Harbour porpoise habitat preferences: robust spatio-temporal inferences from opportunistic data

Saana Isojunno[1],*, **Jason Matthiopoulos[1,2], Peter G. H. Evans[3]**

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

[2]Centre for Research into Environmental and Ecological Modelling, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY16 9LZ, UK

[3]Sea Watch Foundation, Ewyn y Don, Bull Bay, Amlwch, Anglesey, Wales LL68 9SD, UK

ABSTRACT: Statistical habitat modelling is often flagged as a cost-effective decision tool for species management. However, data that can produce predictions with the desired precision are difficult to collect, especially for species with spatially extensive and dynamic distributions. Data from platforms of opportunity could be used to complement or help design dedicated surveys, but robust inference from such data is challenging. Furthermore, regression models using static covariates may not be sufficient for animals whose habitat preferences change dynamically with season, environmental conditions or foraging strategy. More flexible models introduce difficulties in selecting parsimonious models. We implemented a robust model-averaging framework to dynamically predict harbour porpoise *Phocoena phocoena* occurrence in a strongly tidal and topographically complex site in southwest Wales using data from a temporally intensive platform of opportunity. Spatial and temporal environmental variables were allowed to interact in a generalized additive model (GAM). We used information criteria to examine an extensive set of 3003 models and average predictions from the best 33. In the best model, 3 main effects and 2 tensor-product interactions explained 46% of the deviance. Model-averaged predictions indicated that harbour porpoises avoided or selected steeper slopes depending on the tidal flow conditions; when the tide started to ebb, occurrence was predicted to increase 3-fold at steeper slopes.

KEY WORDS: Generalized additive models · *Phocoena phocoena* · Habitat model · Non-linear interactions · Multi-model inference · Model selection · Tidal environments · Wales

INTRODUCTION

Platform-of-opportunity data, collected aboard industrial platforms (oil and gas, renewable energies, seismic exploration), commercial fishing vessels (López et al. 2004), passenger ferries (Williams et al. 2006, Kiszka et al. 2007, MacLeod et al. 2008) and by wildlife tour operators (Hauser et al. 2006, Ingram et al. 2007) are increasingly used in monitoring and for mitigation schemes of marine mammals (Evans & Hammond 2004). To date, such data remain underutilized because their analysis presents several challenges. Opportunistic sampling is not spatially or temporally randomized and may be restricted to a particular route, certain time of the day or phase of tide. This can lead to sparse response data and generate apparent multicollinearities between the corresponding environmental data.

When the track line is predetermined by a commercial design (e.g. fixed shipping routes), the distribution of opportunistic effort may be assumed to be independent of the target species. Relating animal sightings to effort can then be used directly to infer relative distributions (e.g. Kiszka et al. 2007, MacLeod et al. 2008)

*Email: saana.isojunno@gmail.com

and absolute densities within the confines of the area surveyed (Williams et al. 2006). Other opportunistic platforms, such as wildlife tour operators (examined in the present study), determine the track line *in situ*, often according to prevailing environmental conditions (e.g. López et al. 2004, Hauser et al. 2006, Ingram et al. 2007). Such platforms can collect more accurate visual and behavioural data by diverting from the track line to confirm species, gender or group size (Evans & Hammond 2004), but accounting for the response of the platform to the study animals and vice versa can be challenging. Wildlife tour operators in particular may change their course and speed to increase the time spent in the vicinity of the animals during each trip or to visit locations where they believe they have a high probability of finding animals. Thus, the rate of occurrences of the species is not necessarily proportional to observation effort.

Habitat and species distribution models often take the form of statistical regression (Guisan & Zimmermann 2000, Redfern et al. 2006, Aarts et al. 2008, Matthiopoulos & Aarts 2010) that relate species density or occurrence to available static environmental covariates, e.g. sea bottom depth. Often data on prey density is lacking and the models rely on more indirect explanatory variables. Distribution estimates can be highly uncertain for species that live in dynamic environments and have complex behaviour and life histories, especially at finer spatial and temporal scales. Few species have been studied in sufficient detail to quantify the cause and effect of their dynamic responses to environmental change (Guisan & Zimmermann 2000). Collecting such data can be time consuming, while species management and protection requires rigorous estimates of distribution in a timely fashion.

The harbour porpoise *Phocoena phocoena* Linnaeus 1758 is a prime example of a species for which predictive power of distribution and habitat preference has been difficult to obtain across large spatial scales (Embling 2008, Marubini et al. 2009, Embling et al. 2010) because both its environment and its responses to the environment are highly variable. For example, although harbour porpoises are known to eat a wide variety of prey across their entire range, their diet tends to be dominated by only 2 to 4 main prey species in any given area (Santos & Pierce 2003, Santos et al. 2004). There are consistent reports of the associations of harbour porpoises with dynamic, high productivity zones such as upwellings or high energy coastal locations (see for example Watts & Gaskin 1985, Borges & Evans 1997, Weir & O'Brien 2000, Johnston et al. 2005, Tynan et al. 2005, Pierpoint

2008). Such ephemeral aggregations of prey may be particularly important for the harbour porpoise owing to its small body reserves, cold living environment and energetically demanding reproductive schedule (Read & Westgate 1997, Reed et al. 2000, Read 2001, Lockyer 2003, Johnston et al. 2005, Evans et al. 2008).

Preference for short-lived, yet predictable, oceanographic features may go unnoticed in large-scale (typically both extent and resolution are large) surveys that visit a given area only briefly. Such surveys provide a 'snapshot' of the distribution, whereby it is appropriate to average spatial conditions in an area and assume pseudo-equilibrium between the species and its environment (Guisan & Thuiller 2005). Indeed, different study scales may reveal species–environment relationships at different levels in a resource that is structured hierarchically in space, where small, high density patches are nested within large, low density patches (Wiens 1989, Embling 2008, Marubini et al. 2009). Fixed-position land observations have documented that harbour porpoises swim against the prevailing current in association with foraging sea birds at certain tidal states (Evans & Borges 1995, Pierpoint 2008). In salmonid fishes the trade-off between drifting food availability and swimming cost results in a 'sliding spatial niche' across variable flow conditions (Heggenes 1996). To understand whether this also occurs in the harbour porpoise, more temporally intensive distribution surveys (of relatively long duration and frequent sampling periods) are needed at a higher spatial resolution.

We took an empirical approach to explore the extent of dynamic changes in the habitat preferences of the harbour porpoise. We implemented a regression framework for platform-of-opportunity data and used it to dynamically predict porpoise occurrence at a strongly tidal and topographically complex site in southwest Wales. Visual sightings and effort data were collected aboard a commercial passenger boat and restructured to account for the manoevers of the vessel in response to the sightings of harbour porpoises.

MATERIALS AND METHODS

Data collection

Data were collected aboard the Pembrokeshire Dive Charter rigid-inflatable boat (RIB) from 19 July to 20 September 2007. Observations were made by one primary observer (the same person throughout the study period) and one secondary observer (1 of

2 skippers, alternating). In each trip, the vessel approached Skomer Island (51°44′ N, 5°18′ W) from the south side and proceeded around the island on either the west or east side.

Sea bottom depth drops steeply from 20 to 50 m on the west of Skomer Island (area, 2 × 4 km, Fig. 1) but the seabed southeast of the islands is flatter and shallower (~30 m). The island forms 2 narrow sounds in the east, Little Sound and Jack Sound, in-between another smaller island and mainland Pembrokeshire. The sounds connect waters south of Skomer Island to St. Brides Bay in the north, with another steep slope just north of Jack Sound (10 to 40 m depth). Water starts flowing north 2 h before high tide and south 2 h before low tide. Tidal speeds frequently reach over 6 knots in the 2 sounds and the northern tip of the island.

The island was divided into geographically distinct sections of inshore areas (within ~1 km of the shore) that could be identified from the boat (Fig. 1). Effort was recorded every time a section was entered. Each visit to a section, termed 'effort interval', varied in

duration. The boat manoeuvred (0 to 8 knots) within these sections with a variable course. For every effort interval within a section, the time of day, Beaufort sea state, swell height, visibility and boat activity were recorded.

