Addendum to the
Natura Impact Statement

NISA
North Irish Sea Array

Volume 2: Appendices

Appendix A15

MRSea Modelling for Offshore Ornithology

Statkraft

CIP Copenhagen Infrastructure Partners

GoBe APEM Group

ARUP

North Irish Sea Array Windfarm Ltd

North West Irish Sea MRSea Modelling for Offshore Ornithology

North Irish Sea Array Offshore Wind Farm

2026

This document is supplied on and subject to the terms and conditions of the Contractual Agreement relating to this work, under which this document has been supplied.

Confidentiality

This document is confidential.

All information contained within this document is proprietary to GoBe Consultants Ltd and is disclosed in confidence to the specified parties. Information herein may not be reproduced in whole or in part without the express permission from GoBe Consultants Ltd.

www.gobeconsultants.com

Revision	Date	Status	Author:	Checked by:	Approved by:
1.0	December 2025	Draft 1	JS	JC/JM	CC
2.0	February 2026	Draft 2	JS	JM	CC
3.0	April 2026	Draft 3	JS	JM	CC
4.0	June 2026	Final	JS	WH	CC

Contents

1 Introduction 21

1.1 Project Background....21

1.2 MRSea modelling 23

2 Survey Methodology....26

2.1 Data Collection and Survey Design – Digital Aerial Surveys 26

2.2 Quality Assurance Procedure 28

3 Data Analysis....29

3.1 Model-based Abundance Estimates 29

3.2 Modelling Approach Details 29

3.3 Model Specifications....30

3.4 Spatially Explicit Inference....30

3.5 Model Selection 31

3.6 Prediction Grid 33

3.7 Modelling Distribution....35

3.8 Abundance Estimate Apportionment....35

4 Results....37

4.1 Guillemot 37

Abundance Estimates 37

Density Estimate Plots 39

4.2 Razorbill 41

Abundance Estimates 41

Density Estimate Plots 43

4.3 Kittiwake 45

Abundance Estimates 45

Density Estimate Plots 47

4.4 Herring Gull....49

Abundance Estimates 49

Density Estimate Plot....51

4.5 Great Black-backed Gull....53

Abundance Estimates 53

Density Estimate Plots 55

4.6 Gannet 57
Abundance Estimates 57
Density Estimate Plots 59
4.7 Common Tern 61
Abundance Estimates 61
Density Estimate Plots 63
4.8 Roseate Tern 65
Abundance Estimates 65
Density Estimate Plots 67
4.9 Red-throated Diver 69
Abundance Estimates 69
Density Estimate Plots 71
5 Summary 73
5.1 Guillemot 73
5.2 Razorbill 73
5.3 Kittiwake 73
5.4 Herring gull 73
5.5 Great black-backed gull 74
5.6 Gannet 74
5.7 Common tern 74
5.8 Roseate tern 75
5.9 Red-throated Diver 75
6 References 76
A Appendices 78
A.1 MRSea outputs (Original dataset) 78
Guillemot 78
Razorbill 85
A.2 Final Model Summary 92
A.3 Model Setup 94
Guillemot 94
Razorbill 107
Kittiwake 118
Herring Gull 129
Great Black-backed Gull 140

Gannet 151
Common Tern 161
Roseate Tern 172
Red-throated Diver 182
A.4 Final Model Validation 193
Guillemot 193
Razorbill 197
Kittiwake 202
Herring Gull 205
Great Black-backed Gull...208
Gannet 212
Common Tern 216
Roseate Tern 220
Red-throated Diver 223

Figures

Figure A2-1 Ornithological Survey Area (NWIS SPA) and associated DAS transects. 24
Figure A3-1 Ornithological Survey Area, array area, Projected Footprint of Infrastructure, 2km buffers and associated prediction grid utilised for MRSea analysis. 31
Figure A4-1 Model-based density distribution for guillemot within the Survey Area (2024-2025). 37
Figure A4-2 Model-based density distribution for razorbill within the Survey Area (2024-2025). 41
Figure A4-3 Model-based density distribution for kittiwake within the Survey Area (2024-2025). 45
Figure A4-4 Model-based density distribution for herring gull within the Survey Area (2024-2025). .49
Figure A4-5 Model-based density distribution for great black-backed gull within the Survey Area (2024-2025). 53
Figure A4-6 Model-based density distribution for gannet within the Survey Area (2024-2025). 57
Figure A4-7 Model-based density distribution for common tern within the Survey Area (2025). 61
Figure A4-8 Model-based density distribution for roseate tern within the Survey Area (2025). 65
Figure A4-9 Model-based density distribution for red-throated diver within the Survey Area (2024-2025). 69
Figure A6-1 Model-based density distribution for guillemots within the MAC Boundary with a 4km buffer applied, in 2020. 79
Figure A6-2 Model-based density distribution for guillemots within the MAC Boundary with a 4km buffer applied, in 2021. 80
Figure A6-3 Model-based density distribution for guillemots within the MAC Boundary with a 4km buffer applied, in 2022. 81
Figure A6-4 Model-based density distribution for razorbills within the MAC Boundary with a 4km buffer applied, in 2020. 86

Figure A6-5 Model-based density distribution for razorbills within the MAC Boundary with a 4km buffer applied, in 2021. 87
Figure A6-6 Model-based density distribution for razorbills within the MAC Boundary with a 4km buffer applied, in 2022. 88
Figure A6-7 Code excerpt detailing factor level check for categorical variables considered within guillemot MRSea modelling. 91
Figure A6-8 Code excerpt detailing the setup of the initial GLM considered within guillemot MRSea analysis. 91
Figure A6-9 Code excerpt detailing the summary of the initial GLM considered within guillemot MRSea analysis. 92
Figure A6-10 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within guillemot MRSea analysis. 93
Figure A6-11 Code excerpt detailing runs test carried out on the initial GLM considered within guillemot MRSea analysis. 93
Figure A6-12 Runs profiles for explanatory variables considered within the initial GLM considered for guillemot MRSea analysis. 94
Figure A6-13 ACF plot for the initial GLM considered for guillemot MRSea analysis. 95
Figure A6-14 Cumulative residuals of depth plotted for the initial GLM considered for guillemot MRSea analysis. 96
Figure A6-15 Cumulative residuals of distance to colony plotted for the initial GLM considered for guillemot MRSea analysis. 96
Figure A6-16 Cumulative residuals of seabed slope plotted for the initial GLM considered for guillemot MRSea analysis. 97
Figure A6-17 Cumulative residuals of prey density plotted for the initial GLM considered for guillemot MRSea analysis. 97
Figure A6-18 Code excerpt detailing the setup process for the SALSA 1D model considered for guillemot MRSea analysis. 99
Figure A6-19 Code excerpt detailing the summary for the SALSA 1D model for guillemot MRSea analysis. 100
Figure A6-20 Code excerpt for SALSA 2D model setup for guillemot MRSea analysis. 101
Figure A6-21 Code excerpt detailing SALSA 2D summary for guillemot MRSea analysis. 102
Figure A6-22 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for guillemot MRSea analysis. 103
Figure A6-23 Code excerpt detailing factor level check for categorical variables considered within razorbill MRSea modelling. 104
Figure A6-24 Code excerpt detailing the setup of the initial GLM considered within razorbill MRSea analysis. 104
Figure A6-25 Code excerpt detailing the summary of the initial GLM considered within razorbill MRSea analysis. 105
Figure A6-26 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within razorbill MRSea analysis. 106
Figure A6-27 Code excerpt detailing runs test carried out on the initial GLM considered within razorbill MRSea analysis. 106

Figure A6-28 Runs profiles for explanatory variables considered within the initial GLM considered for razorbill MRSea analysis. 107
Figure A6-29 ACF plot for the initial GLM considered for razorbill MRSea analysis. 107
Figure A6-30 Cumulative residuals of depth plotted for the initial GLM considered for razorbill MRSea analysis. 108
Figure A6-31 Cumulative residuals of distance to colony plotted for the initial GLM considered for razorbill MRSea analysis. 108
Figure A6-32 Cumulative residuals of seabed slope plotted for the initial GLM considered for razorbill MRSea analysis. 109
Figure A6-33 Cumulative residuals of prey density plotted for the initial GLM considered for razorbill MRSea analysis. 109
Figure A6-34 Code excerpts detailing the setup process for the SALSA 1D model considered for razorbill MRSea analysis. 110
Figure A6-35 Code excerpt detailing the summary for the SALSA 1D model for razorbill MRSea analysis. 111
Figure A6-36 Code excerpt for SALSA 2D model setup for razorbill MRSea analysis. 112
Figure A6-37 Code excerpt detailing SALSA 2D summary for razorbill MRSea analysis. 113
Figure A6-38 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for razorbill MRSea analysis. 114
Figure A6-39 Code excerpt detailing factor level check for categorical variables considered within kittiwake MRSea modelling. 115
Figure A6-40 Code excerpt detailing the setup of the initial GLM considered within kittiwake MRSea analysis. 115
Figure A6-41 Code excerpt detailing the summary of the initial GLM considered within kittiwake MRSea analysis. 116
Figure A6-42 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within kittiwake MRSea analysis. 116
Figure A6-43 Code excerpt detailing runs test carried out on the initial GLM considered within kittiwake MRSea analysis. 117
Figure A6-44 Runs profiles for explanatory variables considered within the initial GLM considered for kittiwake MRSea analysis. 117
Figure A6-45 ACF plot for the initial GLM considered for kittiwake MRSea analysis. 118
Figure A6-46 Cumulative residuals of depth plotted for the initial GLM considered for kittiwake MRSea analysis. 118
Figure A6-47 Cumulative residuals of distance to colony plotted for the initial GLM considered for kittiwake MRSea analysis. 119
Figure A6-48 Cumulative residuals of seabed slope plotted for the initial GLM considered for kittiwake MRSea analysis. 119
Figure A6-49 Cumulative residuals of prey density plotted for the initial GLM considered for kittiwake MRSea analysis. 120
Figure A6-50 Code excerpt detailing the setup process for the SALSA 1D model considered for kittiwake MRSea analysis. 121
Figure A6-51 Code excerpt detailing the summary for the SALSA 1D model for kittiwake MRSea analysis. 122

