Fine-scale harbour seal at-sea usage mapping around Orkney and the North coast of Scotland

Scottish Marine and Freshwater Science Vol 7 No 27

E L Jones, S Smout, C Blight, C Sparling and B McConnell

Fine-scale harbour seal at-sea usage mapping around Orkney and the North coast of Scotland

Scottish Marine and Freshwater Science Report Vol 7 No 27

Esther L. Jones, Sophie Smout, Clint Blight, Carol Sparling and Bernie McConnell

ISSN: 2043-7722

DOI: 10.7489/1876-1

Marine Scotland Science is the directorate of the Scottish Government responsible for the integrated management of Scotland's seas. Marine Scotland Science (formerly Fisheries Research Services) provides expert scientific and technical advice on marine and fisheries issues. Scottish Marine and Freshwater Science is a series of reports that publishes the results of research and monitoring carried out by Marine Scotland Science. It also publishes results of marine and freshwater scientific work that has been carried out for Marine Scotland under external commission. These reports are not subject to formal external peer-review.

This report presents the results of marine and freshwater scientific work carried out for Marine Scotland under external commission. The work was funded through the Scottish Government's Contract Research Fund and by Scottish Natural Heritage.

Fine-scale harbour seal at-sea usage mapping around Orkney and the North coast of Scotland

Esther L. Jones, Sophie Smout, Clint Blight, Carol Sparling and Bernie McConnell.

Sea Mammal Research Unit
Scottish Oceans Institute,
University of St Andrews, St Andrews, Fife, KY16 8LB, UK

Sea Mammal Research Unit

Scottish Natural Heritage
Dualchas Nàdair na h-Alba
All of nature for all of Scotland
Nàdar air fad airson Alba air fad

Executive summary

The report describes how fine-scale harbour seal usage maps around Orkney and the north coast of Scotland, can be used and interpreted, as well as the caveats and limitations, and methodology used to produce them (Appendix A).

1. Harbour seal movement data from telemetry tagged seals, collected between 2003 and 2015, were combined with terrestrial count data. The most recent year that count data were collected from each onshore location was used, which ranged from 2008 to 2015. Population-level species distribution maps and associated confidence intervals were produced around Orkney and the north coast of Scotland at a resolution of 0.6 km x 0.6 km.
2. At the time of publication, the usage maps are available to download in geo-referenced formats from the Marine Scotland Information website at the following address (http://marine.gov.scot/information/fine-scale-harbour-seal-sea-usage-mapping-around-orkney-and-north-coast-scotland).
3. Seasonal usage was also investigated (Appendix B). There were a number of data and numerical constraints that meant robust seasonal usage maps could not be produced. To address these constraints, recommendations of telemetry and terrestrial count data collection and further analysis are made.

Contents

Background

Harbour seals (*Phoca vitulina*) are one of two resident breeding seal species around the UK. While the UK harbour seal population has remained approximately stable, in some areas there have been dramatic declines since 2000. Animals within the Orkney and the North Coast management area have been particularly affected, with numbers decreasing by 78% between 2000 and 2013 (Duck *et al.* 2014). Concern around the status of the population coupled with the uncertainty surrounding the risk of collisions between tidal turbines and seals has led to the need for more information on the real risks presented to this species by tidal turbines. A key element of models for assessing collision risk is determining the number of animals that use the area close to turbines.

Currently, broad scale seal usage maps generated for the Scottish Government by the Sea Mammal Research Unit (SMRU) can be used for the above purpose, where site-specific survey work has not resulted in the generation of accurate and high resolution density, or passage rate estimates for the turbine location(s) (http://www.smru.st-andrews.ac.uk/smrudownloader/uk_seal_usage_of_the_sea). These maps are based on long-term telemetry data collected from tagging studies conducted at many different sites around the UK. In areas where few telemetry data are available, a 'null model' is used. This model, based on a general analysis of the telemetry data, predicts the distribution of seals based on distance to coast and haulout site. These usage maps are produced at a 5 km x 5 km resolution, which is appropriate for defining seal usage over large areas, but is likely to mask important heterogeneity within specific areas of interest to the tidal energy industry.

This report addresses these issues by:

- characterising harbour seal usage using a fine-scale resolution with 0.6 km x 0.6 km grid cells more appropriate to the scale of commercial tidal developments
- incorporating environmental covariates within the 'null model' to provide more realistic predictions of seal distribution
- presenting an efficient GAM-GEE modelling framework for the development of the null model that uses all available movement data
- including recent telemetry (movement) data available in the study area
- providing estimates of uncertainty in usage

This report also investigates whether the available data support the production of seasonal usage maps.

Introduction

The maps shown in this report represent at-sea distributions of harbour seals around Orkney and the North coast of Scotland at 0.6 km x 0.6 km resolution. They can be used in assessments where local seal distribution needs to be taken into account, such as when considering the impact of offshore tidal stream or wave marine renewable developments. Harbour seal movement data from 2003 to 2015 were collected through telemetry tags attached to individual animals that provide information on the at-sea and haulout site locations over the months that tags remain attached to individuals. Through a series of data processing protocols to correct for locational error and sampling bias, the locations were transformed into smoothed density surfaces (Movement data, section 1.2). Only a sample of animals within the population were tagged and so, to provide species distribution maps, locational information was required within the study area to characterise areas where animals were known to be located (through terrestrial counts) but where none of the tagged animals went. Movement data from the tagged animals were combined with environmental data (to characterise habitat important to harbour seals) to model usage in these areas, termed 'null modelling' (Null modelling, section 1.7). Smoothed density surfaces (derived from the movement data) were aggregated with the null modelled surfaces to produce seal distribution surfaces. These were combined with terrestrial count data from the most recent survey year (between 2008 and 2015) to produce population-level species distribution maps (Terrestrial counts, section 1.3). Uncertainty was propagated throughout the analysis to produce maps of associated upper and lower 95% confidence intervals. Details of the methodology are presented in Appendix A and details of the seasonal analysis are presented in Appendix B.

Usage Maps

Figure 1 shows at-sea distribution of harbour seals around Orkney and the North coast of Scotland. The map can be interpreted as the estimated mean number of seals present in a 0.6 km x 0.6 km grid cell. The confidence interval maps (Figures 2 and 3) show lower and upper 95% confidence intervals and can also be interpreted as the upper and lower bounds on the estimated number of seals in each grid cell. Harbour seal at-sea usage across the whole map is estimated as 2420 (95%CI 809, 4103)[1]. These maps can be used as inputs into various impact assessment processes which require spatially explicit estimates of animal density such as collision risk models. Figures 1 to 3 are for illustrative purposes only. Usage map data are available to download from Marine Scotland Information http://marine.gov.scot/information/fine-scale-harbour-seal-sea-usage-mappingaround- orkney-and-north-coast-scotland .

[1] Note that these aggregated confidence intervals are likely to be inflated. See caveat 2 in section 5.

Seasonality was investigated and results are presented in Appendix B. A number of constraints prevented the production of seasonal usage maps. First, the timing of tag deployments allowed telemetry data to be split into spring/summer and autumn/winter. However, dividing telemetry data between seasonal maps increased the size of confidence intervals in the resulting usage maps. This was due to a larger proportion of usage arising from the null model (for areas where telemetry data did not exist but terrestrial count data were available). Second, locations of tag deployments differed seasonally. This may affect predictions of usage if at-sea usage by animals was strongly influenced by tagging location. This is less problematic for non-seasonal maps where telemetry data were aggregated over time.

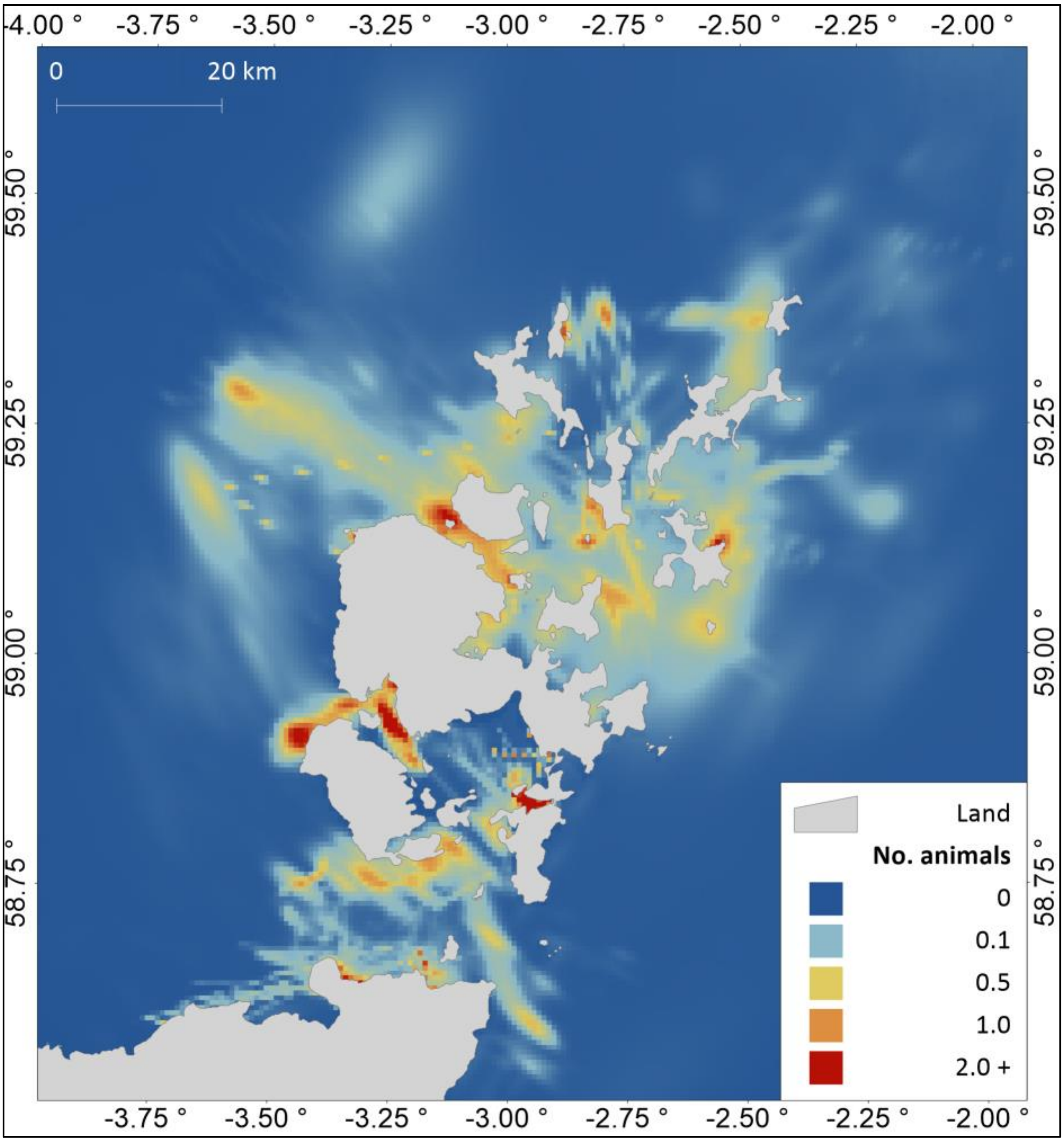


Figure 1. At-sea harbour seal usage showing number of animals per 0.6 km x 0.6 km grid cell.

