

1 **Running title:** Spatial models for distance sampling
2 **Number of words:** ~??
3 **Number of tables:** ?
4 **Number of figures:** ?
5 **Number of references:** ?

6 **Spatial models for distance sampling data:**
7 **recent developments and future directions**

8 **David L. Miller^{1*}, Louise Burt², Eric Rexstad²,**
9 **Len Thomas².**

- 10 *1. Department of Natural Resources Science, University of Rhode Island,*
11 *Kingston, Rhode Island 02881, USA*
12 *2. Centre for Research into Ecological and Environmental Modelling,*
13 *The Observatory, University of St. Andrews, St. Andrews KY16 9LZ,*
14 *Scotland*

15 ***Correspondence author. dave@ninepointeightone.net**

16

Summary

17

Since the initial work by Hedley & Buckland (2004), there have been

18

many advances to the methodology for density surface modelling in

19

distance sampling. This review aims to describe some of the recent

20

work, in particular from spatial smoothing. We offer a comparison of

21

the various options for the practitioner as well as an examples and

22

software.

23

Keywords: Distance sampling; spatial modelling; generalized additive mod-

24

els; Poisson processes; abundance estimation.

25 Introduction

26 When surveying biological populations it is increasingly common to record
27 spatially referenced data; for example: coordinates of observations, bathy-
28 metry or chlorophyll A levels. Mapping the spatial distribution of a pop-
29 ulation can be extremely useful for practitioners, especially when commu-
30 nicating results to non-experts. Spatial models allow for the vast databases
31 spatially-referenced data to be harnessed, allowing for interactions between
32 environmental covariates and population densities to be investigated. Includ-
33 ing spatial covariates into the model (for example, latitude and longitude)
34 can account for spatial autocorrelation. Recent advances in both methodo-
35 logy and software have made spatial modelling readily available to the non-
36 specialist (e.g. Wood (2006), Rue *et al.* (2009)). Note that here we use the
37 term “spatial model” to include any model which includes spatially referenced
38 covariates, not just those which contain smooths of location.

39 This article concerns combining spatial modelling techniques with dis-
40 tance sampling (Buckland *et al.* (2001), Buckland *et al.* (2004)). Distance
41 sampling takes simple strip sampling and extends it to the case where detec-
42 tion is not certain, for example when animals are cryptic.

43 Observers travel along transect centre lines or stand at points and record
44 the perpendicular distance from the centre line or point to the object of in-
45 terest (y). These distances are used to estimate the *detection function* ($g(y)$)
46 by modelling the decrease in detectability with increasing distance from the
47 line or point. The detection function may also include animal/observer spe-
48 cific covariates (Marques *et al.* (2007)). From the fitted detection function,

the probability of detection can be calculated, this gives the probability that an animal within the truncation distance is detected, which can then be used to calculate density and abundance (Buckland *et al.* (2001), Chapter 3).

In a distance sampling analysis one assumes that the objects of interest are distributed according to some process (Buckland *et al.* (2001), Section 2.1). If the objects' locations are not dependent on any spatially varying covariates (such as location, distance from coast, depth, etc) a homogenous process is assumed; so with respect to the line, the objects are distributed uniformly. It is often possible to design surveys such that this assumption holds (for example, ensuring that transect lines run perpendicular to geographical features that would attract or repel animals) or by post-stratification (Buckland *et al.* (2001), Section 3.7).

Hedley & Buckland (2004) were the first to address spatial modelling of distance sampling data, allowing for a relaxation of the homogeneity of the point process, by including a rate parameter which is a function of spatially varying covariates. Thinking of the underlying placement of the objects as an inhomogeneous point process allows us to think of the detection process as a “thinning” (Cox & Isham (1980), Section 4.3) of the process, resulting in another inhomogeneous point process. By assuming the object placement and detection processes are independent, it is possible to separate these two processes (placement and thinning) in the likelihood.

Modelling the spatial process not only permits the use of spatially referenced data, it also gives practitioners the opportunity to use data from opportunistic surveys, for example “incidental” data arising from “ecotourism” cruises can be included in analyses (Williams *et al.* (2006)). Although with

74 such non-random designs, spatial placement is less important than placement
75 with respect to the range of covariate values expected to be encountered
76 within the area of interest.

77 The rest of the article is structured as follows: we describe two methods
78 which take the point process approach before going on to describe the two-
79 stage approach of Hedley & Buckland (2004). We then describes recent ad-
80 vances, along with some practical advice regarding the model fitting, formula-
81 tion and checking. Throughout this article a motivating data set is used to il-
82 lustrate the methods. These data are from a combination of several shipboard
83 surveys conducted on pan-tropical spotted dolphins in the Gulf of Mexico.
84 These data consist of 47 observations of groups of dolphins. The group size
85 was recorded, as well as the Beaufort sea state at the time of the observation.
86 Coordinates for each observation and depth at a series of points over the pre-
87 diction area were also available as covariates for the analysis. A complete ex-
88 ample analysis can be found at <http://www.github.com/dill/dsm/wiki/>.

89 Direct modelling of the process

90 From the point process description, two modelling procedures arise. One
91 approach is to directly model the point process, estimating the observation
92 process as the thinning of that point process (Niemi & Fernández (2010),
93 Johnson *et al.* (2010)). A second approach consists of performing a distance
94 analysis and using the fitted detection function as part of spatial model
95 (Hedley & Buckland (2004)).

96 Johnson *et al.* (2010) propose a point process-based model for distance
 97 sampling data (henceforth referred to as DSpat). They first assume that
 98 the locations of all individuals in the survey area (not just those which were
 99 observed) are a realisation of an inhomogeneous Poisson process which is a
 100 function of space. The authors then take the novel approach of allowing
 101 for separate (disjoint) regions of the survey area to have different detection
 102 functions associated with them. The sum of these detection functions is then
 103 used as a thinning of the Poisson process. The parameters are then found via
 104 standard maximum likelihood methods for point processes (see, e.g. Badde-
 105 ley & Turner (2000)). In contrast to Hedley & Buckland (2004), parameters
 106 are estimated jointly so uncertainty from both the spatial pattern and the
 107 observation process is incorporated into variance estimates for the abund-
 108 ance. Concurrent estimation of the parameters also ensures that interactions
 109 between the thinning and underlying point process are estimated correctly.
 110 The authors also address the issue of overdispersion (commonly a symptom
 111 of animals or groups clustering), unmodelled by spatial covariates in a man-
 112 ner similar to that for GLMs (see *Recent Developments*, below, for another
 113 approach).