Because of the slow speed and variable course of the boat, the same individuals could be observed repeatedly. Instead of attempting to keep track of individuals indefinitely, we assumed that porpoises were stationary enough to stay in the geographical section for the duration of the effort interval. Individuals were counted within the section so that we only recorded those individuals as separate sightings that added to the existing count of total animals in the effort interval. Once we entered a new section, i.e. new effort interval started, the count could start again assuming new animals. Because the passenger boat often stayed with the animals, we could observe animals at close ranges (<150 m) and for long durations (>10 min). *In situ* counting of individuals was also made easier because of variation in individual size and because, atypically for the spe-
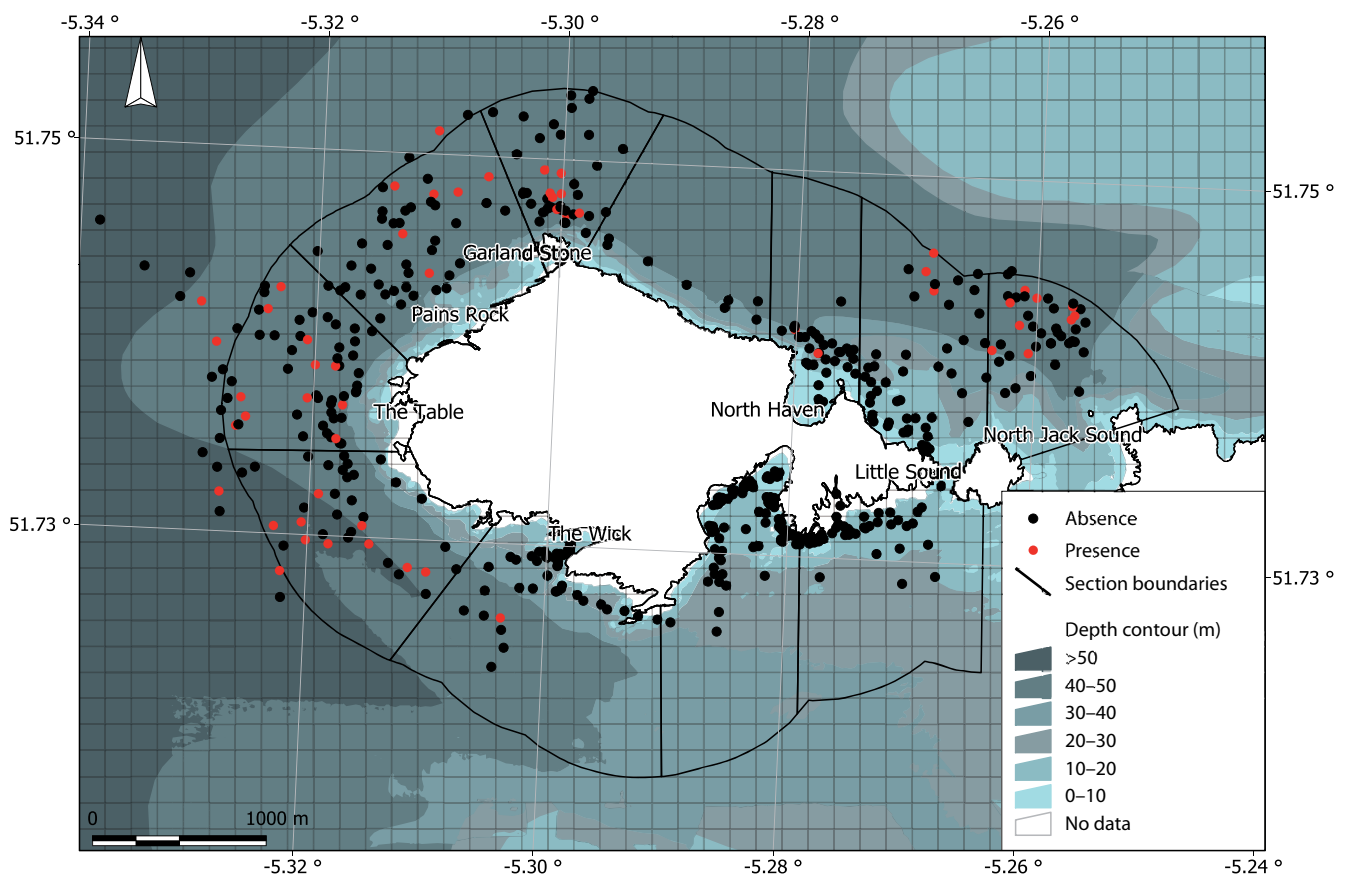

Fig. 1. Skomer Island survey sections, overlaid with a grid of 150 × 150 m cells, analysed data set of visits to grid cells and the integrated surface of depth values. Red points represent the grid cell visits that included a harbour porpoise sighting and black points represent visits with absence of sightings. Data points outside the section boundaries were assigned to the closest section

cies, some individuals carried distinct white colour patterns on their backs and flanks.

GPS data were available on vessel position, speed and bearing for every survey minute. For each sighting we also recorded time of day, geographical position, bearing, estimated sighting distance, total number of adults and calves, animal behaviour, animal heading and bird and ocean current associations. Although the detection distance could be more than 200 to 300 m, sightings were not logged on GPS until the animals were approached to within ~150 m.

Environmental data used in the analysis

No prey data were available at the required temporal scale within the study area. Indirect variables may provide even better predictors than prey density alone at small spatial scales and in complex topography, especially in environments that constrain the predator physiologically or energetically, such as swimming in strong currents (Guisan & Zimmermann 2000, Torres et al. 2008). Indirect exploratory variables were chosen according to data availability and relevance to the oceanographic conditions that potentially affect both prey distribution and porpoise swimming cost. Unfortunately, no data encompassing the study area were available on current strength.

Bathymetry

The Countryside Council for Wales provided a bathymetric surface from a multibeam survey (pixel size, 5 × 5 m) (Longdin & Browning Ltd 2002). Missing depth values outside the range of the multibeam survey were obtained from bathymetric elevation and aspect surfaces (pixel size, 120 × 120 m) (n = 57, 11% of all analysed data of visits to grid cells) obtained as ASCII grids via EDINA Marine Digimap Service (©Crown Copyright / SeaZone Solutions Ltd 2008) (Fig. 1). Slope (0 to 90°) and aspect (i.e. the azimuthal direction in which a tangent plane faces, −180 to 180°) were calculated from a 3 × 3 pixel window from this surface (Slope and Aspect functions in Manifold 8.0, CDA International).

Tide

Measured tidal heights (accurate to 0.1 cm) were obtained every 15 min from the Milford Haven tide gauge (51° 42′ N, 5° 3′ W) courtesy of the British Oceanographic Data Centre. The rate of change in tide was calculated as $-\sin(2\pi m_i / \Delta t_i) \times \Delta h_i$, where m_i is the number of minutes from the previous high tide, Δt_i is the time difference (in minutes) between the high and low tide, and Δh_i is the difference in tidal height of high and low tides in the tidal cycle i. The index can be thought of as a circular transformation of the minutes from previous high tide weighted by the range of tidal heights in that cycle.

Weather data

Hourly sea surface temperature (SST) readings (accurate to 0.1°C), wind direction (degrees, °), wind speed (knots) and sea level pressure (hPa) from the Turbot bank buoy (51° 36′ N, 5° 6′ W) were provided by the Met Office Marine Automatic Weather Station network. Sea state was recorded *in situ* with the standard Beaufort scale. The swell was visually approximated to 0.5 m resolution.

Restructuring the opportunistic data

Sections were used to partition the study area in real time (Fig. 1). For finer spatial analysis, the effort interval was not used as the unit of porpoise occurrence. Instead, a grid of 150 × 150 m squares was overlaid on the area (Fig. 2) and GPS data within the effort intervals were used to estimate observation effort and presence of a sighting for each grid cell. Each visited grid cell was marked as presence or absence only once during a visit to a section. Grid cells with one or more sightings were marked as presence, and all other visited grid cells as absence. The time stamp for each grid cell visit was taken as the first visit to that grid cell. Using effort interval as the unit of data restructuring ensured that there were no duplicate presence or absence data within an effort interval and that increased duration or number of visits to a grid cell with sightings could not introduce bias in the analysis.

Presence or absence was considered instead of the full abundance data to reduce biases in the opportunistic platform. The boat operator responded to porpoise sightings by changing the course and speed of the vessel. The vessel was more likely to approach and manoeuvre in the vicinity of a greater number of animals foraging rather than travelling, when the sea state was calmer, and earlier rather than later during a trip. The vessel approached feeding gannets *Morus bassanus* because harbour porpoises were
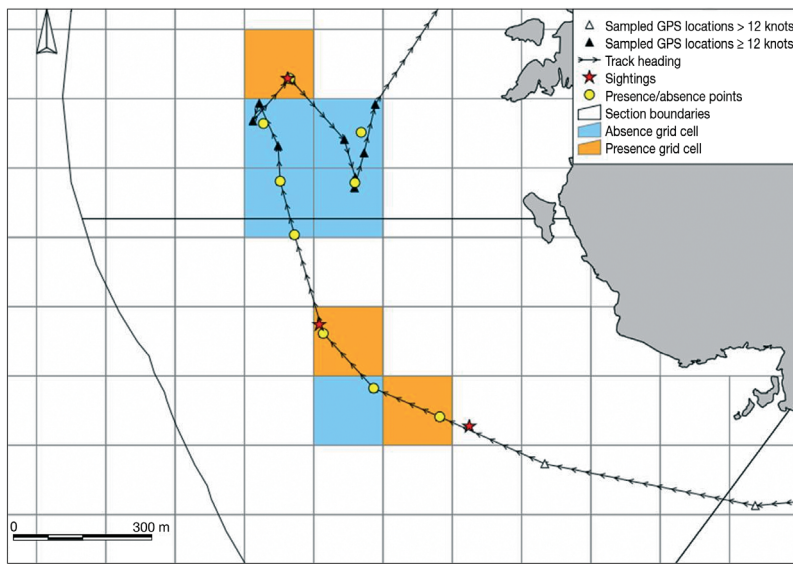
Fig. 2. Restructuring the opportunistic data. Vessel track locations (sampled every minute) are shown as triangles and sighting locations of harbour porpoises *Phocoena phocoena* as stars. All track data collected above vessel speed of 12 knots were discarded. Sightings were first appended to the track locations to give presence/absence per minute of track line. The track locations were then grouped by grid cells in each visit to a section so that the location of the new presence/absence point was taken as the average track location. Therefore, in a visit to the section each grid cell was visited only once in the restructured data. Presence was defined as a visited grid cell with at least one presence (shown in orange). Sections were used as practical boundaries (see 'Materials and methods: data collection') within which individual harbour porpoises were counted; therefore, a visit to the same grid cell in different sections could result in a double absence (in the picture), but not a double presence