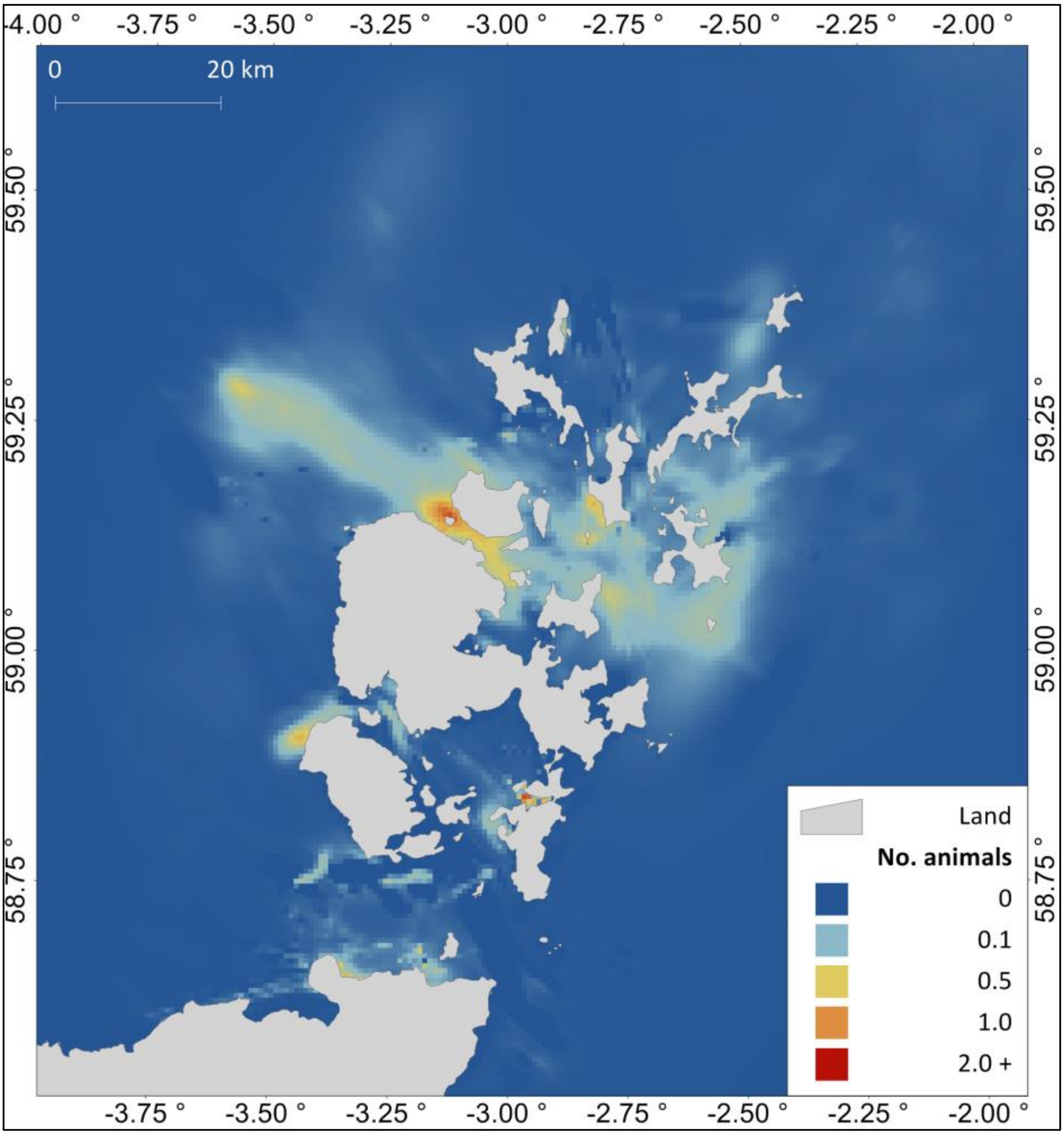


Figure 2. Lower 95% confidence intervals for at-sea harbour seal usage showing number of animals per 0.6 km x 0.6 km grid cell.

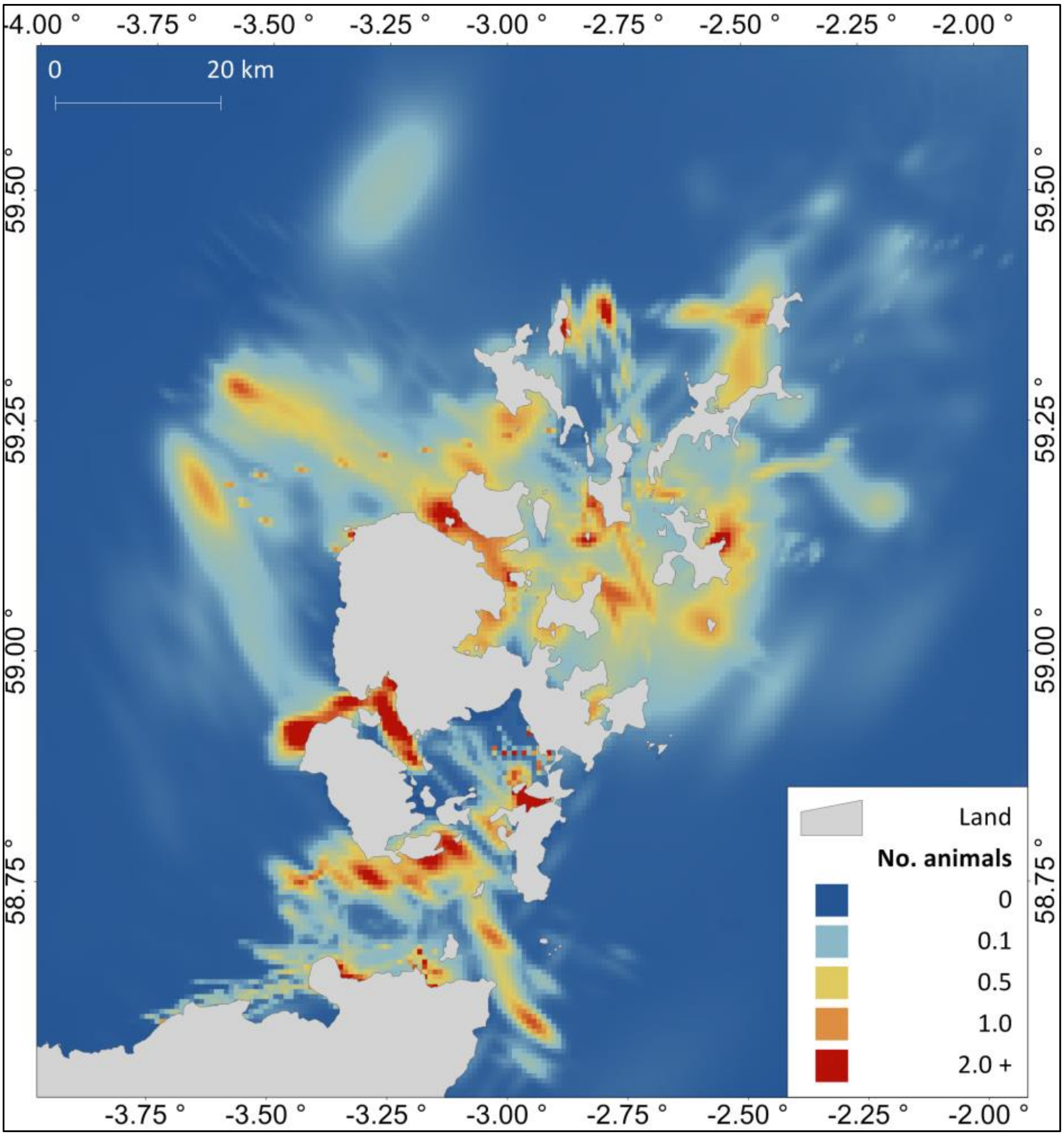


Figure 3. Upper 95% confidence intervals for at-sea harbour seal usage showing number of animals per 0.6 km x 0.6 km grid cell.

Frequently asked questions

Q. What are the scale and projection of the usage maps?

A. Maps are gridded as 0.6 km x 0.6 km cells on a Universal Transverse Mercator Zone 30(North) World Geodetic System 1984 datum (UTM30N WGS84) projection. GIS files are provided in UTM30N WGS84 GeoTIFF and shapefile formats.

Q. How should the maps be interpreted?

A. The maps can be interpreted as the estimated mean number of seals in each 0.6 km x 0.6 km grid square. The upper and lower confidence intervals relate to 95% confidence in this mean. The maps show usage in the ‘at-sea’ (marine) environment. To convert usage in each grid cell to the number of animals per 1 km x 1 km, divide by 0.36.

Q. Should these maps be used in place of the 5 km x 5 km usage maps?

A. These 0.6 km x 0.6 km maps were produced as a response to the requirement to assess the risks of proposed commercial tidal developments around Orkney to the local harbour seal population. The higher resolution was chosen to reflect the finest-scale that the underlying movement data would support. Maps at 5 km x 5 km resolution should be used for all other parts of the UK, or where the area of interest is much larger than the scale of the fine-scale usage analysis, (e.g. offshore wind farm developments). The overall estimated mean number of animals will be different between the original and fine-scale usage analyses because the original maps used all observations to estimate the population in 2013 population estimates whereas the fine-scale maps are scaled to the most contemporary population estimates available.

Q. Do these maps reflect current population estimates?

A. The maps are scaled using the most recent terrestrial count data available at the time of map production (Figure 6). Harbour seal at-sea usage across the whole map is estimated as 2420 (95%CI 809, 4103).

Q. How much usage is reflected by the telemetry data and how much by the null modelling?

A. Telemetry data comprised 92% of usage, null modelling constituted 8% of usage.

Q. Can population estimates be summed over more than one grid cell?

A. The total usage over larger, aggregated areas can be estimated by summing the means for the grid cells it contains. Confidence intervals can also be aggregated (see Caveats and limitations).

Caveats and limitations

1. The estimates of seal numbers were calculated by grid square, therefore the numbers will depend on the size of the grid squares used. Larger cell sizes will result in larger numbers. This must be kept in mind when comparing the results of the present study with previous usage maps calculated at lower resolution.

2. Where developments are likely to impact areas that include several cells on the fine-scale grid, estimated means can be summed over these cells to estimate the total abundance of seals in an area. In many areas, the uncertainty in estimates for neighbouring cells cannot be considered independent of spatial smoothing within the model. A conservative approach to quantify the associated uncertainty is to use the sum of the lower bounds and upper bounds for individual cells as the lower and upper bounds for the whole area. This is likely to overestimate the size of the confidence interval because it ignores both the different data points contributing to the usage estimate of each grid cell and the gradual decay, with distance, in the spatial correlation implicit in the movement data (i.e. high usage in one cell increases the probability that an adjacent cell will also have high usage). Consequently, the degree of overestimation will increase as the scale over which this summing occurs increases.
3. The analysis presented here does not distinguish between habitat that may be important for specific events such as foraging or breeding and areas that might be used as ‘commuting corridors’ between such sites. Usage is displayed over all types of activity. Nor does the analysis take into account patterns of residency and site turnover of animals. For example, mean usage (the estimated number of animals present in a grid cell) does not distinguish between usage by many individuals using an area a small amount, or a small number of individuals using an area a lot. The number of individuals exposed to collision risk is likely to be different between these two situations. This issue is true of any static density inputs into collision risk models.
4. Usage maps are scaled to population estimates using terrestrial count data from August, which does not account for intra-annual movement of the population outwith the study area, or redistribution of animals between haulout sites within the study area.
5. The usage maps account for seal movement in two-dimensions (longitude and latitude) and do not provide information about how animals use depth.

Acknowledgements

Historical telemetry data used in this analysis were funded by Marine Scotland (MS), Department for Energy and Climate Change (DECC), Scottish Natural Heritage (SNH), and Natural Environmental Research Council (NERC). We are also grateful to Professor Paul Thompson for allowing access to data from three tagged animals that were funded by the Moray Firth Marine Mammal Monitoring Plan (Marine Scotland, Beatrice Offshore Windfarm Ltd (BOWL), Moray Offshore Renewables Ltd (MORL), Highlands and Islands Enterprise (HIE) and the Crown Estate (TCE)).

Appendix A – Analysis

Figure 4 shows a schematic flowchart of the analytical processes used to estimate usage of harbour seals.