114 Niemi & Fernández (2010) also use Poisson processes but incorporate it
 115 into a fully Bayesian approach. Their intensity function takes the form of a
 116 product of a parametric function of the covariates and a mixture of Gaussian
 117 kernels as a spatial smooth. An appropriate degree of smoothing could be
 118 selected by putting prior distributions on the number and locations of the
 119 “knots” of the spatial smooth (the means of the Gaussian kernels) and then
 120 using reversible jump MCMC (Green (1995)). However, because the authors

only include a single precision parameter for all of the kernels, small and large scale variation cannot both be accommodated. As in Johnson *et al.* (2010), the detection function was used as a thinning of the process, although (unlike DSpat) only one detection function was used across the whole region with known parameters. This means that unlike DSpat (but similar to the count model, above), the uncertainty in the detection function is not incorporated in the spatial model.

Both of the above Poisson process models do not account for group size, both stating that this could be included by considering a marked point process (Cox & Isham (1980), Section 5.5). Both methods offer direct modelling of the point process, although with some drawbacks compared to the methodology of Hedley & Buckland (2004). It should be noted that the loss of efficiency from using a two-stage approach is not large (Buckland *et al.* (2004), p. 313). For these reasons, the article focuses on method of Hedley & Buckland (2004) and the advances which can be applied to their methodology.

Density surface modelling

We refer to the approach of Hedley & Buckland (2004) as *density surface modelling* (DSM), this is used as a rather general description for modelling distance sampling data using spatially referenced data. The approach is incorporated into the popular software package Distance (Thomas *et al.* (2010)). Rather than modelling the point process directly, DSM uses a spatial model for the survey area using the counts, abundance (of individuals

144 or groups) or observation density as response. The principle is simple: just
 145 as conventional and multiple covariate distance sampling (CDS and MCDS,
 146 respectively) extend strip transect sampling to the case where detection is
 147 not guaranteed, DSM extends a spatial model for strip transects to line and
 148 point transects.

149 First, consider conducting a strip transect survey. Strips are divided into
 150 contiguous *segments* (indexed by j), which are of length l_j ; small enough
 151 such that the density does not vary a lot in the segment. For each of these
 152 segments, the number of individuals observed (n_j) is used as the response
 153 (see *Practical advice*, below, for how to deal with size bias in grouped popu-
 154 lations). This count can then be modelled as a function of spatial and envir-
 155 onmental covariates (the \mathbf{z}_{jk} for k indexing the covariates: e.g. location, sea
 156 surface temperature, weather conditions) using a generalized additive model
 157 (GAM; e.g. Wood (2006)). The covered area enters the model as an offset
 158 (the area of segment j , $A_j = 2wl_j$, where w is the truncation distance). We
 159 can model the counts as a function of covariates measured for each segment:

$$\mathbb{E}(n_j) = \exp \left[\log_e (A_j) + \beta_0 + \sum_k f_k (\mathbf{z}_{jk}) \right], \quad (1)$$

160 where the f_k s are smooth functions of the covariates in the GAM case and
 161 β_0 is an intercept term. The distribution of n_j is modelled as quasi-Poisson
 162 in Distance but other options are possible (see discussion of the Tweedie
 163 distribution, below).

165 If perpendicular distance is recorded and a CDS analysis is performed, we
166 replace A_j by $A_j \hat{P}_a$ in eqn 1, where \hat{P}_a is the probability of detection, making
167 the offset the effective area of the segment. Modelling then operates in two
168 stages, first a detection function is fitted to the distance data to obtain \hat{P}_a ,
169 then the following model is fitted:

$$\mathbb{E}(n_j) = \exp \left[\log_e \left(A_j \hat{P}_a \right) + \beta_0 + \sum_k f_k(z_{jk}) \right], \quad (2)$$

170 This formulation can also be used for point transects by setting $A_j = w\pi^2$,
171 $\forall j$. The above definition of the smooth terms is rather general because several
172 covariates could be included in single smooth terms via tensor products of
173 univariate bases (see Wood (2006), Section 4.1.8) or via multivariate spline
174 bases (e.g. thin plate regression splines; Wood (2003)). A typical use of a
175 bivariate spline in this setting is to smooth with respect to spatial coordinates
176 by including the centroid of the j^{th} segment or point. Basis choice for spatial
177 smooths is covered below. Note that even if location is not used, the model
178 is still spatial (in some sense), because the covariates used in the GAM are
179 spatially referenced.

180 If animals occur in groups or clusters, then the response variable in equa-
181 tion 2 could be either the number of groups to estimate group abundance
182 or, if group size has been recorded, then the response variable could be the
183 number of individuals per segment to estimate the individual abundance.

184 Figure 1 (top panel) shows the raw observations from the dolphin data,
185 along with the transect lines, overlaid on the depth data. Figure 2 shows

186 a GAM fitted to the dolphin data, the top panel shows predictions from
 187 a model where the counts were models as a smooth function of depth, the
 188 bottom panel shows predictions where a smooth of spatial location was also
 189 included. Further discussion of the plots follows in *Practical advice*, below.

190 DSM WITH COVARIATES AT THE OBSERVATION LEVEL

191 The above model only considers the case where the covariates are meas-
 192 ured only at the segment/point level (which we refer to environmental or
 193 spatially-referenced covariates). Often covariates (ζ_{ij} , for individual/group i
 194 in segment j) are collected on the level of individuals (or groups); for example
 195 sex, length or observer identity. A multiple covariate distance sampling ana-
 196 lysis (MCDS; Marques & Buckland (2003), Marques *et al.* (2007)) can then
 197 be performed and the probability of detection estimated as a function of the
 198 individual level covariates $\hat{P}_a(\zeta_i)$. Individual level covariates can be incor-
 199 porated into the model by making the response the Horvitz-Thompson es-
 200 timator of per-segment abundance and altering the offset term to be covered
 201 area rather than the effective area:

$$\mathbb{E}(\hat{N}_j) = \exp \left[\log_e(A_j) + \beta_0 + \sum_k f_k(z_{jk}) \right], \quad (3)$$

202 for the multiple covariate case it is simply a case of estimating \hat{N}_j for each
 203 segment via the usual Horvitz-Thompson-type estimator (Thompson (2002):

$$\hat{N}_j = \sum_{i=1}^{n_j} \frac{1}{\hat{P}_a(\zeta_{ij})}.$$

205 Our aims in a DSM analysis are usually two-fold: estimating overall abund-
206 ance and investigating the relationship between abundance and environ-
207 mental covariates.

