Wildlife Biology

WLB-00719

Boulanger, J., Poole, K. G., Gunn, A., Adamczewski, J. and Wierzchowski, J. 2020. Estimation of trends in zone of influence of mine sites on barren-ground caribou populations in the Northwest Territories, Canada, using new methods. – Wildlife Biology 2020: wlb.00719

Appendix 1

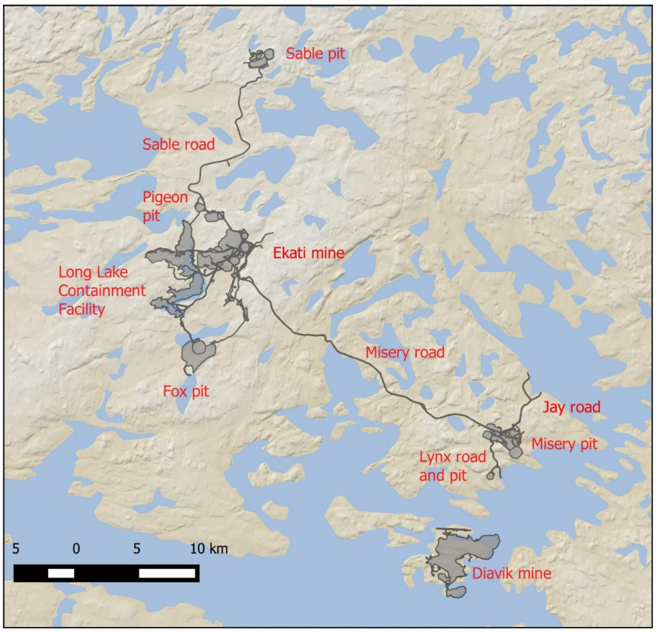
Timelines for Ekati and Diavik mines footprints and measurement of distance from mines and derivation of NDVI metrics

Timing of changes to the Ekati and Diavik footprints and the source of digital layers are provided in Table A1. Distance from mine was measured from footprints (including mine roads), and centroids of mine areas. Comparison of model fit with distance measured from centroid or footprint for collar suggested models with distance from footprint were more supported likely because points that fall within the footprint would receive a 0 distance from mine and therefore selection for mine footprint would be estimated as part of the intercept term. In contrast, a location that fell within a footprint would get assigned a distance using centroids. It is likely that habitat selection is universally low within a footprint so using the footprint basically pools habitat selection for footprint areas which may improve model fit compared to centroids. The difference between centroids and footprints is not as substantial for aerial surveys given that most transect cells do not fall in footprints and 91% of cells were closer to roads than mine centroids.

Table A1. Mine timelines used for measurement of distance from Ekati and Diavik along with a map of footprint features

Feature	Year in	
Main Diavik footprint	<2000	
Main Ekati footprint	<2000	
Misery road and	2000	

associated roads		
Long Lake Containment Facility	2000	
Fox pit	2003	
Pigeon pit	2014	
Lynx road	2015	
Lynx pit	2015	
Sable road	2016	
Jay road	2017	



NDVI vegetation indices were downloaded from MODIS data strips (Didan 2015, Didan et al. 2015) from the USGS web site (<https://earthexplorer.usgs.gov/>) and summarized in 8 × 8 km cells that covered the focal collar and aerial study area. Reliability indices were used to screen out tiles that were dominated by cloud cover. NDVI Data frames corresponding to 15 day periods were then summarized by average NDVI score to allow an overall index of seasonality for the time periods of the analysis (Fig. A1).

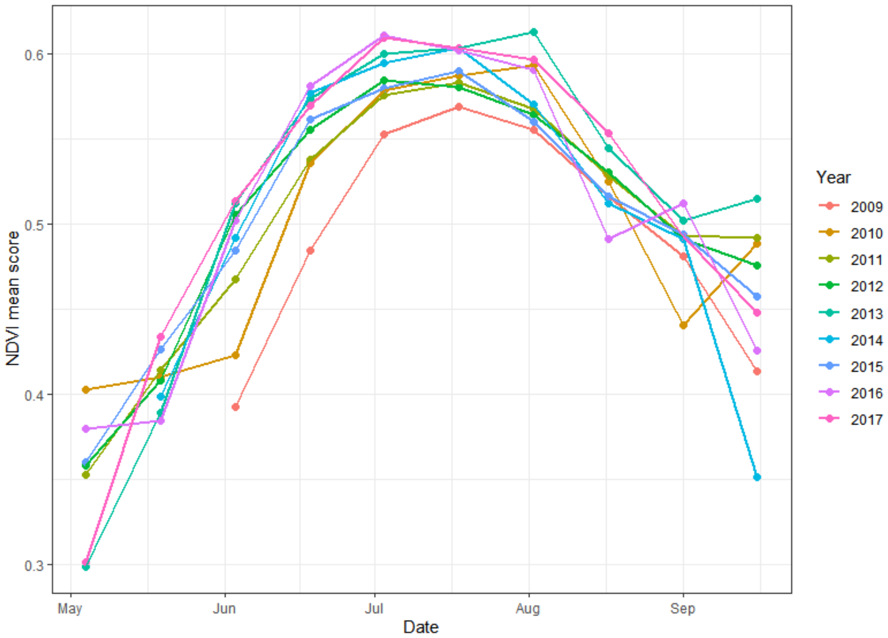


Figure A1. Mean NDVI scores for the entire study area used to model seasonality of habitats for the collar analysis (2009–2017). The main period of the analysis began in July through October.

Appendix 2

Supplementary information for collar analysis (2009–2017)

All captures of wildlife, including caribou, in the NWT are carried out with an approved Wildlife Research Permit. Any animal handling procedures are reviewed by a multi-partner Animal Care Committee. There is a Standard Operating Procedure for caribou captures. Chase times are limited and every effort is made to limit the handling time. Collar programming has varied over the years as the technology has changed. A common pattern is three daily locations and downloads every second day or every day at key periods. Most collars are set up to operate for about three years or in some cases for four years, depending on programming and battery size. All collars have a programmed breakaway mechanism that parts the collar at a time before the battery runs out. Caribou captures are carried out by helicopter using a netgun fired to envelop the caribou and enable handling. Selection of caribou to collar is opportunistic; the helicopter will target a small group of caribou and strive to isolate one caribou from the group and then capture it. Captured animals are of various ages, as assessed from the incisors. We believe the collared caribou are generally representative of the herd in the sense that there is no selection for caribou of a particular age or condition. Field conditions and limited chase times do not allow for selectivity.

Table A2. Sample sizes of collared caribou relative to mine areas as a function of season and year. Total collars pertain to all collars available in a given year. A collared caribou was classified as having encountered a mine area if at least one of its locations was within the 95th percentile of movement distances for the given year. The breakdown of encounter by season is also given.