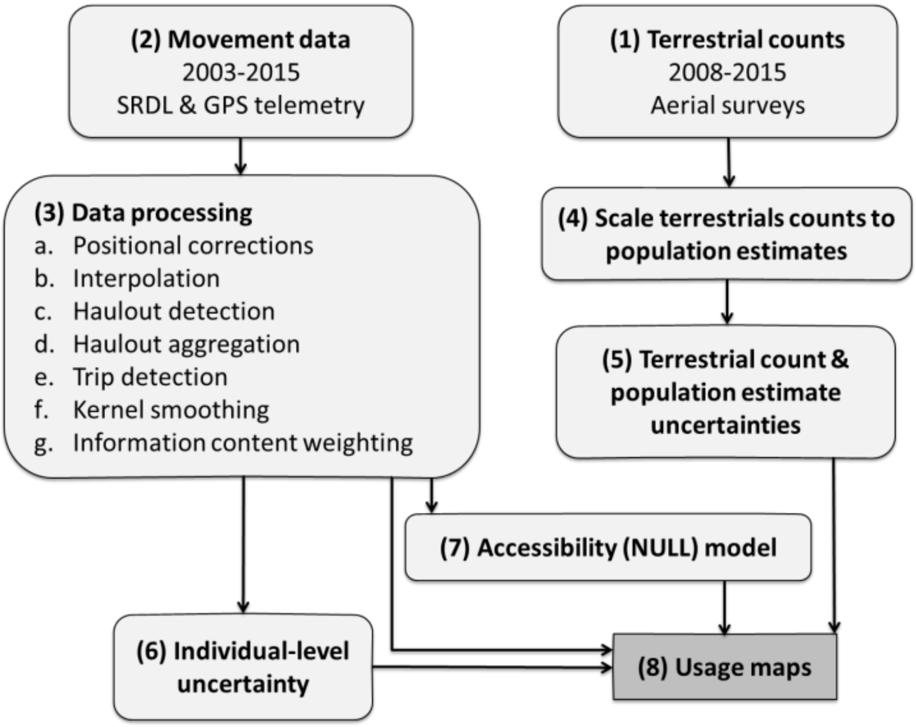


Figure 4. Flowchart showing the analytical methodology used to produce harbour seal usage maps. Adapted from Figure 1 in Jones et *al.* (2015).

1.1 Study area

A study area centred on Orkney was delineated from 58.52°N to 59.66°N and 3.98°W to 1.88°W, to include the majority of telemetry data from the surrounding area (Figure 5). A larger analytical area (referred to as the spatial extent of the analysis in Figure 5) was delineated to capture movement data from animals that spent time at-sea within the study area (but spent time on land outside of the study area). This ensured that usage around the outer edges of the study area was not underestimated. An appropriate spatial resolution of 0.6 km x 0.6 km was determined by the movement data and estimated from the median distance (median = 0.64 km; variance = 2.7 km) between each location of an individual. Analyses were conducted using R 3.3.1 (R Core Team, 2016) and GIS software Manifold 8.0 (Manifold Software Limited, 2013) and all maps were projected using Universal Transverse Mercator 30 North, World Geodetic System 1984 datum (UTM30N WGS84). Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) shoreline data from NOAA were used in all figures where the shoreline is represented, available from www.ngdc.noaa.gov/mgg/shorelines/gshhs.html.

1.2 Movement data

Sixty adult (defined as any animal more than one-year-old) animals, tagged between 2003 and 2015 around Orkney, the North coast of Scotland and the Moray Firth, spent time within the study area. Between 2003 and 2005, Satellite Relay Data Loggers (SRDL) were deployed that used the Argos satellite system for data transmission (Argos 2011). Between 2011 and 2015, GPS phone tags that used the GSM mobile network with a Fastloc© hybrid protocol were deployed (McConnell *et al.* 2004). Telemetry data were processed through a set of data-cleansing protocols to remove null and missing values, and duplicated records from the analysis.

SRDL positional error was corrected using a Kalman filter (a statistical method for improving estimates of position) and data were used to estimate positions at two-hourly intervals (Royer & Lutcavage 2008; Jones *et al.* 2015). The majority of GPS locations have an expected error of ≤ 55 m (Dujon *et al.* 2014), however occasional outliers were excluded using thresholds of residual error and number of satellites, and then straight-line interpolated to regularise to the same two-hourly intervals as the SRDL data (Jones *et al.* 2015). Three animals (pv1-foxy-03, pv18a-Izzy-06, pv59-09-15) had few locations within the study area, and three animals (pv44-013-12, pv9-Chris-04, pv9-Gabe-04) did not have any haulout records, so these six animals were excluded, bringing the total number of animals used in the analysis to 54 (Table 1).

Continuous spatial surfaces to represent the proportion of time animals spent in different areas were derived by kernel-smoothing the telemetry data. The *KS* library in R (Duong 2016) was used to estimate spatial bandwidth of the 2D kernel applied to each animal/haulout site map. Mean and variance were scaled to population size by combining each one with the population mean and variance estimates of each haulout site, and these were aggregated to the total usage map.

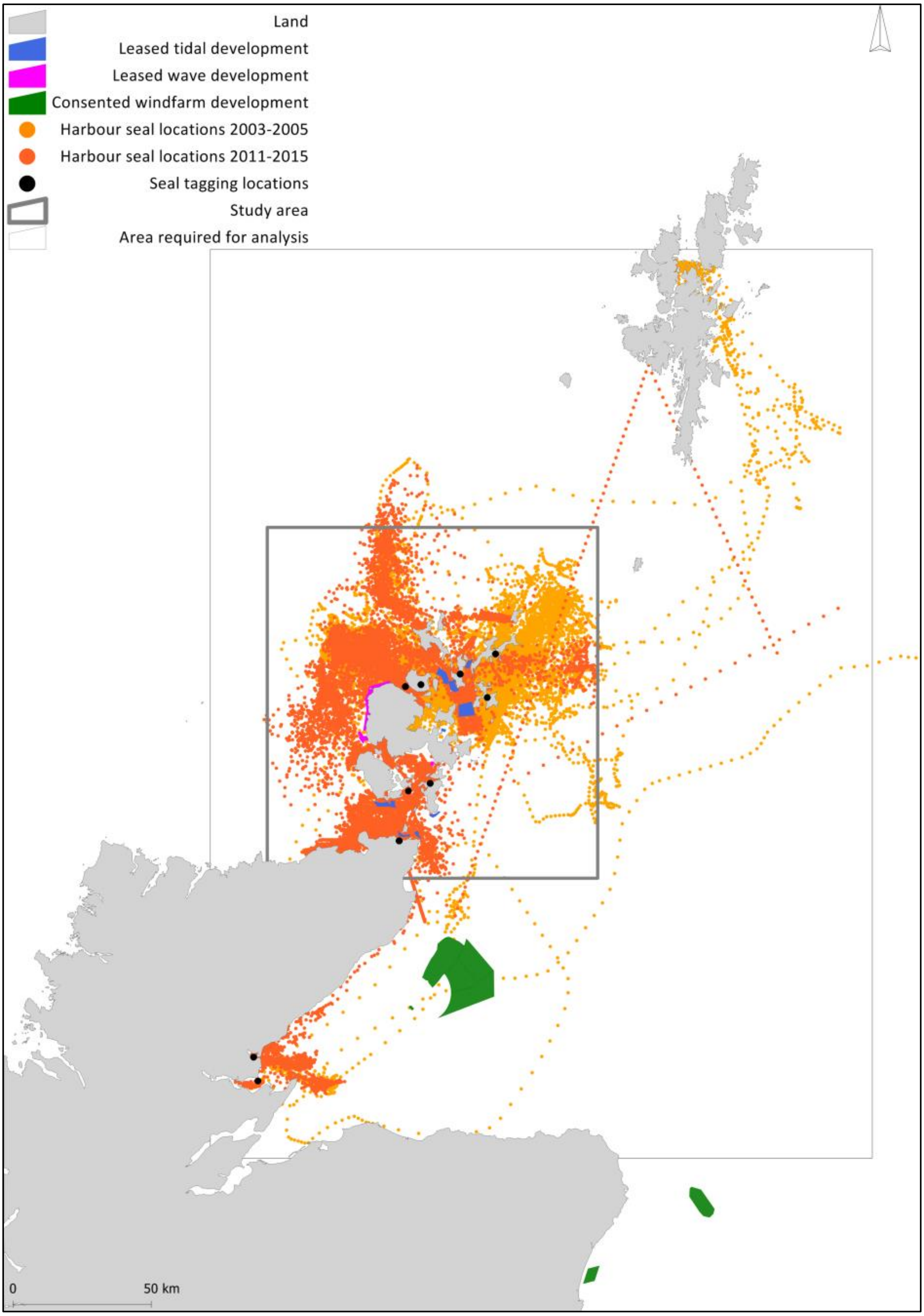


Figure 5. Map showing the tracks of 54 animals included in the analysis (orange circles), their tagging locations (black circles), proposed offshore marine renewable developments (blue, pink, green areas), and the spatial extents of the analysis (outer rectangle), and final study area (inner rectangle).

Table 1. Animals included in the analysis showing animal reference number, tag type, age-class, region and location of tagging, original funding, sex, and year tagged. Starred (*) rows indicate animals that were not included in the original 5km x 5km usage maps (Jones *et al.* 2015) but were used in the fine-scale usage maps. (**) column related to the primary seasonality of the animal; AW = autumn/winter (post-moult), SS = spring/summer (pre-moult). Funding codes acronyms are: Department of Energy and Climate Change (DECC), Scottish Natural Heritage (SNH), Marine Scotland (MS), Natural Environmental Research Council (NERC), and Moray Firth Marine Mammal Monitoring Plan (MFMMMP).

#	Animal reference	Tag type	age	Management region	Tagging location	Funding	Sex	Year	Season**
1	pv1-ali-03	SRDL	1+	Orkney & N coast	Sanday	DECC	F	2003	AW
2	pv1-Arnie-03	SRDL	1+	Orkney & N coast	Eynhallow	DECC	M	2003	AW
3	pv1-bo-03	SRDL	1+	Orkney & N coast	Sanday	DECC	F	2003	AW
4	pv1-Bob-03	SRDL	1+	Orkney & N coast	Eynhallow	DECC	M	2003	AW
5	pv1-cat-03	SRDL	1+	Orkney & N coast	Sanday	DECC	F	2003	AW
6	pv1-dot-03	SRDL	1+	Orkney & N coast	Sanday	DECC	F	2003	AW
7	pv1-erin-03	SRDL	1+	Orkney & N coast	Rousay	DECC	F	2003	AW
8	pv6-Ken-04	SRDL	1+	Orkney & N coast	Stronsay	DECC	M	2004	SS
9	pv6-Len-04	SRDL	1+	Orkney & N coast	Stronsay	DECC	M	2004	SS
10	pv6-Max-04	SRDL	1+	Orkney & N coast	Rousay	DECC	M	2004	SS
11	pv6-Oli-04	SRDL	1+	Orkney & N coast	Eynhallow	DECC	M	2004	SS
12	pv6-pat-04	SRDL	1+	Orkney & N coast	Stronsay	DECC	F	2004	SS
13	pv6-Pete-04	SRDL	1+	Orkney & N coast	Eynhallow	DECC	M	2004	SS
14	pv6-queenie-04	SRDL	1+	Orkney & N coast	Rousay	DECC	F	2004	SS
15	pv6-sally-04	SRDL	1+	Orkney & N coast	Eynhallow	DECC	F	2004	SS
16	pv11-James-05	SRDL	1+	Moray Firth	Dornoch	DECC	M	2005	SS
17	pv11-Kath-05	SRDL	1+	Moray Firth	Dornoch	DECC	F	2005	SS
18	pv24-112-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	AW
19	pv24-148-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	AW
20	pv24-150-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	F	2011	AW
21	pv24-151-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	AW
22	pv24-153-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	F	2011	AW
23	pv24-155-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	AW
24	pv24-165-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
25	pv24-394-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
26	pv24-541-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
27	pv24-580-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	F	2011	SS
28	pv24-590-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
29	pv24-598-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	F	2011	SS
30	pv24-622-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
31	pv24-x625-11	GPS	1+	Orkney & N coast	Pentland	SNH, MS	M	2011	SS
32	pv44-003-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	F	2012	SS
33	pv44-004-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	F	2012	SS
34	pv44-005-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, SNH, MS	M	2012	SS