208 To calculate an abundance estimate for some region of interest, the ne-
209 cessary covariates (those included in the model) must be available for the
210 whole of the region, and they must also be available at the required resolu-
211 tion (using prediction grid cells that are smaller than the resolution of the
212 spatially referenced data will not have an effect on abundance/density estim-
213 ates). Having acquired the relevant data and calculated the associated areas
214 of the prediction cells, predictions can be made for the particular covariate
215 levels and abundance estimates calculated from summing predicted values
216 over the prediction grid cells.

217 As with any predictions which are outside of the range of the data, one
218 should heed the usual warnings regarding extrapolation. For example, in an
219 offshore study the effect of a continental shelf maybe cause significant issues
220 if there was not search effort on both sides of the shelf. Frequently, maps
221 of abundance or density are required and any spurious predictions can be
222 visually assessed, as well as by plotting a histogram of the predicted values.
223 A sensible definition of the region of interest is required to avoid prediction
224 outside the range of the data.

225 Abundance estimation is not the only information contained in these mod-
226 els. By looking at plots of marginal smooths of the spatially referenced
227 covariates, one can begin to understand the relationships between the covari-

228 ates and abundance. Going back to the dolphin data, we can see the effect
 229 of depth on abundance in Figure 3. There we can see that there is a large
 230 depth effect between 0 and 500m which then seems to level off (a straight line
 231 could be drawn inside the confidence band (dashed line)), indicating that the
 232 dolphins prefer water deeper than 500m. Note that the y axis in such plots
 233 is on the scale of the link function (log in this case), so care should be taken
 234 in their interpretation.

235 VARIANCE ESTIMATION

236 Estimating the variance of abundances calculated using DSM is not straight
 237 forward as uncertainty from the estimated parameters of the detection func-
 238 tion must be incorporated into the spatial model. A second consideration is
 239 that in a line transect survey, adjacent segments are likely to be correlated;
 240 failing to account for this spatial autocorrelation will lead to artificially low
 241 variance estimates and hence misleadingly narrow confidence intervals.

242 *Resampling-based methods*

243 Hedley & Buckland (2004) describe a method of calculating the variance in
 244 the abundance estimates using a parametric bootstrap, resampling from the
 245 residuals of the fitted model. The bootstrap then follows the following steps:

246 Denote the fitted values for the model to be $\hat{\boldsymbol{\eta}}$. For $b = 1, \dots, B$ (where
 247 B is the number of resamples required):

- 248 1. Resample (with replacement) the per-segment residuals, store the val-
 249 ues in \mathbf{r}_b .

- 250 2. Refit the model but with the response set to $\hat{\boldsymbol{\eta}} + \mathbf{r}_b$ (where $\hat{\boldsymbol{\eta}}$ are the
251 fitted values from the original model).
- 252 3. Take the predicted values for the new model and store them.

253 From the predicted values stored in the last step, the per-location and abund-
254 ance variance can be calculated in the usual manner. The total variance of
255 the abundance estimate can then be found by combining the variance es-
256 timate from the bootstrap procedure with the variance of the probability of
257 detection from the detection function model (using the delta method; Seber
258 (1982)). This assumes that the two components of the variance are independ-
259 ent and the method does not take into account spatial autocorrelation
260 (the individual segments are treated as independent).

261 The above procedure assumes that there is no correlation in space between
262 segments and that residuals can be swapped around. Clearly if many animals
263 are observed in a segment then we would expect there to be a relatively high
264 level in the next segment (especially because the segments are defined after
265 the survey). A moving block bootstrap (MBB) can account for some of
266 the spatial autocorrelation in the variance estimation. The segments are
267 grouped together into overlapping blocks, (so if the block size is 5, block
268 one is segments 1, ..., 5, the second block is segments 2, ..., 6, and so on).
269 Then, at step (2) above, resamples are taken of the blocks (i.e. groups of
270 segments together) rather than individual segments within the transects.
271 Using blocks should account for some of the autocorrelation between the
272 segments, inflating the variances accordingly.

273 Williams *et al.* (2006) use a slight variation on the MBB, resampling

274 either days or trips such that the total segment length was approximately
 275 the same as that in the original survey. The authors use a jackknife (Efron
 276 (1979)), removing one day (or trip) in turn and refitting the model to the
 277 remaining data. Predictions from the fitted model could be used to calcu-
 278 late a variance and from that confidence intervals (assuming that abundance
 279 estimates are log-normally distributed; Buckland *et al.* (2001), Section 3.6)
 280 can be calculated. By calculating variances for both day and trip, the au-
 281 thors also propose an informal test of between-day correlation: if adjacent
 282 days are independent then the variance estimates for trip and day should be
 283 similar, on the other hand if the adjacent days are autocorrelated then it
 284 would be expected that the trip variance would be lower (and the confidence
 285 intervals narrower). This test could then be used to decide which of the two
 286 resampling units should be used to calculate the abundance variance (if there
 287 was evidence of autocorrelation then trip should be used). The authors also
 288 used the jackknife approach to produce maps of the study area showing how
 289 the surface changed when different parts of the data were removed.

290 The methods detailed above account only for variability in the spatial part
 291 of the model, not the uncertainty in the detection function. The above mov-
 292 ing block bootstrap can be modified to take into account detection function
 293 uncertainty by generating new distances from the fitted detection function
 294 and then re-calculating the offset by fitting a detection function to the new
 295 data. The (new) procedure works as follows:

296 For $b = 1, \dots, B$ (where B is the number of resamples required):

- 297 1. Resample (with replacement) the per-block residuals, store the values
 298 in \mathbf{r}_b .