Table 1. Candidate covariates for models to predict occurrence of harbour porpoise in southwest Wales

Variable	Description
Depth	Average depth of the visited grid cell, meters (m)
Slope	Slope of the sea bottom in the range of 0 to 90° calculated from a pixel window of 3 × 3 (slope function in Manifold)
Aspect	Aspect of the sea bottom in the range of −180 to 180° (aspect function in Manifold)
Height	Height of the tide, meters (m)
tideFlow	Rate of change of the tidal height in meters (m)
Sunrise	Minutes from local time of sunrise
Day	Number of days from the beginning of the year
SST	Sea surface temperature in tenths of degrees Celsius (°C)
DD	Wind direction, degrees (°)
FF	Wind speed, knots
MSLP	Sea level pressure in tenths of hectopascals (hPa)
SS	Beaufort state of the sea surface, induced by the local winds (7 categories)
Swell	Swell that is not induced by the local winds

often found to forage in association with them. Thus, search effort for porpoises increased after their detection, reducing the degrees of freedom in the relative abundance data. Restructuring the data re-

moved the effect of increased effort at a grid cell location after detection but not the availability bias of foraging versus travelling porpoises.

Overview of modelling

The occurrence of a harbour porpoise sighting in each on-effort grid cell was modelled as a generalized additive model (GAM) with a binary response within the mgcv (1.4-0) library in R (Wood 2004, 2006a, 2008). We allowed for non-linear interactions in the model structure in order to capture any changes in preferences for one covariate as a function of another covariate. We first modelled occurrence with no interactions and then compared the best model with models that included interaction terms. Because of a large number of possible combinations of univariate and interaction covariates, there was a large uncertainty in which model structure to choose. This uncertainty was addressed by weighted averaging of the best models.

Model without interaction terms

Models were fitted with univariate smooths of each candidate variable (Table 1). The duration of the grid cell visit was not included as an offset because the vessel could respond to the presence of animals by staying longer in the grid cell. It was assumed that existing animals would be detected immediately after the boat moved into a new grid cell and those animals would not enter an empty cell after the boat. This assumption would be violated if porpoises avoided the vessel. We believe avoidance effects were comparatively small in this opportunistic survey. Harbour porpoises appeared habituated to the wildlife watching RIBs (rigid-hulled inflatable boats) that visit the survey area up to 4 times a day during the summer. Most times the engine of the vessel was turned off in the vicinity of the animals and their behaviour could be

observed without disruption for extended periods of time (varying from 10 to 60 min). The number of GPS locations logged per cell ranged from 1 to 24, of which 72.5% contained only one GPS location. If the vessel followed a straight course at its average speed (4 knots) during single visits, it would take a minimum time of 70 s to cross a 150 m grid cell. Because all of the 55 first detections were made within the first minute of the visit to the cell, we could assume that the likelihood of presence did not increase with visit duration.

In the mgcv library, the degree of smoothness (or inversely, the complexity or wiggliness of the non-linear relationship) of model terms is estimated as part of model fitting. The trade-off between fit and smoothness is controlled by the smoothing parameter λ. Thus, models with different degrees of freedom are not nested and step-wise model selection is inappropriate. Instead, Wood (2006b) recommends shrinkage smoothers as an alternative approach to automated model selection. These incorporate a shrinkage component into the smoothing penalties that forces all the parameters of the smooth to zero when the model finds no signal in the data. Cyclic cubic regression splines (Wood 2006a) were fitted to cyclic covariates, such as wind direction and minutes from high tide, to ensure continuity of the response at the end points of the covariate's range.

In binary models, the unbiased risk estimator (UBRE) can be minimized to estimate the smoothing parameter λ and compare models (Wood 2006a). A probit model was chosen over other link functions because it explained more of the deviance and appeared to retain smaller UBRE scores regardless of whether the model included interaction terms. To assess how well the model fitted the data, the explained deviance for binary data (Nagelkerke 1991, Faraway 2006) was used.

Model with interaction terms

In a GAM, smooths of several covariates allow for non-linear interaction terms. We used tensor products, which can be constructed of any type of low rank smooths to represent interactions of covariates that have been measured at different scales or quantities, for example tidal height versus wind direction (Wood 2006a,b).

Automatic selection of interaction terms was necessary because there was little previous knowledge of their relative importance. Since the data set was rather small, only 1 or 2 products of second-order

(2-dimensional) interactions were considered. Latitude and longitude cannot have a direct causal relationship with the underlying process of habitat selection, but they may correlate spatially with environmental variables that do. We wanted to avoid the possibility of latitude and longitude competing in model selection with causal covariates and we therefore did not include them in model selection.

We explored all 3003 models that included all the main effects and up to 2 unique pairs of tensor products. Tensors were allowed to include any of the candidate main effects. Shrinkage smoothers were used for automated model selection of each full model. Therefore, no separate model runs were necessary for models excluding each of the main effects or either of the tensors.

Restructuring of the data did not remove the possibility of observing the same animals across sections, which might have resulted in some degree of serial correlation. An autocorrelation structure could not be included in the model with the shrinkage smoothers; autocorrelation was therefore checked after model selection. The best main effects model and interaction model were fitted with a continuous first-order autoregressive structure (AR1, lag = 1) with trip as the grouping factor. This autocorrelation structure did not improve the main effects model, and the estimated value for the autocorrelation parameter (φ) was small and negative (−0.034) for the best interaction model. The best interaction model was therefore interpreted without an autocorrelation structure.

Average predictions from a confidence set of models

In situations, such as ours, where multiple models are being considered and a large number (*R*) of them are similarly supported by the data, Burnham & Anderson (2002) recommend Akaike weights as a relative measure of model support. The Akaike weight for the *i*th model is essentially its Akaike information criterion (AIC), normalized by all the AICs from the confidence set of *R* models:

$$
w_i = \frac{\exp(-\tfrac{1}{2}\Delta\mathrm{AIC}_i)}{\sum_{r=1}^{R}\exp(-\tfrac{1}{2}\Delta\mathrm{AIC}_r)} \tag{1}
$$

where the likelihood of model g_i is proportional to $\exp(-\tfrac{1}{2}\Delta\mathrm{AIC}_i)$ and $\Delta\mathrm{AIC}_i$ and $\Delta\mathrm{AIC}_r$ are the differ-

ence between the AIC of the best model and the ith and rth model structure, respectively.

To select the confidence set of models, Burnham & Anderson (2002) suggested a cut-off value of 0.05 for the evidence ratios. Evidence ratios compare the likelihood of the model g_i, given the data, to the likelihood of the best model:

$$\frac{L(g_i \mid x)}{L(g_{\min} \mid x)} > \text{cut-off} \qquad (2)$$

Because of the large number of explored models, we took a more inclusive cut-off value of 0.00028 that was determined visually as the point where the cumulative sum of the Akaike weights, in decreasing order, stabilized. We also included 2 models whose UBRE scores were lower than the lowest UBRE score in the set, but whose Akaike weights were less than this cut-off value (Models 2 and 24 in Table 3).

Model predictions were averaged across the confidence set weighted by the Akaike weights w_i (Burnham & Anderson 2002) as

$$\hat{\bar{\theta}} = \sum_{i=1}^{R} w_i \hat{\theta} \qquad (3)$$

where $\hat{\theta}$ is the estimated probability of occurrence. Confidence intervals and coefficients of variation (CV) for the averaged estimates of occurrence were obtained by using an unconditional variance estimator derived by Burnham & Anderson (2002). 95% CIs were constructed assuming a t-multiplier:

$$95\% \text{ CI} = \hat{\bar{\theta}} \pm t(0.025, R)\sqrt{Var(\hat{\bar{\theta}}) / R} \qquad (4)$$

RESULTS

A total of 540 grid cell visits were used in the analyses. Of these, 55 had sightings (Fig. 1).

Model without interaction terms

The best model without interaction terms was used as a baseline for comparison with the single best interaction model and the model-averaged predictions. The model retained depth, minutes from sunrise and SST (Fig. 3, Table 2) and explained 33.38% of the deviance. Harbour porpoise occurrence on the water surface increased in deeper waters and in the earlier hours of daylight (Fig. 3a,b,d). Porpoises were seen more in the middle of the observed range of SSTs (Fig. 3c).

The single best model with interaction terms

Out of 3003 possible models, 52 did not converge, probably because of their poor ability to fit the data, and were therefore excluded from further analyses. The best interaction model with the lowest UBRE score (−0.501) retained depth and minutes from sunrise as main effects, and 2 tensor products: slope–tidal flow and tidal height–wind speed (Model 1; Tables 2 & 3). This model had an AIC weight of 0.112 (Eq. 1) and accounted for 45.25% of the deviance.