35	pv44-007-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	F	2012	SS
36	pv44-011-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, SNH, MS	M	2012	SS
37	pv44-014-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, SNH, MS	M	2012	AW
38	pv44-017-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	M	2012	SS
39	pv44-018-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	M	2012	SS
40	pv44-020-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	F	2012	SS
41	pv44-021-12	GPS	1+	Orkney & N coast	Eday	NERC, SNH, MS	F	2012	SS
42	pv47-392-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, MS	M	2012	AW
43	pv47-427-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, MS	M	2012	AW
44	pv47-539-12	GPS	1+	Orkney & N coast	Eday	NERC, MS	M	2012	AW
45	pv47-583-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, MS	M	2012	AW
46	pv47-585-12	GPS	1+	Orkney & N coast	Eday	NERC, MS	M	2012	AW
47	pv47-588-12	GPS	1+	Orkney & N coast	Eynhallow	NERC, MS	M	2012	AW
48	pv57-197-14	GPS	1+	Orkney & N Coast	St Margarets	SNH, MS, NERC	F	2014	AW
49	pv57-199-14	GPS	1+	Orkney & N Coast	Switha	SNH, MS, NERC	M	2014	AW
50	pv57-200-14	GPS	1+	Orkney & N Coast	St Margarets	SNH, MS, NERC	F	2014	AW
51	pv57-913-14	GPS	1+	Orkney & N Coast	St Margarets	SNH, MS, NERC	F	2014	AW
*52	pv59-05-15	GPS	1+	Moray Firth	Loch Fleet	MFMMMP	F	2015	SS
*53	pv59-07-15	GPS	1+	Moray Firth	Loch Fleet	MFMMMP	F	2015	SS
*54	pv59-12-15	GPS	1+	Moray Firth	Loch Fleet	MFMMMP	F	2015	SS

1.3 Terrestrial counts

Harbour seals are surveyed during their moult in August when the greatest number of animals haul out on land for an extended period. During aerial surveys all seals along a specified coastline are counted and coordinates are recorded to an accuracy of approximately 50 m. Surveys take place within two hours of low tide when low tide is between 12:00 and 18:00 hours (Thompson *et al.* 2005). Survey effort is variable between locations (Figure 6). Surveyed coastline was gridded to 0.6 km x 0.6 km and the most recent available count (ranging from 2008 and 2015) was used in each grid cell (Figure 7). Grid cells that were surveyed but in which no animals were located were given a value of 0 (Table 2). A single population estimate was produced and associated uncertainty attributed to each haulout grid cell. Full details of this method are available from Jones et *al.* (2015); Supplementary information (http://www.int-res.com/articles/suppl/m534p235_supp.pdf).

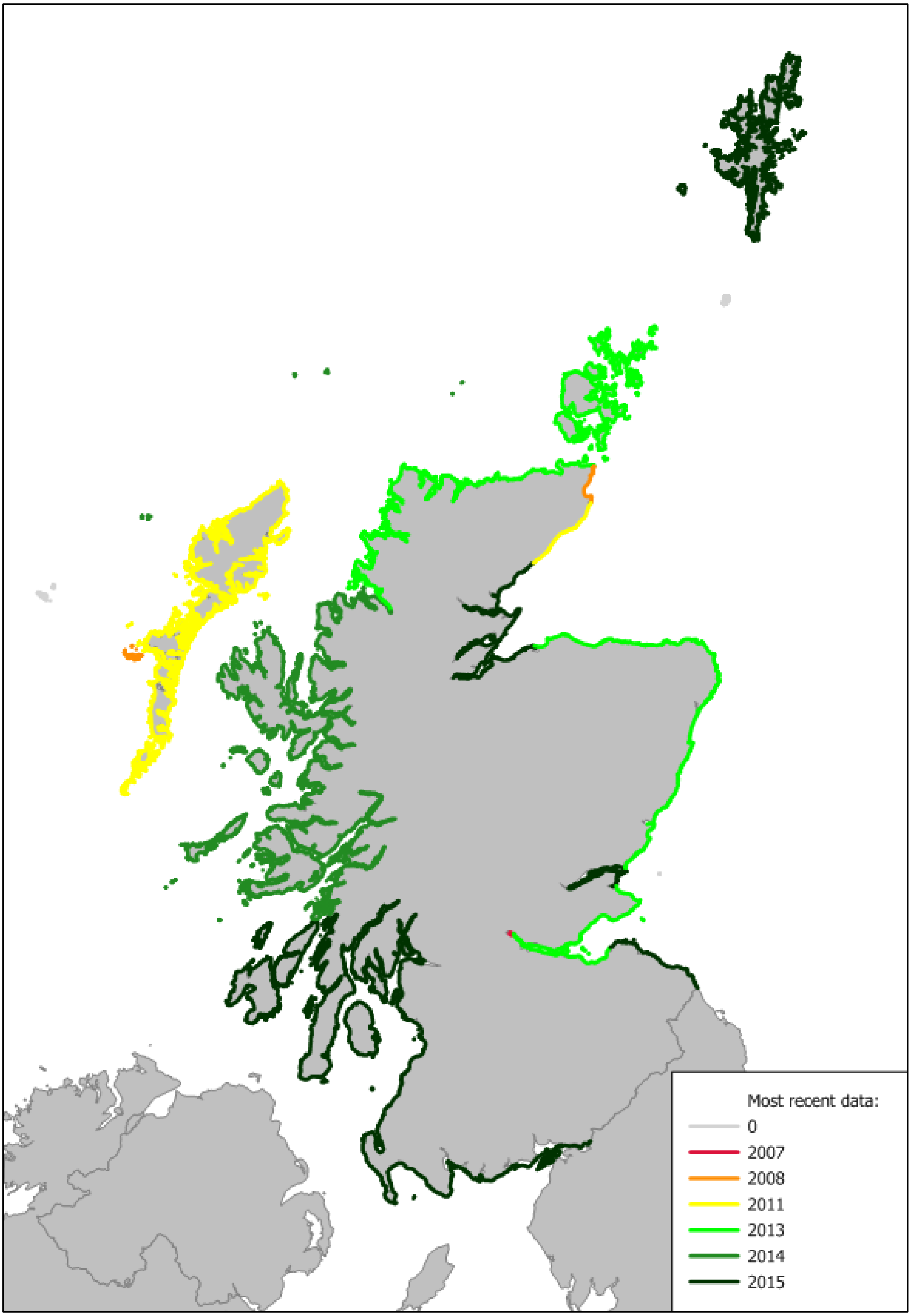


Figure 6. Map showing the most recent terrestrial surveys around the coast of Scotland. Grey lines represent 0 terrestrial counts.

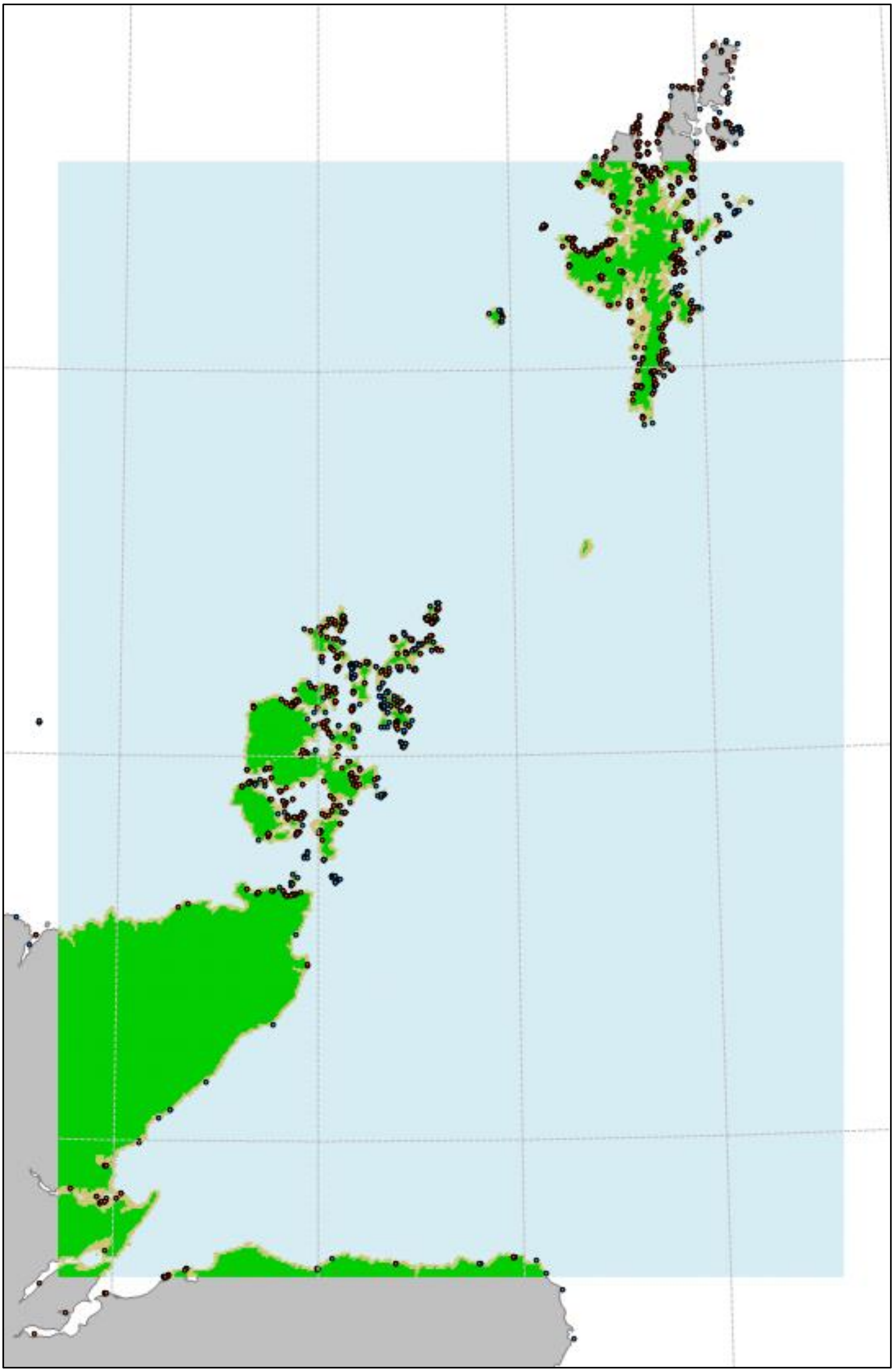


Figure 7. Locations of most recent aerial survey (non-0) counts (ranging from 2008 to 2015 for individual locations) within the analytical study area.

Table 2. Most recent harbour seal terrestrial counts available for each section of surveyed coastline up to 2015.

Coastline surveyed	Year	Count
Shetland	2015	2837
Orkney	2013	1865
North Coast of Scotland	2013	59
Moray Firth (Duncansby Head to Wick)	2008	1
Moray Firth (Wick to Helmsdale)	2011	0
Moray Firth (Helmsdale to Findhorn)	2015	468
Moray Firth (North Grampian)	2013	39
Total		**5269**

1.4 Environmental data

Harbour seals are central-place foragers, regularly hauling out on land in between spending time at sea travelling and foraging. Therefore, their distribution is likely to be strongly linked to their haulout locations. Geodesic (shortest at-sea) distance, seabed depth (bathymetry), and seabed sediment have been shown to effectively characterise UK seal habitat preference in the North Sea (Aarts *et al.* 2008). Like any predator, seals most likely respond dynamically to their environment with regards to the location of their prey species. However, fish distributional data were not available at the scale of this analysis and so a number of environmental covariates were used to predict seal distribution: geodesic distance and seabed depth characterised seal movement; sediment, annual mean tidal power, and peak flow were used as a proxy for prey distributions (Table 3 and Figure 8).

1. **Geodesic distance** By definition, central-place foragers have a home-range (they move to other locations before returning to their original departure point). For marine animals, this is represented by the shortest distance between a haulout site and an at-sea location. By contrast, Euclidean distance is the straight-line distance from one point to another, and ignores barriers to movement such as land. Around complex coastlines such as the UK, Euclidean and geodesic distances can vary widely. Geodesic distance was calculated using the R library *gdistance* (van Etten 2015) at a resolution of 0.6 km x 0.6 km for each haulout site to determine the distance between each seal location and the corresponding haulout site.