- 299 2. Let $n_b = \hat{\boldsymbol{\eta}} + \mathbf{r}_b$, rounding to the nearest integer.
- 300 3. Generate n_b new distances from the fitted detection function, refit a
301 new detection function (with the same key function and adjustment
302 terms and selecting the number of adjustments using AIC, if required).
- 303 4. Calculate \hat{P}_a and hence a new offset.
- 304 5. Refit the spatial model (with the same covariates but allowing the
305 smoothing parameter to be selected), to the new response ($\hat{\boldsymbol{\eta}} + \mathbf{r}_b$)
306 with the new offset.
- 307 6. Take the predicted values for the new model and store them.

308 By refitting the detection function in each bootstrap resample should
309 account for the uncertainty in the detection function much much better than
310 using the delta method to combine the variances.

311 *Variance propagation*

312 Rather than using the bootstrap methods above, Williams *et al.* (2011) cal-
313 culate the variance without having to refit the model many times. Their
314 method incorporates the uncertainty in the estimation of the detection func-
315 tion into the variance of the spatial model, albeit only in the case where
316 covariates are measured at a point/segment level only. Their procedure is as
317 follows:

- 318 1. Fit the model described in eqn 2.

- 319 2. Re-fit the model with an additional random effects term. This term
320 characterises the uncertainty in the estimation of the detection function
321 (via the uncertainty of the probability of detection, \hat{P}_a).
- 322 3. Variance estimates of the abundance calculated (via the method given
323 in Wood (2006), page 245) from the model will include uncertainty
324 from estimation of the detection function.

325 We consider propagating the uncertainty in this manner not only to be more
326 computationally efficient but also preferable from a technical perspective.
327 The bootstrap methods described above do not fully account for spatial
328 autocorrelation, this failure to account for spatial autocorrelation will lead
329 to wider confidence intervals for the abundance (or density).

330 *Visualising uncertainty*

331 There are several ways to visualise the uncertainty measures calculated above.
332 For the bootstrap methods, if at each round of the bootstrap the predicted
333 values are stored per prediction grid cell, the coefficient of variation can be
334 calculated per cell and then displayed. Figure 4 shows maps of the coefficient
335 of variation for the model which includes both location and depth covariates.
336 The top panel shows the result of running 1000 bootstrap replications in-
337 cluding detection function uncertainty as above. The bottom panel shows
338 the same plot but using the variance propagation method.

339 Recent developments

340 EDGE EFFECTS

341 Recent work (Ramsay (2002), Wang & Ranalli (2007), Wood *et al.* (2008),
342 Scott-Hayward et al (in prep) and Miller and Wood (submitted)) has high-
343 lighted the need to take care when smoothing over areas with complicated
344 boundaries; for example, if the survey area includes rivers, peninsulae or
345 islands. If two parts of the domain (either side of a peninsula, say) are inap-
346 propriately linked by the model (the distance between the points is measured
347 “as the crow flies”, rather than “as the fish swims”) then the boundary feature
348 can be “smoothed across” leading to incorrect inference. Ensuring that a real-
349 istic spatial model has been fit to the data (and, for example, that whales
350 have not been estimated to dwell on land) is essential for valid inference.
351 The soap film smoother of Wood *et al.* (2008) is particularly appealing as
352 the model jointly estimates boundary conditions for a complex study area
353 along with the “interior” smooth. This can be particularly helpful when
354 uncertainty is estimated via a bootstrap as the model helps avoid large, un-
355 realistic predictions which can plague other smoothers (Bravington & Hedley
356 (2009)).

357 Even if the study area does not have a complicated boundary, edge effects
358 can still be problematic. Miller et al (in prep.) show that when using global
359 smoothers, smoothing towards the plane can cause the fitted surface to “curl-
360 up” as predictions move further away from the data. They suggest the use of
361 *Duchon splines* (a generalisation of thin plate regression splines) to alleviate
362 the problem by smoothing toward the intercept.

364 The Tweedie distribution offers a very flexible alternative to the quasi-Poisson
 365 distribution is the usual response distribution when modelling count data
 366 (Candy (2004)). Through the parameter p , many common distributions
 367 arise; varying p between 1 (Poisson) and 2 (gamma) leads to a random vari-
 368 able which is a sum of M gamma variables where M is Poisson distributed
 369 (Jørgensen (1987)). Although it is possible to perform optimization to find
 370 p , this is generally seen as unnecessary as the distribution does not change
 371 appreciably when p is changed by less than 0.1 (therefore trial and error is
 372 usually reasonable). Mark Bravington (pers. comm.) suggested plotting the
 373 square root of the absolute value of the residuals and if this plot is flat a
 374 “correct” p has been found. Additionally he suggests a value of 1.5/1.6 for p
 375 for fisheries and 1.2 marine mammal work is generally acceptable.

376 Practical advice

377 Figure 5 shows a flow diagram of the modelling process for creating a density
 378 surface model for distance sampling data. The diagram shows which methods
 379 are compatible with each other and what the options are for modelling a
 380 particular data set.

381 In the experience of the authors, it is sensible to start with a detection
 382 function without covariates and a simple smooth of spatial location and then
 383 add in more complicated features such as covariates in the detection function,
 384 or using a soap film smoother (perhaps afterwards dropping the location
 385 term). Model discrimination can be performed for the detection function

386 using goodness-of-fit tests (Buckland *et al.* (2004) and AIC. For the spatial
387 model, generalized cross validation (GCV) score and percentage deviance
388 explained are useful metrics, we also highly recommend the use of standard
389 GAM diagnostic plots. An example of such plots is given in Figure 6 along
390 with a description of their uses.

391 In the dolphin analysis, we include a smooth of location. This not only
392 doubles the percentage deviance explained (27.3% to 52.7%), it also allows us
393 to account for spatial autocorrelation (in a primitive way). One can see this
394 when comparing the two plots in Figure 2 and the plot of the depth in Figure
395 1, the plot of the smooth of depth alone looks very similar to the raw plot of
396 the depth data. A smooth of an environmental-level covariate such as depth
397 can be very useful for assessing the relationships between abundance/density
398 and the covariate, but estimates of abundance/density from such models may
399 be misleading.

400 In the analysis we have converted from latitude and longitude to metres
401 from the point (27.01, -88.3). This is because the bivariate smoother which we
402 use (the thin plate spline, Wood (2003)) is isotropic: it treats the wigglyness
403 of the smoother in each direction as equal: a move of 1 degree in latitude is
404 not the same as a move of 1 degree in longitude, the move to meters from
405 the centre of the study area is sensible (using SI units removes the need for
406 conversion later).