Year	total	Collars encountering mine area			
		All seasons	Summer	Fall	Winter
2009	12	9	9	4	4
2010	17	13	13	9	0
2011	16	5	5	3	0
2012	20	10	10	9	3
2013	12	6	5	0	3
2014	17	13	12	12	1
2015	45	24	24	8	6
2016	41	31	30	6	6
2017	50	42	42	32	22
total	230	153	150	83	45

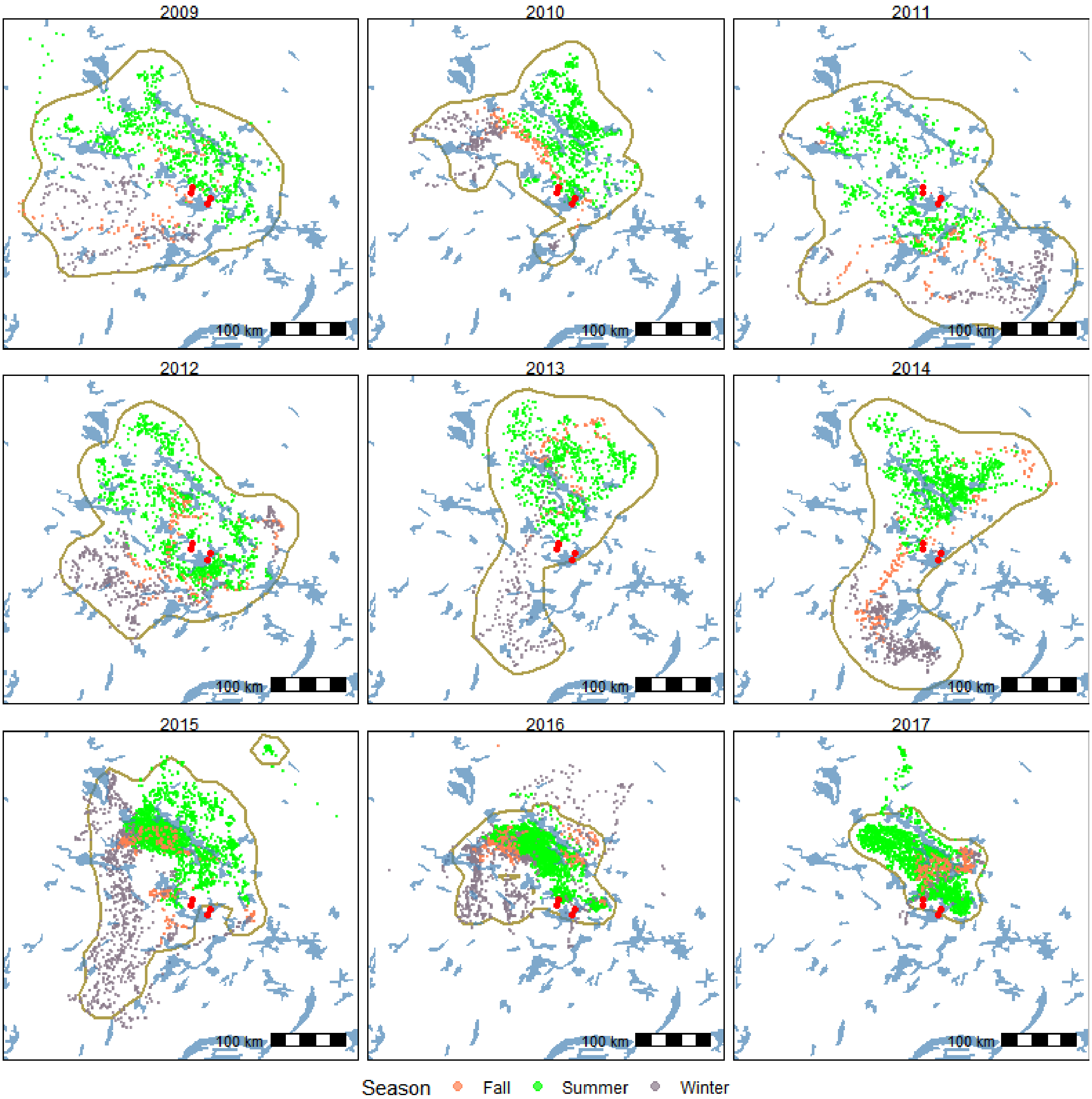


Figure A2. Extent of collared Bathurst caribou on summer, fall and early winter ranges from 2009-2017. Locations are coded by season. Mine feature centroids are delineated in red. The Lupin and Jericho mine sites to the north are shown along with Ekati and Diavik mines to the south. The brown polygon displays 95% kernel utilization distributions.

Table A3. Distances moved per day for male and female Bathurst caribou from season and year-specific estimates. The 95th percentile was used to define available areas. The sample size is the number of year and season combinations for males and females.

Statistic	mean	SD	min	max	n
<u>Females</u>					
95th percentile	29.61	7.27	16.69	51.68	45
Median	8.75	3.93	4.60	19.08	45
Mean	10.50	3.69	5.84	18.99	45
<u>Males</u>					
95th percentile	27.60	10.30	13.32	59.01	17
Median	7.83	3.65	3.04	16.83	17
Mean	9.38	3.63	5.09	20.30	17

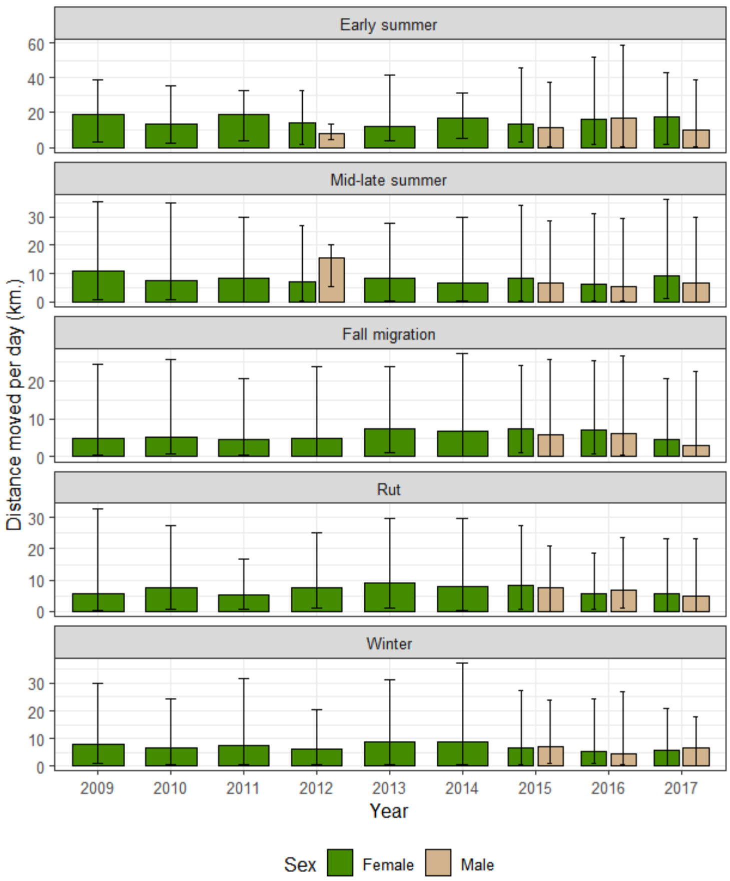


Figure A3. Summary of seasonal median distances moved per day for male and female collared Bathurst caribou. Error bars represent the 5th and 95th percentile of movement rates.

Table A4. Base RSF model for Bathurst caribou collar data from 2009–2017 based on conditional logistic regression. Robust standard errors and confidence limits are given as well as tests for parameter significance. Interaction terms are indented from main terms.