As with the model without interaction terms, the occurrence of harbour porpoises increased in deeper waters and in the earlier hours of daylight (Fig. 4a,b). The probability tended to be lower when tides were flooding, particularly in regions of steep sea bed (Fig. 4c,d). The effect of stronger winds depended on the level of tidal height: a negative effect occurred at medium-high tidal heights, and a positive effect occurred at low tidal height (Fig. 4e,f). Overall, the interaction terms improved the UBRE scores (Table 2).

Average predictions from a confidence set of models

We chose a set of 33 models with evidence ratios ranging from only 0.00028 to 0.194 with respect to the best model (see corresponding section in 'Materials and methods'; Table 3). Together the 33 models accounted for 23.9% of all Akaike weights.

The relative importance of each covariate was measured by summing the Akaike weights (Eq. 3) of all the models in which the covariate was retained by shrinkage (Burnham & Anderson 2002) (Table 4). The 4 most important covariates were those included in the single best interaction model: depth, minutes from sunrise, and the 2 tensor products, slope–tidal flow and tidal height–wind speed. Both depth and minutes from sunrise were equally retained (in 75.8% of the models each) and important (relative importance for depth was 0.85, and 0.82 for minutes from sunrise). The 2 tensor products, by contrast, were retained only in a minority of the cases (39.4% slope–tidal flow and 9.1% tidal height–wind speed) but accounted for a large amount of the Akaike weights (the relative variable importances were 0.79 and 0.49, respectively). Other important covariates were SST, tidal height alone and tidal height–minutes from high tide interaction. Those covariates were present in more mod-
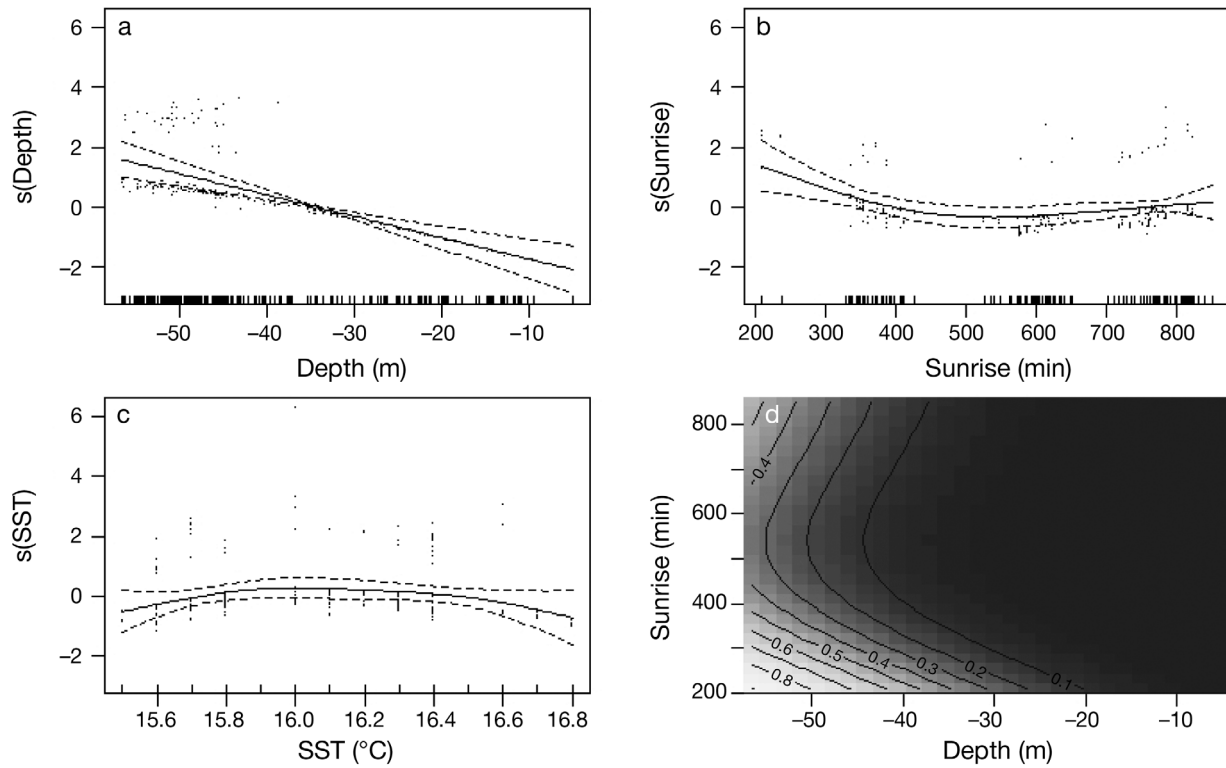
Fig. 3. Component smooth functions of (a) Depth, (b) Sunrise and (c) sea surface temperature (SST) on the scale of the linear predictor. Solid lines show the smooth estimate, and dashed lines the estimate ±2 SE. (d) 2-dimensional fitted values on the response scale for minutes from sunrise (Sunrise) and Depth in the best main effects model

Table 2. Summary of the best main effects and interaction models to predict occurrence of harbour porpoise in southwest Wales. AIC: Akaike's information criterion; UBRE: unbiased risk estimator; s: smooth function; te: tensor product. See Table 1 for description of variables

Model	Main effects model	The best interaction model
Formula	s(Depth) + s(Sunrise) + s(SST)	s(Depth) + s(Sunrise) + te(Slope, tideFlow) + te(Height, FF)
Estimated degrees of freedom for each covariate	1.00, 2.68, 2.83	1.03, 1.42, 12.22, 45.14
UBRE / AIC	−0.479 / 269.07	−0.501 / 257.36
Explained deviance	33.38%	45.25%
Autocorrelation	−0.012	−0.034
Weight	2951 models	0.0003	0.1121

els (45.5, 24.2 and 33.3%, respectively) than tidal height–wind speed interaction, but accounted for lower Akaike weights (0.21, 0.15 and 0.12, respectively) (Table 4).

Averaged predicted occurrence (Eq. 3) was plotted throughout the range of each covariate, given the mean values for all other variables (Fig. 5). Harbour porpoise occurrence was predicted to be more likely at greater depths (range, 0 to 60 m), medium to low slope values, higher tidal heights, stronger ebbing tides or earlier in the day. The occurrence of a por-

poise sighting tended to decrease with flooding tide, but the decrease was more pronounced (up to ¼) at higher slope values. The SE values for this interaction term were higher for extreme (low and high) fitted values. The predictions of the tidal height–wind speed term were more uncertain at higher values of tidal height (Fig. 6d).

Spatial predictions were made from the confidence set of models and averaged across different stages of the tidal cycle. To predict from realistic values of tidal height and tidal flow, their observed values were

Table 3. Confidence set of models used to average predictions for the occurrence of harbour porpoise in southwest Wales. AIC: Akaike's information criterion; UBRE: unbiased risk estimator

Model	UBRE	ΔAIC	exp(−0.5 × ΔAIC)	Weights | 2951 models	Weights | 33 models	Evidence ratios
1	−0.501	0.0	1.000	0.11207	0.46952	1.00000
2	−0.496	15.4	0.000	0.00005	0.00022	0.19414
3	−0.495	3.3	0.194	0.02176	0.09115	0.09139
4	−0.492	4.8	0.091	0.01024	0.04291	0.08527
5	−0.492	6.7	0.035	0.00398	0.01666	0.07543
6	−0.492	4.9	0.085	0.00956	0.04004	0.06527
7	−0.492	7.0	0.031	0.00343	0.01435	0.06372
8	−0.491	5.2	0.075	0.00845	0.03542	0.05239
9	−0.491	5.5	0.065	0.00731	0.03064	0.04905
10	−0.491	5.5	0.064	0.00714	0.02992	0.04096
11	−0.491	7.8	0.020	0.00230	0.00962	0.03915
12	−0.490	8.6	0.013	0.00150	0.00629	0.03680
13	−0.490	8.1	0.018	0.00197	0.00827	0.03548
14	−0.490	12.0	0.002	0.00028	0.00116	0.03401
15	−0.490	5.9	0.052	0.00587	0.02460	0.03057
16	−0.490	6.0	0.049	0.00550	0.02303	0.02461
17	−0.490	8.6	0.014	0.00154	0.00644	0.02413
18	−0.490	7.8	0.020	0.00225	0.00942	0.02049
19	−0.489	11.8	0.003	0.00031	0.00131	0.02006
20	−0.489	8.6	0.013	0.00150	0.00629	0.01803
21	−0.489	6.4	0.041	0.00459	0.01923	0.01762
22	−0.489	6.5	0.039	0.00439	0.01838	0.01450
23	−0.488	6.6	0.037	0.00412	0.01728	0.01447
24	−0.488	16.4	0.000	0.00003	0.00013	0.01371
25	−0.488	6.8	0.034	0.00381	0.01597	0.01339
26	−0.488	8.5	0.015	0.00163	0.00681	0.01339
27	−0.488	7.4	0.024	0.00270	0.01133	0.01338
28	−0.487	8.5	0.014	0.00162	0.00679	0.01310
29	−0.487	7.4	0.025	0.00276	0.01156	0.00933
30	−0.486	8.6	0.013	0.00150	0.00628	0.00279
31	−0.486	8.7	0.013	0.00147	0.00615	0.00246
32	−0.486	9.3	0.009	0.00105	0.00438	0.00046
33	−0.486	8.0	0.018	0.00202	0.00846	0.00028

averaged over 2 h intervals of both the neap and spring tidal cycle (no direct measurements of current speeds were available) (Figs S1 to S12 in the Supplement at www.int-res.com/articles/suppl/m448 p155_supp.pdf). Harbour porpoises were most likely to be encountered at the northwest slope of the island along the 50 m bathymetric contour during the turn of high tide (up to 2 h after high tide, Fig. 7). During spring tides, the probability of occurrence was higher and extended over a longer period of time, 1 h before and 2 h after high tide. An increase in occurrence was predicted in deeper waters before high tide and over a steeper sloping seabed after high tide. No reliable predictions could be made in deeper waters at low tide or a turning tide, where CV values exceeded 4. Uncertainty was at its lowest in shallow, low density areas (CV < 1) (Fig. 7, Figs S1 to S12 in the Supplement).