2. **Seabed depth** The bathymetric metadata and Digital Terrain Model data products were derived from the European Marine Observation and Data Network (EMODnet) Bathymetry portal (http://www.emodnet-bathymetry.eu) released August/September 2015. The seabed depth data had a resolution of

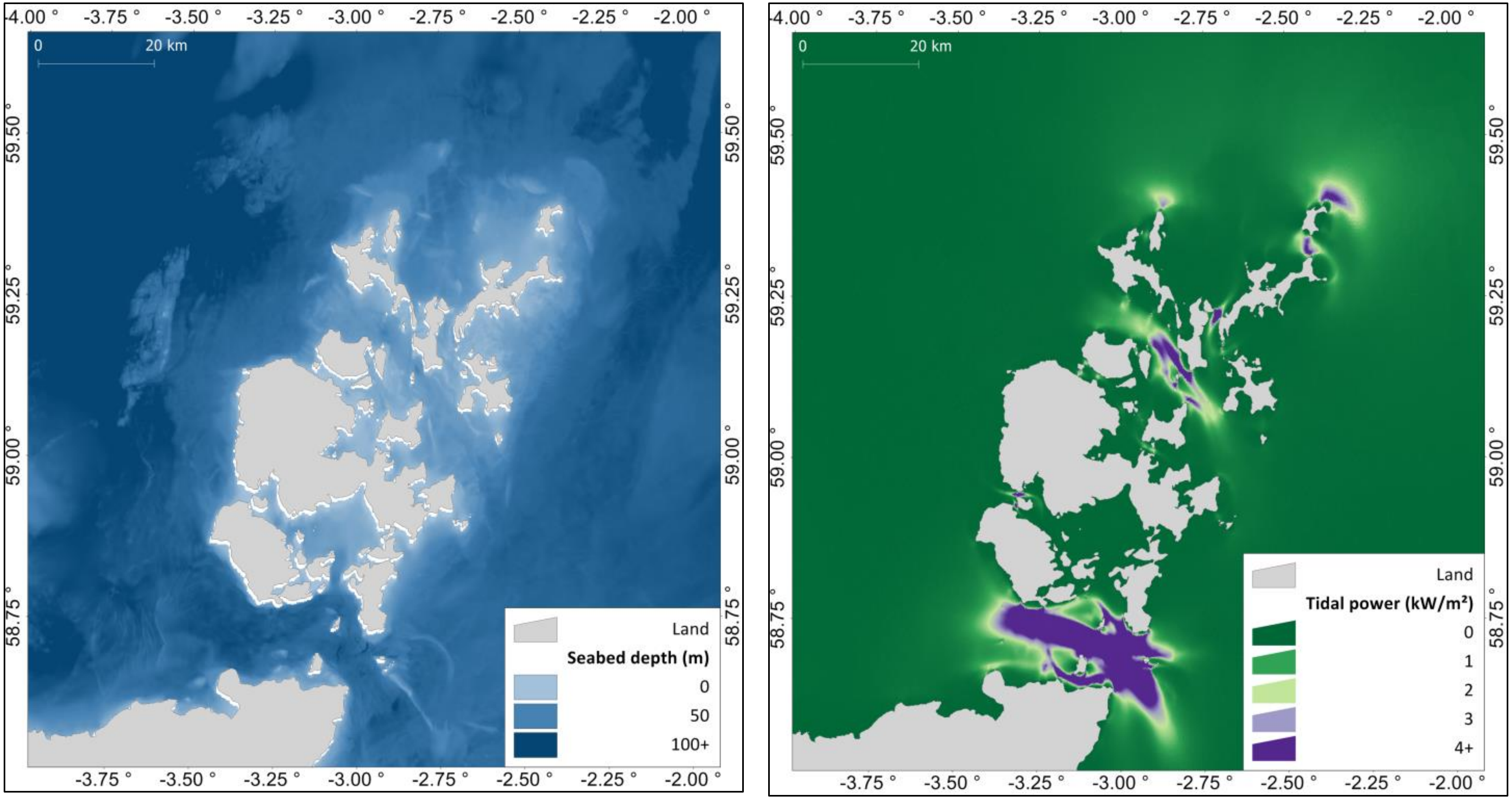
1/8 minutes (~230 m) and are based on the seabed depth at the Lowest Astronomical Tide (LAT).

3. **Tidal power and peak flow** Annual mean tidal power (kWm^{-2}), peak flow for mean spring tide (ms^{-1}), and peak flow for mean neap tide (ms^{-1}) were produced by the Pentland Firth and Orkney Waters Hydrodynamic Model (PFOW), courtesy of Marine Scotland (Price *et al.* 2016). The covariates had unstructured grid resolution of 1 km at the coast and 10 km at the shelf edge. The flow covariates represent the spatial distribution of an average of ebb and flood tides and therefore integrate over tidal states.

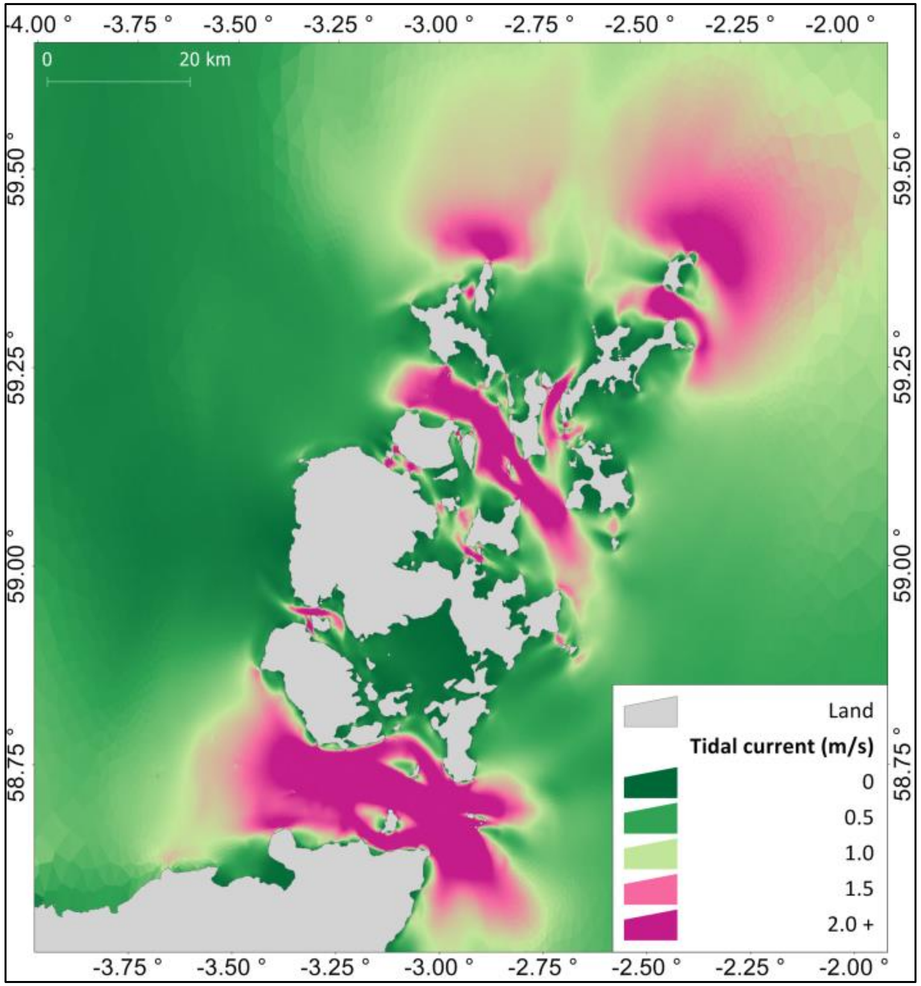
4. **Proportion of sand/gravel/mud** Sediment type was derived from the British Geological Survey (http://www.bgs.ac.uk/products/digbath250/home.html), obtained from core samples spaced 5 km apart on average. A simplified Folk classification system (Folk 1954) was applied to derive variables containing proportions of sand, gravel, and mud. Data were given as a percentage-by-weight of gravel (particles > 2.0 mm in diameter), sand (0.0625 - 2.0 mm in diameter), and mud (particles < 0.0625 mm in diameter). Spatial autocorrelation between the three covariates was calculated by randomly sub-sampling the cores to calculate semi-variograms (Isaaks & Srivastava 1990). Each sediment covariate was kriged at a 1 km resolution using the semi-variograms and the resultant local estimates were normalised to 100% (Aarts *et al.* 2008). These covariates do not take into account other substrate (such as underlying rock) that may be present on the seabed, or biotope information.

Table 3. Environmental covariates used in 'null modelling' to predict usage where movement data were limited.

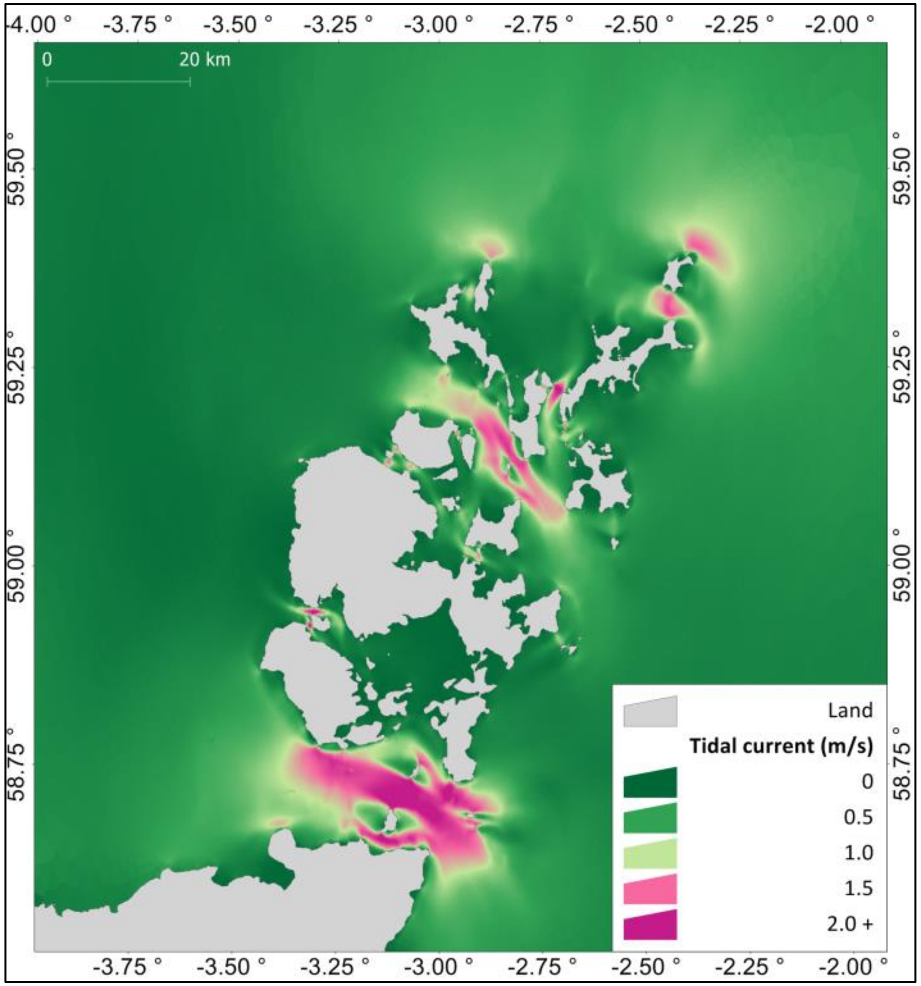
Description	Data source	Original scale and projection
Geodesic distance to haulout (km)	User defined	0.6 km x 0.6 km grid squares, UTM30N WGS84
Seabed depth (m)	EMODnet	Vector 1/8’ resolution, longitude-latitude WGS84
Annual mean tidal power (kWm^{-2})	PFOW Hydrodynamic model	Unstructured grid resolution of 1 km at coast and 10 km at shelf edge, UTM30N WGS84
Peak flow for mean spring tide (ms^{-1})	PFOW Hydrodynamic model	Unstructured grid resolution of 1 km at coast and 10 km at shelf edge, UTM30N WGS84
Peak flow for mean neap tide (ms^{-1})	PFOW Hydrodynamic model	Unstructured grid resolution of 1 km at coast and 10 km at shelf edge, UTM30N WGS84
Proportion of sediment type including CO_3 concentration (%)	British Geological Survey (Digbath 250)	UTM30N WGS84