407 If animals occur in groups rather than individually a size bias can occur
408 due to larger groups being more visible than smaller groups. The expected
409 group size can be obtained from a regression of probability of detec-
410 tion against the logarithm of group size. Having calculated the expected

411 group size, this can be used calculate the per-segment abundance, rather
412 than simply summing the number of observed individuals.

413 Discussion

414 The field is quickly evolving to allow modelling of more complex data how-
415 ever the basic principle remains as in Hedley & Buckland (2004), albeit with
416 various additions to the modelling process. We expect to see large advances
417 two areas: temporal inferences and the handling of spatial autocorrelation.
418 These should become more mainstream as modern spatio-temporal model-
419 ling techniques are adopted. Petersen *et al.* (2011) provide a very basic
420 framework for temporal modelling; their model includes extra smooth terms
421 for their spatial and depth smooth terms after the construction of an off-
422 shore windfarm which are included via an indicator. Spatial autocorrelation
423 can be accounted for via approaches that explicitly introduce correlations
424 such as generalized estimating equations (GEEs; Hardin & Hilbe (2003)) or
425 via mechanisms such as that of Skaug (2006), which allows observations to
426 cluster according to one of several states (e.g. “feeding” or “transit”) taking
427 into account short-term agglomerations (“hot spots”).

428 Acknowledgments

429 DLM wishes to thank Mark Bravington and Sharon Hedley for their help
430 and patience in explaining and providing code for their variance propagation
431 method and alerting him to the existence of the Markov modulated Poisson

432 process.

433 LEN: Do we need to say something about the Navy funding me
434 here?

435 References

- 436 Baddeley, A. & Turner, R. (2000) Practical maximum pseudolikelihood for spatial
 437 point patterns. *Australian & New Zealand Journal of Statistics*, **42**, 283–322.
 438 URL [http://onlinelibrary.wiley.com/doi/10.1111/1467-842X.00128/](http://onlinelibrary.wiley.com/doi/10.1111/1467-842X.00128/abstract)
 439 abstract
- 440 Bravington, M. & Hedley, S.L. (2009) Antarctic minke whale abundance estimates
 441 from the second and third circumpolar IDCR/SOWER surveys using the
 442 SPLINTR model.
 443 URL [http://www.iwcoffice.org/_documents/sci_com/sc61docs/](http://www.iwcoffice.org/_documents/sci_com/sc61docs/SC-61-IA14.pdf)
 444 SC-61-IA14.pdf
- 445 Buckland, S.T., Anderson, D., Burnham, K.P., Laake, J.L., Borchers, D.L. &
 446 Thomas, L. (2001) *Introduction to Distance Sampling*. Oxford University Press.
- 447 Buckland, S.T., Anderson, D., Burnham, K.P., Laake, J.L., Borchers, D.L. &
 448 Thomas, L. (2004) *Advanced Distance Sampling*. Oxford University Press.
- 449 Candy, S. (2004) Modelling catch and effort data using generalised linear models,
 450 the Tweedie distribution, random vessel effects and random stratum-by-year
 451 effects. *Ccamlr Science*, **11**, 59–80.
 452 URL http://www.ccamlr.org/ccamlr_science/Vol-11-2004/04candy.pdf
- 453 Cox, D.R. & Isham, V. (1980) *Point Processes*. Monographs on Applied Probability
 454 and Statistics. Chapman and Hall. ISBN 9780412219108.
- 455 Efron, B. (1979) Bootstrap methods: another look at the jackknife. *The Annals of*
 456 *Statistics*, **7**, 1–26.
- 457 Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and
 458 Bayesian model determination. *Biometrika*, **82**, 711–732.
- 459 Hardin, J. & Hilbe, J. (2003) *Generalized Estimating Equations*. Chapman and
 460 Hall/CRC, London, UK.
- 461 Hedley, S.L. & Buckland, S.T. (2004) Spatial models for line transect sampling.
 462 *Journal of Agricultural, Biological, and Environmental Statistics*, **9**, 181–199.
- 463 Johnson, D.S., Laake, J.L. & Ver Hoef, J.M. (2010) A model-based approach for
 464 making ecological inference from distance sampling data. *Biometrics*, **66**, 310–
 465 318.
- 466 Jørgensen, B. (1987) Exponential dispersion models. *Journal of the Royal Statist-*
 467 *ical Society. Series B, Statistical Methodology*, **49**, 127–162.

- 468 Marques, F. & Buckland, S.T. (2003) Incorporating covariates into standard line
469 transect analyses. *Biometrics*, **59**, 924–935.
- 470 Marques, T.A., Thomas, L., Fancy, S. & Buckland, S.T. (2007) Improving estimates
471 of bird density using multiple-covariate distance sampling. *The Auk*, **124**, 1229–
472 1243.
- 473 Niemi, A. & Fernández, C. (2010) Bayesian Spatial Point Process Modeling of Line
474 Transect Data. *Journal of Agricultural, Biological, and Environmental Statistics*,
475 **15**, 327–345.
- 476 Petersen, I.K., MacKenzie, M., Rexstad, E., Wisz, M.S. & Fox, A.D. (2011) Com-
477 paring pre- and post-construction distributions of long-tailed ducks *Clangula*
478 *hyemalis* in and around the Nysted offshore wind farm, Denmark: a quasi-
479 designed experiment accounting for imperfect detection, local surface features
480 and autocorrelation. 2011-1.
- 481 Ramsay, T. (2002) Spline smoothing over difficult regions. *Journal of the Royal*
482 *Statistical Society. Series B, Statistical Methodology*, pp. 307–319.
- 483 Rue, H., Martino, S. & Chopin, N. (2009) Approximate Bayesian inference for
484 latent Gaussian models by using integrated nested Laplace approximations. *J.*
485 *R. Statist. Soc. B*, **71**, 319–392.
- 486 Seber, G.A.F. (1982) *The Estimation of Animal Abundance and Related Paramet-*
487 *ers*. Blackburn Pr. ISBN 9781930665552.
488 URL [http://books.google.com/books?id=bnGaPQAACAAJ&dq=seber&cd=](http://books.google.com/books?id=bnGaPQAACAAJ&dq=seber&cd=10&source=gbs_api)
489 [10&source=gbs_api](http://books.google.com/books?id=bnGaPQAACAAJ&dq=seber&cd=10&source=gbs_api)
- 490 Skaug, H.J. (2006) Markov modulated Poisson processes for clustered line transect
491 data. *Environmental and Ecological Statistics*, **13**, 199–211.
- 492 Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley,
493 S.L., Bishop, J.R., Marques, T.A. & Burnham, K.P. (2010) Distance software:
494 design and analysis of distance sampling surveys for estimating population size.
495 *Journal of Applied Ecology*, **47**, 5–14.
- 496 Thompson, S.K. (2002) *Sampling*. Wiley, 2nd edn. ISBN 9781118162965.
497 URL [http://books.google.com/books?id=qukULxJ--QAC&printsec=](http://books.google.com/books?id=qukULxJ--QAC&printsec=frontcover&dq=intitle:sampling+inauthor:thompson&cd=1&source=gbs_api)
498 [frontcover&dq=intitle:sampling+inauthor:thompson&cd=1&source=gbs_](http://books.google.com/books?id=qukULxJ--QAC&printsec=frontcover&dq=intitle:sampling+inauthor:thompson&cd=1&source=gbs_api)
499 [api](http://books.google.com/books?id=qukULxJ--QAC&printsec=frontcover&dq=intitle:sampling+inauthor:thompson&cd=1&source=gbs_api)
- 500 Wang, H. & Ranalli, M. (2007) Low-rank smoothing splines on complicated do-
501 mains. *Biometrics*, **63**, 209–217.