DISCUSSION

We set out to predict the habitat preferences of a highly mobile species in a dynamic environment using temporally intensive opportunistic data. Our modelling approach acknowledged the rapid changeability of habitat quality and the fact that the measured environmental variables were merely proxies for the underlying drivers of harbour porpoise distribution. Data were restructured to account for biases in opportunistic effort. Functionally flexible models were constructed and explored in an extensive search of the model space. Uncertainty related to the specification of more complex models was addressed by model averaging. We compared this with less flexible (no interaction terms) and less robust (single model) approaches. We showed that the best interaction models outperformed the best main effects model (Tables 2
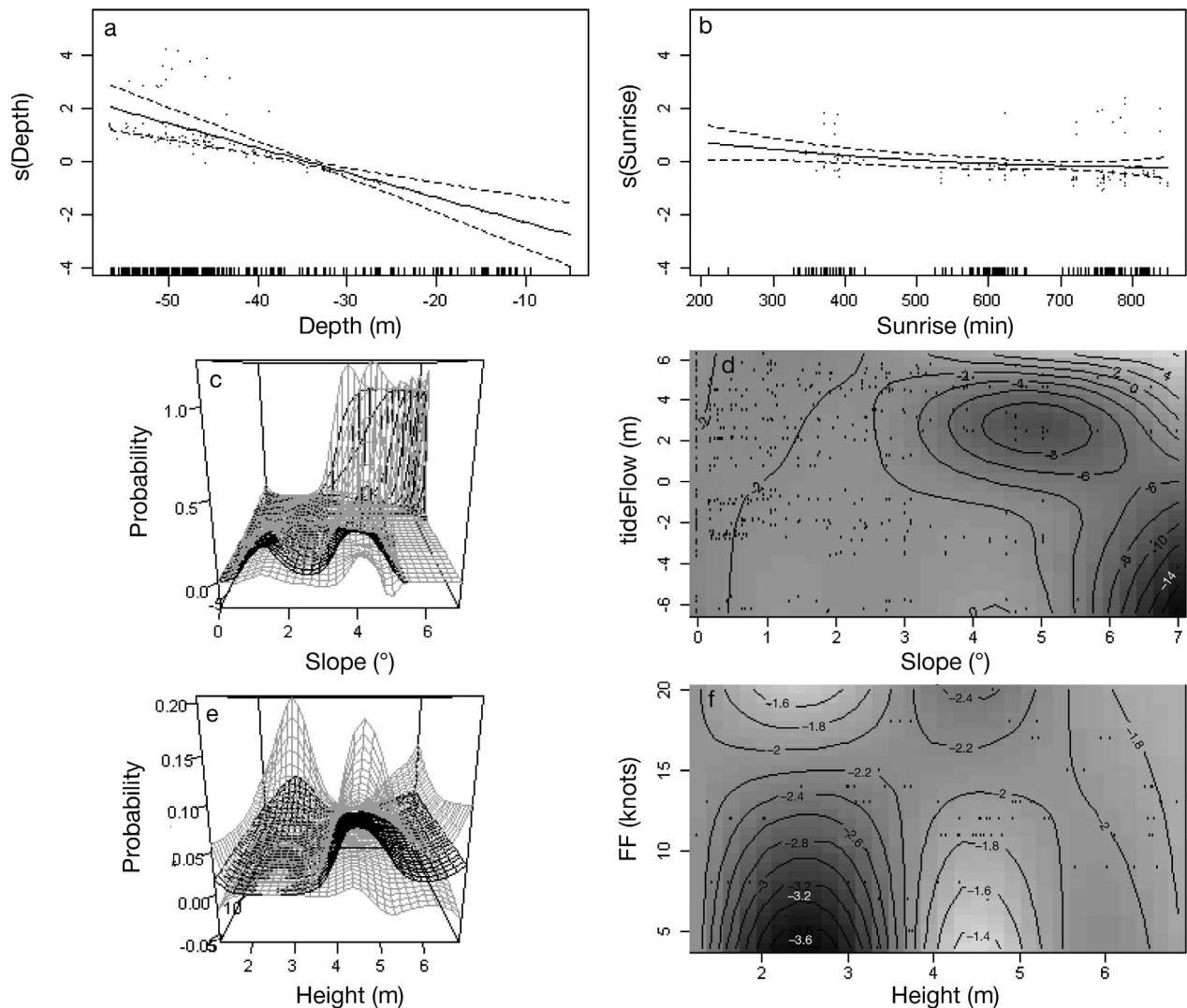
Fig. 4. Component smooth functions of (a) Depth and (b) Sunrise on the scale of the linear predictor, and 2-dimensional fitted values on (c,e) the response scale or (d,f) the link scale for the best interaction model without an autocorrelation structure. Points in (a,b) are the partial residuals, and the points in (d,f) show the data coverage. Dashed lines in (a,b) show the smooth estimate ± 2 SE. Grey surfaces in (c,e) show the fitted values ± SE. Paler shading indicates higher predicted value (d,f). See Table 1 for description of variables

& 3) and that covariates that were able to explain the data but excluded from the best main effects model could be incorporated by model averaging (Table 4).

The covariates with the highest relative importance were depth, minutes from sunrise and interactions between slope–tidal flow and tidal height–wind speed (Table 4). Of the 7 top covariates (Akaike weights higher than 87% of the models), all but depth were dynamic, and 4 included a tidal variable. Given the short study period (3 mo) and limited spatial extent, the importance and shape of response of each variable should be interpreted in the context of fine spatio-temporal scale and limited range of the environmental data. Further, interpreting the predicted

occurrence of harbour porpoises on water surface as 'habitat preference' assumes that a constant proportion of porpoises were detected across all combinations of environmental conditions. Although the models were informed by observation conditions (wind, tide, sea state, swell), the availability of porpoises for observation on the water surface might also vary with their own behavioural state, dive duration and swimming speed (perception bias, Buckland et al. 2001).

Depth appeared to be the single most important predictor in the data. All models in the confidence set included depth (Table 4). Porpoises were more likely to occur in the deeper ranges (>60 m) of the study area, with a 3-fold difference in occurrence evident