-4.00 ° -3.75 ° -3.50 ° -3.25 ° -3.00 ° -2.75 ° -2.50 ° -2.25 ° -2.00 °
0 20 km
59.50 ° 59.25 ° 59.00 ° 58.75 °
Land
Seabed depth (m)
0
50
100+
Land
Tidal power (kW/m²)
0
1
2
3
4+

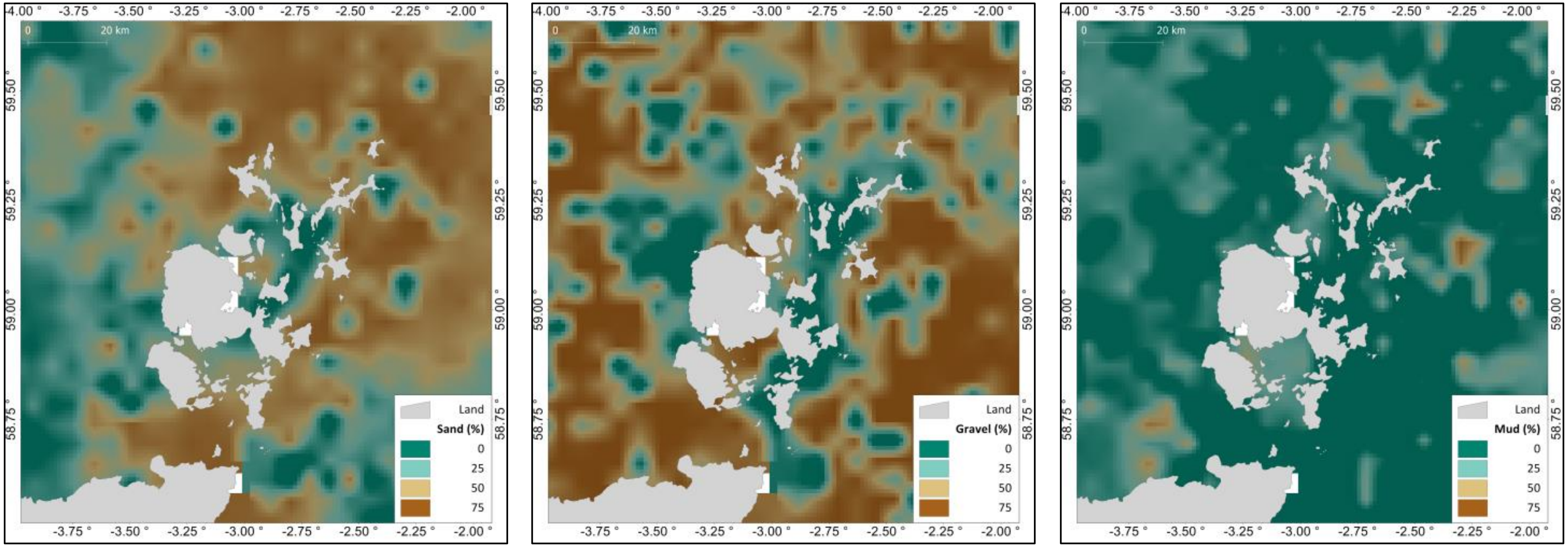
a **b**


-4.00 °
-3.75 °
-3.50 °
-3.25 °
-3.00 °
-2.75 °
-2.50 °
-2.25 °
-2.00 °
0
20 km
59.50 °
59.25 °
59.00 °
58.75 °
Land
Tidal current (m/s)
0
0.5
1.0
1.5
2.0 +

c


-4.00 °
-3.75 °
-3.50 °
-3.25 °
-3.00 °
-2.75 °
-2.50 °
-2.25 °
-2.00 °
0
20 km
59.50 °
59.25 °
59.00 °
58.75 °
Land
Tidal current (m/s)
0
0.5
1.0
1.5
2.0 +

d



e **f** **g**

Figure 8. Environmental covariates (except geodesic distance) used for model selection in 'null modelling', (a) Seabed depth, (b) Annual mean tidal power, (c) Peak flow for mean spring tide, (d) Peak flow for mean neap tide, (e) Proportion of sand in seabed sediment, (f) Proportion of gravel in seabed sediment, (g) Proportion of mud in seabed sediment. White areas in the map are those with no data.

1.5 Haulout aggregation

A 0.6 km x 0.6 km grid cell was identified as a haulout grid cell either from the telemetry data when tagged animals moved onto land that was within the grid cell, and/or from the terrestrial count data where animals were counted in that grid cell when hauled out. Haulout grid cells were aggregated for three reasons: (a) the granularity of a 0.6 km x 0.6 km grid cell may not be consistent with animal behaviour and space-use if more than one haulout grid cell forms part of a connected aggregation (e.g. seals may return to a haulout grid cell close to the departure haulout grid cell); (b) using single haulout grid cells maximised the number of haulout grid cells defined by the terrestrial count data but which did not have telemetry data directly associated with them. This would have resulted in inflated uncertainty as the null model would contribute more usage to the analysis than would be necessary; and (c) using single haulout grid cells that were associated with telemetry data but where the terrestrial count was 0 reduced the importance of those telemetry data.

Haulout grid cells were aggregated using a clustering algorithm based on the distances between them (taking account of the complex coastline in the area). An investigation into the appropriate spatial scale for clustering was carried out. Hierarchal cluster analysis with a centroid agglomeration method was implemented to generate aggregated haulout cells ranging from shortest separation of 0.6 km (no clustering) to 15 km (maximum clustering) in increments of 0.6 km (reflecting the scale of the analysis) (Everitt *et al.* 2011). *Telemetry clusters* were defined as having telemetry data from at least one tagged animal associated with any haul out grid cell in the cluster. *Null clusters* were defined as having at least one non-0 count from the terrestrial count data but no tagged animals visited any haul out grid cells within the cluster. To identify the most appropriate clustering size, four factors were considered: (1) total number of clusters, given computational constraints; (2) number of telemetry clusters where the sum of survey counts was 0; (3) largest aggregated survey count associated with a cluster, and number of telemetered animals that used the cluster with the largest aggregated survey count. This was to ensure that the population estimate from one large haul out grid cell (where only a few tagged animals were associated with it) did not unduly influence the analysis; (4) approximate contribution of telemetry usage in the final usage maps (defined using the terrestrial count associated with telemetry clusters). Balancing these considerations gave an optimum cluster size of 4.2 km, which resulted in 218 telemetry clusters and 103 null clusters. Of the telemetry clusters, 164 had no terrestrial count data associated with them. To keep these clusters in the analysis, their terrestrial counts were changed to 1. To ensure the overall terrestrial count in the analysis was consistent, a count of 164 animals was then subtracted from telemetry clusters where the original count was greater than 1 by rescaling the counts to sum of the counts minus 164.

1.6 Modelling review

Existing seal usage maps use an integrated framework of density estimation and regression modelling to provide mean estimates and uncertainty of seal usage scaled by local population estimates around the UK (Jones *et al.* 2015). Predictions of usage are required for locations at which seals are known to haul out but for which no telemetry data are available. Models are fitted using telemetry data from other sites and (environmental) covariates to estimate seal usage. The model can be used to predict usage from a null haulout grid cell based on the values of covariates (such as bathymetry) in the local area around that cell.

Telemetry data have intrinsic properties that need to be accounted for. Firstly, they provide information about an individual observed at a specific location and time, which is based partly on where the animal chooses to be and partly on where it had been a short-time previously (spatial and temporal dependencies, termed as autocorrelation). To account for autocorrelation, data are often 'thinned' systematically until autocorrelation is effectively removed. However, discarding data can be problematic, for example, due to small sample size or study area relative to the scale of animal movement. Secondly, telemetry data, by their nature are presence-only locations (i.e. they show where animals were observed, but in contrast to survey data do not contain locations at which animals are known to be absent). There are few presence-only methods for habitat modelling, such as climate envelope modelling (Tsoar *et al.* 2007); the majority of methods also require absence data. Absences can be created artificially by generating points within the study area, termed *pseudo-absences*. Comparative analyses have shown that even where presence-only data are available, models that use presence-absence methods outperform presence-only methods (Elith *et al.* 2006). Regression modelling techniques such as Generalised Linear Models (GLM) or Generalised Additive Models (GAM) are generally robust to pseudo-absences that are selected randomly (Wisz & Guisan 2009), given there are a sufficient number of presence points (Barbet-Massin *et al.* 2012). However, there is evidence to suggest that thinning presence-only data in order to deal with autocorrelation may adversely affect the ability of the model to accurately predict in analyses where data are limited. Therefore, we augmented the GLM regression modelling approach taken in (Jones *et al.* 2015) and implemented a Generalised Additive Modelling – Generalised Estimating Equation (GAM-GEE) modelling framework while retaining all the available data in the analysis. This methodology allows all telemetry data to be used because it explicitly deals with spatio-temporal autocorrelation (Pirotta *et al.* 2011).

1.7 Null modelling

To predict usage in areas of the study area where terrestrial counts were available but movement data were not, a null model was built using all available telemetry data to predict probability of presence based on covariates. Five pseudo-absences were associated with each presence point by repeatedly selecting at-sea points within the study area so that a representative range of underlying environmental covariates could be associated with the pseudo-absence points. Multicollinearity between the covariates was tested using variation inflation factor analysis with the R library *car* (Fox & Weisberg 2011). The threshold for high multicollinearity (termed Variance Inflation Factor – VIF) was taken to be 5 (Sokal & Rohlf, 2012). Peak flows for mean spring and neap tides were highly correlated so these covariates were not included in the same model during model selection. All other covariates had a VIF score between 1.5 and 3.7.

The *geepack* R library (Højsgaard *et al.* 2006) was used to fit binomial GAM-GEEs with a logit link function and an independent working correlation structure to account for any residual autocorrelation within defined subsets of data (Pirotta *et al.* 2011). Weightings were used so that each presence had five-times more weighting than each pseudo-absence (thus giving presence and pseudo-absences the same importance). As per Russell *et al.* (2016), autocorrelation was estimated separately for presences and pseudo-absences to avoid underestimating the autocorrelation within the presences of an individual. Each pseudo-absence was assumed to be independent. Year was included as a factor and geodesic distance was kept as a linear covariate term within the linear predictor. The *splines* library was used to implement cubic *β*-splines that allowed all other covariates to vary as a function of one-dimensional smooth terms within the linear predictor (4 degrees of freedom) with one internally positioned knot at the mean of each covariate (Pirotta *et al.* 2011).

Model selection was carried out using five-fold cross-validation (Wiens *et al.* 2008) (with spatial blocks based on haulout cluster) using between 10 and 20 equal-area bins with a moving window. This allowed robust estimation of the precision associated with the spatial predictions because, by using the independent working correlation structure, we accounted for any residual autocorrelation within defined panels of data (Pirotta *et al.* 2011). Forwards model selection was used by adding in each covariate (linear and spline terms separately), then using cross-validation to select the highest scoring covariate until there was no improvement in scores. The selected model predicted usage for each null haulout cluster.

1.8 Population-scaling and uncertainty

Within-haul out variance was modelled using data-rich haulout clusters (determined experimentally to be those sites which had ≥ 7 tagged animals associated with them). Variance was estimated using linear models with explanatory covariates of sample size (number of tagged animals in the haulout cluster) and mean usage by seals. The models predicted variance for data-poor and null usage clusters (where terrestrial count data existed but movement data did not). Within-haulout cluster variance was estimated for null clusters by setting the sample size of the uncertainty model to zero.

Density estimation was used to generate usage maps for those haulout sites where telemetry data were available. Individual-level weightings were applied to account for differences in the magnitude of data collected by an animal over its tag lifespan and for variation in the operational settings of the tag itself (Jones *et al.* 2015). The harbour seal population in each haulout cluster was estimated from terrestrial count data, which were rescaled to allow for the proportion of animals that were at sea when surveys were carried out. Using mean haul out probabilities over all available months and their variances, a distribution of population estimates was derived ranging from the value of each terrestrial count (minimum population size) to three times the count (maximum population size). A likelihood distribution was sampled using parametric bootstrapping 500 times per count to produce a distribution of estimates. A single population estimate and variance for each haulout cluster was calculated from the bootstrapped estimates (Jones *et al*. 2015; Supplementary Information).