- 502 Williams, R., Hedley, S.L., Branch, T.A., Bravington, M.V., Zerbini, A.N. & Find-
503 lay, K.P. (2011) Chilean blue whales as a case study to illustrate methods to
504 estimate abundance and evaluate conservation status of rare species. *Conserva-*
505 *tion Biology*, **25**, 526–535.
- 506 Williams, R., Hedley, S.L. & Hammond, P. (2006) Modeling distribution and
507 abundance of Antarctic baleen whales using ships of opportunity. *Ecology and*
508 *Society*, **11**, 1.
- 509 Wood, S.N. (2003) Thin plate regression splines. *Journal of the Royal Statistical*
510 *Society. Series B, Statistical Methodology*, **65**, 95–114.
- 511 Wood, S.N. (2006) *Generalized Additive Models: An introduction with R*. Chapman
512 & Hall/CRC.
- 513 Wood, S.N., Bravington, M.V. & Hedley, S.L. (2008) Soap film smoothing. *Journal*
514 *of the Royal Statistical Society. Series B, Statistical Methodology*, **70**, 931–955.

Figures

Fig. 1 Top: the survey area, transect centrelines and observations with size of circle corresponding to the group size overlaid onto depth data; bottom left, histogram of observed distances with fitted detection function; bottom right, plot of distance versus group size with linear trend showing the relation between distance and group size.

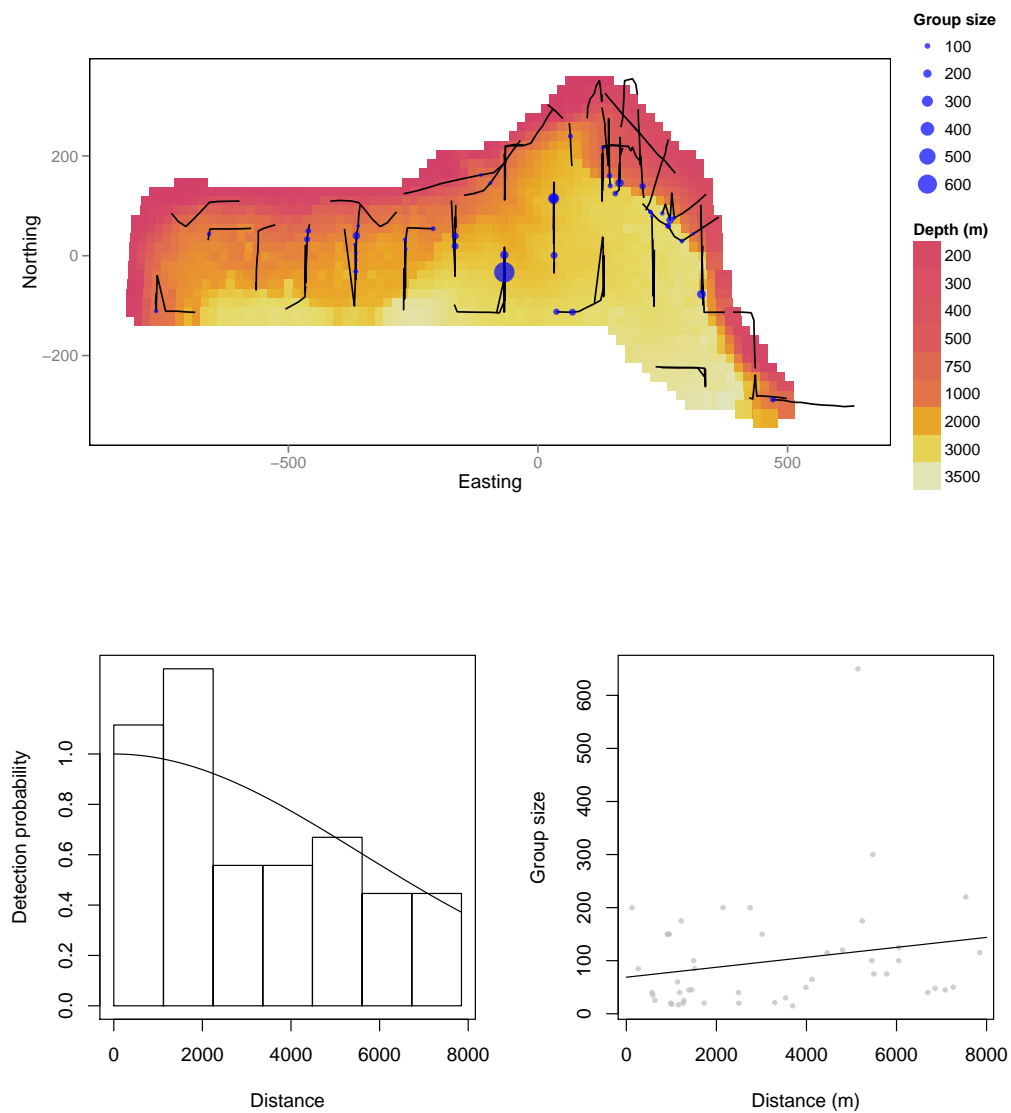


Fig. 2 Predictions for the dolphin data. Top: Predictions from the model using only depth as an explanatory variable, bottom: the model using both depth and location.