Parameter	β	SE	Confidence Interval		Z-score	p-value
Boulder/bedrock	-0.129	0.073	-0.273	0.015	-1.749	0.080
×Summer	0.173	0.074	0.028	0.318	2.332	0.020
×Winter	-0.083	0.071	-0.222	0.056	-1.174	0.240
Forest	-0.681	0.179	-1.032	-0.329	-3.797	<0.001
summer	-0.024	0.185	-0.385	0.338	-0.128	0.898
winter	0.547	0.178	0.198	0.897	3.069	0.002
Within seasonal range	2.043	0.165	1.719	2.367	12.352	<0.001
Low shrub	-0.043	0.055	-0.151	0.064	-0.790	0.429
×summer	-0.150	0.057	-0.261	-0.038	-2.639	0.008
×winter	-0.190	0.056	-0.300	-0.081	-3.423	0.001
Moss/lichen×Drought	0.064	0.014	0.037	0.091	4.623	<0.001
Sedge^2	-0.300	0.044	-0.387	-0.214	-6.799	<0.001
Fall	-0.109	0.045	-0.197	-0.021	-2.416	0.016
Summer	-0.226	0.048	-0.320	-0.131	-4.691	<0.001
Winter	-0.277	0.076	-0.425	-0.128	-3.657	<0.001
NDVI	0.027	0.006	0.015	0.039	4.394	<0.001
Treeherb^2	0.000	0.000	-0.001	0.000	-1.528	0.126
Tundra	-0.243	0.064	-0.368	-0.119	-3.828	<0.001
summer	-0.111	0.065	-0.237	0.016	-1.712	0.087
winter	-0.358	0.086	-0.526	-0.190	-4.170	<0.001
Tussock-summer	-0.103	0.047	-0.194	-0.012	-2.213	0.027
winter	0.029	0.037	-0.043	0.102	0.790	0.430
water	-0.894	0.051	-0.994	-0.793	-17.405	<0.001
water^2	-0.286	0.021	-0.328	-0.245	-13.459	<0.001
water × drought	0.122	0.038	0.048	0.196	3.249	0.001

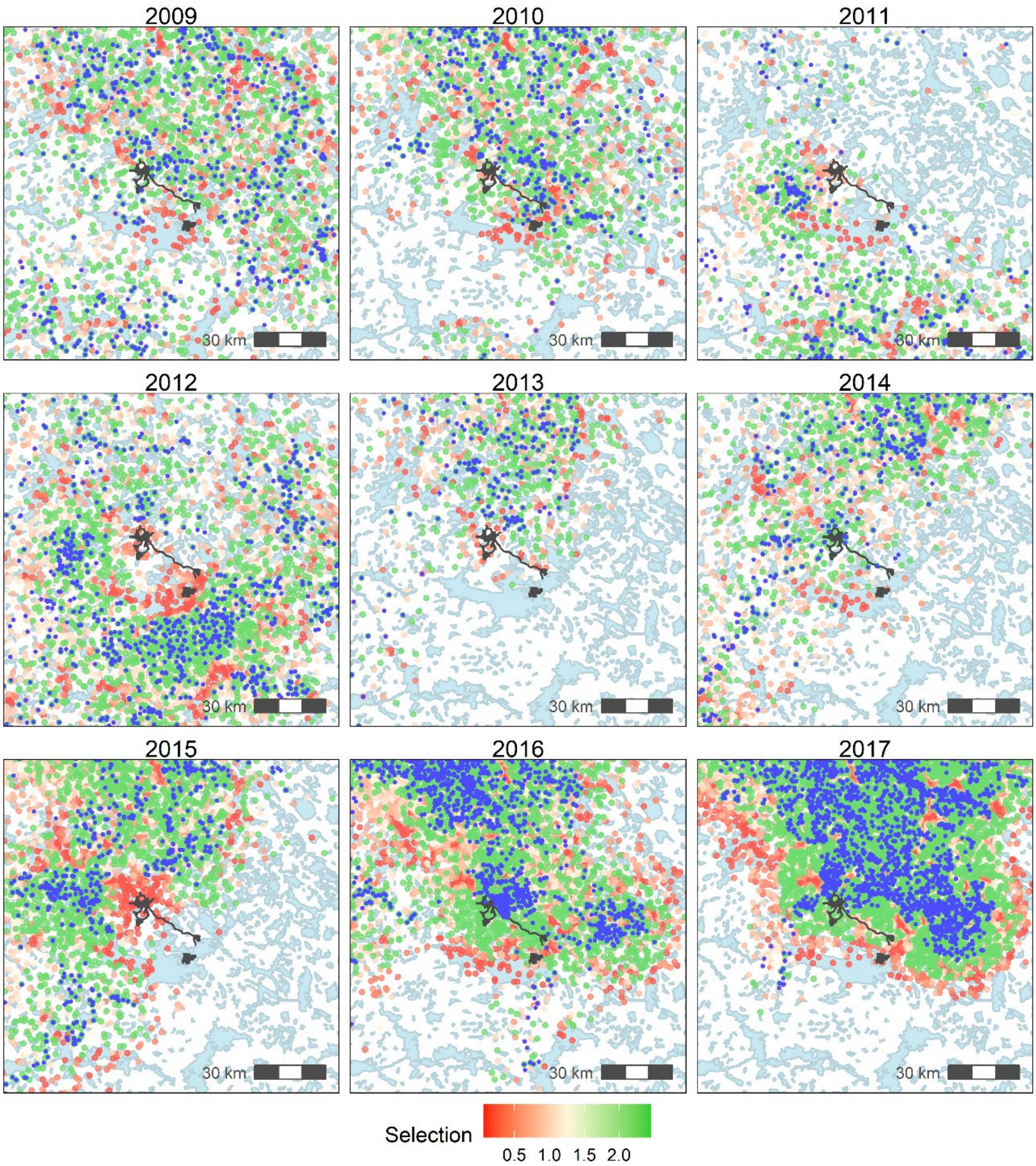


Figure A4. Spatial predictions of segmented model with used caribou locations indicated in blue. Red areas are areas of avoidance with green areas being selected.

Appendix 3

Cross validation of base model results

For the cross-validation analysis we randomly subdivided the data into training and testing datasets based on Huberty's rule of thumb (Huberty 1994). The goodness of fit of a model developed with the training data set was then compared with the testing dataset. We estimated the Pearson correlation (Zar 1996) of successive resource selection function (RSF) score bins with the frequency of used locations in each bin (adjusted for availability area of each bin). If the model fit the data then the RSF bin score and area-adjusted frequencies should be positively correlated (Boyce et al. 2002). We also estimated expected frequencies of each bin based on RSF score and compared these to observed frequencies using regression methods (Johnson et al. 2006). These tests were conducted for both overall model fit and the fit of yearly data. The k-fold selection process was repeated 100 times, which allowed an estimate of precision across a range of potential cross validation samples.

Cross validation statistics are given in Table A5. Interested readers should consult Boyce et al. (2002) and Johnson et al. (2006) for details on each test.

The mean correlation of predicted and observed data was 0.98 which was significant in all resamplings of the data set. Of greater interest was goodness of fit of yearly data sets. As shown in Fig. A5, cross-validation results suggested reasonable fit with observed and expected frequencies of RSF bins occurring for all 100 resamplings of the data. Fit was slightly reduced for 2011, 2013 and 2014 as indicated by a larger spread of points around the line of agreement. One additional factor that influenced fit was the sample size of points. Bins with lower sample sizes (red points in Fig. A5) often showed greater spread than bins with higher sample sizes. The fit for 2014 was potentially compromised by drought conditions during this year and subsequent changes in habitat selection. However, correlations between bin frequencies and odds ratios scores were significant in 90% or more of the resamplings.

Table A5. Cross-validations scores based upon (Boyce et al. 2002) and Johnson et al. (2006) based upon 100 resamplings of data.