Individual and population-level variances were combined to form uncertainty estimates for the usage maps. Maps for all haulout clusters were normalised and then scaled according to the number of animals observed there in the surveys, also accounting for the mean proportion of time animals spent at sea. Telemetry-based maps were aggregated with the predictions of the null model to create total usage map(s) for the whole area of the study.

1.9 Results

Geodesic distance, proportion of sand, tidal power, and seabed depth were chosen in the selected model. The selected model had polynomial (spline) terms for proportion of sand and seabed depth, using 4 degrees of freedom, whilst geodesic distance and tidal power were linear. Geodesic distance and seabed depth had negative coefficients, indicating that usage decreases with increasing distance and depth from haulout sites. The relationships found between usage and geodesic distance and seabed depth are expected for central-place foragers, and corroborates other literature showing this finding (Aarts *et al.* 2008). There was a slight preference in sediment type for 55% sand, and mean tidal power exhibited a negative coefficient, meaning that usage declined with increasing power. Figure 9 shows the probability of seal presence modelled as functions of each covariate with accompanying 95% confidence intervals derived through parametric bootstrapping. There is no simple or reliable method for providing a goodness of fit metric using these techniques.

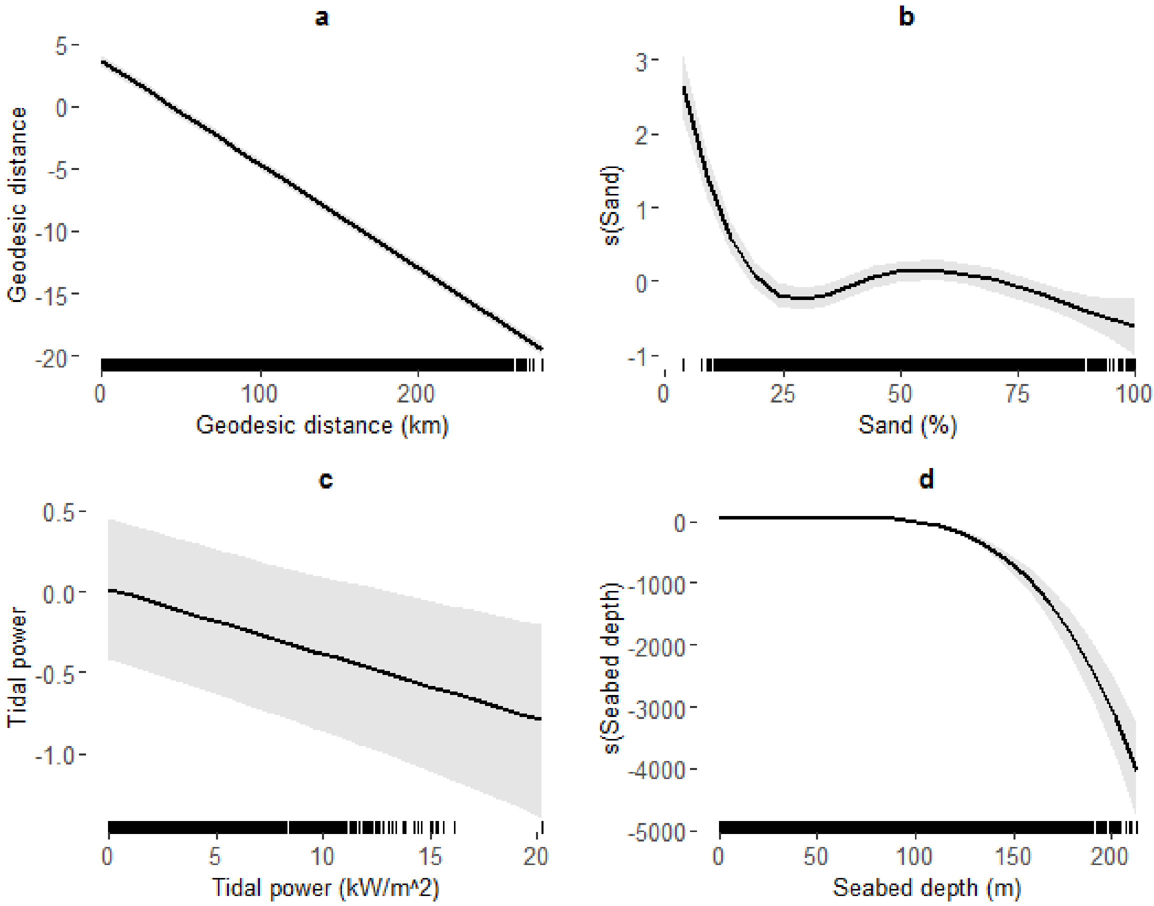


Figure 9. Marginal relationships between covariates and probability of seal presence on the link scale of the model (logit). Modelling of seal presence was carried out as a function of (a) Geodesic distance to haulout, (b) Proportion of sand in sediment (c) Mean tidal power, (d) Seabed depth. The shaded areas represent 95% confidence intervals (using parametric bootstrapping). A rug plot with actual data values are shown on the x-axis of each plot.

Appendix B – Seasonal analysis

1.10 Methods

Seasonality within the telemetry data was investigated. The timing of tag deployments supported a pre-moult/post-moult temporal split, roughly analogous to spring/summer and autumn/winter (Figure 10). There was seasonal split between tag deployments: 31 tags were deployed pre-moult spanning Julian days 50 to 212 (mid-Feb to July) in years 2004, 2005, 2011, 2012, 2014, and 2015, and 23 tags were deployed post-moult spanning days 275 to 450 (October to March) in years 2003 (with the exception of the two tags pv1-cat-03, pv1-dot-03 that had deployment days greater than 450 days), 2011, 2012, 2014, and 2015. Figures 11 and 12 show telemetry data and tagging locations for spring/summer and autumn/winter respectively.

1.11 Results

A number of constraints prevented the production of seasonal usage maps. Telemetry tags were placed onto animals either before (March) or after (September) their annual moult in August. Therefore, telemetry data from a single animal were either collected in spring/summer (February to July) inclusive of breeding and moulting seasons, or autumn/winter (October to March) when animals were not breeding or moulting. Splitting these data during spatial smoothing meant that inevitably the null model (used in areas where telemetry data did not exist but terrestrial count data were available) contributed a higher proportion of usage (spring/summer = 45.4%; autumn/winter = 15.7%) than was the case for the non-seasonal usage map (8.0%). This widened the confidence intervals for each seasonal map. Second, locations and quantity of tag deployments differed seasonally. This may affect predictions of usage if at-sea usage by animals was strongly influenced by tagging location. Third, terrestrial count data were collected during August. Scaling seasonal maps to population levels using these data would give each map the same population estimate. This means that, like non-seasonal maps, seasonal maps do not account for intra-annual changes to the population (animals entering or leaving the study area, or redistribution of animals at haulout sites) but assume the population remains constant in terms of distribution and abundance.

To further investigate seasonality and produce seasonal usage maps, additional telemetry data collection is required. This would ensure that confidence intervals in the resulting usage maps could be reduced by minimising the usage contributed by the null model. Any potential correlation of tagging location and at-sea usage (at the scale of the analysis), could be also accounted for. In order to scale usage maps to the population present in any given time of year, terrestrial count data would also need to be collected seasonally. However, count surveys are costly and non-moult haulout probability is more variable therefore population census surveys are restricted to August. Finally, additional exploratory analysis and modelling should be undertaken to ensure any seasonal usage estimates are robust and defensible.

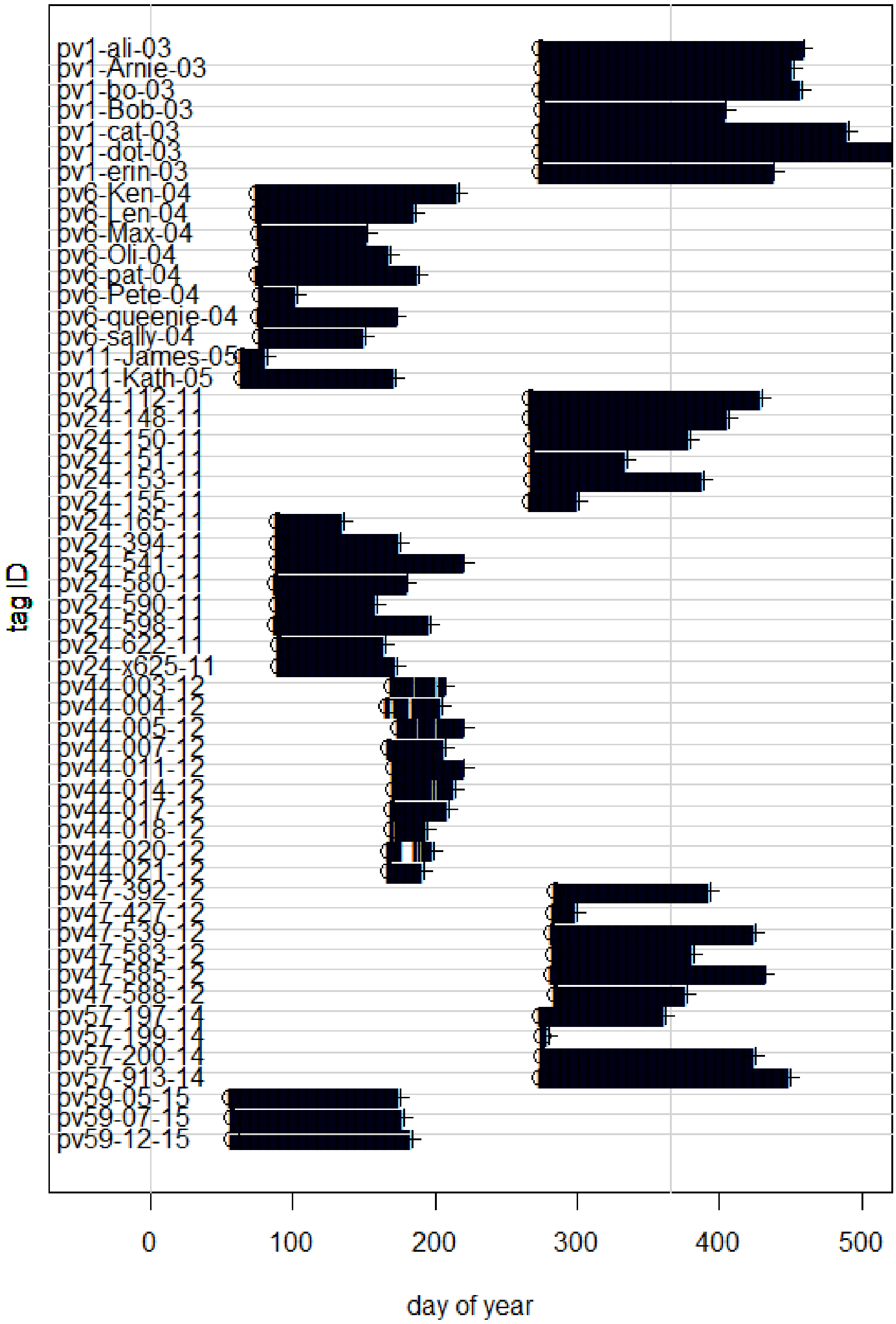


Figure 10. Temporal extent (day of year) of movement data by animal.