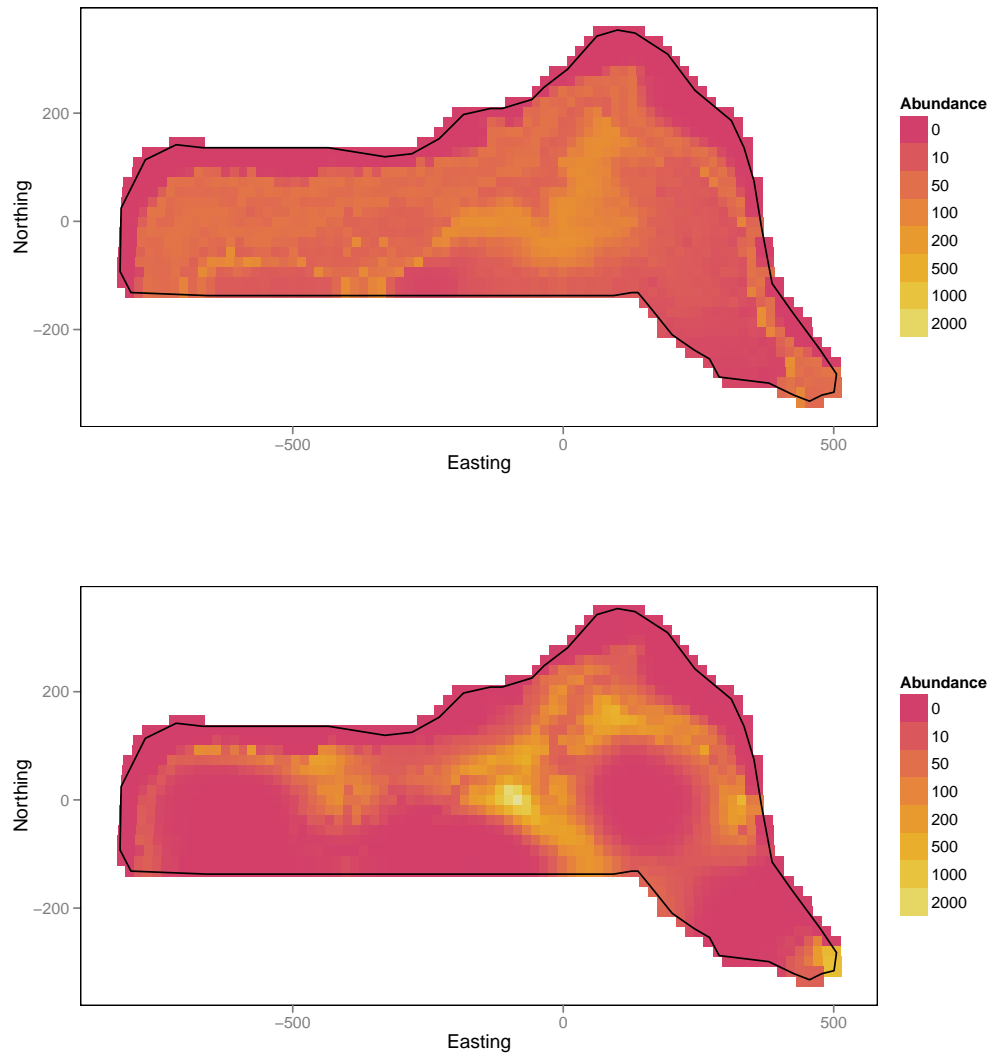


Fig. 3 Plot of the effect on the response of depth, note that it is possible to draw a straight line between 750m and 3000m within the confidence band, so the wiggles in the smooth may not be indicative of any relationship. What is clear is that there is some effect up to about 500m. The number in brackets on the y axis indicates the effective degrees of freedom of the smooth term. The rug ticks at the bottom of the plot indicate where the data were collected.