Correlation mean bin score and area-adjusted frequencies (Boyce et al. 2002)								
No.	Year	Mean	Median	Min	Max	5th percentile	97.5 percentile	Proportion significant
1	2009	0.900	0.916	0.655	0.982	0.769	0.967	1
2	2010	0.931	0.942	0.797	0.991	0.842	0.988	1
3	2011	0.759	0.791	0.317	0.967	0.387	0.950	0.89
4	2012	0.787	0.799	0.594	0.934	0.658	0.890	1
5	2013	0.672	0.715	0.208	0.969	0.326	0.925	0.73
6	2014	0.589	0.601	0.326	0.830	0.381	0.779	0.65
7	2015	0.930	0.940	0.795	0.994	0.807	0.987	1
8	2016	0.910	0.915	0.755	0.995	0.781	0.977	1
9	2017	0.963	0.973	0.804	0.998	0.878	0.996	1
PropExp vs PropUsed-Slopes (Johnson et al. 2006)								
No.	Year	Mean	Median	Min	Max	5th percentile	97.5 percentile	
1	2009	0.972	0.977	0.685	1.272	0.778	1.151	
2	2010	0.888	0.892	0.672	1.039	0.764	1.021	
3	2011	0.889	0.890	0.559	1.181	0.700	1.077	
4	2012	0.807	0.809	0.652	0.926	0.702	0.916	
5	2013	0.840	0.830	0.533	1.299	0.596	1.115	
6	2014	1.013	1.008	0.801	1.280	0.870	1.179	
7	2015	0.976	0.969	0.839	1.158	0.883	1.114	
8	2016	0.932	0.933	0.838	1.018	0.858	1.004	
9	2017	1.033	1.034	0.899	1.137	0.927	1.119	
PropExp vs PropUsed-Intercepts (Johnson et al. 2006)								
No.	Year	Mean	Median	Min	Max	5th percentile	97.5 percentile	
1	2009	0.003	0.002	-0.027	0.032	-0.015	0.022	
2	2010	0.011	0.011	-0.004	0.033	-0.002	0.024	
3	2011	0.011	0.011	-	0.044	-0.008	0.030	

				0.018			
4	2012	0.019	0.019	0.007	0.035	0.008	0.030
5	2013	0.016	0.017	-0.030	0.047	-0.012	0.040
6	2014	-0.001	-0.001	-0.028	0.020	-0.018	0.013
7	2015	0.002	0.003	-0.016	0.016	-0.011	0.012
8	2016	0.007	0.007	-0.002	0.016	0.000	0.014
9	2017	-0.003	-0.003	-0.014	0.010	-0.012	-0.012

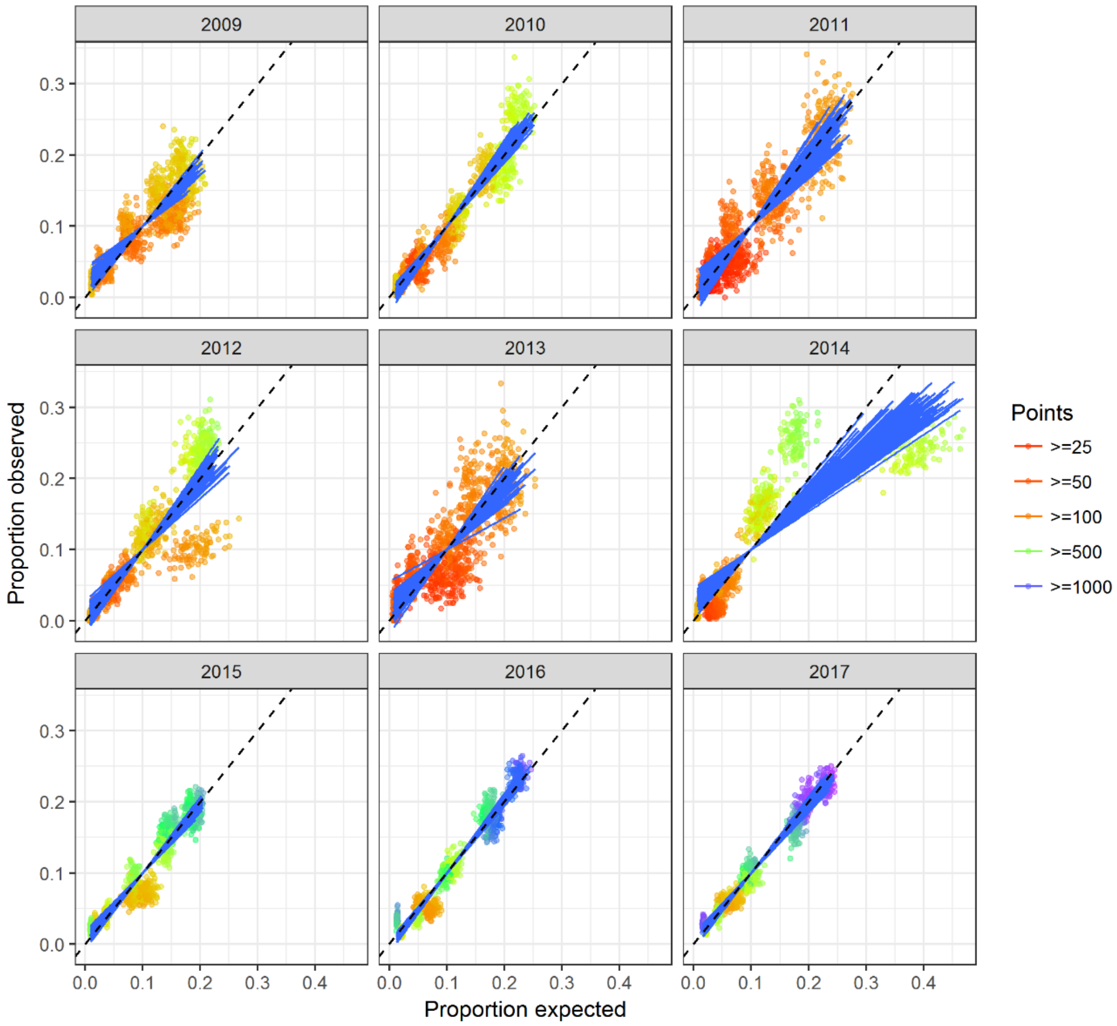


Figure A5. Cross validation estimates for base model (Table A5). If models fit then the proportions of expected frequencies should be higher correlated with observed frequencies of RSF bins. A larger spread indicates lesser fit. The slope of the regression curves should also be close to 1 as delineated by the hatched line. Slopes from each resampling are indicated by blue lines.

Appendix 4

GAM analysis of larger scale gradients

Generalized additive modelling (GAM) analysis suggested relatively even gradients of habitat selection by caribou for distances up to 50–100 km from the Ekati and Diavik mines followed by decreases in 2009, 2013 and 2014 (Fig. A6). In contrast, selection increased at further distances in 2015–2017.