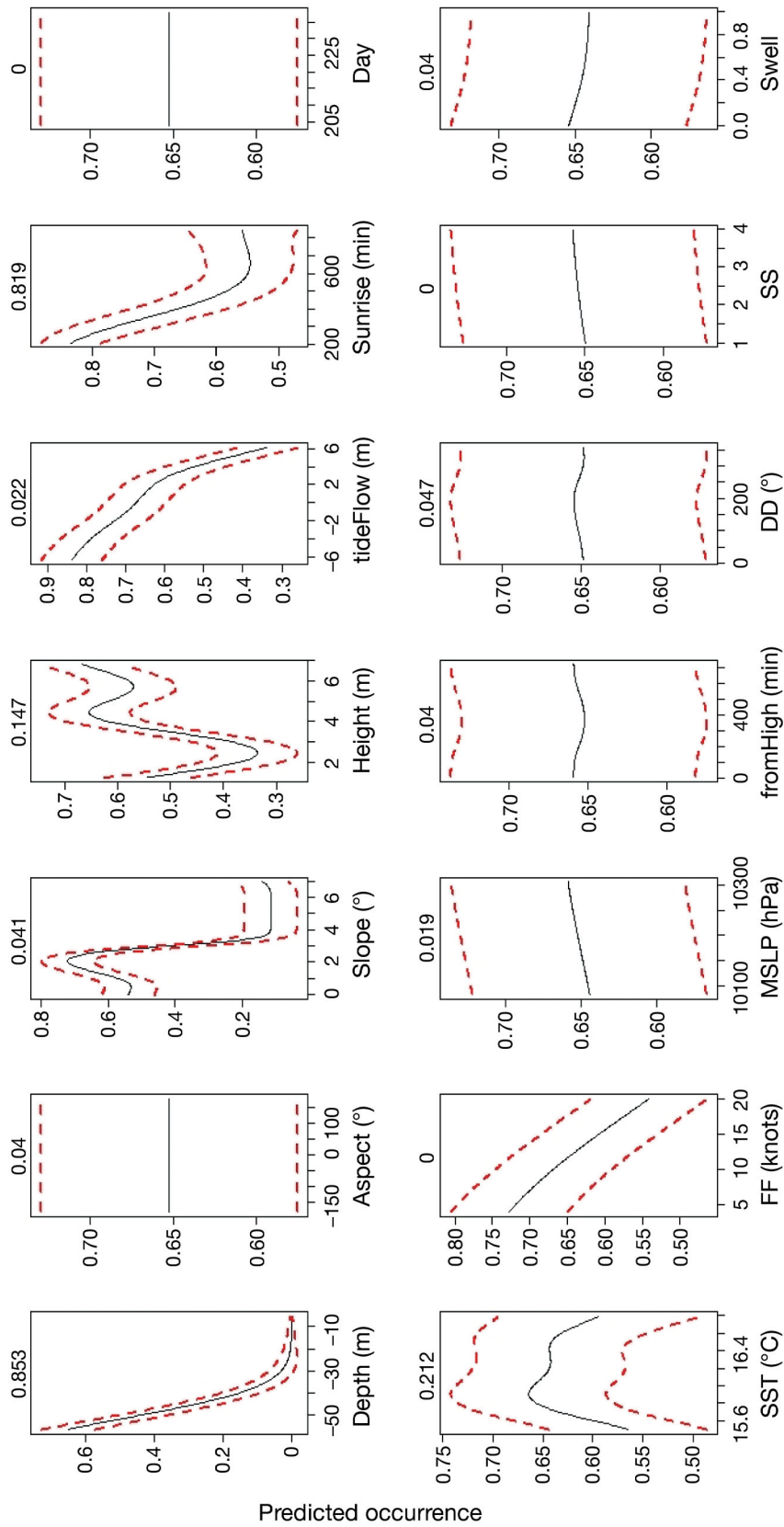Table 4. Covariate combinations, their presence (%) in the 33 confidence set of models and their importance (for each covariate, $\sum_{i=1}^{R} w_i \times 100$ where w_i is the Akaike weight of that model). See Table 1 for description of variables

Covariate	Presence (%)	100Σw_i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
Depth	75.76	85.32	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					x	x	x	x	x	x	x	x
Sunrise	75.76	81.90	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x	x						x	x	x	x	x	x		
Slope × tideFlow	39.39	79.70	x	x	x	x		x	x	x	x	x	x	x	x	x	x					x		x		x	x								
Height × FF	9.09	48.58	x												x					x					x			x							
SST	45.45	21.19					x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x							x					x
Height	24.24	14.65		x		x		x								x	x				x	x			x		x							x	
Height × fromHigh	33.33	11.88							x		x			x	x			x	x		x	x	x						x	x	x				x
Height × Sunrise	3.03	9.12			x																														
DD	12.12	4.67																					x			x	x	x	x						
Slope	9.09	4.13																						x					x		x				
Aspect	3.03	4.00						x																											
Swell	3.03	4.00						x																											
Depth × SS	3.03	4.00						x																											
fromHigh	9.09	3.98		x			x					x	x																						
Height × DD	6.06	3.97																x																	
Depth × Height	3.03	3.54								x																									
SST × DD	12.12	3.08																										x			x	x			
Sunrise × DD	3.03	2.99										x																							
Depth × SST	9.09	2.98																						x		x		x							
tideFlow	9.09	2.23																		x														x	x
MSLP	3.03	1.92																					x												
Sunrise × Swell	6.06	1.79																			x														x
Height × tideFlow	6.06	1.76																											x			x			
Depth × Sunrise	3.03	1.73																							x										
Depth × MSLP	3.03	1.60																									x								
Depth × DD	3.03	0.83													x																				
Height × MSLP	6.06	0.63																								x					x				
Aspect × Sunrise	3.03	0.44																																x	
Slope × Swell	3.03	0.44																																x	
Sunrise × Day	3.03	0.12														x																			
SST × FF	3.03	0.02		x																															

Fig. 5. Averaged (solid line) predictions for each covariate in the confidence set of models and their 0.95 (95%) confidence limits (red dashed lines; with t-multiplier) when all other variables are fixed on their mean value. The relative importance of each variable is shown on top of each figure. See Table 1 for description of variables
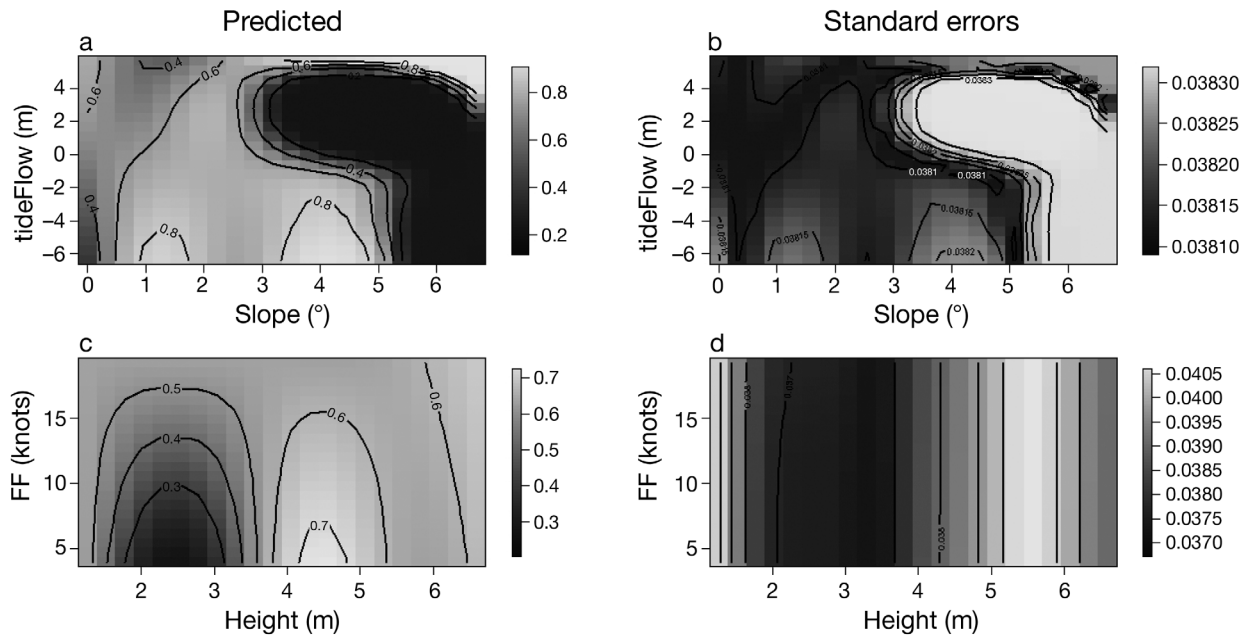
Fig. 6. (a,c) Average predictions on the response scale and (b,d) associated standard error values for interaction terms. See Table 1 for description of variables
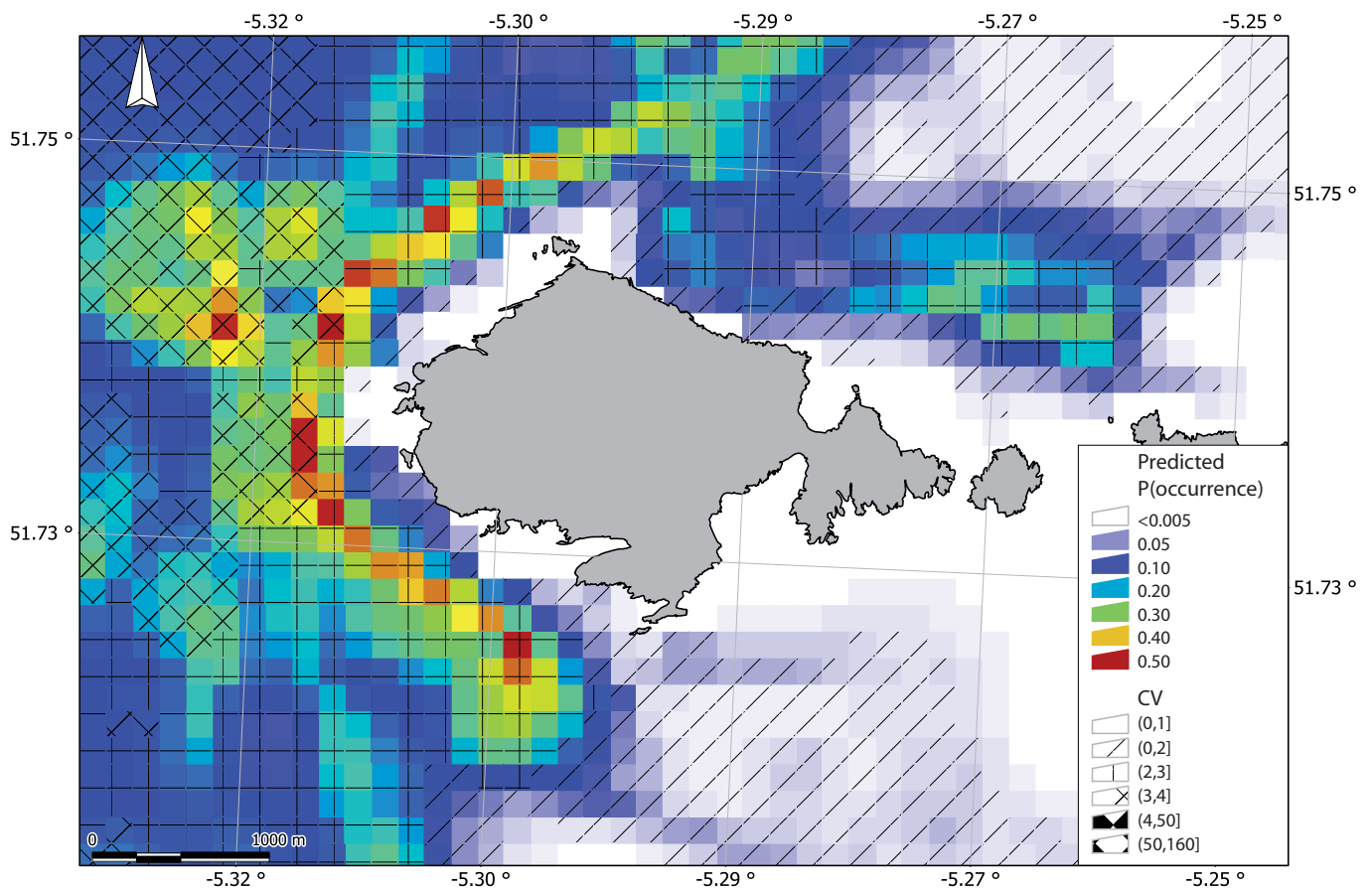
Fig. 7. *Phocoena phocoena*. Example spatial predictions (means and confidence interval, CV) for probability of harbour porpoise occurrence (P[occurrence]) at 2 to 4 h after high tide during spring tide. Predictions for other phases of the tide are shown in Figs S1 to S12 in the Supplement at www.int-res.com/articles/suppl/m448p155_supp.pdf

between 20 and 50 m. Importance of depth with low preference for shallow (<20 m) waters is consistent with the literature (20 to 200 m on shelf waters, with lower preference for waters exceeding 100 m) (Carretta et al. 2001, Northridge et al. 1995, Weir et al. 2000, Evans et al. 2003, Reid et al. 2003, Marubini et al. 2009, Embling et al. 2010). Watts & Gaskin (1985) suggested that the animals may in fact be actively avoiding turbulence in shallow (<10 m) areas.

Tidal variables were present in all of the best 33 models, but no single tidal variable appeared to fully explain the effects of tide on harbour porpoise occurrence (Table 4). Similarly, tidal variables are supported in the literature, but the preferred tidal phase or speed appears to vary across studies (Johnston et al. 2005, Pierpoint 2008, Marubini et al. 2009, Embling et al. 2010). Instead of particular tidal conditions or depth per se, porpoises are likely to select a range of current regimes and topography that enhance relative vorticity (Johnston et al. 2005) and concentrate prey in available high quality patches (Borges & Evans 1997). Including predicted or measured current speed and direction could capture such regimes more directly; unfortunately such data were not available for the present study.

The data indicated preference for both medium to low (1 to 2°) and high (3 to 5°) slope values, but the occurrence at higher slopes was reduced by a factor of 4 at flood tide. In the study area, this coincides with a south-flowing tide and a turning tide from south- to north-flowing tide. The presence of an island in the middle of the study area may have augmented this relationship. Wind- and tide-induced upwelling and stirring at the steeply sloping sides of an island will probably enhance local primary productivity, given a sufficient life span for the feature (Arístegui et al. 1997). Islands can also act as flow obstacles that generate nutrient uplifting eddies in their wake (Arístegui et al. 1997, Coutis & Middleton 2002). Johnston et al. (2005) showed that focal regions of tagged animals tracked by satellite overlapped with localized fronts where herring *Clupea harengus* and euphausiid zooplankton aggregate, and that a large eddy and a frontal system extended to these areas (see also Watts & Gaskin 1985, Read & Westgate 1997). In Pembrokeshire, foraging hotspots in the vicinity of islands (Pierpoint 2001, 2008) may indicate a similar process, as was found within Mousa Sound, southeast Shetland Islands (Borges & Evans 1997).

An increase in harbour porpoise occurrence at the beginning and end of the day may be due to increasing foraging effort in the day–night transition phase or an increase in activity at night. The latter hypothesis is supported by Westgate et al. (1995) and Carlström (2005) who demonstrated activity maxima at night, which was also revealed in neighbouring Cardigan Bay, west Wales, by acoustic monitoring (Pesante et al. 2008). SST explained variability in the data better than did season. There was probably not enough data to distinguish between SST and season owing to the study period being confined to the summer months. Seasonal effects would have been expected to be important in late summer when harbour porpoise numbers peak in this region (Evans et al. 2003) and in the study area where more than 20% of our sightings included calves.

Contrary to the results from most systematic transect surveys (e.g. Evans & Hammond 2004, Tynan et al. 2005, Embling 2008, Marubini et al. 2009), sea state and swell did not appear to affect harbour porpoise presence. The slower speed and responsiveness of the opportunistic platform to the first cues of the animals may have enhanced detections relative to observation conditions. The variability in detections with sighting conditions may have also been better captured by the tidal height–wind speed interaction term; such interaction may reflect changes in perception bias (Buckland et al. 2001) owing to the behaviour of the porpoises and variability in ability to sight the animals as a function of different levels of tidal height and wind speed. Averaged predictions from the interaction term showed an increased detection of porpoises with wind speed at low tidal heights and a reverse, i.e. negative effect at high tidal heights. In shallow tidal waters in southwest Wales, waves vary semi-diurnally so that short period waves (those that affect sea state) respond to current refraction over the tidal cycle (Jones 2000). In particular, it appeared that our detections in the field declined rapidly when wind direction was opposite to the direction of tidal waves.

Three important covariates were excluded from the single best model (SST, tidal height and tidal height–minutes from high tide). They were retained in more models in the confidence set than the slope–tidal flow interaction that was retained in the best model (Table 4). This is due to the uncertainty related to more complex models, and because no single covariate could completely capture the effect of tide; facts that highlight the need for model averaging. Because we had a high threshold for discarding potential model structures for the confidence set, we were more inclusive in selecting models to be used in our averaged predictions. The top 4 models had 31 to 3590 more AIC weight in predictions than the bottom 4 in the confidence set (Table 3).

This study shows that biologically feasible inter-actions can be objectively selected from a set of all possible 2-dimensional interactions, and that a large number (>98%) of interaction terms that fail to ex-plain data can be discarded based on this selection. The UBRE score of the best interaction model de-creased by 0.022 units (~12 units of AIC) from the main effects model, and the explained deviance in-creased from 33.4 to 45.25%. The weight of the best main effects model was too small and UBRE was too high to be included in the confidence set of models.

Burnham & Anderson (2002) warned against data dredging and recommended building the set of pos-sible models based on biological hypotheses. How-ever, very few cetacean species have been studied in sufficient detail to develop specific hypotheses about the ecological processes that determine distributions (Redfern et al. 2006). Similarly, had we restricted our-selves to a predetermined set of models, the lack of previous knowledge on non-linear interactions might have resulted in some biologically important rela-tionships being left out. Since the same data were used for both model specification and prediction, the hypotheses generated by our exhaustive search should be tested for the population from an indepen-dent data set. This can be seen as an iterative process of both prediction- and hypothesis-driven modelling, as recommended by Redfern et al. (2006).

Since the AIC for GAMs in mgcv is based on an approximation of model likelihood (Wood 2006a), model averaging based on AIC weights might have introduced artificial model selection uncertainty. The confidence set was chosen so that it included both the lowest UBRE and AIC scores (Table 3); however, within the confidence set, the ranking of models from the 2 criteria was not identical. For example, the model with the second best UBRE score had the second worst AIC score in the set. Consequently, the importance of the retained vari-ables (a function of the Akaike weights) in that model was artificially reduced. These biases could be addressed by emerging methodologies such as UBRE weighted estimates.

Our results highlight the view that incorporating flow conditions can improve model accuracy at small spatial scales. Further research is needed to explore the importance of flow conditions at larger scales. The necessary temporal resolution for such dynamic models is challenging in terms of cost and effort, but to this end, dedicated survey data could be supple-mented with platform-of-opportunity observations. We have shown here that the inherent biases of such opportunistic data can be accounted for in a robust approach to give estimates of harbour porpoise habi-tat preference and spatial usage.

