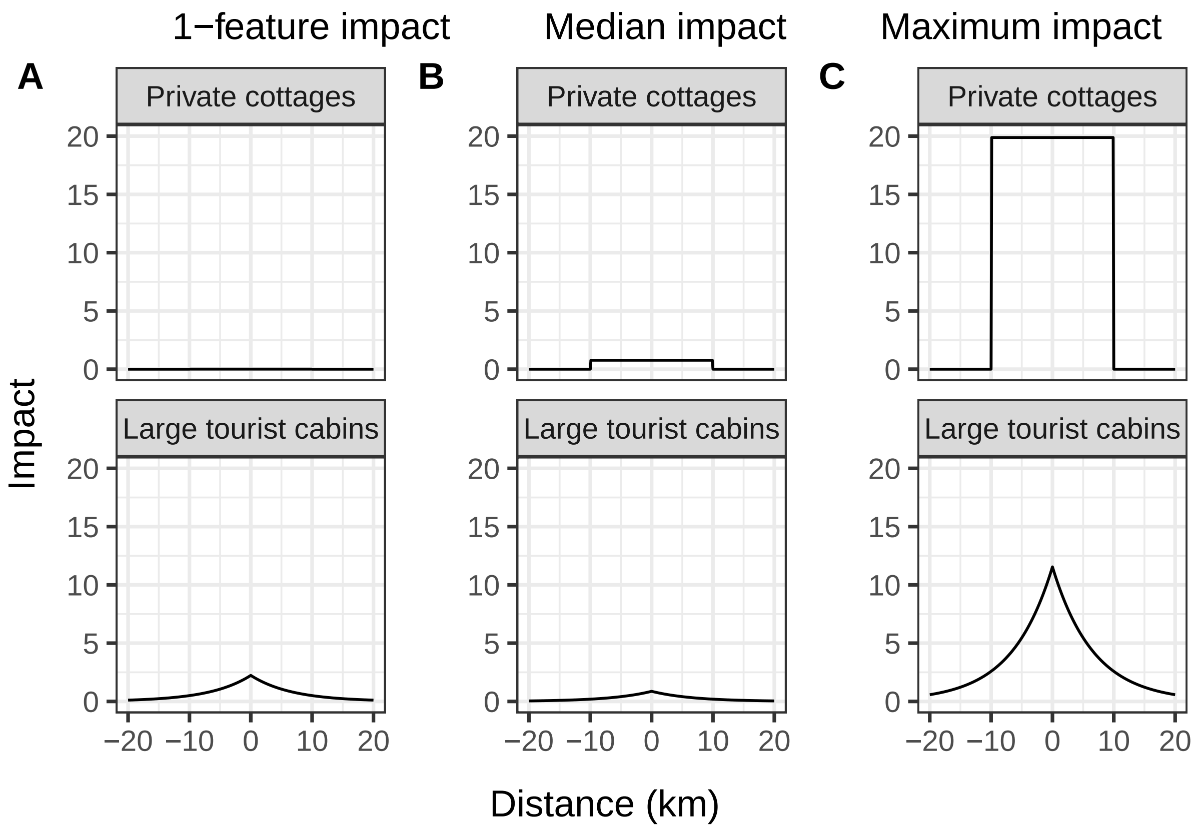

Figure C5: Impact of private cottages and tourist cabins considering only 1 feature, the median number of features (103 for private cottages, 0.38 for tourist cabins), and the maximum number of each type of feature (2664 for cottages, 5 for cabins), given their respective influence functions and ZoI. The impact presented here is the multiplication between the magnitude of the impact (the model coefficients) and the cumulative influence variable (eq. 4). The impact of only one private cottage is negligible. At their median values, the impacts of private and public cabins are comparable, while at their maximum the cumulative impact of private cottages might be higher than that of tourist cabins.

When cumulative impacts of infrastructure are spatialized by multiplying the magnitude of the impacts by the cumulative influence measures, we see how the relative impact of private cottages and large tourist cabins change across space (see Fig. 4 in the main text). As a consequence, since reindeer avoided high densities of both infrastructure types at relatively large extents, areas of high habitat suitability for reindeer correspond to those in which the cumulative influence of both infrastructure is low – what matches with the locations used by reindeer, indicated through the GPS data.

References

Bakkestuen, V., Erikstad, L., & Halvorsen, R. (2008). Step-less models for regional environmental variation in Norway. Journal of Biogeography, 35(10), 1906–1922. https://doi.org/10.1111/j.1365-2699.2008.01941.x

Burnham, K. P., & Anderson, D. R. (2002). Model selection and multimodel inference: A practical information-theoretic approach (2nd ed). Springer.

Evans, A. L., Lian, M., das Neves, C. G., Os, Ø., Andersen, R., Aanes, R., Strand, O., Tryland, M., & Arnemo, J. M. (2013). Physiologic evaluation of medetomidine-ketamine anesthesia in free-ranging Svalbard (Rangifer tarandus platyrhynchus) and wild Norwegian Reindeer (Rangifer tarandus tarandus). Journal of Wildlife Diseases, 49(4), 1037–1041. https://doi.org/10.7589/2013-03-049

Fieberg, J., Signer, J., Smith, B., & Avgar, T. (2021). A 'How to' guide for interpreting parameters in habitat-selection analyses. Journal of Animal Ecology, 90(5), 1027–1043. https://doi.org/10.1111/1365-2656.13441

Huais, P. Y. (2018). multifit: An R function for multi-scale analysis in landscape ecology. Landscape Ecology, 33(7), 1023–1028. https://doi.org/10.1007/s10980-018-0657-5

Kutner, M. H. (Ed.). (2005). Applied linear statistical models (5th ed). McGraw-Hill Irwin.

Laforge, M. P., Vander Wal, E., Brook, R. K., Bayne, E. M., & McLoughlin, P. D. (2015). Process-focussed, multi-grain resource selection functions. Ecological Modelling, 305, 10–21. https://doi.org/10.1016/j.ecolmodel.2015.03.003

Panzacchi, M., Van Moorter, B., Strand, O., Loe, L. E., & Reimers, E. (2015). Searching for the fundamental niche using individual-based habitat selection modelling across populations. Ecography, 38(7), 659–669. https://doi.org/10.1111/ecog.01075

Therneau, T. M. (2020). A Package for Survival Analysis in R. https://CRAN.R-project.org/package=survival

Therneau, T. M., & Grambsch, P. M. (2000). Modeling Survival Data: Extending the Cox Model. Springer.