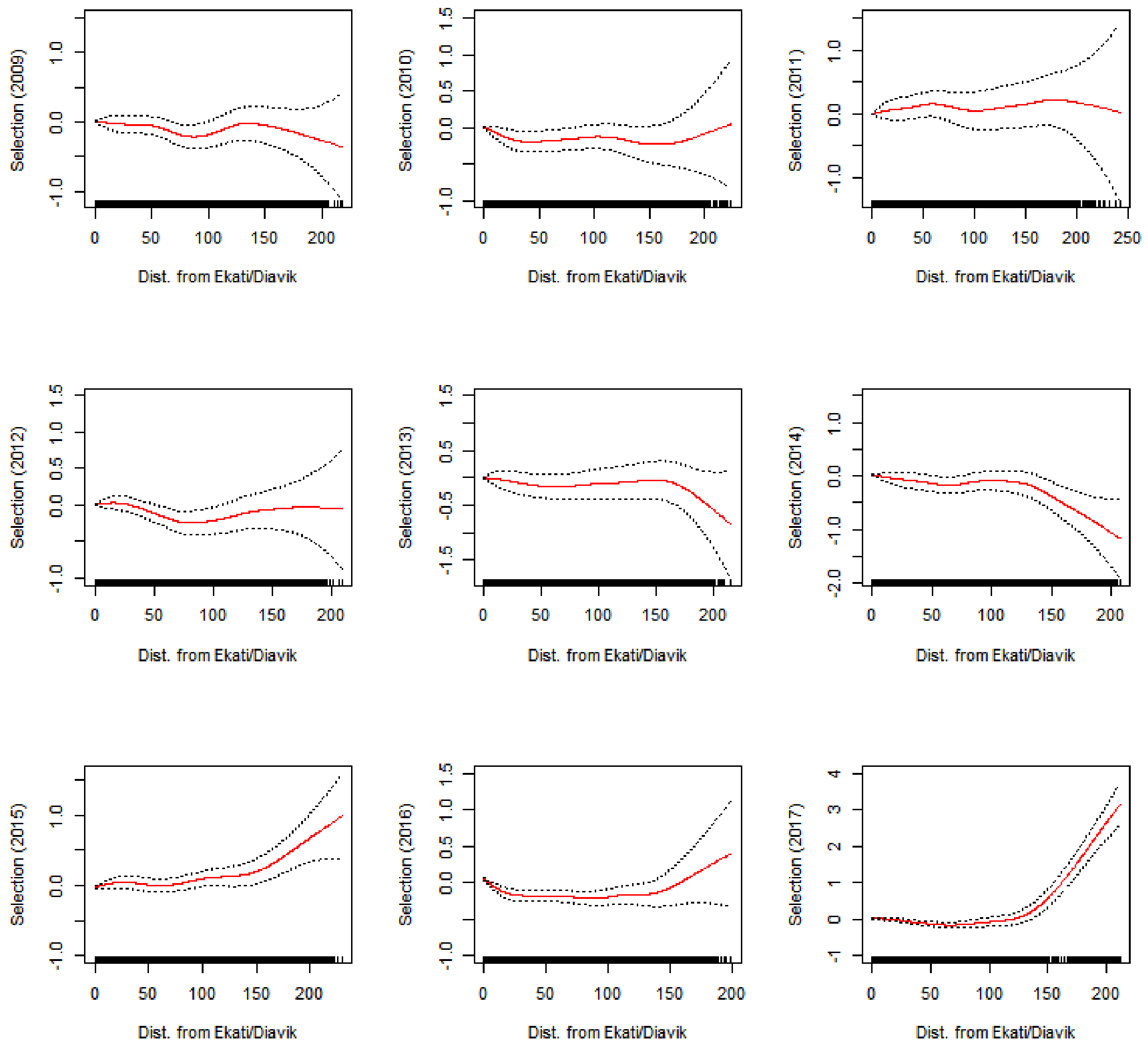


Figure A6. Large scale selection gradients by caribou relative to the Ekati and Diavik mines from 2009–2017 as estimated by generalized additive models.

Appendix 5

Re-analysis of 1996–2008 collar data set

The 1996–2008 caribou collar data set originally analyzed in Boulanger et al. (2012) was re-analysed using the segmented R package. The sample size of collars in the Bathurst herd from 1996–2006 was low and as a result there were not enough data to test for and estimate year-specific zones of influence (Table A6).

Table A6. Sample sizes of collared caribou for the Ekati and Diavik analysis.

Collars	Year											
	1996	1997	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Encountered Mine	9	7	14	13	9	11	10	4	18	14	19	10
Total collars	9	7	14	13	12	11	10	6	18	14	19	12

In addition, the fix interval was longer in earlier years (due to use of satellite collars and not more recent GPS-satellite technology) therefore reducing the sample sizes of points (Fig. A7) as discussed in Boulanger et al. (2012). The highest sample sizes occurred in 2006 to 2008 which was also when satellite collar fix intervals were shorter in duration.

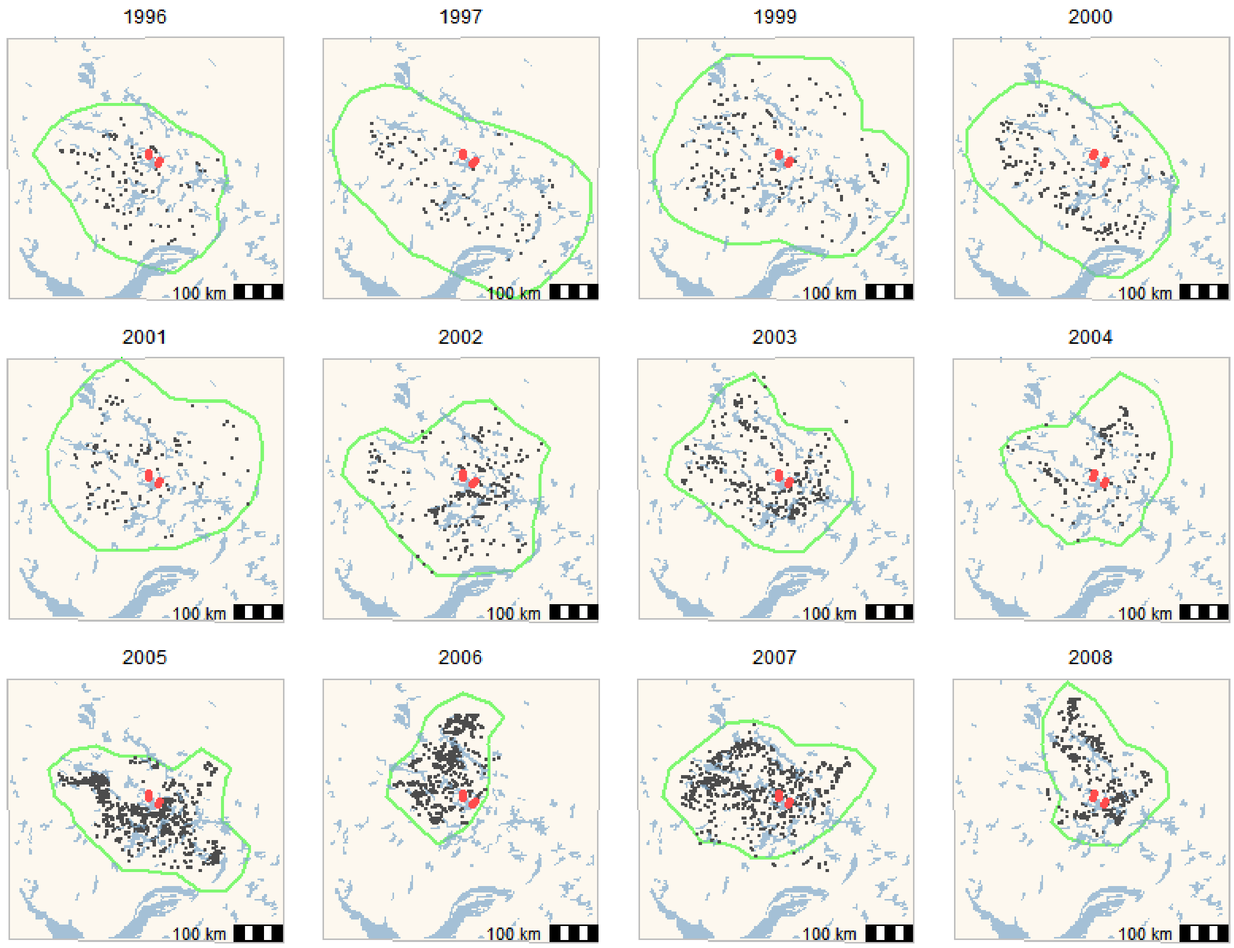


Figure A7. Distribution of collared caribou points (black dots) and seasonal range as estimated by a 95% kernel estimator (green line) for 1996–2008. The Ekati/Diavik mine is indicated by the red points.