LITERATURE CITED

Aarts G, Mackenzie ML, McConnell BJ, Fedak MA, Matthiopoulos J (2008) Estimating space use and envi-ronmental preference from wildlife telemetry data. Eco-graphy 31:140–160

Arístegui J, Tett P, Hernandez-Guerra A, Basterretxea G and others (1997) The influence of island-generated eddies on chlorophyll distribution: a study of mesoscale variation around Gran Canaria. Deep-Sea Res I 44:71–96

Borges L, Evans PGH (1997) Spatial distribution of the har-bour porpoise and fish prey and their associations in southeast Shetland, N. Scotland. Eur Res Cetaceans, 10: 262–265

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borcher DL, Thomas L (2001) Introduction to distance sampling. Oxford University Press, New York, NY

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information–theoretic approach. Springer, New York, NY

Carlström J (2005) Diel variation in echolocation behaviour of wild harbour porpoises. Mar Mamm Sci 21:1–12

Carretta JV, Taylor BL, Chivers SJ (2001) Abundance and depth distribution of harbor porpoise (*Phocoena phoco-ena*) in northern California determined from a 1995 ship survey. Fish Bull 99:29–39

CDA International (2007) Manifold system 8012316 ultimate edition. www.manifold.net

Coutis PF, Middleton JH (2002) The physical and biological impact of a small island wake in the deep ocean. Deep-Sea Res 49:1341–1361

Embling C (2008) Predictive models of cetacean distribu-tions off the west coast of Scotland. PhD dissertation, University of St Andrews, St Andrews

Embling CB, Gillibrand PA, Gordon J, Shrimpton J, Stevick PT, Hammond PS (2010) Using habitat models to identify suitable sites for marine protected areas for harbour por-poises (*Phocoena phocoena*). Biol Conserv 143:267–279

Evans PGH, Borges L (1995) Associations between por-poises, seabirds and their prey in south-east Shetland, N. Scotland. Eur Cetacean Soc 9:173–178

Evans PGH, Hammond PS (2004) Monitoring cetaceans in European waters. Mammal Rev 34:131–156

Evans PGH, Anderwald P, Baines ME (2003) UK cetacean status review. Report to English Nature and Countryside Council for Wales. Sea Watch Foundation, Oxford

Evans PGH, Lockyer CH, Smeenk C, Addink M, Read AJ (2008) Harbour porpoise *Phocoena phocoena*. In: Harris S, Yalden DW (eds) Mammals of the British Isles. The Mammal Society, Southampton

Faraway JJ (2006) Extending the linear model with R: gen-eralized linear, mixed effects and nonparametric regres-

sion models. Chapman & Hall/CRC, Boca Raton, FL

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. Ecol Lett 8: 993–1009

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135:147–186

Hauser DDW, VanBlaricom GR, Holmes EE, Osborne RW (2006) Evaluating the use of whalewatch data in determining killer whale (*Orcinus orca*) distribution patterns. J Cetacean Res Manag 8:273–281

Heggenes J (1996). Habitat selection by brown trout (*Salmo trutta*) and young Atlantic salmon (*S. salar*) in streams: static and dynamic hydraulic modelling. Regul Rivers Res Manag 12:155–169

Ingram SN, Walshe L, Johnston D, Rogan E (2007) Habitat partitioning and the influence of benthic topography and oceanography on the distribution of fin and minke whales in the Bay of Fundy, Canada. J Mar Biol Assoc UK 87:149–156

Johnston DW, Westgate AJ, Read AJ (2005) Effects of fine-scale oceanographic features on the distribution and movements of harbour porpoises *Phocoena phocoena* in the Bay of Fundy. Mar Ecol Prog Ser 295:279–293

Jones B (2000) A numerical study of wave refraction in shallow tidal waters. Estuar Coast Shelf Sci 51:331–347

Kiszka J, Macleod K, Van Canneyt O, Walker D, Ridoux V (2007) Distribution, encounter rates, and habitat characteristics of toothed cetaceans in the Bay of Biscay and adjacent waters from platform-of-opportunity data. ICES J Mar Sci 64:1033–1043

Lockyer C (2003) Harbour porpoises (*Phocoena phocoena*) in the North Atlantic: biological parameters. In: Haug T, Desportes G, Víkingsson G, Witting L (eds) NAMMCO Sci Pub 5:71–90

Longdin & Browning Ltd (2002) Habitat and feature distribution in Pembrokeshire marine SAC: acoustic habitat survey. Countryside Council for Wales Science report 514

López A, Pierce GJ, Valeiras X, Santos MB, Guerra A (2004) Distribution patterns of small cetaceans in Galician waters. J Mar Biol Assoc UK 84:283–294

MacLeod CD, Mandleberg L, Schweder C, Bannon SM, Pierce GJ (2008) A comparison of approaches for modelling the occurrence of marine animals. Hydrobiologia 612:21–32

Marubini F, Gimona A, Evans PGH, Wright PJ, Pierce GJ (2009) Habitat preferences and interannual variability in occurrence of the harbour porpoise *Phocoena phocoena* off northwest Scotland. Mar Ecol Prog Ser 381:297–310

Matthiopoulos J, Aarts G (2010) The spatial analysis of marine mammal abundance. In: Boyd IL, Bowen D, Iverson S (eds) Marine mammal ecology and conservation: a handbook of techniques. Oxford University Press, Oxford

Nagelkerke N (1991) A note on the general definition of the coefficient of determination. Biometrika 78:691–692

Northridge SP, Tasker ML, Webb A, Williams JM (1995) Distribution and relative abundance of harbour porpoises (*Phocoena phocoena* L.), white-beaked dolphins (*Lagenorhynchus albirostris* Gray), and minke whales (*Balaenopteua acutorostrata* L.) around the British Isles. ICES J Mar Sci 52:55–66

Pesante G, Evans PGH, Baines ME, McMath M (2008) Abundance and life history parameters of bottlenose dolphin in Cardigan Bay: monitoring 2005–2007. CCW Mar Monit Rep 61. Countryside Council for Wales, Bangor

Pierpoint C (2001) Harbour porpoise distribution in the coastal waters of SW Wales. A report to the International Fund for Animal Welfare. IFAW, Yarmouth Port, MA

Pierpoint C (2008) Harbour porpoise (*Phocoena phocoena*) foraging strategy at a high energy near-shore site in south-west Wales UK. J Mar Biol Assoc UK 88:1167–1173

Read AJ (2001) Trends in the maternal investment of harbour porpoises are uncoupled from the dynamics of their primary prey. Proc R Soc Lond B Biol Sci 268:573–577

Read AJ, Westgate AJ (1997) Monitoring the movements of harbor porpoises (*Phocoena phocoena*) with satellite telemetry. Mar Biol 130:315–322

Redfern JV, Ferguson MC, Becker EA, Hyrenbach KD and others (2006) Techniques for cetacean habitat modelling. Mar Ecol Prog Ser 310:271–295

Reed JZ, Chambers C, Hunter CJ, Lockyer C, Kastelein R, Fedak MA, Boutilier RG (2000) Gas exchange and heart rate in the harbour porpoise, *Phocoena phocoena.* J Comp Physiol B 170:1–10

Reid JB, Evans PGH, Northridge SP (2003) Atlas of cetacean distribution in north-west European waters. JNCC (Joint Nature Conservation Committee), Peterborough

Santos MB, Pierce GJ (2003) The diet of harbour porpoise (*Phocoena phocoena*) in the Northeast Atlantic. Oceanogr Mar Biol Annu Rev 41:355–390

Santos MB, Pierce GJ, Learnmonth JA, Reid RJ and others (2004) Variability in the diet of harbour porpoises (*Phocoena phocoena*) in Scottish waters 1992–2003. Mar Mamm Sci 20:1–27

Torres LG, Read AJ, Halpin P (2008) Fine-scale habitat modelling of a top marine predator: Does prey data improve predictive capacity? Ecol Appl 18:1702–1717

Tynan CT, Ainley DG, Barth JA, Cowles TJ, Pierce SD, Spear LB (2005) Cetacean distributions relative to ocean processes in the northern California Current system. Deep-Sea Res 52:145–167

Watts P, Gaskin DE (1985) Habitat index analysis of the harbour porpoise (*Phocoena phocoena*) in the southern coastal Bay of Fundy, Canada. J Mammal 66:733–744

Weir CR, O'Brien SH (2000) Association of the harbour porpoise (*Phocoena phocoena*) with the western sea front. Eur Cetacean Soc 14:61–65

Weir CR, Pollock C, Cronin C, Taylor S (2001) Cetaceans of the Atlantic Frontier, north and west of Scotland. Cont Shelf Res 21:1047–1071

Westgate AJ, Read AJ, Berggren P, Koopman HN, Gaskin DE (1995) Diving behaviour of harbour porpoises, *Phocoena phocoena.* Can J Fish Aquat Sci 52:1064–1073

Wiens JA (1989) Spatial scaling in ecology. Funct Ecol 3: 385–397

Williams R, Hedley SL, Hammond PS (2006) Modeling distribution and abundance of Antarctic baleen whales using ships of opportunity. Ecol Soc 11:1. www.ecologyandsociety.org/vol11/iss1/art1/

Wood SN (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models. J Am Stat Assoc 99:673–686

Wood SN (2006a) Generalized additive models: an introduction with R. Chapman & Hall/CRC, Boca Raton, FL

Wood SN (2006b) Low-rank scale-invariant tensor product smooths for generalized additive mixed models. Biometrics 62:1025–1036

Wood SN (2008) Fast stable direct fitting and smoothness selection for generalized additive models. J R Stat Soc B 70:495–518