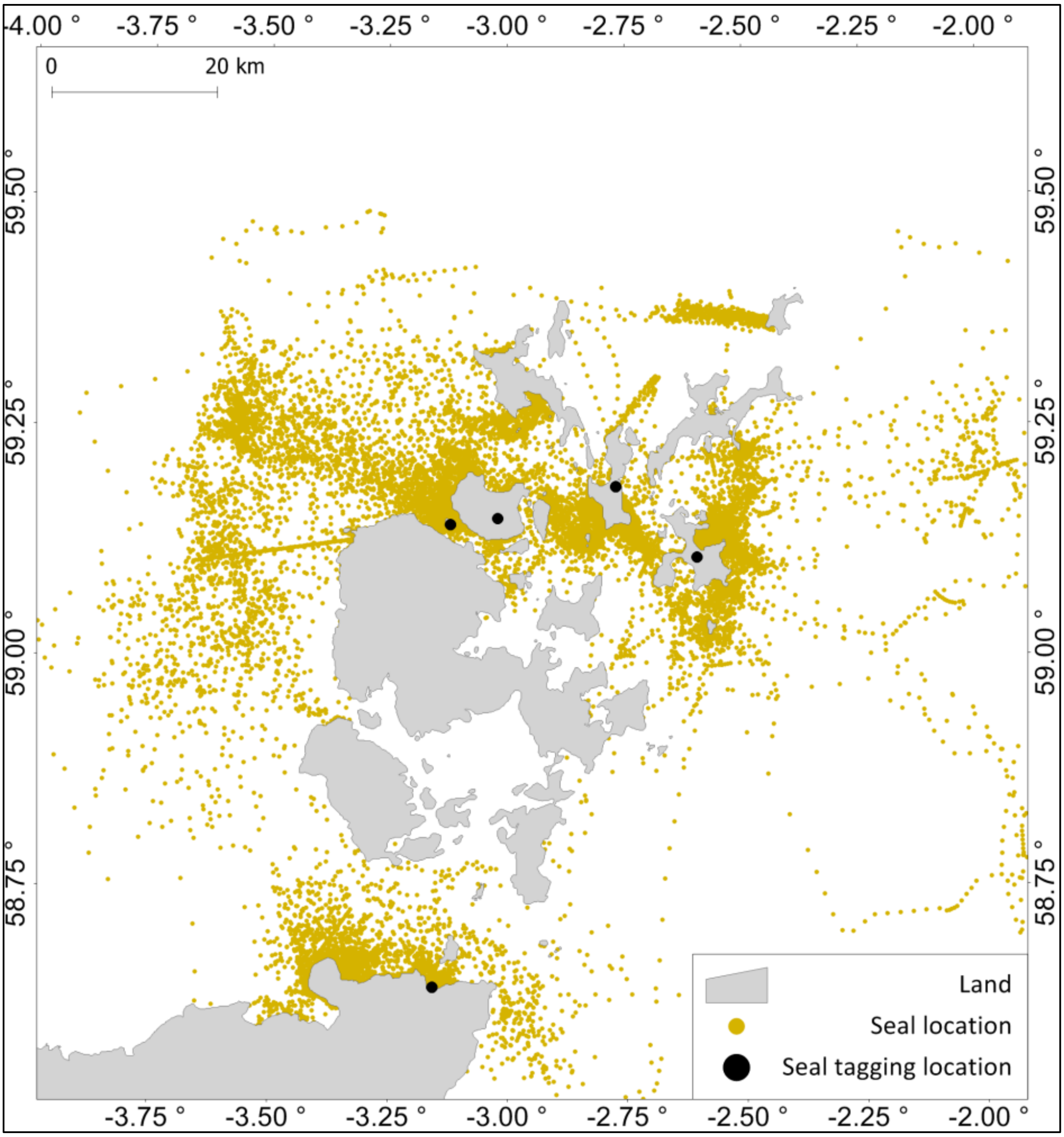


Figure 11. Spatial distribution of movement data available from 31 animals in spring/summer.

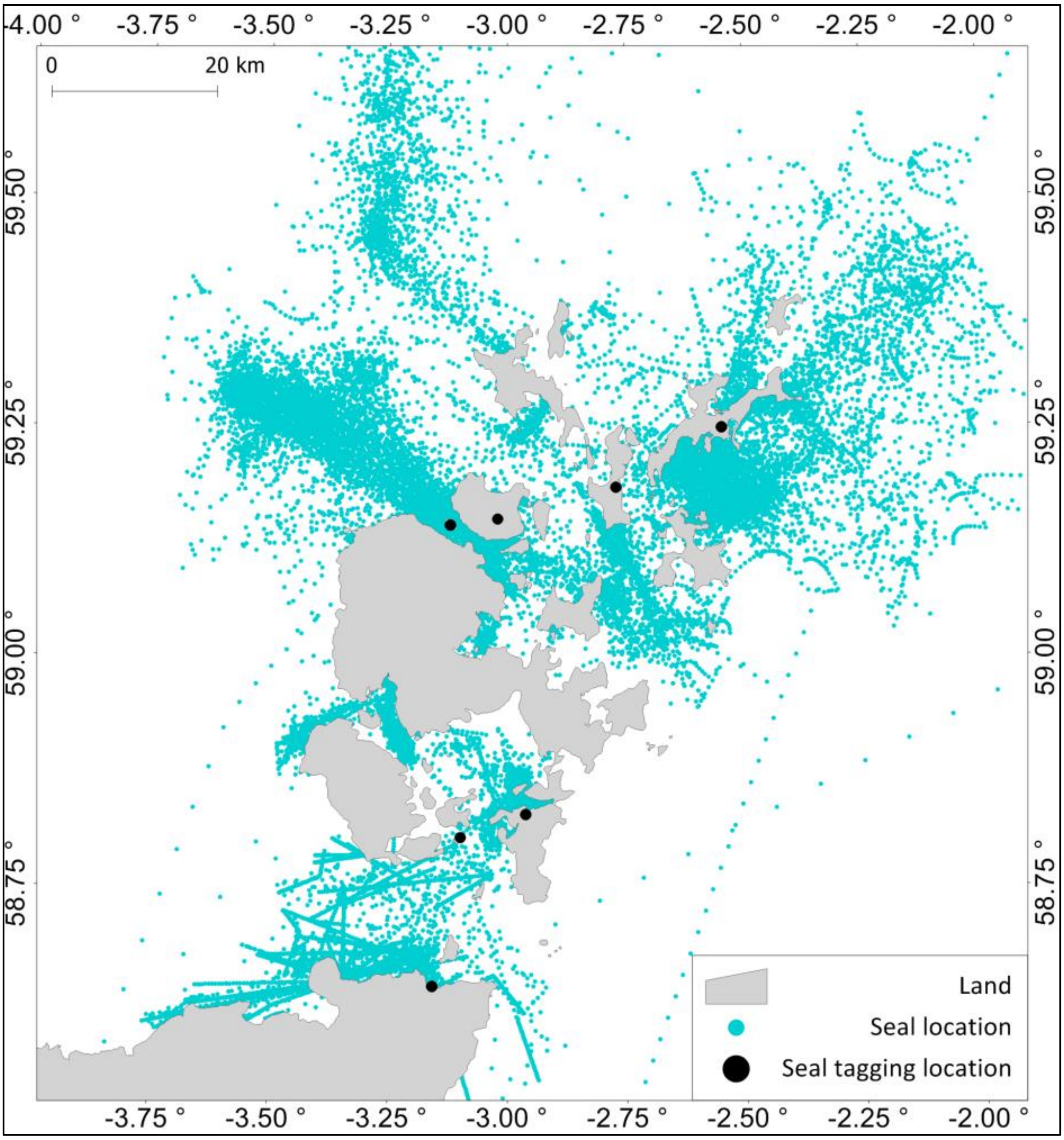


Figure 12. Spatial distribution of movement data available from 23 animals in autumn/winter.

References

Aarts, G., Mackenzie, M.L., McConnell, B.J., Fedak, M. & Matthiopoulos, J. (2008). Estimating space-use and habitat preference from wildlife telemetry data. *Ecography*, **31**, 140–160.

Argos. (2011). Argos User's Manual 2007-2011. *CLS*.Barbet-Massin, M., Jiguet, F., Albert, C.H. & Thuiller, W. (2012). Selecting pseudo-absences for species distribution models: How, where and how many? *Methods in Ecology and Evolution*, **3**, 327–338.

Duck, C.D., Morris, C.D. & Thompson, D. (2014). *The status of UK harbour seal populations in 2013, including summer counts of grey seals*. SCOS Briefing Paper 14/03 (report available at http://www.smru.st-andrews.ac.uk/research-policy/scos/).

Dujon, A.M., Lindstrom, R.T. & Hays, G.C. (2014). The accuracy of Fastloc-GPS locations and implications for animal tracking. *Methods in Ecology and Evolution*, **5**, 1162–1169.

Duong, T. (2016). ks: Kernel Smoothing. R package version 1.10.4. https://CRAN.R-project.org/package=ks

Elith, J., Graham, C.H., Anderson, R.P., Dudik, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B. A, Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S., Zimmermann, N.E., Dudík, M., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Li, J., Lohmann, L.G., Loiselle, B. A., Manion, G., Moritz, C., Nakazawa, Y., Overton, J.M., Peterson, A.T., Steven, J., Richardson, K., Scachetti-pereira, R., Schapire, R.E., Soberón, J., Dudik, M., Ferrier, S., Nakamura, M., Townsend, A., Phillips, S.J., Sobern, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

van Etten, J. (2015). R package gdistance: Distances and Routes on Geographical Grids.

Everitt, B.S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster Analysis*, 5th edn. Wiley.

Folk, R.L. (1954). The Distinction between Grain Size and Mineral Composition in Sedimentary-Rock Nomenclature. *The Journal of Geology*, **62**, 344–359.

Fox, J. & Weisberg, S. (2011). *An R Companion to Applied Regression*, 2nd edn. Sage, Thousand Oaks CA.

Højsgaard, S., Halekoh, U. & Yan, J. (2006). The R Package geepack for Generalized Estimating Equations. *Journal of Statistical Software*, **15**, 1–11.

Isaaks, E.H. & Srivastava, R.M. (1990). *Applied Geostatistics*. Oxford University Press.

Jones, E.L., McConnell, B.J., Smout, S., Hammond, P.S., Duck, C.D., Morris, C.D., Thompson, D., Russell, D.J.F., Vincent, C., Cronin, M., Sharples, R.J. & Matthiopoulos, J. (2015). Patterns of space use in sympatric marine colonial predators reveal scales of spatial partitioning. *Marine Ecology Progress Series*, **534**, 235–249.

Manifold Software Limited. (2013). Manifold System Ultimate Edition 8.0.28.0.

McConnell, B.J., Beaton, R., Bryant, E., Hunter, C., Lovell, P. & Hall, A.J. (2004). Phoning home - A new GSM mobile phone telemetry system to collect mark-recapture data. *Marine Mammal Science*, **20**, 274–283.

Price, D., Stuiver, C., Johnson, H., Gallego, A. & O'Hara Murray, R. (2016) The Scottish Shelf Model. Part 2: Pentland Firth and Orkney Waters Sub-Domain. *Scottish Marine and Freshwater Science*, **7**, 248.

Pirotta, E., Matthiopoulos, J., Mackenzie, M.L., Scott-Hayward, L. & Rendell, L. (2011). Modelling sperm whale habitat preference : a novel approach combining transect and follow data. *Marine Ecology Progress Series*, **436**, 257–272.

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Royer, F. & Lutcavage, M. (2008). Filtering and interpreting location errors in satellite telemetry of marine animals. *Journal of Experimental Marine Biology and Ecology*, **359**, 1–10.

Russell, D.J.F., Hastie, G.D., Thompson, D., Janik, V.M., Hammond, P.S., Scott-Hayward, L.A.S., Matthiopoulos, J., Jones, E.L. & McConnell, B.J. (2016). Avoidance of windfarms by harbour seals is limited to pile driving. *Journal of Applied Ecology*.

Sokal, R.R. & Rohlf, F.J. (2012) *Biometry*, 4th ed. W H Freeman.

Thompson, D., Lonergan, M. & Duck, C. (2005). Population dynamics of harbour seals Phoca vitulina in England : monitoring growth and catastrophic declines. *Journal of Applied Ecology*, **42**, 638–648.

Tsoar, A., Allouche, O., Steinitz, O., Rotem, D. & Kadmon, R. (2007). A comparative evaluation of presence-only methods for modelling species distribution. *Diversity and Distributions*, **13**, 397–405.

Wiens, T.S., Dale, B.C., Boyce, M.S. & Kershaw, G.P. (2008). Three way k-fold cross-validation of resource selection functions. *Ecological Modelling*, **212**, 244–255.

Wisz, M.S. & Guisan, A. (2009). Do pseudo-absence selection strategies influence species distribution models and their predictions? An information-theoretic approach based on simulated data. *BMC Ecology*, **9**, 8.