A preliminary analysis that attempted to estimated ZOI from collars based on pooled year data (i.e. every two years pooled) was unsuccessful with non-convergence of the segmented models. As a result, we used the pooling of years from Boulanger et al. (2012) which was based on phases of mine activity, with an iteration zone of 50 km. Estimates of ZOI suggested a significant ZOI for 2003–2008 when both mines were in operation, but with a non-significant βzoi (at α=0.05) (Table A7).

Table A7. Estimates of ZOI (km) from analysis of Ekati and Diavik (1996–2008)

Years	Phase	Zone of influence					Odds ratios			Significance βzoi	
		ZOI	St.Err	Conf. Int		CV	OR	Conf. Int		Z	p-value
1996-99	Ekati construction	2.52	5.79	-8.83	13.87	230.0%	1.65	0.33	2.03	-0.43	0.667
2000-02	Ekati operation Diavik construction	36.94	33.19	-28.10	101.99	89.8%	2.16	0.95	1.01	-1.22	0.222
2003-08	Ekati & Diavik operation	15.50	4.51	6.66	24.34	29.1%	2.04	0.90	1.02	-1.39	0.164

A plot of estimates with segmented predictions demonstrate low sample sizes for the 1996-1999 and 2000–2002 periods (Fig. A8). Sample size is increased for the 2003–2008 period with a trend of lower densities of used points in the proximity of the Ekati and Diavik mines. As a result, a ZOI of 15.5 km was estimated, similar to that obtained by Boulanger et al. (2012; 11 km, CI=1–17 km).

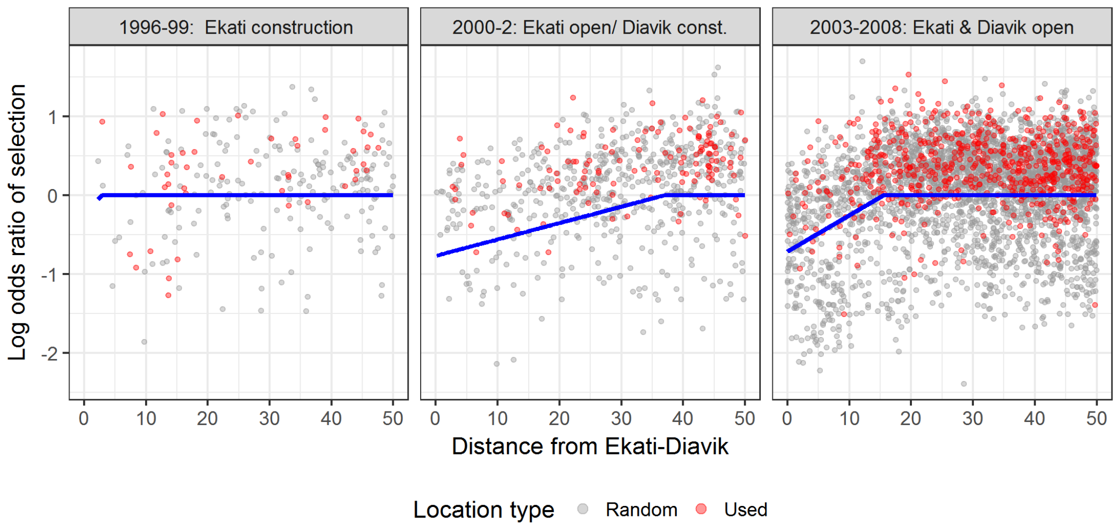


Figure A8. Estimates of raw section predictions and segmented regression predictions from the Ekati and Diavik 1996–2008 ZOI analysis. The blue line indicates the predicted ZOI curve from program segmented.

Appendix 6

Aerial survey analysis details

The development of the base habitat model and other statistical details on the analysis of the aerial survey data set from 1998–2008 are given in Boulanger et al. (2012). The main difference in the analysis presented in this paper is the addition of the 2009 and 2012 data sets and the use of the segmented package to estimate ZOI. Table A8 summarizes the aerial survey data set.

Table A8. Summary of yearly sample sizes for ZOI aerial survey

Year	Number of surveys		Cells (1 km) surveyed	Cells (1 km)with caribou	
	Ekati	Diavik		Count	Proportion (%)
1998	17		6,715	268	4.0
1999	18		7,110	410	5.8
2000	12		4,740	120	2.5
2001	11		4,345	448	10.3
2002	8	8	5,416	339	6.3
2003	9	9	6,093	260	4.3
2004	9 combined		6,093	168	2.8
2005	10 combined		6,770	446	6.6
2006	10	8	10,432	311	3.0
2007	9	10	13,157	332	2.5
2008	10	10	12,643	206	1.6
2009	13 combined		8,971	178	2.0
2012	11 combined		8,129	149	1.8

The design of aerial surveys are shown in Fig. A9.

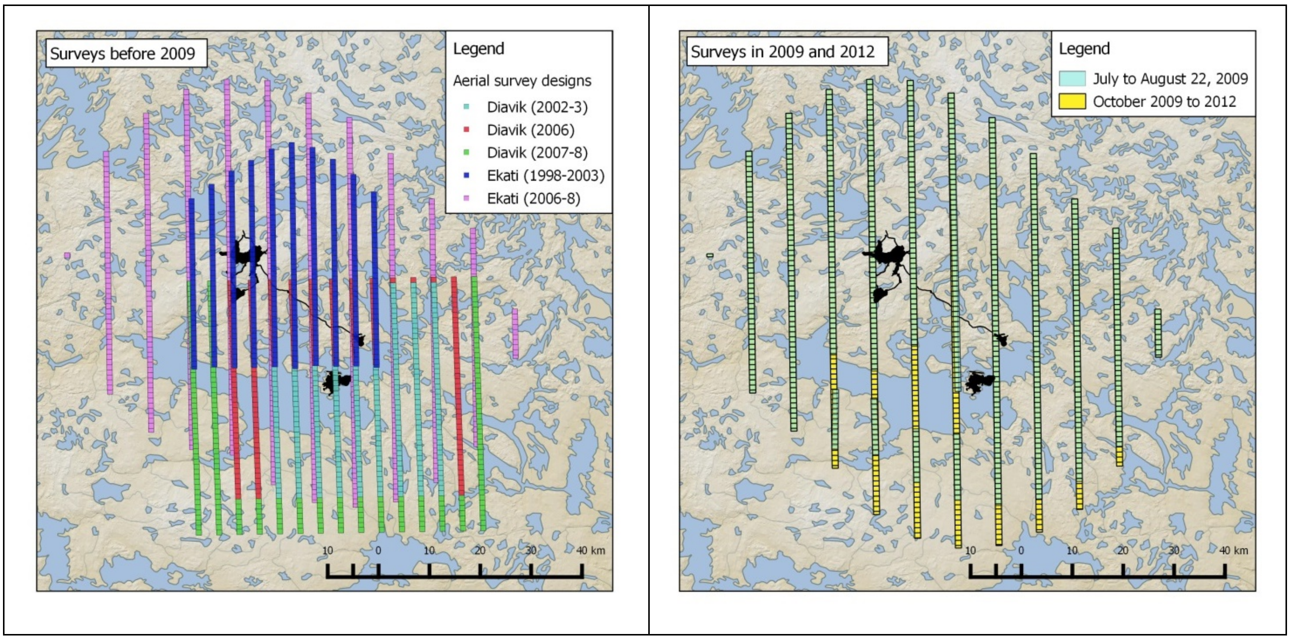


Figure A9. Survey designs flown prior to 2009 (left) and in 2009–2012 (right). Some surveys in the prior to 2009 period were flown concurrently between Diavik and Ekati. The Diavik and Ekati mine roads and overall footprints are shown in black.

The base model developed by Boulanger et al. (2012) was used in analyses given that the same time series (except for the addition of 2009 and 2012) was analyzed. Goodness of fit test suggested acceptable fit for all years (Table A9). For this analysis NDVI data were obtained from 1 km MODIS data strips with values made compatible with the Boulanger et al. (2012) analysis. Analyses of distance to disturbance were conducted using distance to mine centroids for Diavik, the main Ekati site including the Long Lake Containment Facility (LLCF), the Fox Pit, and all main roads (defined as roads located inside and outside of the main mine sites and pits with significant levels of daily traffic; e.g. the Fox and Misery roads) for all years of the analysis (Supplementary information). Distances to the perimeter of footprints were also considered, however the difference in distances was negligible (correlation=0.998) given that 91% of cells were closer to roads then centroids.

Table A9. Base habitat model for aerial survey analysis for the Ekati and Diavik mine area aerial surveys. Standardized slope estimates are given for habitat variables (from Boulanger et al. 2012).

Parameter	Estimate	SE	CI	χ^2	p
Intercept	-3.33	0.04	-3.40 - -3.26	8737.26	<0.0001
Esker	0.04	0.02	0.01 - 0.07	5.52	0.0188
Relative occupancy	0.58	0.01	0.56 - 0.61	2656.08	<0.0001
Lowshrub^2	-0.06	0.03	-0.11 - -0.01	6.28	0.0122
Sedge wet	0.15	0.04	0.08 - 0.23	15.71	<0.0001
Tundra^2	-0.10	0.02	-0.14 - -0.06	28.18	<0.0001
Tundra × NDVI	0.49	0.25	0.00 - 0.97	3.87	0.0492
Water	-0.14	0.08	-0.29 - 0.02	2.97	0.0848
Water^2	-0.23	0.05	-0.32 - -0.14	25.70	<0.0001

Naïve generalized linear models, which did not account for autocorrelation in the data set, were run to provide initial estimates of ZOI in *segmented*. These estimates were then used as starting points in the more complex generalized estimating equation analyses where models were fit based upon minimizing the standardized Pearson residuals which indexes fit of the model and observed data (Appendix 7). Figure A10 demonstrates predicted odds ratios and ZOI for 2008 and 2009 from the segmented analysis. Full results are given in Table 4 and Fig. 5 in the main manuscript.

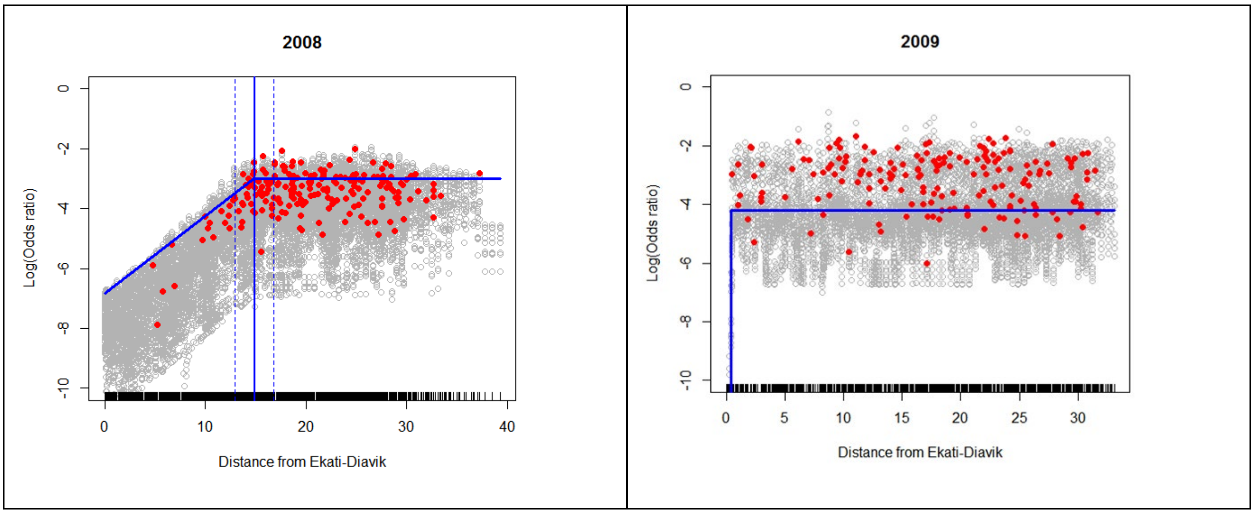


Figure A10. Prediction plots of segmented models for 2008 and 2009. The red dots indicate transect cells where caribou were detected and the grey dots indicate cells where caribou were not detected. The blue line is the mean odds ratio score with ZOI indicated by vertical blue lines (with confidence limits as hashed lines).

Appendix 7

Use of segmented R package to estimate ZOI

An understanding of how *segmented* parameterizes breakpoints is needed to be able to use the segmented package to estimate ZOI. The segmented package is meant to fit a variety of segmented relationships including ones that include more than one breakpoint. For ZOI analyses we are interested in a particular relationship (Fig. A11) where a non-zero slope occurs up to the ZOI after which the slope is 0 (meaning the mine has no effect on habitat selection). The basic parameterization of the segmented regression analysis (in the *segmented* package) is $\alpha Z+\beta(Z-\psi)$ where ψ is the threshold ZOI distance, α is the non-zero slope of the left hand curve, and β is the difference in slopes between the right hand curve and left hand curve (Muggeo 2003), and Z is the distance from mine x-axis variable. The slope of the right hand curve should be 0 if a ZOI exists (Fig. A11). This constraint on the analysis can be imposed by setting the distance from mine x-axis variable used in the analysis (Z) *to a negative value*. This causes the equation ($\alpha Z+\beta(Z-\psi)$) to search iteratively for a breakpoint ψ where $\alpha=\beta$ (with α describing the slope of the left hand curve) therefore fitting the intended ZOI relationship (Muggeo 2008).

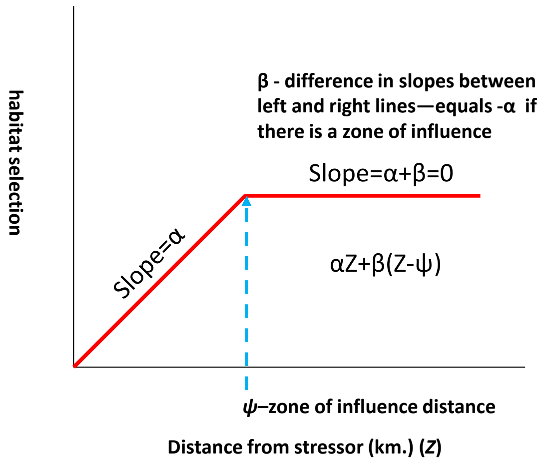


Figure A11. The parameterization of ZOI in program segmented.

The R code for ZOI estimation for the 2003-2008 period for Ekati/Diavik is provided below. The first analysis uses a *glm* model object as the baseline habitat model which demonstrates the full features of segmented. An analysis that uses a generalized estimation equation (GEE) model, which accounts for repeated sampling of segments, is then demonstrated.

A segmented analysis using the *glm* R object

1. An R data set (ac2) contains the transect data for aerial surveys from 2003–2008. A hit binary variable (1-caribou present in segment, 0-not present in segment) is used with

additional habitat variables as described in Boulanger et al. (2012). An additional variable in this data set is d_ekdi which is the distance from mine/road infrastructure as detailed in Boulanger et al. (2012).

```
ac2$neg.d_ekdi<- -ac2$d_ekdi
```

2. A baseline model from Boulanger et al. (2012) is fit in R using the *glm* method—Logistic regression (family=”binomial”(“logit”) is specified. This analysis does not contain a distance from mine variable.

```
out.airsurvglm<-glm(hit~reloccupan + ESKER_HA+sedgewet + water + I(water^2) +
I(lowshrub^2) +I(tundra^2)+ I((tundra * ndvi)), data=ac2,family=binomial("logit"))
```

3. The segmented command is then used to estimate the cutpoint using the *out.airsurvglm* object. The glm object is listed, followed by the distance from mine to search for a threshold, ~neg.d_ekdi. A starting point for the search is given (*psi=-18*). A bootstrap method is used to assess various values of psi. A breakpoint at 13.99 is estimated.

```
> zoi_20038_glm<-segmented(out.airsurvglm,~neg.d_ekdi,psi=-18)

> summary(zoi_20038_glm)

        ***Regression Model with Segmented Relationship(s)***

Call:
segmented.glm(obj = out.airsurvglm, seg.Z = ~neg.d_ekdi, psi = -18)

Estimated Break-Point(s):
  Est. St.Err
-13.99   0.87

Meaningful coefficients of the linear terms:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  -3.37668    0.05227 -64.606  < 2e-16 ***
reloccupan    0.58607    0.01639  35.768  < 2e-16 ***
ESKER_HA     -0.02448    0.02377  -1.030   0.3032
sedgewet      0.06929    0.02712   2.555   0.0106 *
```

```
water              -0.67214    0.05566 -12.075  < 2e-16 ***
I(water^2)         -0.22634    0.05146  -4.398 1.09e-05 ***
I(lowshrub^2)      -0.01342    0.02452  -0.547   0.5841
I(tundra^2)        -0.10550    0.02427  -4.347 1.38e-05 ***
I((tundra * ndvi))  0.20122    0.10900   1.846   0.0649 .
U1.neg.d_ekdi      -0.08833    0.01042  -8.476       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
(Dispersion parameter for binomial family taken to be 1)

Null     deviance: 15338  on 55187  degrees of freedom
Residual deviance: 13378  on 55177  degrees of freedom
AIC: 13400

Convergence attained in 2 iterations with relative change 1.359137e-05
```

4. A confidence interval on the ZOI is then estimated.

```
> confint(zoi_20038_glm,rev.sgn=TRUE)
$neg.d_ekdi
  Est. CI(95%).l CI(95%).u
 13.99     12.29      15.7
```

5. A plot of the estimated ZOI relationship can also be obtained via the plot command (Fig. A12). The confidence limits on the relationship are displayed. A rug plot details the density of distance measurements. The y units are in the logit scale used in the analysis.

```
plot(zoi_20038_glm,rev.sgn=TRUE,dens.rug=T,conf.level=0.95,shade=T)
```

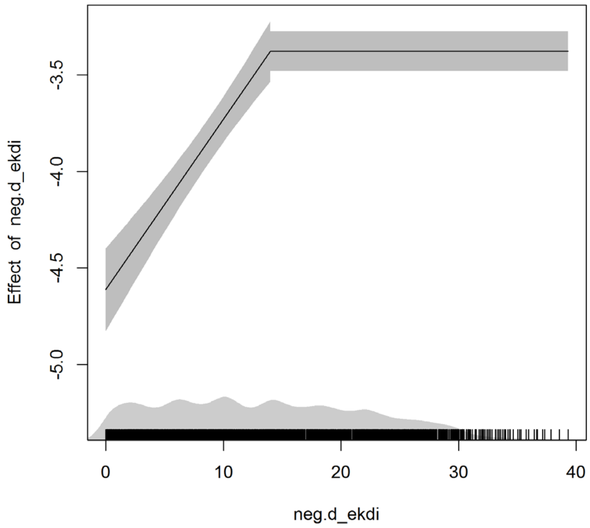


Figure A12. The estimated ZOI relationship for Ekati-Diavik 2003–2008 using the segmented package. A rug plot along the x-axis indexes the relative sample size of distance measurements used in the analysis.

One issue with using the *glm* method is that it does not adequately account for repeated observations at segments. A generalized estimation equation (GEE) method or other approach can be used to model repeated measures. The analysis outlined above can be further fitted using the *geepack* package in R. However, the *segmented.default* command then has to be used to estimate the threshold and a function that assesses model fit has to be specified for *segmented.default*. Vito Muggeo (per. comm.) suggested that the raw residuals could be used. Below is an R function that is used with segmented.default.

```
f<-function(x){
  raw.res<- x$y-x$fitted
  v.raw.res<-x$fitted*(1-x$fitted)
  r<-sum(raw.res^2/v.raw.res)
  r}
```

The *geeglm* command for the base habitat model is as follows. The *geeglm* package is required for this analysis. An id variable (jbs_trans) identifies the transect segments that are repeatedly measured. The data set is sorted by the jbs_trans variable for this analysis. An exchangeable correlation matrix is used to model correlations between repeated transect measurements.

```
out.airsurvgeeglm<-geeglm(hit~reloccupan +  ESKER_HA+sedgewet + water + I(water^2)
```

```
+ I(lowshrub^2) +I(tundra^2)+ I((tundra * ndvi)), data=ac2,id=jbs_tran, family="binomial",
corstr="exchangeable")
```

The accompanying segmented command is below. Note that an added seg.control line is used to refer segmented default to the function to optimize/search for the ZOI.

```
ogeeglm0308<-segmented.default(out.airsurvgeeglm0308,seg.Z=~neg.d_ekdi,psi=-
18,control=seg.control(fn.obj="f(x)"))
```

This approach yields a similar estimate of ZOI of 14.5 km (CI= 12.8–16.4) as detailed in Boulanger et al. (2016). Precision of estimates from the *segmented.default* method is reduced slightly due to adjustment of variances by the geeglm method.

The collar analysis uses the same general approach with clogit used for conditional logistic regression with caribou id as strata and each id and year combination as a cluster (Prima et al 2017). A robust variance estimation method is used.

```
Clogit_collars<-clogit(bincar~
        yearfact:rgt30F+
         bedbould + bedbould:season
        +forest  + forest:season
        +lowshrub
        +I(sedgewet*ndvi)+ sedgewet:season +I(sedgewet*sedgewet)
        +I(treeherb*treeherb)
        +tundra+ tundra:season
        +Tussock:summwint
        +water +I(water*water)+I(water*julydrought)
        +I(mosslich*julydrought)
        +lowshrub:season +inrange
    +strata(id)+cluster(cluster),  data=collars0817yrF,robust=T,
method="efron",model=TRUE)
```

The resulting clogit_collars object is then used with the segmented command to obtain ZOI estimates as with the aerial survey analyses. A model matrix approach can be used to obtain yearly ZOI estimates using the same underlying base habitat model. For collar analyses it is suggested that the number of bootstrap starting values is increased to at least 100 to ensure estimate stability.

References

Boyce, M. S. et al. 2002. Evaluating resource selection functions. – Ecol. Model. 157: 231–300.

Didan, K. 2015. MOD13A2 MODIS/Terra Vegetation Indices 16-Daly L3 Global 1km SIN Grid V006. NASA EOSDIS Land Processes DAAC. – <http://doi.org/10.5067/MODIS/MOD12A2.006>.

Didan, K. et al. 2015. Modis vegetation index user's guide (MOD13 Series), ver. 3.0. – Univ. of Arizona.

Huberty, C. J. 1994. Applied discriminant analysis. – Wiley.

Johnson, C. J. et al. 2006. Resource selection functions based on use-availability data: theoretical motivation and evaluation methods. – J. Wildl. Manage. 70: 347–357.

Muggeo, V. M. R. 2003. Estimating regression models with unknown break-points. – Stat. Med. 22: 3055–3071.

Muggeo, V. M. 2008. segmented: an R package to fit regression models with broken line relationships. – R News 8: 20–25.