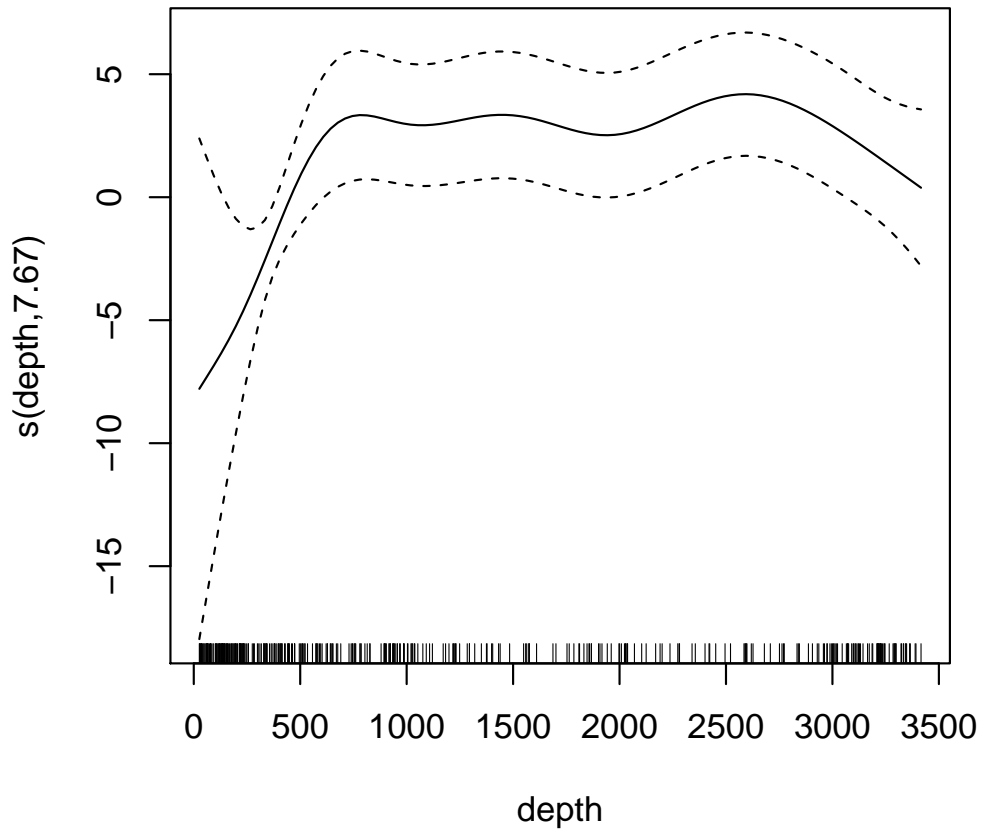


Fig. 4 Plot of coefficient of variation maps, showing the uncertainty in the fitted model. The top panel shows the estimate using the moving block bootstrap incorporating detection function uncertainty, the bottom panel shows the same plot using the variance propagation method. The bootstrap plot seems far more noisy.

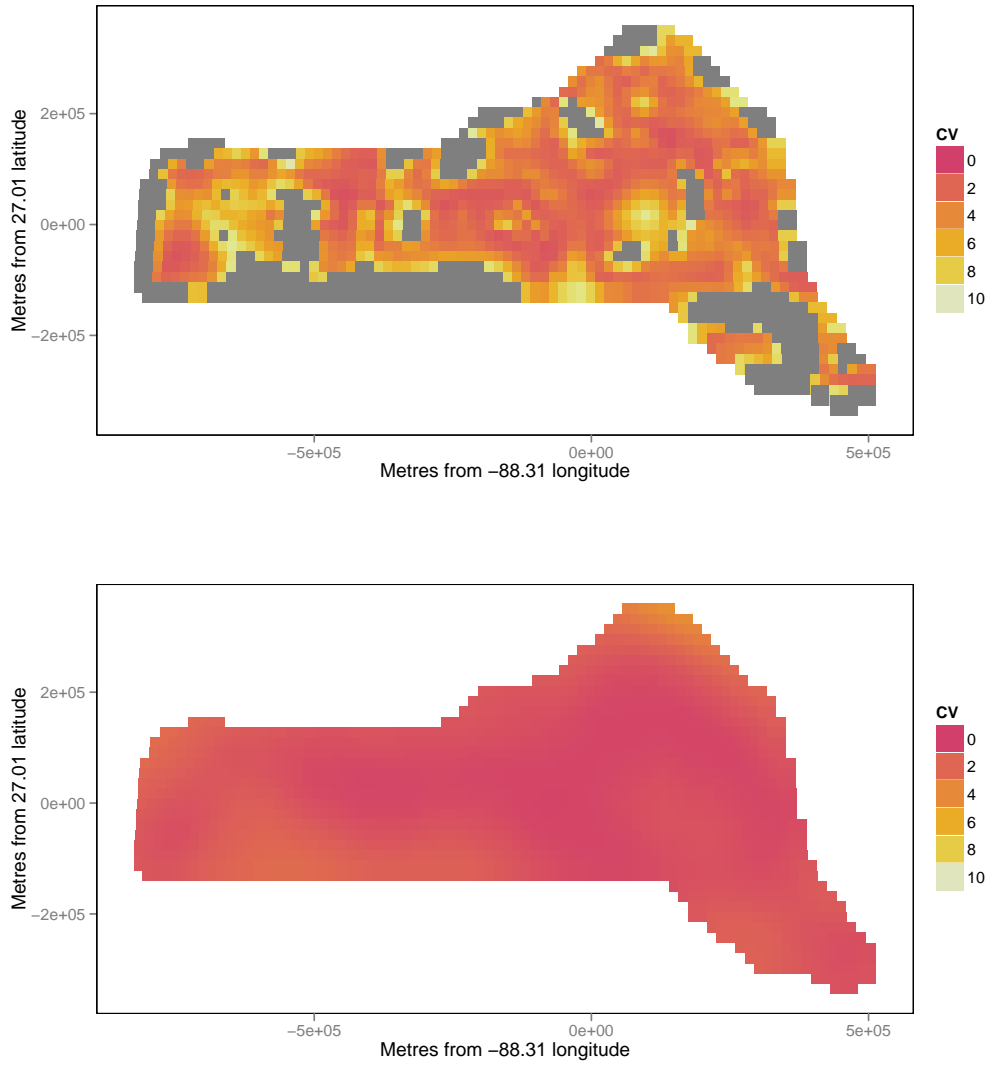


Fig. 5 Flow diagram showing the modelling process for creating a density surface model.

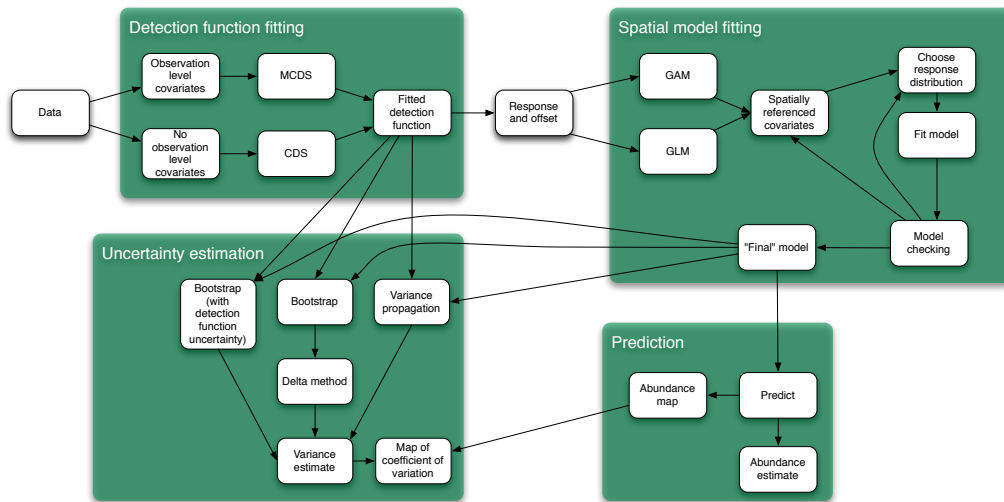


Fig. 6 Example of model diagnostics for the model which included both location and depth covariates for the dolphin data. From top left clockwise:

