Advanced Distance Sampling

S. T. BUCKLAND
University of St Andrews

D. R. ANDERSON
Colorado Cooperative Fish and Wildlife Research Unit

K. P. BURNHAM
Colorado Cooperative Fish and Wildlife Research Unit

J. L. LAAKE
National Marine Mammal Laboratory, Seattle

D. L. BORCHERS
University of St Andrews

L. THOMAS
University of St Andrews

OXFORD
UNIVERSITY PRESS
Contents

1 Introduction to advanced distance sampling 1
   S. T. Buckland and D. R. Anderson

2 General formulation for distance sampling 6
   D. L. Borchers and K. P. Burnham
   2.1 Introduction 6
   2.2 CDS revisited 7
      2.2.1 Conventional line transect estimator 8
      2.2.2 Conventional point transect estimator 9
   2.3 Horvitz-Thompson: a versatile estimator 9
      2.3.1 Animals that occur as individuals 10
      2.3.2 Animals that occur in clusters 11
      2.3.3 CDS estimators 11
   2.4 Maximum likelihood estimation 11
      2.4.1 ‘Covered’ animals 12
      2.4.2 Random detection with known probability 13
      2.4.3 CDS likelihoods 16
   2.5 Summary so far and preview of advances 17
      2.5.1 Summary 17
      2.5.2 Preview of advances 18
   2.6 Advanced methods for detection function estimation 19
      2.6.1 Multiple covariate distance sampling 19
      2.6.2 Mark-recapture distance sampling 21
      2.6.3 Estimation when π(y) is unknown 23
   2.7 Estimating animal density surfaces 24
      2.7.1 The count method 25
      2.7.2 The waiting distance method 25
      2.7.3 Cluster size surface estimation 26
   2.8 Survey design 26
      2.8.1 Likelihood-based inference 26
      2.8.2 Design-based inference 27
      2.8.3 Adaptive distance sampling 27
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.9 Model selection</td>
<td>28</td>
</tr>
<tr>
<td>2.10 Summary</td>
<td>29</td>
</tr>
<tr>
<td>3 Covariate models for the detection function</td>
<td>31</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>31</td>
</tr>
<tr>
<td>3.2 A conditional likelihood framework for distance sampling</td>
<td>32</td>
</tr>
<tr>
<td>3.3 Line transect sampling</td>
<td>33</td>
</tr>
<tr>
<td>3.3.1 The conditional likelihood</td>
<td>33</td>
</tr>
<tr>
<td>3.3.2 Incorporating covariates into semiparametric models for the detection function</td>
<td>35</td>
</tr>
<tr>
<td>3.3.3 Abundance estimation</td>
<td>38</td>
</tr>
<tr>
<td>3.4 Point transect sampling</td>
<td>43</td>
</tr>
<tr>
<td>3.5 Example</td>
<td>45</td>
</tr>
<tr>
<td>3.6 Discussion</td>
<td>47</td>
</tr>
<tr>
<td>4 Spatial distance sampling models</td>
<td>48</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>48</td>
</tr>
<tr>
<td>4.2 Spatial line transect models</td>
<td>49</td>
</tr>
<tr>
<td>4.2.1 Deriving a likelihood</td>
<td>50</td>
</tr>
<tr>
<td>4.2.2 A likelihood based on inter-detection distances</td>
<td>52</td>
</tr>
<tr>
<td>4.2.3 Clustered populations</td>
<td>54</td>
</tr>
<tr>
<td>4.3 Practical implementations of spatial line transect models</td>
<td>55</td>
</tr>
<tr>
<td>4.3.1 A waiting distance model</td>
<td>55</td>
</tr>
<tr>
<td>4.3.2 A count model</td>
<td>57</td>
</tr>
<tr>
<td>4.4 Spatial distribution of Antarctic minke whales</td>
<td>60</td>
</tr>
<tr>
<td>4.5 Spatial point transect models</td>
<td>63</td>
</tr>
<tr>
<td>4.5.1 Deriving a likelihood</td>
<td>63</td>
</tr>
<tr>
<td>4.5.2 A point transect count model</td>
<td>65</td>
</tr>
<tr>
<td>4.6 Discussion</td>
<td>66</td>
</tr>
<tr>
<td>5 Temporal inferences from distance sampling surveys</td>
<td>71</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>71</td>
</tr>
<tr>
<td>5.2 Concepts</td>
<td>73</td>
</tr>
<tr>
<td>5.2.1 Sampling and population variation</td>
<td>73</td>
</tr>
<tr>
<td>5.2.2 Sampling covariance</td>
<td>75</td>
</tr>
<tr>
<td>5.2.3 Empirical and process models</td>
<td>76</td>
</tr>
<tr>
<td>5.2.4 Trend</td>
<td>76</td>
</tr>
<tr>
<td>5.2.5 Abundance as a fixed or random quantity</td>
<td>77</td>
</tr>
</tbody>
</table>
5.3 Trend estimation from global abundance estimates 79
5.3.1 Graphical exploration 79
5.3.2 Linear trend models 79
5.3.3 Smoothing 84
5.3.4 Trend estimation when samples covary 88
5.4 Spatio-temporal analysis 91
5.4.1 Transect-level models of trend 91
5.4.2 Spatio-temporal modelling 93
5.5 Process models 93
5.5.1 State-space models 94
5.5.2 Generalizing state-space models 96
5.6 Other analysis methods 98
5.6.1 Time series methods 98
5.6.2 Quality control methods 98
5.7 Survey design 99
5.7.1 Repeating transects 99
5.7.2 Sample size 101
5.7.3 Planning long-term studies 105

6 Methods for incomplete detection at distance zero 108
J. L. Laake and D. L. Borchers

6.1 Introduction 108
6.2 Likelihood and Horvitz–Thompson 111
6.2.1 Constant detection probability 112
6.2.2 Detection probability changing with distance 114
6.2.3 Independence issues 117
6.2.4 Multiple covariates 121
6.2.5 Unobserved heterogeneity 122
6.3 State and observation models 123
6.3.1 State models 124
6.3.2 Observation models 128
6.3.3 Observation configurations 137
6.4 Example data 140
6.5 Estimation for independent configuration 142
6.5.1 Distance only 142
6.5.2 Distance and covariates 154
6.6 Estimation for trial configuration 159
6.6.1 Distance and covariates 159
6.6.2 Distance, covariates, and responsive movement 162
6.7 Estimation for removal configuration 166
6.8 Dealing with availability bias 168
6.8.1 Static availability 168
6.8.2 Hazard-rate models for dynamic availability 170
CONTENTS

6.8.3 Discrete availability: animal-based 171
6.8.4 Discrete availability: cue-based 174
6.8.5 Intermittent availability 175
6.8.6 Design-based availability estimation 178

6.9 Special topics 180
6.9.1 Uncertain duplicate identification 180
6.9.2 When should double-observer methods be used? 182

6.10 Field methods 185
6.10.1 Marked animals 185
6.10.2 Observation configuration 186
6.10.3 Data collection and recording 188

7 Design of distance sampling surveys and Geographic Information Systems 190
S. Strindberg, S. T. Buckland, and L. Thomas

7.1 The potential role of GIS in survey design 190
7.2 Automated survey design 191
7.2.1 Point transect design 192
7.2.2 Line transect design using lines of fixed length 199
7.2.3 Line transect design using lines that span the full width of the survey region 206
7.2.4 Zigzag samplers 211

7.3 Estimation for uneven coverage probability designs 224
7.3.1 Objects that occur singly 225
7.3.2 Objects that occur in clusters 226
7.3.3 Variance estimation 226

7.4 Choosing between survey designs by simulation 226

8 Adaptive distance sampling surveys 229
J. H. Pollard and S. T. Buckland

8.1 Introduction 229
8.2 Design-unbiased adaptive point transect surveys 230
8.2.1 Survey design 230
8.2.2 Estimation 233
8.2.3 Simulated example 237
8.2.4 Discussion 239

8.3 Design-unbiased adaptive line transect surveys 240
8.3.1 Survey design 240
8.3.2 Estimation 241
8.3.3 Discussion 246
8.4 Fixed-effort adaptive line transect surveys 247
   8.4.1 Survey design 247
   8.4.2 Estimation 247
   8.4.3 Simulation 254
   8.4.4 Discussion 259

9 Passive approaches to detection in distance sampling 260
   P. M. Lukacs, A. B. Franklin, and D. R. Anderson
   9.1 Introduction 260
   9.2 Trapping webs 262
      9.2.1 Density estimation 262
      9.2.2 Including data from recaptures 263
      9.2.3 Design of trapping webs 266
      9.2.4 An example 266
      9.2.5 A critique of the trapping web 269
   9.3 Trapping line transects 270
      9.3.1 Density estimation 271
      9.3.2 Including data from recaptures 271
      9.3.3 Design of trapping line transects 272
      9.3.4 An example 272
      9.3.5 A critique of the trapping line transect 278
   9.4 Discussion and summary 278

10 Assessment of distance sampling estimators 281
   R. M. Fewster and S. T. Buckland
   10.1 Introduction 281
      10.1.1 Notation 281
   10.2 Estimation framework 282
   10.3 Model and design 283
      10.3.1 Model-based inference 283
      10.3.2 Design-based inference 284
      10.3.3 Distance sampling: a composite approach 286
   10.4 Simulation framework 286
      10.4.1 Testing the design 287
      10.4.2 Testing the model 288
      10.4.3 Testing the full line transect estimation procedure 291
   10.5 Example: testing the design 292
      10.5.1 Testing equal coverage designs for var(\hat{N}) 293
      10.5.2 A design without equal coverage probability 297
10.6 Example: non-uniformity within the strip 298
10.6.1 Estimation of $N_c$ 298
10.6.2 Asymptotic result when $\theta$ is estimated 301
10.7 Example: full estimation procedure 302
10.8 Trial by simulation: a completely model-based approach 302
10.9 Summary 306

11 Further topics in distance sampling 307
K. P. Burnham, S. T. Buckland, J. L. Laake, D. L. Borchers,
T. A. Marques, J. R. B. Bishop, and L. Thomas

11.1 Distance sampling in three dimensions 307
11.1.1 Three-dimensional line transect sampling 307
11.1.2 Three-dimensional point transect sampling 309

11.2 Conventional distance sampling: full likelihood examples 312
11.2.1 Line transects: simple examples 312
11.2.2 Point transects: simple examples 318
11.2.3 Some numerical confidence interval comparisons 321

11.3 Line transect surveys with random line length 327
11.3.1 Introduction 327
11.3.2 Line transect sampling with fixed $n$ and random $L$, under Poisson object distribution 328
11.3.3 Technical comments 331
11.3.4 Discussion 333

11.4 Models for the search process 335
11.4.1 Continuous hazard-rate models 335
11.4.2 Discrete hazard-rate models 340
11.4.3 Further modelling of the detection process 343

11.5 Combining mark-recapture and distance sampling surveys 350

11.6 Combining removal methods and distance sampling 352
11.6.1 Introduction 352
11.6.2 Combining removal methods with distance sampling 354

11.7 Point transect sampling of cues 356
11.7.1 Introduction 356
11.7.2 Estimation 357

11.8 Migration counts 359
11.8.1 Background 359
11.8.2 Modelling migration rate 360
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.8.3</td>
<td>Modelling detection probabilities</td>
<td>360</td>
</tr>
<tr>
<td>11.8.4</td>
<td>An example: gray whales</td>
<td>364</td>
</tr>
<tr>
<td>11.9</td>
<td>Estimation with distance measurement errors</td>
<td>371</td>
</tr>
<tr>
<td>11.9.1</td>
<td>Conventional distance sampling: ( g(0) = 1 )</td>
<td>371</td>
</tr>
<tr>
<td>11.9.2</td>
<td>Independent multiplicative measurement errors</td>
<td>373</td>
</tr>
<tr>
<td>11.9.3</td>
<td>Mark-recapture distance sampling: ( p(0) &lt; 1 )</td>
<td>375</td>
</tr>
<tr>
<td>11.9.4</td>
<td>Maximum likelihood vs pdf correction approach</td>
<td>377</td>
</tr>
<tr>
<td>11.10</td>
<td>Relating object abundance to population abundance for indirect sampling</td>
<td>377</td>
</tr>
<tr>
<td>11.10.1</td>
<td>Introduction</td>
<td>377</td>
</tr>
<tr>
<td>11.10.2</td>
<td>Discrete-time modelling</td>
<td>378</td>
</tr>
<tr>
<td>11.10.3</td>
<td>Continuous-time modelling</td>
<td>378</td>
</tr>
<tr>
<td>11.10.4</td>
<td>Conclusions</td>
<td>384</td>
</tr>
<tr>
<td>11.11</td>
<td>Goodness of fit tests and q-q plots</td>
<td>385</td>
</tr>
<tr>
<td>11.11.1</td>
<td>Quantile-quantile plots</td>
<td>385</td>
</tr>
<tr>
<td>11.11.2</td>
<td>Kolmogorov-Smirnov test</td>
<td>387</td>
</tr>
<tr>
<td>11.11.3</td>
<td>Cramér-von Mises test</td>
<td>388</td>
</tr>
<tr>
<td>11.11.4</td>
<td>The Cramér-von Mises family of tests</td>
<td>388</td>
</tr>
<tr>
<td>11.12</td>
<td>Pooling robustness</td>
<td>389</td>
</tr>
</tbody>
</table>

**References**

**Index**