Figure A6-52 Code excerpt for SALSA 2D model setup for kittiwake MRSea analysis. 123
Figure A6-53 Code excerpt detailing SALSA 2D summary for kittiwake MRSea analysis. 124
Figure A6-54 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for kittiwake MRSea analysis. 125
Figure A6-55 Code excerpt detailing factor level check for categorical variables considered within herring gull MRSea modelling. 126
Figure A6-56 Code excerpt detailing the setup of the initial GLM considered within herring gull MRSea analysis. 126
Figure A6-57 Code excerpt detailing the summary of the initial GLM considered within herring gull MRSea analysis. 127
Figure A6-58 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within herring gull MRSea analysis. 128
Figure A6-59 Code excerpt detailing runs test carried out on the initial GLM considered within herring gull MRSea analysis. 128
Figure A6-60 Runs profiles for explanatory variables considered within the initial GLM considered for herring gull MRSea analysis. 128
Figure A6-61 ACF plot for the initial GLM considered for herring gull MRSea analysis. 129
Figure A6-62 Cumulative residuals of depth plotted for the initial GLM considered for herring gull MRSea analysis. 130
Figure A6-63 Cumulative residuals of distance to colony plotted for the initial GLM considered for herring gull MRSea analysis. 130
Figure A6-64 Cumulative residuals of seabed slope plotted for the initial GLM considered for herring gull MRSea analysis. 131
Figure A6-65 Cumulative residuals of prey density plotted for the initial GLM considered for herring gull MRSea analysis. 131
Figure A6-66 Code excerpt detailing the setup process for the SALSA 1D model considered for herring gull MRSea analysis. 133
Figure A6-67 Code excerpt detailing the summary for the SALSA 1D model for herring gull MRSea analysis. 134
Figure A6-68 Code excerpt for SALSA 2D model setup for herring gull MRSea analysis. 135
Figure A6-69 Code excerpt detailing SALSA 2D summary for herring gull MRSea analysis. 135
Figure A6-70 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for herring gull MRSea analysis. 136
Figure A6-71 Code excerpt detailing factor level check for categorical variables considered within great black-backed gull MRSea analysis. 137
Figure A6-72 Code excerpt detailing the setup of the initial GLM considered within great black-backed gull MRSea analysis. 137
Figure A6-73 Code excerpt detailing the summary of the initial GLM considered within great black-backed gull MRSea analysis. 138
Figure A6-74 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within great black-backed gull MRSea analysis. 138
Figure A6-75 Code excerpt detailing runs test carried out on the initial GLM considered within great black-backed gull MRSea analysis. 139

Figure A6-76 Runs profiles for explanatory variables considered within the initial GLM considered for great black-backed gull MRSea analysis. 139
Figure A6-77 ACF plot for the initial GLM considered for great black-backed gull MRSea analysis. ..140
Figure A6-78 Cumulative residuals of depth plotted for the initial GLM considered for great black-backed gull MRSea analysis. 140
Figure A6-79 Cumulative residuals of distance to colony plotted for the initial GLM considered for great black-backed gull MRSea analysis. 141
Figure A6-80 Cumulative residuals of seabed slope plotted for the initial GLM considered for great black-backed gull MRSea analysis. 141
Figure A6-81 Cumulative residuals of prey density plotted for the initial GLM considered for great black-backed gull MRSea analysis. 142
Figure A6-82 Code excerpt detailing the setup process for the SALSA 1D model considered for great black-backed gull MRSea analysis. 143
Figure A6-83 Code excerpt detailing the summary for the SALSA 1D model for great black-backed gull MRSea analysis. 144
Figure A6-84 Code excerpt for SALSA 2D model setup for great black-backed gull MRSea analysis. 145
Figure A6-85 Code excerpt detailing SALSA 2D summary for great black-backed gull MRSea analysis. 146
Figure A6-86 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for great black-backed gull MRSea analysis....147
Figure A6-87 Code excerpt detailing factor level check for categorical variables considered within gannet MRSea modelling....148
Figure A6-88 Code excerpt detailing the setup of the initial GLM considered within gannet MRSea analysis....148
Figure A6-89 Code excerpt detailing the summary of the initial GLM considered within gannet MRSea analysis....149
Figure A6-90 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within gannet MRSea analysis....149
Figure A6-91 Code excerpt detailing runs test carried out on the initial GLM considered within gannet MRSea analysis. 150
Figure A6-92 Runs profiles for explanatory variables considered within the initial GLM considered for gannet MRSea analysis. 150
Figure A6-93 ACF plot for the initial GLM considered for gannet MRSea analysis. 151
Figure A6-94 Cumulative residuals of depth plotted for the initial GLM considered for gannet MRSea analysis....152
Figure A6-95 Cumulative residuals of distance to colony plotted for the initial GLM considered for gannet MRSea analysis. 152
Figure A6-96 Cumulative residuals of seabed slope plotted for the initial GLM considered for gannet MRSea analysis. 153
Figure A6-97 Cumulative residuals of prey density plotted for the initial GLM considered for gannet MRSea analysis. 153
Figure A6-98 Code excerpt detailing the setup process for the SALSA 1D model considered for gannet MRSea analysis. 154

Figure A6-99 Code excerpt detailing the summary for the SALSA 1D model for gannet MRSea analysis. 155
Figure A6-100 Code excerpt for SALSA 2D model setup for gannet MRSea analysis. 156
Figure A6-101 Code excerpt detailing SALSA 2D summary for gannet MRSea analysis. 156
Figure A6-102 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for gannet MRSea analysis. 157
Figure A6-103 Code excerpt detailing factor level check for categorical variables considered within guillemot MRSea modelling. 158
Figure A6-104 Code excerpt detailing the setup of the initial GLM considered within common tern MRSea analysis. 158
Figure A6-105 Code excerpt detailing the summary of the initial GLM considered within common tern MRSea analysis. 159
Figure A6-106 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within common tern MRSea analysis. 159
Figure A6-107 Code excerpt detailing runs test carried out on the initial GLM considered within common tern MRSea analysis. 160
Figure A6-108 Runs profiles for explanatory variables considered within the initial GLM considered for common tern MRSea analysis. 160
Figure A6-109 ACF plot for the initial GLM considered for common tern MRSea analysis. 161
Figure A6-110 Cumulative residuals of depth plotted for the initial GLM considered for common tern MRSea analysis. 162
Figure A6-111 Cumulative residuals of distance to colony plotted for the initial GLM considered for common tern MRSea analysis. 162
Figure A6-112 Cumulative residuals of distance to coast plotted for the initial GLM considered for common tern MRSea analysis. 163
Figure A6-113 Cumulative residuals of seabed slope plotted for the initial GLM considered for common tern MRSea analysis. 163
Figure A6-114 Cumulative residuals of prey density plotted for the initial GLM considered for common tern MRSea analysis. 164
Figure A6-115 Code excerpt detailing the setup process for the SALSA 1D model considered for common tern MRSea analysis. 165
Figure A6-116 Code excerpt detailing the summary for the SALSA 1D model for common tern MRSea analysis. 166
Figure A6-117 Code excerpt for SALSA 2D model setup for common tern MRSea analysis. 167
Figure A6-118 Code excerpt detailing SALSA 2D summary for common tern MRSea analysis. 167
Figure A6-119 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for common tern MRSea analysis. 168
Figure A6-120 Code excerpt detailing factor level check for categorical variables considered within roseate tern MRSea modelling. 169
Figure A6-121 Code excerpt detailing the setup of the initial GLM considered within roseate tern MRSea analysis. 169
Figure A6-122 Code excerpt detailing the summary of the initial GLM considered within roseate tern MRSea analysis. 169

Figure A6-123 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within roseate tern MRSea analysis. 170
Figure A6-124 Code excerpt detailing runs test carried out on the initial GLM considered within roseate tern MRSea analysis. 170
Figure A6-125 Runs profiles for explanatory variables considered within the initial GLM considered for roseate tern MRSea analysis. 171
Figure A6-126 ACF plot for the initial GLM considered for roseate tern MRSea analysis. 172
Figure A6-127 Cumulative residuals of depth plotted for the initial GLM considered for roseate tern MRSea analysis. 172
Figure A6-128 Cumulative residuals of distance to colony plotted for the initial GLM considered for roseate tern MRSea analysis. 173
Figure A6-129 Cumulative residuals of seabed slope plotted for the initial GLM considered for roseate tern MRSea analysis. 173
Figure A6-130 Cumulative residuals of prey density plotted for the initial GLM considered for roseate tern MRSea analysis. 174
Figure A6-131 Code excerpt detailing the setup process for the SALSA 1D model considered for roseate tern MRSea analysis. 175
Figure A6-132 Code excerpt detailing the summary for the SALSA 1D model for roseate tern MRSea analysis. 176
Figure A6-133 Code excerpt for SALSA 2D model setup for roseate tern MRSea analysis. 177
Figure A6-134 Code excerpt detailing SALSA 2D summary for roseate tern MRSea analysis. 177
Figure A6-135 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for roseate tern MRSea analysis. 178
Figure A6-136 Code excerpt detailing factor level check for categorical variables considered within red-throated diver MRSea modelling. 179
Figure A6-137 Code excerpt detailing the setup of the initial GLM considered within red-throated diver MRSea analysis. 179
Figure A6-138 Code excerpt detailing the summary of the initial GLM considered within red-throated diver MRSea analysis. 179
Figure A6-139 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within red-throated diver MRSea analysis. 180
Figure A6-140 Code excerpt detailing runs test carried out on the initial GLM considered within red-throated diver MRSea analysis. 180
Figure A6-141 Runs profiles for explanatory variables considered within the initial GLM considered for red-throated diver MRSea analysis. 181
Figure A6-142 ACF plot for the initial GLM considered for red-throated diver MRSea analysis. 182
Figure A6-143 Cumulative residuals of depth plotted for the initial GLM considered for red-throated diver MRSea analysis. 182
Figure A6-144 Cumulative residuals of distance to coast plotted for the initial GLM considered for red-throated diver MRSea analysis. 183
Figure A6-145 Cumulative residuals of distance to colony plotted for the initial GLM considered for red-throated diver MRSea analysis. 183
Figure A6-146 Cumulative residuals of seabed slope plotted for the initial GLM considered for red-throated diver MRSea analysis. 184

Figure A6-147 Cumulative residuals of prey density plotted for the initial GLM considered for red-throated diver MRSea analysis. 184
Figure A6-148 Code excerpt detailing the setup process for the SALSA 1D model considered for red-throated diver MRSea analysis. 186
Figure A6-149 Code excerpt detailing the summary for the SALSA 1D model for red-throated diver MRSea analysis. 187
Figure A6-150 Code excerpt for SALSA 2D model setup for red-throated diver MRSea analysis. 188
Figure A6-151 Code excerpt detailing SALSA 2D summary for red-throated diver MRSea analysis. .188
Figure A6-152 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for red-throated diver MRSea analysis. 189
Figure A6-153 Cumulative residual plot of depth within the final SALSA 2D model selected for guillemot MRSea analysis. 190
Figure A6-154 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for guillemot MRSea analysis. 191
Figure A6-155 Cumulative residual plot of prey density within the final SALSA 2D model selected for guillemot MRSea analysis. 191
Figure A6-156 Cumulative residual plot of seabed slope within the final SALSA 2D model selected for guillemot MRSea analysis. 192
Figure A6-157 Observed vs fitted values plot for the final SALSA 2D model selected for guillemot MRSea analysis. 192
Figure A6-158 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for guillemot MRSea analysis. 193
Figure A6-159 Mean-variance plot for the final SALSA 2D model selected for guillemot MRSea analysis. 194
Figure A6-160 Cumulative residual plot of depth within the final SALSA 2D model selected for razorbill MRSea analysis. 195
Figure A6-161 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for razorbill MRSea analysis. 196
Figure A6-162 Cumulative residual plot of prey density within the final SALSA 2D model selected for razorbill MRSea analysis. 196
Figure A6-163 Observed vs fitted values plot for the final SALSA 2D model selected for razorbill MRSea analysis. 197
Figure A6-164 Scaled Pearsons Residuals vs fitted values plot for the final SALSA 2D model selected for razorbill MRSea analysis. 197
Figure A6-165 Mean-variance plot for the final SALSA 2D model selected for razorbill MRSea analysis. 198
Figure A6-166 Cumulative residual plot of depth within the final SALSA 2D model selected for kittiwake MRSea analysis. 199
Figure A6-167 Observed vs fitted values plot for the final SALSA 2D model selected for kittiwake MRSea analysis. 200
Figure A6-168 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for kittiwake MRSea analysis. 200
Figure A6-169 Mean-variance plot for the final SALSA 2D model selected for kittiwake MRSea analysis. 201

Figure A6-170 Cumulative residual plot of depth within the final SALSA 2D model selected for herring gull MRSea analysis. 202
Figure A6-171 Observed vs fitted values plot for the final SALSA 2D model selected for herring gull MRSea analysis. 203
Figure A6-172 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for herring gull MRSea analysis. 203
Figure A6-173 Mean-variance plot for the final SALSA 2D model selected for herring gull MRSea analysis. 204
Figure A6-174 Cumulative residual plot of depth within the final SALSA 2D model selected for great black-backed gull MRSea analysis. 205
Figure A6-175 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for great black-backed gull MRSea analysis. 206
Figure A6-176 Observed vs fitted values plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis. 206
Figure A6-177 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis. 207
Figure A6-178 Mean-variance plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis. 208
Figure A6-179 Cumulative residual plot of depth within the final SALSA 2D model selected for gannet MRSea analysis. 209
Figure A6-180 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for gannet MRSea analysis. 210
Figure A6-181 Observed vs fitted values plot for the final SALSA 2D model selected for gannet MRSea analysis. 210
Figure A6-182 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for gannet MRSea analysis. 211
Figure A6-183 Mean-variance plot for the final SALSA 2D model selected for gannet MRSea analysis. 212
Figure A6-184 Cumulative residual plot of depth within the final SALSA 2D model selected for common tern MRSea analysis. 213
Figure A6-185 Cumulative residual plot of distance to coast within the final SALSA 2D model selected for common tern MRSea analysis. 214
Figure A6-186 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for common tern MRSea analysis. 214
Figure A6-187 Observed vs fitted values plot for the final SALSA 2D model selected for common tern MRSea analysis. 215
Figure A6-188 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for common tern MRSea analysis. 215
Figure A6-189 Mean-variance plot for the final SALSA 2D model selected for common tern MRSea analysis. 216
Figure A6-190 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for roseate tern MRSea analysis. 217
Figure A6-191 Observed vs fitted values plot for the final SALSA 2D model selected for roseate tern MRSea analysis. 218

Figure A6-192 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for roseate tern MRSea analysis. 218
Figure A6-193 Mean-variance plot for the final SALSA 2D model selected for roseate tern MRSea analysis. 219
Figure A6-194 Cumulative residual plot of distance to coast within the final SALSA 2D model selected for red-throated diver MRSea analysis. 220
Figure A6-195 Observed vs fitted values plot for the final SALSA 2D model selected for red-throated diver MRSea analysis. 221
Figure A6-196 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for red-throated diver MRSea analysis. 221
Figure A6-197 Mean-variance plot for the final SALSA 2D model selected for red-throated diver MRSea analysis. 222

Tables

Table A1-1 RFI submissions from the An Coimisiún Pleanála relevant to the MRSea Modelling 18
Table A1-2 A summary of the species considered for MRSea analysis and the months used on a species-by-species basis. 21
Table A3-1 DAS survey information included within MRSea analysis. 28
Table A3-2 Final model used in species-specific analysis. 29
Table A3-3 Environmental variables, and associated data sources, considered for MRSea analysis. ..32
Table A4-1 The raw count of individual guillemot recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 35
Table A4-2 The raw count of individual guillemot recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 35
Table A4-3 The raw count of individual razorbill recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 39
Table A4-4 The raw count of individual razorbill recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 39
Table A4-5 The raw count of individual kittiwake recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 43
Table A4-6 The raw count of individual kittiwake recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates

(CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 43

Table A4-7 The raw count of individual herring gull recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 47

Table A4-8 The raw count of individual herring gull recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 47

Table A4-9 The raw count of individual great black-backed gull recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 51

Table A4-10 The raw count of individual great black-backed gull recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. 51

Table A4-11 The raw count of individual gannet recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 55

Table A4-12 The raw count of individual gannet recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 55

Table A4-13 The raw count of individual common tern recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 59

Table A4-14 The raw count of individual common tern recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 59

Table A4-15 The raw count of individual roseate tern recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 63

Table A4-16 The raw count of individual roseate tern recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 63

Table A4-17 The raw count of individual red-throated diver recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 67

Table A4-18 The raw count of individual red-throated diver recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model. 67

Table A6-1 The raw count of individual guillemots recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer. 76

Table A6-2 The raw count of individual guillemots recorded within each survey, alongside apportioned, and corrected for availability bias, design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer. 77

Table A6-3 The raw count of individual razorbills recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer. 83

Table A6-4 The raw count of individual razorbills recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer. 84

Table A6-5 Final MRSea model selections for each species. 89

Acronyms

Term	Definition
ACF	Autocorrelation function
ANOVA	Analysis of Variance
CI	Confidence interval
CREEM	Centre for Research into Ecological and Environmental Modelling
CReSS	Complex Region Spatial Smoother
CTR	Control zone
DAS	Digital aerial survey
DBE	Design-based estimate
ECC	Export Cable Corridor
EIAR	Environmental Impact Assessment Report
GAM	Generalised Additive Model
GVIF	Generalised variance inflation factors
LOD	Limit of Deviation
MAC	Maritime Area Consent
MRSea	Marine Renewables Strategic Environmental Assessment
NISA	The North Irish Sea Array
NWIS	North-west Irish Sea
OWF	Offshore wind farm
PFI	Projected Footprint of Infrastructure
QA	Quality assurance
SALSA	Spatially Adaptive Local Smoothing Algorithm
SPA	Special Protection Area

1 Introduction

1.1.1 North Irish Sea Array Windfarm Ltd (NISA, hereafter referred to as ‘the Developer’) has been considering the Request for Further Information (RFI) issued by An Bord Pleanála (now An Coimisiún Pleanála) as well as the third-party submissions received following public consultation. At An Coimisiún Pleanála’s behest, the Developer has also continued to consult with stakeholders in respect of the 2024 planning application throughout 2024-2026. The Developer has refined elements of the design to respond to the third-party submissions, the continued public and stakeholder consultation and the RFI. Full details of consultation undertaken can be found in Appendix A1.2 Consultation Report.

1.1.2 Any cross reference to a chapter, section, table, image, figure or appendix within this document is to another location within the Addendum to the EIAR unless explicitly stated otherwise. Any cross reference to anything included in the 2024 EIAR will be clearly labelled as such.

1.1.3 The sections relevant to MRSea Modelling are included below.

Table A1-1 RFI submissions from the An Coimisiún Pleanála relevant to the MRSea Modelling

RFI Section	RFI	Relevance to Chapter
1b	The scientific information provided as part of the planning application documentation should be based on up-to-date survey reports and data. Accordingly, the applicant is requested to confirm/provide justification/verification that the information submitted in support of the planning application remains relevant and appropriate at the point of submitting further information or to update same as required.	The timeframes associated with the RFI have necessitated a review of the datasets previously used in the 2024 EIAR to ensure any necessary updates to the baseline environment are captured. This report ensures that additional data has fed into the ornithological assessment and therefore remains valid and representative of up to date information.
1c	The applicant is requested to confirm whether any on-going or additional surveying has been carried out since the application was lodged and, if so, the applicant is invited to submit any further survey data results and analysis and update the planning application documentation, as appropriate.	Additional ornithological surveys have been undertaken since submission of the Application. These additional surveys are described in Section 15.2.5 of Chapter 15: Offshore Ornithology.

1.1 Project Background

1.1.1 This document has been prepared by GoBe Consultants Limited (GoBe) on behalf of the Developer to accompany Chapter 15: Offshore and Intertidal Ornithology. It is a technical report that sits alongside the technical baseline (Appendix A15.1 Offshore and Intertidal Ornithology Technical Baseline; hereafter referred to as the ‘Technical Baseline’) detailing the methodology used to derive model based abundance estimates and distributions from raw observations obtained from Digital Aerial Surveys (DAS).

1.1.2 The North Irish Sea Array (NISA) Offshore Wind Farm (OWF) (hereafter the ‘proposed development’) is proposed for construction 11.3 km off the east coast of Ireland (at the nearest point to the mainland). The proposed development will consist of offshore wind turbine generators (WTGs), an offshore substation platform (OSP), inter-array cables and export cables. The area considered in the context of offshore ornithological receptors includes the entire proposed development array area, covering 88.5 km^2, an asymmetric 2 km buffer surrounding the array area (201.6 km^2), and the offshore Export Cable Corridor (ECC) covering a further 67.9 km^2.

1.1.3 Since the original 2024 EIAR submission, and in response to the RFI, the Developer has refined the WTG layout (refer to Appendix A5.1: Design Refinements for more information) which has resulted in a reduction in spatial extent of the layout. For the purposes of the ornithology assessment only this refined layout is represented by a Projected Footprint of Infrastructure (PFI)[1], which includes both project option layouts, and the limit of deviation (LOD), and covers an area of 57.7 km^2 (Figure A3-1). However, it should be noted only one layout will be constructed.

1.1.4 This equates to a 70% reduction from the original Maritime Area Consent (MAC) boundary. The Developer has also undertaken an additional 12 months of DAS covering the entire North-west Irish Sea (NWIS) special protection area (SPA) in response to several Sections of the RFI, including but not limited to:

- 8. (a) iii. …*It is requested that the applicant considers undertaking targeted surveys covering one winter period, with two surveys per month undertaken in critical months for wintering Red-throated Diver.*
- 8. (c) iii. …*The applicant should consider displacement effects up to 10 km from the proposed array area during operation.*
- 8. (c) vi. *Baseline Data Vintage: The Board queries the age and relevance of the submitted aerial and boat-based survey data used in the application… …The applicant is requested to provide justification that the original digital aerial surveys and boat-based data remains relevant and appropriate at the point of submitting additional information to support the proposed development.*
- 8. (c) ix. *The applicant is requested to apply more appropriate regional population estimates to these species and revise the baseline and assessment accordingly.*
- 8 (e) vi. *North-west Irish Sea cSPA Common Guillemot: …The applicant is requested to justify interpretation of the data in relation to Common Guillemot and, where appropriate, re-evaluate the data and re-interpret the consequences for the impacts on the Conservation Objectives of the NWIS cSPA, having regard to the observation from the DAU.*

---

[1] The PFI represents a defined ornithology study sub-area within the overall array area, where all permanent offshore array infrastructure will be located. The PFI has been developed following design refinements to minimise spatial overlap with sensitive receptors and occupies just 57.7 km^2, representing 2.5% of the NWIS cSPA.

1.1.5 The Developer considers that the additional 12 months DAS dataset addresses the RFI by providing robust, recent, site-specific data on the abundance and distributions across the full NWIS SPA as well as specific regions of interest within it. This is a powerful dataset that informs the assessment, most notably, the 'population size' and 'change in distribution' attribute of the NWIS SPA conservation objectives (NPWS, 2023).

1.1.6 Also within this report are the updated abundances clipped to the PFI calculated from the original 29 months of DAS using the original MRSea model for guillemot and razorbill. These can be found in Section A.1. Only the abundances have been updated by being clipped to the PFI, and 2km buffer, as opposed to the array area and associated buffer - full details of the methodology can be found in the original Appendix 15.2 MRSea Modelling for Offshore Ornithology of the 2024 EIAR.

1.2 MRSea modelling

1.2.1 This report summarises analysis of DAS undertaken at the NWIS SPA (hereafter 'NWIS DAS'). The survey programme consisted of 12 surveys, with one per month conducted between September 2024 and August 2025 inclusive. Full details of survey design and specifications can be found in the Technical Baseline.

1.2.2 GoBe were tasked with providing model-based abundance and density estimates for several species using the Marine Renewables Strategic Environmental Assessment (MRSea) modelling framework, created specifically for offshore wind development (Scott-Hayward *et al.*, 2013). A single model run was completed for each species, and included all behaviours, to estimate seabird receptor distribution across all surveys.

1.2.3 The aim of this analysis was to support the impact assessment of the proposed development and to provide comparison to the design-based estimates (abundance estimates used within displacement modelling) within the EIAR. MRSea was carried out for the NWIS SPA over a period of 12 months in addition to the pre-existing modelling (guillemot and razorbill only) carried out for 29 months across the MAC Boundary plus 4km buffer. While the outputs of these two analyses have not been combined within this appendix, MRSea outputs from both sources have been clipped to the new PFI boundary to provide abundance and density estimates. For the 12-month dataset, these outputs are presented in section 4. The updated abundance and density estimates, clipped to the PFI and associated 2km buffer, from the original 29-month DAS dataset can be found in Appendix A15.1.

1.2.4 The distribution of seabird species within the PFI, estimated from models, can be compared with the distributions predicted from raw observations to get a better understanding of area usage in relation to environmental variables. Where necessary, information about 'hotspots' (i.e. areas that regularly support high densities of birds) can be used to help inform the EIA process in terms of understanding area usage relative to the location of the proposed development. This analysis is additional to the standard assessment guidance and was undertaken in acknowledgement that the proposed development now sits within the NWIS SPA, which was not the case when its MAC was awarded. Therefore, a robust assessment and verification of outcomes on impacts to qualifying interests is of high importance.

1.2.5 Species that were frequently detected within raw observation data in varying densities throughout all months of the survey period (September 2024 – August 2025), were prioritised for MRSea analysis. This included guillemot, razorbill, kittiwake, herring gull and great black-backed gull. Following this, species with high numbers of observations within core months of biological importance (e.g. breeding periods) were then analysed. This included gannet, common tern, roseate tern, red-throated diver and common scoter. All species analysed are also qualifying interests of the NWIS SPA and potentially sensitive to displacement. A summary of species considered for MRSea analysis, and the months used on a species-by-species basis, is provided below (Table A1-2).

1.2.6 Raw observations for some species (e.g. Manx Shearwater) exceeded the MRSea guidance threshold in certain months, but these occurrences were limited and not sustained across the dataset. Incorporating these sporadic peaks led to model instability and non-convergence, resulting in an unreliable model fit. Consequently, MRSea analyses were not undertaken for these species. The full list of species excluded from MRSea analysis is provided in Table A1-2.

1.2.7 Common scoter were excluded from MRSea analysis due to failed model convergence. The raw observations were highly clustered near shore and collected over only a few months, providing insufficient spatial and temporal coverage for reliable model smoothing or prediction. This is required for accurate, and ecologically meaningful, model estimates to be created. Therefore, design-based estimates have been used for this species.

Table A1-2 A summary of the species considered for MRSea analysis and the months used on a species-by-species basis.

Species	Months Used in MRSea Analysis
Arctic Skua	Excluded from MRSea analysis
Arctic tern	Excluded from MRSea analysis
Black guillemot	Excluded from MRSea analysis
Black-headed gull	Excluded from MRSea analysis
Common gull	Excluded from MRSea analysis
Common scoter	Excluded from MRSea analysis
Common tern	May 2025 – August 2025
Cormorant	Excluded from MRSea analysis
Fulmar	Excluded from MRSea analysis
Gannet	September 2024 – November 2024 & February 2025 – August 2025
Great black-backed gull	All survey months.
Great crested grebe	Excluded from MRSea analysis
Great northern diver	Excluded from MRSea analysis
Great skua	Excluded from MRSea analysis
Guillemot	All survey months.

Herring gull	All survey months.
Kittiwake	All survey months.
Little tern	Excluded from MRSea analysis
Oystercatcher	Excluded from MRSea analysis
Puffin	Excluded from MRSea analysis
Razorbill	All survey months.
Red-throated diver	September 2024 – May 2025
Roseate tern	May 2025 – August 2025

2 Survey Methodology

2.1 Data Collection and Survey Design – Digital Aerial Surveys

2.1.1 The NWIS DAS data were collected via monthly transect surveys across the NWIS SPA (2,333 km^2) shown in Figure A2-1.

2.1.2 The original survey design ('Plan A') for the NWIS SPA (referred to herein as the 'Survey Area') comprised contiguous east-west transects across the Survey Area, targeting 18% coverage. Due to restricted access to the Dublin Airport Control Zone (CTR), an alternative design ('Plan B') was implemented, maintaining alignment with Plan A where possible while avoiding the CTR and adding two north-south transects. For Plan B there was a target coverage of 15%. When no flightlines within the CTR were possible, Plan C, also providing 15% coverage, was implemented. Further details of surveys used within MRSea analysis, and their associated plan type are included in section 3.4.

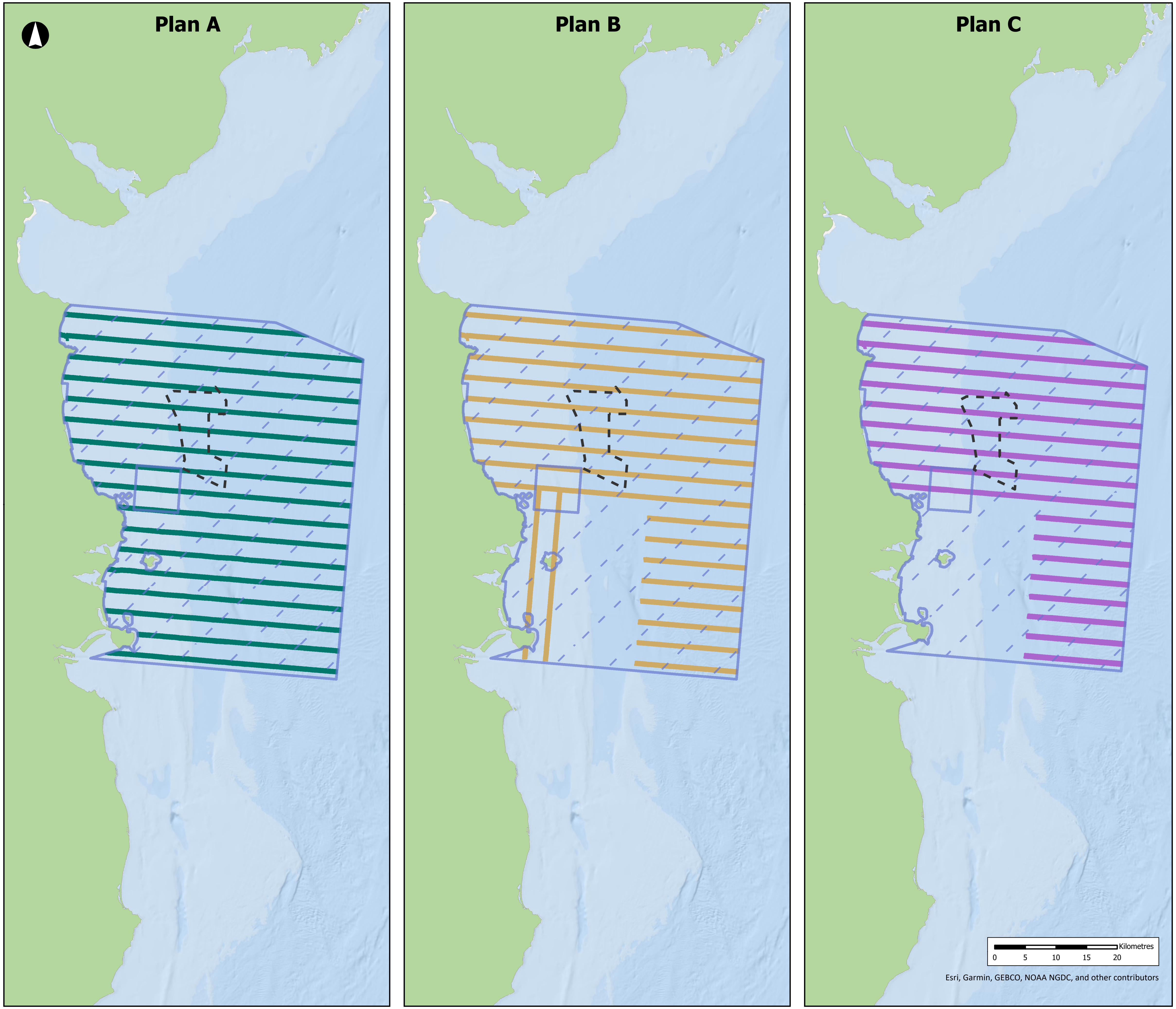
2.1.3 The data collected were 1.5 cm ground survey distance digital still images, by a twin-engine aircraft flying at an altitude of 1,300 feet at a speed of approximately 120 knots, with GPS-linked flight management ensuring accurate track adherence. Images were collected along 17-19 transects across the Survey Area, with a minimum of 15% of the sea surface covered during each survey. Surveys were also conducted under the following environmental conditions:

- Cloud base: >1,700 feet;
- Visibility: >5 km;
- Windspeed: <30 knots;
- Sea state: 4 or less (Beaufort 5 or 6); and
- No icing conditions.

2.1.4 Seabirds were identified to species level during image analysis. Where identification to species level was not possible, individuals were classified to the lowest taxonomic level. To ensure the accuracy of the species identification, the survey contractor (APEM Ltd) conducted internal quality assurance which guaranteed any missed individuals were included in the data. The data collected during the NWIS DAS provided the following information:

- Date and time of each seabird and recorded during a survey;
- Corresponding coordinates for each seabird recorded;
- Age, sex and moult status of seabirds, where possible;
- Additional behavioural information whether a bird is sitting, flying, or diving; and
- Estimated flight heights, where possible.

2.1.5 Further details regarding data collection and survey design are available in Appendix A15.1 Offshore and Intertidal Ornithology Technical Baseline.



Esri, Garmin, GEBCO, NOAA NGDC, and other contributors

NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

NWIS SPA DAS Survey Extent (2024 - 2025)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:600,000 @A3	**A2-1**
Status: Issue	

2.2 Quality Assurance Procedure

2.2.1 A standard internal Quality Assurance (QA) procedure was carried out for each survey. Images were assessed in batches with a different staff member responsible for each batch. Each bird image was reviewed and checked by APEM's own dedicated QA manager, ensuring that 100% of birds found in the images were subject to internal QA. The QA manager, an experienced ornithologist, is responsible for maintaining and updating the image library and provides advice and guidance to the image processing staff. Images containing no birds were removed and kept separately for further QA. Of these 'blank' images, 20% were randomly selected for the QA process.

3 Data Analysis

3.1 Model-based Abundance Estimates

3.1.1 To provide more detail for seabird species within the Survey Area, a model-based approach was used to determine statistically robust, spatially distributed, population estimates. Using model-based techniques means that environmental variables can be included within the analysis to help predict abundance and density distributions within the Survey Area or any geographical subset (e.g. the PFI). MRSea analysis was used to generate estimates of seabird distribution and abundance, underpinned by raw observation data recorded in the NWIS DAS imagery (Scott-Hayward *et al.,* 2014).

3.1.2 MRSea is a statistical analysis software package specifically designed by the Centre for Research into Ecological and Environmental Modelling (CREEM) for both baseline analysis (site characterisation) and, where data are available, pre- and post-construction analysis (Scott-Hayward *et al.*, 2013). The latter can be used for investigating potential changes in distributions of birds following an offshore wind farm's (OWF) development. The model uses a "Complex Region Spatial Smoother" spatial modelling technique with a "Spatially Adaptive Local Smoothing Algorithm" (CReSS-SALSA) to estimate bird distributions in a Generalised Additive Model (GAM) framework. This approach was developed to analyse spatial abundance data following environmental change, such as the construction of a wind farm, and allows spatially auto-correlated and zero-inflated data to be modelled in a robust way (Onoufriou *et al.*, 2021).

3.2 Modelling Approach Details

3.2.1 The steps used to fit an MRSea model are described below in general terms. Actual model fit and the environmental variables included in the final outputs varied on a species-by-species basis. Full final model validation plots are presented in Appendix A.4.

3.2.2 The model was based on a spatially adaptive GAM to permit nonlinear relationships on the link scale for each candidate environmental covariate. Generalised variance inflation factors (GVIF) were used to test for collinearity between variables within each model (Fox & Weisberg, 2015). Variables with a GVIF value >5 were removed from further analysis at this stage on a model-by-model basis. Correlation between variables was also assessed by visual examination using correlation plots provided as part of the suite of diagnostics produced within the MRSea package.

3.2.3 The X and Y coordinates were included as a two-dimensional spatial smoother. Survey was included as a factor variable within the model which also included an interaction term between the survey variable and the smoothed spatial term (X and Y coordinates). This allowed for the knot coefficients and positions to vary between surveys if appropriate.

3.3 Model Specifications

3.3.1 Due to the nature of seabird count data being highly variable, it generally displays properties of an over-dispersed Poisson distribution. Models were assessed visually to determine the most appropriate error terms to account for high variability. It was also important to consider that data collected along transects and within repeated surveys may lead to spatial and/or temporal correlation. This was assessed as part of the model process with runs tests and autocorrelation function plots being visually inspected.

3.3.2 A CReSS method was used to fit the spatial density surface for each model, due to its ability to handle complex geographic features such as coastlines and islands (Scott-Hayward *et al.* 2014). Model flexibility was determined by both the number of 'knots' used (i.e. anchor points) for the model and the effective range (r) of the basis associated with each knot. SALSA2D was used to determine the number/location of knots and the r parameters with quantile regression Bayesian Information Criterion (QBIC) as the selection criterion.

3.4 Spatially Explicit Inference

3.4.1 The data used within the modelling process were collected from a DAS programme, across the Survey Area, conducted from September 2024 to August 2025 (inclusive). Information regarding survey number, plan type and if fishing vessel activity was present is also provided (Table A3-1). One survey per month was carried out. Following current guidance, where 10 or less observations of a species were present within a month, that survey was removed from analysis to avoid issues when running spatial elements of the modelling process (NatureScot, 2023).

3.4.2 When locations are close together geographically (and sometimes temporally), they tend to have similar counts. If the model does not include environmental variables that explain these patterns, the residuals (differences between observed and predicted values) will show spatial correlation. This violates a key assumption in many statistical models, such as GAMS, which require residuals to be independent. If this assumption is broken, estimates of precision can become unreliable. To address this, when residual correlation was detected, robust standard errors were applied. This adjusts for autocorrelation and provides more realistic estimates of uncertainty.

3.4.3 As part of the model setup process, a blocking structure was chosen such that the residuals were permitted to be correlated within blocks but could be deemed independent between blocks. Survey ID (month as a numeric variable), concatenated with Transect ID, was specified as the blocking structure. This ensured that the model should treat data from within each transect of each survey as correlated but be independent between different transects and surveys. These independence assumptions were validated by visual assessment of autocorrelation function (ACF) plots (See Appendix A.3).

Table A3-1 DAS survey information included within MRSea analysis.

DAS							
Survey No.	Survey Month 2024	Survey Plan	Fishing Vessel Activity Present	Survey No.	Survey Month 2025	Survey Plan	Fishing Vessel Activity Present
1	September	Plan B	No	5	January	Plan B	Yes
2	October	Plan B	No	6	February	Plan B	Yes
3	November	Plan A	Yes	7	March	Plan B	Yes
4	December	Plan B	Yes	8	April	Plan C	Yes
				9	May	Plan C	Yes
				10	June	Plan B	Yes
				11	July	Plan B	Yes
				12	August	Plan B	Yes

3.5 Model Selection

3.5.1 The MRSea modelling process starts with a basic GAM and then adds flexibility using SALSA in one and two dimensions. In the one-dimensional SALSA (SALSA1D) step, an initial knot is placed at the median of each variable. During model optimisation, additional knots are added where extra flexibility is most needed within the covariate range. A selection procedure is also included at this point which automatically determines if covariates should be smoothed or linear or removed from the model entirely.

3.5.2 For the two-dimensional SALSA (SALSA2D) step, initial knots are positioned to provide good coverage across the spatial area (on a survey-by-survey basis), and these can move as the model selects the best configuration. The Quasi-Bayesian Information Criterion (QBIC) is used to decide how much flexibility to allow in the spatial models, including whether to add or remove knots based on areas that require more adaptability.

3.5.3 Model fits were assessed via review of levels of residual autocorrelation using autocorrelation functions and run tests. Additionally, model selection was completed using an Analysis of Variance (ANOVA), enabling p-values for each term to be scrutinised. The two-dimensional relationship was plotted and checked for sensibility and an F-test was used for cumulative residual plots to check if covariates were modelled appropriately.

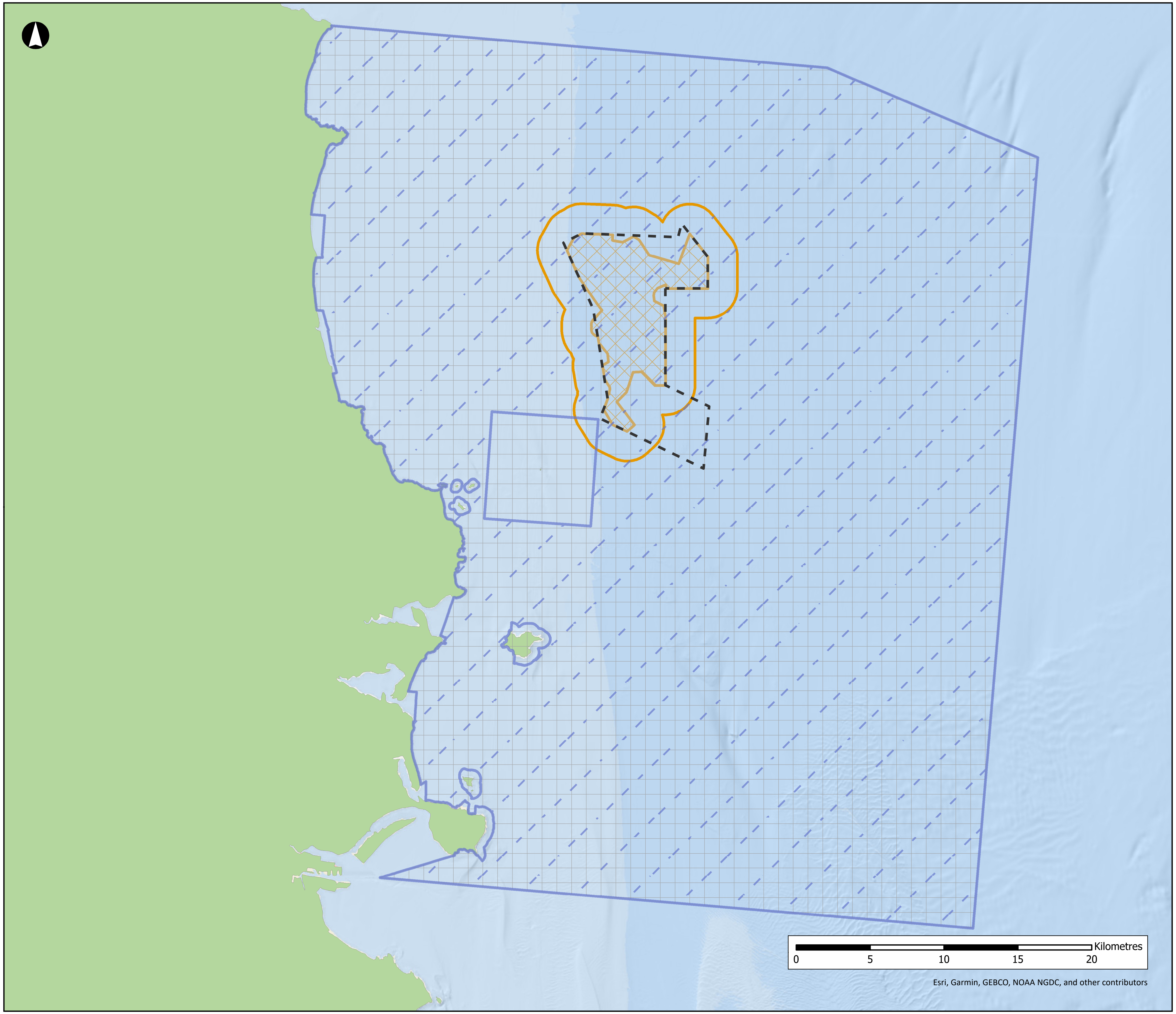
3.5.4 The final models used for each analysis are summarised in Table A3-2. A full summary of each model and its associated setup and diagnostic plots are presented in Appendices A.2, A.3 & A.4.

Table A3-2 Final model used in species-specific analysis.

Model	Model Covariates
Guillemot	count ~ survey + depth + distance to colony + seabed slope + prey density + survey:LFR.g + offset
Razorbill	count ~ survey + depth + distance to colony + prey density + offset
Kittiwake	count ~ survey + depth + survey:LFR.g + offset
Herring gull	count ~ survey + depth + survey:LFR.g + offset
Great black-backed gull	count ~ survey + depth + distance to colony + survey:LFR.g + offset
Gannet	count ~ survey + depth + distance to colony + survey:LFR.g + offset
Common tern	count ~ survey + depth + distance to colony + distance to coast + survey:LFR.g + offset
Roseate tern	count ~ survey + distance to colony + survey:LFR.g + offset
Red-throated diver	count ~ survey + distance to coast + survey:LFR.g + offset

3.6 Prediction Grid

3.6.1 A prediction grid was constructed by clipping a grid of 1 km^2 grid cells to the shapefile of the Survey Area to provide a framework which allowed the inclusion of environmental variables and visualisation of modelled species density outputs (Figure A3-1). Environmental variables were extracted from various data sources and clipped to the spatial extent of the Survey Area using geospatial tools (QGIS Development Team, 2024). A full list of environmental variables considered within analysis, and their associated data sources, are presented in Table A3-3.



Esri, Garmin, GEBCO, NOAA NGDC, and other contributors

⬚ Array Area

▭ North-West Irish Sea SPA

▨ Projected Footprint of Infrastructure

▭ Projected Footrprint of Infrastructure +2km Buffer

□ Prediction Grid used for MRSea Analysis (1km^2) Grid

NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Project Terminology Map (NWIS SPA, Array Area, Projected Footprint of Infrastructure, +2km Buffer, MRSea Analysis Prediction Grid)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:250,000 @A3	**A3-1**
Status: Issue	

Table A3-3 Environmental variables, and associated data sources, considered for MRSea analysis.

Environmental Variable	Data Source
Depth	Extracted from publicly available bathymetry data (GEBCO Compilation Group, 2024).
Seabed slope	Extracted from publicly available bathymetry data (GEBCO Compilation Group, 2024).
Distance to coast	Extracted from publicly available coastline shapefile data (ArcGIS, 2022).
Distance to colony	Extracted from publicly available SPA shapefile data (JNCC, 2024 & National Parks & Wildlife Service, 2024)
Prey density	Extracted from publicly available lesser sandeel (*Ammodytes marinus*) species distribution data (Langton *et al.*, 2021).

3.7 Modelling Distribution

3.7.1 MRSea analysis was undertaken for the entirety of the Survey Area (2,333 km^2), the output of the model-based abundance estimates for each species was also clipped to the PFI (57.7 km^2) and PFI plus 2km buffer (160.2 km^2). For each area of interest species abundance, the density per km^2, and lower and upper confidence intervals (CIs) were calculated. The 95% CIs were generated from the 1,000 replicates of a parametric bootstrap created during the modelling process. Density values were used to visually present species distribution in the Survey Area, PFI and the PFI plus a 2km buffer.

3.7.2 Within Appendix A.1, model-based abundance and density estimates from the original 29-month DAS data have been clipped from the original survey area (MAC boundary plus a 4km buffer) to provide new outputs for the PFI and PFI plus a 2km buffer. The full details of model selection and validation remain within the original application (Appendix A15.2 MRSea Modelling for Offshore Ornithology).

3.8 Abundance Estimate Apportionment

3.8.1 Bird sightings that could not be identified to species level were apportioned by calculating an apportionment abundance derived from design-based estimates (DBEs). Specifically, the unapportioned DBE estimate was subtracted from the apportioned DBE estimate to quantify the additional number of birds attributable to the MRSea apportionment process, and this value was then allocated accordingly.

3.8.2 For species where detectability was also influenced by availability bias (guillemot, razorbill and red-throated diver), a further correction was applied by multiplying the apportioned estimates by species-specific availability correction factors. Correction factors were derived from region appropriate values cited in Dunn *et al*. (2024), where published rates were available, and more generalist values from Thaxter *et al*. (2010) where not.

3.8.3 Apportioned estimates are likely to be consistently higher because the apportioning process redistributes birds detected within the Survey Area to represent the full population, rather than only those directly observed. Even where availability bias does not apply, apportioning scales the raw counts to reflect the proportion of the population expected to be present, which inherently inflates the estimate relative to the unadjusted totals.

4 Results

4.1 Guillemot

4.1.1 The following section presents model-based abundance estimates and density estimate plots derived from guillemot MRSea analysis. Details regarding model summary, setup and final model validation plots can be found in the appendices (A.2, A.3 & A.4).

4.1.2 Diagnostic plots indicated that the final model selected for guillemot MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.1.3 Guillemot were recorded in all surveys. The peak raw count within the Survey Area was recorded in August 2025, with 47,814 individuals. Raw counts for each month are presented in Table A4-1.

4.1.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-1 and Table 4-2. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in July 2025 at 41,109 guillemot. Within the PFI alone, the peak unapportioned abundance was estimated at 16,284 guillemot for the same period. Across the Survey Area, the peak estimated abundance was 287,894 guillemot in August 2025 (Table A4-1).

4.1.5 The use of the Plan C survey method in April 2025 resulted in lower raw guillemot counts relative to adjacent months. This limitation should be considered when interpreting the modelled abundance estimates.

4.1.6 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in July 2025 at 48,941 guillemot. Within the PFI alone, the peak apportioned abundance was estimated at 19,382 guillemot for the same period. Across the Survey Area, the peak estimated abundance was 328,669 guillemot in August 2025 (Table A4-2).

Table A4-1 The raw count of individual guillemot recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})
01	September	20,618	124,691 (79,829 – 171,198)	53.44	126,815 (72,505 – 685,254)	54.35	4,843 (2,873 – 8,504)	84.01	13,753 (8,168 – 24,085)	85.87
02	October	10,704	64,696 (42,536 – 87,715)	27.73	66,383 (38,526 – 535,345)	28.45	1,863 (1,189 – 3,053)	32.31	5,435 (3,410 – 9,046)	33.94
03	November	1,568	7,849 (5,338 – 10,705)	3.36	8,711 (5,150 – 35,270)	3.73	425 (279 – 671)	7.36	1,141 (742 – 1,829)	7.13
04	December	1,074	6,537 (3,027 – 10,388)	2.80	6,217 (3,637 – 21,832)	2.66	560 (358 – 933)	9.72	1,427 (912 – 2,370)	8.91
05	January	1,035	6,253 (3,378 – 9,321)	2.68	6,454 (3,787 – 26,640)	2.77	228 (146 – 365)	3.96	639 (408 – 1,034)	3.99
06	February	1,391	8,285 (4,887 – 11,661)	3.55	8,403 (4,703 – 43,570)	3.60	301 (192 – 489)	5.22	830 (521 – 1,381)	5.18
07	March	2,465	14,870 (8,658 – 21,759)	6.37	14,933 (9,135 – 65,582)	6.40	826 (520 – 1,379)	14.33	2,273 (1,455 – 3,747)	14.19
08	April	1,214	7,800 (5,107 – 10,669)	3.34	8,165 (4,682 – 61,942)	3.50	188 (114 – 320)	3.25	552 (334 – 952)	3.45
09	May	2,779	18,165 (11,635 – 26,908)	7.78	19,494 (9,359 – 121,530)	8.35	340 (209 – 570)	5.90	961 (574 – 1,644)	6.00
10	June	11,867	25,006 (16,360 – 37,786)	10.72	105,454 (48,751 – 931,517)	45.19	132 (67 – 270)	2.29	530 (265 – 1,077)	3.31
11	July	38,021	217,658 (153,274 – 297,362)	93.28	239,441 (135,042 – 1,505,282)	102.61	16,284 (7,730 – 37,364)	282.46	41,109 (20,820 – 88,370)	256.68
12	August	47,814	286,607 (172,694 – 426,872)	122.83	287,894 (165,025 – 1,758,617)	123.38	12,518 (8,311 – 19,362)	217.13	34,512 (22,155 – 54,907)	215.49

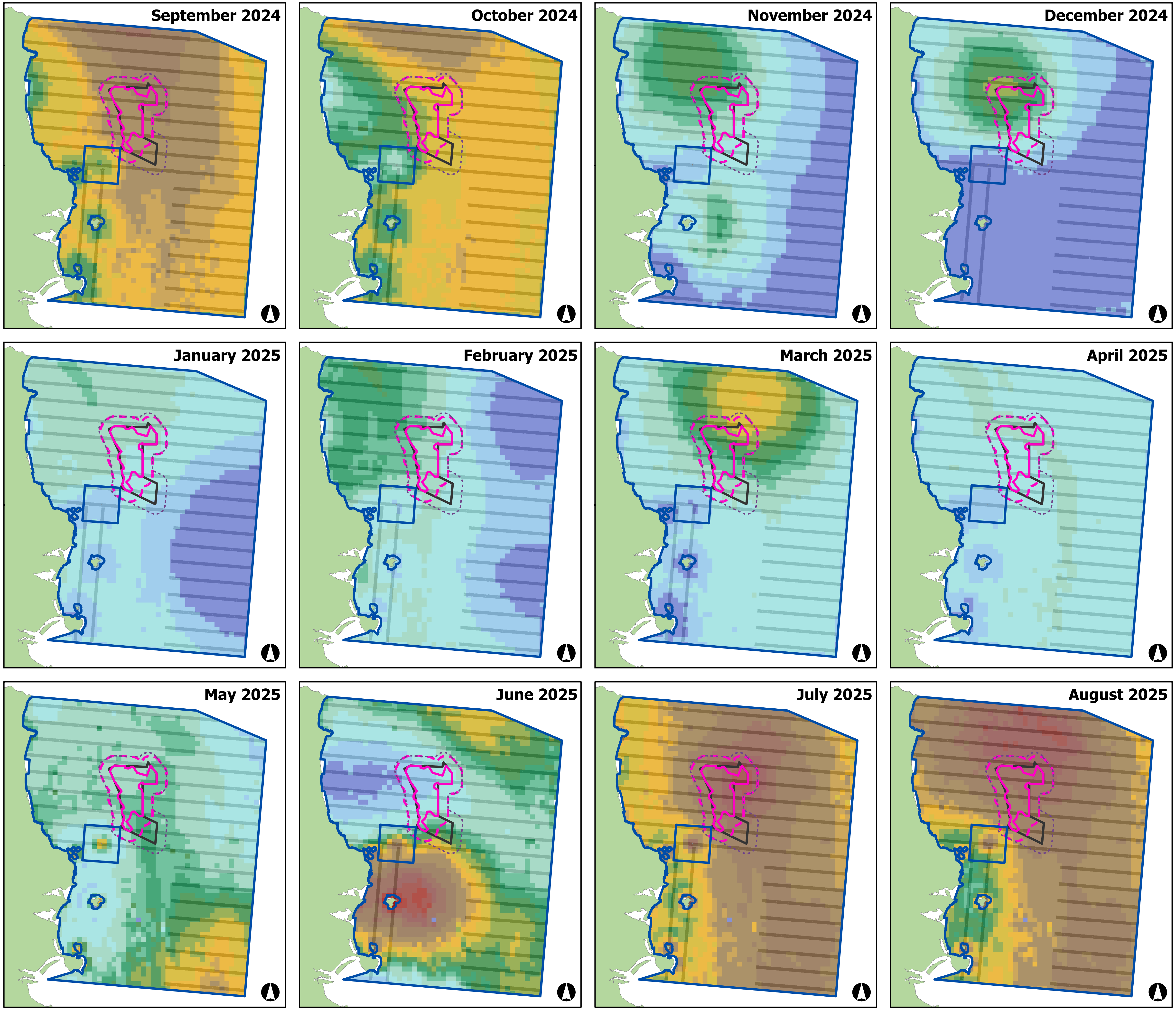
Table A4-2 The raw count of individual guillemot recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned and Corrected for Availability Bias)	
			Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})	Population Estimate	Density Estimate (birds/km^{2})
01	September	20,618	155,433 (100,616 – 209,799)	65.16	157,795 (97,397 – 778,834)	67.62	6,279 (4,088 – 10,350)	108.92	18,025 (11,814 – 29,515)	112.55
02	October	10,704	91,762 (64,172 – 119,858)	38.47	93,745 (61,003 – 644,946)	40.17	2,277 (1,485 – 3,675)	39.49	6,768 (4,388 – 11,013)	42.26
03	November	1,568	14,064 (10,050 – 18,131)	5.90	15,119 (10,763 – 47,602)	6.48	737 (559 – 1,039)	12.78	2,092 (1,603 – 2,933)	13.06
04	December	1,074	13,183 (6,684 – 20,380)	5.53	12,733 (9,108 – 34,677)	5.46	1,008 (724 – 1,532)	17.48	2,750 (2,026 – 4,076)	17.17
05	January	1,035	12,665 (7,082 – 18,684)	5.31	12,952 (9,152 – 41,702)	5.55	519 (402 – 713)	9.00	1,571 (1,242 – 2,134)	9.81
06	February	1,391	14,694 (9,345 – 19,902)	6.16	14,863 (9,584 – 65,045)	6.37	493 (339 – 762)	8.56	1,413 (971 – 2,199)	8.82
07	March	2,465	23,906 (14,527 – 34,221)	10.02	23,997 (15,800 – 95,606)	10.28	1,370 (938 – 2,152)	23.77	3,745 (2,589 – 5,829)	23.39
08	April	1,214	12,410 (8,589 – 16,516)	5.20	12,932 (8,366 – 83,433)	5.54	301 (205 – 474)	5.22	812 (526 – 1,336)	5.07
09	May	2,779	25,243 (16,637 – 36,987)	10.58	28,194 (14,908 – 161,964)	12.08	484 (311 – 786)	8.39	1,334 (827 – 2,231)	8.33
10	June	11,867	36,097 (24,060 – 53,165)	15.13	142,719 (68,382 – 1,225,688)	61.16	210 (124 – 390)	3.64	759 (412 – 1,476)	4.74
11	July	38,021	261,549 (186,625 – 355,476)	109.65	286,626 (166,929 – 1,737,945)	122.83	19,382 (9,574 – 43,551)	336.19	48,941 (25,678 – 103,127)	305.58
12	August	47,814	326,565 (198,103 – 481,810)	136.90	328,669 (192,842 – 1,954,496)	140.85	14,086 (9,435 – 21,652)	244.33	38,987 (25,327 – 61,533)	243.43

Density Estimate Plots

4.1.7 Modelled guillemot density estimates show strong seasonal variation across the Survey Area. Densities are consistently low during winter (November to February), with birds mainly inshore and only low densities present within the PFI footprint (Figure A4-1). Densities increase through spring (March - May), becoming more widespread and reaching moderate levels within and around the PFI. Where surveys included Lambay Island (e.g. in June), very clear hotspots were recorded around the island. The same distribution patterns are assumed to have occurred in May; however, due to airspace restrictions, neither Plan A nor Plan B surveys could be carried out.

4.1.8 Peak densities occur in the post-breeding season (July to September), when guillemots disperse from breeding colonies. There is a relatively even distribution throughout this period with some distinct hotspots, particularly south and northeast of the PFI (Figure A4-1). Overall, guillemot presence is most concentrated and spatially structured in early summer, with minimal clustering during winter.



- Array Area
- 2km Buffer from Array Area
- Projected Footprint of Infrastructure
- 2km Buffer from Projected Footprint of Infrastructure
- North-West Irish Sea SPA
- Survey Transect Layout

Guillemot Density (birds per km^2)

- 0 - 1
- 1 - 2
- 3 - 4
- 5 - 6
- 7 - 8
- 9 - 10
- 10 - 15
- 15 - 20
- 20 - 30
- 30 - 40
- 40 - 50
- 50 - 100
- 100 - 200
- 200 - 300
- 300 - 400
- 400 - 500
- 500 - 1000
- >1000

Kilometres
0 5 10 15 20 25

NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Guillemot (km^2)

Job No: 281240

Datum: WGS84 **Projection:** UTM30N

Date: February 2026

Figure No:

Scale: 1:750,000 @A3

Status: Issue

A4-1

4.2 Razorbill

4.2.1 The following section presents model-based abundance estimates and density estimate plots derived from razorbill MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.2.2 Diagnostic plots indicated that the final model selected for razorbill MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.2.3 Razorbill were recorded in all surveys. The peak raw count within the Survey Area was recorded in August 2025, with 19,785 individuals. Raw counts for each month are presented in Table A4-3.

4.2.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-3 and Table A4-4. The peak unapportioned abundance in PFI plus 2 km buffer was estimated in August 2025 at 18,086 razorbill. Within the PFI alone, the peak unapportioned abundance was estimated at 6,235 razorbill for the same period. Across the Survey Area, the peak estimated abundance was 139,078 razorbill for the same period (Table A4-3).

4.2.5 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 20,833 razorbill. Within the PFI alone, the peak apportioned abundance was estimated at 7,211 razorbill for the same period. Across the Survey Area, the peak estimated abundance was 161,053 razorbill for the same period (Table A4-4).

Table A4-3 The raw count of individual razorbill recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	8,859	52,338 (25,639 – 82,960)	22.43	65,166 (32,762 – 714,354)	27.93	914 (512 – 1,667)	15.86	2,973 (1,689 – 5,307)	18.56
02	October	1,468	8,759 (4,357 – 13,054)	3.75	10,025 (5,210 – 39,962)	4.30	315 (171 – 622)	5.46	934 (526 – 1,786)	5.83
03	November	1,091	5,485 (4,168 – 6,878)	2.35	6,807 (3,765 – 34,407)	2.92	170 (96 – 309)	2.95	515 (298 – 918)	3.21
04	December	1,203	7,133 (3,507 – 10,982)	3.06	7,853 (4,882 – 21,529)	3.37	219 (124 – 391)	3.79	659 (402 – 1,109)	4.12
05	January	1,710	10,216 (5,336 – 16,135)	4.38	11,784 (6,588 – 83,654)	5.05	413 (230 – 776)	7.16	1,204 (700 – 2,158)	7.52
06	February	995	6,001 (4,055 – 8,086)	2.57	7,992 (4,223 – 76,889)	3.42	130 (71 – 240)	2.26	398 (223 – 722)	2.49
07	March	870	5,210 (4,006 – 6,381)	2.23	6,512 (3,502 – 36,161)	2.79	163 (89 – 313)	2.82	493 (275 – 919)	3.08
08	April	624	4,042 (2,625 – 5,674)	1.73	5,009 (2,550 – 27,577)	2.15	125 (61 – 255)	2.17	379 (192 – 760)	2.37
09	May	236	1,535 (817 – 2,562)	0.66	1,916 (808 – 12,667)	0.82	48 (20 – 118)	0.83	145 (63 – 351)	0.90
10	June	2,356	2,047 (1,136 – 3,268)	0.88	31,791 (11,131 – 677,391)	13.62	8 (2 – 40)	0.14	35 (10 – 150)	0.22
11	July	4,965	29,763 (21,569 – 38,704)	12.76	35,322 (18,532 – 171,321)	15.14	706 (340 – 1,475)	12.25	2,142 (1,071 – 4,330)	13.37
12	August	19,785	118,391 (77,624 – 162,314)	50.74	139,078 (79,087 – 565,931)	59.60	6,235 (3,407 – 11,202)	108.14	18,086 (10,117 – 32,253)	112.92

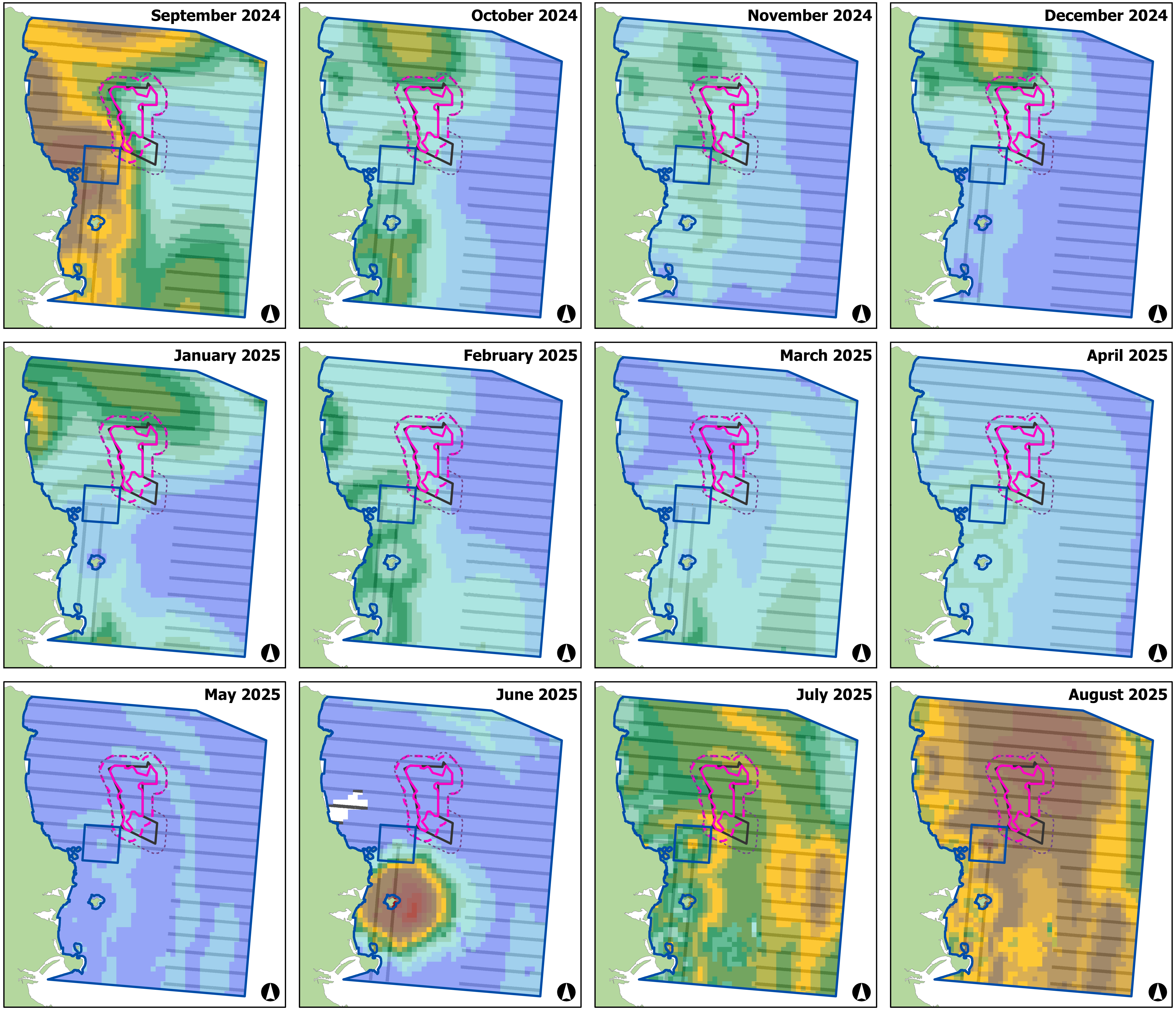
Table A4-4 The raw count of individual razorbill recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned and Corrected for Availability Bias)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	8,859	65,744 (32,376 – 102,729)	27.56	85,396 (46,639 – 861,844)	36.60	1,154 (673 – 2,055)	20.03	3,874 (2,338 – 6,666)	24.19
02	October	1,468	12,059 (6,078 – 18,282)	5.06	14,849 (8,830 – 52,270)	6.36	436 (256 – 820)	7.56	1,307 (796 – 2,372)	8.16
03	November	1,091	9,473 (7,336 – 11,693)	3.97	12,001 (8,107 – 47,331)	5.14	337 (242 – 515)	5.84	1,028 (750 – 1,543)	6.42
04	December	1,203	11,657 (6,403 – 17,529)	4.89	14,007 (10,204 – 31,513)	6.00	395 (274 – 616)	6.85	1,224 (895 – 1,799)	7.64
05	January	1,710	17,719 (10,412 – 26,274)	7.43	20,484 (13,833 – 112,484)	8.78	683 (450 – 1,148)	11.85	2,157 (1,512 – 3,378)	13.47
06	February	995	9,159 (6,226 – 12,183)	3.84	12,100 (7,275 – 100,295)	5.19	188 (113 – 328)	3.27	597 (372 – 1,011)	3.72
07	March	870	7,273 (5,549 – 8,991)	3.05	9,250 (5,396 – 47,203)	3.96	267 (172 – 459)	4.63	700 (421 – 1,246)	4.37
08	April	624	5,889 (3,949 – 8,156)	2.47	6,837 (3,972 – 33,129)	2.93	160 (85 – 311)	2.77	474 (256 – 918)	2.96
09	May	236	1,971 (1,033 – 3,263)	0.83	2,427 (1,137 – 14,953)	1.04	57 (25 – 138)	0.99	171 (75 – 411)	1.07
10	June	2,356	2,814 (1,519 – 4,487)	1.18	37,394 (13,324 – 789,517)	16.03	9 (2 – 47)	0.16	41 (13 – 176)	0.26
11	July	4,965	34,031 (24,619 – 44,343)	14.27	42,861 (23,302 – 201,300)	18.37	925 (499 – 1,820)	16.05	2,700 (1,452 – 5,249)	16.86
12	August	19,785	131,615 (86,295 – 179,984)	55.18	161,053 (93,314 – 643,046)	69.02	7,211 (4,017 – 12,819)	125.07	20,833 (11,835 – 36,830)	130.08

Density Estimate Plots

4.2.6 The modelled razorbill density estimates show pronounced seasonal variability across the Survey Area, with low and spatially limited densities through winter (October to February), during which birds are largely restricted to inshore waters and only very low densities occur within the PFI footprint (Figure A4-2).

4.2.7 Peak densities occur in the summer (June - August), when a distinct high-density hotspot develops to the south of the PFI and elevated densities extend across much of the southern portion of the Survey Area (Figure A4-2). Although densities remain relatively high in July, they become more diffuse, and by August spatial dispersion is increasingly prevalent. This trend is in alignment with expected razorbill post-breeding (July onwards) dispersal patterns.



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Razorbill (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:750,000 @A3	**A4-2**
Status: Issue	

4.3 Kittiwake

4.3.1 The following section presents model-based abundance estimates and density estimate plots derived from kittiwake MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.3.2 Diagnostic plots indicated that the final model selected for kittiwake MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.3.3 Kittiwake were recorded in all surveys. The peak raw count within the Survey Area was recorded in August 2025, with 2,759 individuals. Raw counts for each month are presented in Table A4-5.

4.3.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-5 and Table A4-6. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 1,933 kittiwake. For the same period, within the PFI and Survey Area, the peak unapportioned abundance was estimated at 710 and 16,153 kittiwake respectively (Table A4-5).

4.3.5 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 1,943 kittiwake. Within the PFI alone, the peak apportioned abundance was estimated at 710 kittiwake for the same period. Across the Survey Area, the peak estimated abundance was 16,996 kittiwake in August 2025 (Table A4-6).

Table A4-5 The raw count of individual kittiwake recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	824	4,887 (2,566 – 8,045)	2.09	5,110 (2,676 – 9,473)	2.19	175 (91 – 318)	3.04	480 (249 – 880)	3.00
02	October	841	5,121 (2,364 – 8,571)	2.19	4,407 (2,464 – 7,616)	1.89	114 (76 – 168)	1.98	328 (215 – 493)	2.05
03	November	2,177	10,917 (6,488 – 16,031)	4.68	10,899 (4,654 – 28,074)	4.67	187 (112 – 316)	3.24	663 (404 – 1,122)	4.14
04	December	91	542 (342 – 793)	0.23	564 (343 – 912)	0.24	19 (12 – 29)	0.34	53 (33 – 81)	0.33
05	January	469	2,809 (2,233 – 3,455)	1.20	2,905 (2,007 – 4,255)	1.24	100 (73 – 136)	1.73	273 (198 – 376)	1.70
06	February	349	1,807 (1,346 – 2,326)	0.77	2,161 (1,409 – 3,325)	0.93	74 (51 – 107)	1.29	203 (139 – 295)	1.27
07	March	569	3,420 (2,562 – 4,324)	1.47	3,536 (2,259 – 5,528)	1.52	121 (82 – 176)	2.11	332 (221 – 486)	2.07
08	April	362	2,334 (1,642 – 3,108)	1.00	2,972 (1,779 – 5,042)	1.27	35 (23 – 52)	0.60	100 (65 – 150)	0.62
09	May	246	1,088 (673 – 1,601)	0.47	1,709 (918 – 3,330)	0.73	4 (1 – 18)	0.07	29 (11 – 85)	0.18
10	June	433	1,861 (1,249 – 2,588)	0.80	3,435 (2,060 – 5,790)	1.47	5 (2 – 18)	0.09	25 (11 – 69)	0.16
11	July	1,570	8,621 (5,799 – 12,260)	3.69	9,353 (5,535 – 15,737)	4.01	403 (258 – 629)	6.99	1,180 (743 – 1,865)	7.36
12	August	2,759	14,786 (10,375 – 19,425)	6.34	16,153 (10,497 – 25,322)	6.92	710 (531 – 944)	12.31	1,933 (1,409 – 2,644)	12.07

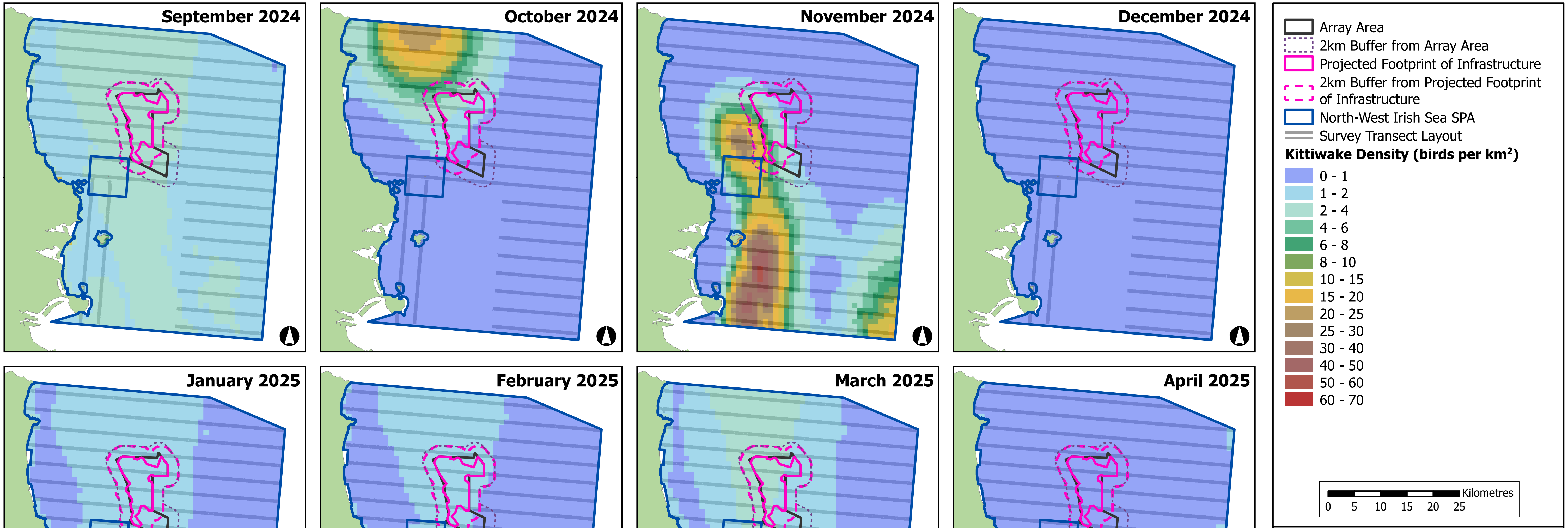
Table A4-6 The raw count of individual kittiwake recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	824	5,098 (2,775 – 8,239)	2.14	5,321 (7,997 – 17,469)	2.28	175 (267 – 584)	3.04	480 (249 – 880)	3.00
02	October	841	5,568 (2,667 – 9,173)	2.33	4,854 (7,318 – 14,933)	2.08	119 (195 – 364)	2.06	337 (224 – 502)	2.10
03	November	2,177	12,799 (7,854 – 18,396)	5.37	12,781 (17,435 – 45,509)	5.48	191 (303 – 619)	3.31	694 (435 – 1,153)	4.34
04	December	91	681 (395 – 1,043)	0.29	703 (1,047 – 1,959)	0.30	19 (32 – 61)	0.34	57 (37 – 85)	0.36
05	January	469	4,041 (2,680 – 6,238)	1.69	4,137 (6,144 – 10,399)	1.77	100 (173 – 309)	1.73	287 (212 – 390)	1.79
06	February	349	2,223 (1,636 – 2,815)	0.93	2,577 (3,987 – 7,312)	1.10	74 (126 – 232)	1.29	232 (168 – 324)	1.45
07	March	569	3,564 (2,698 – 4,477)	1.49	3,680 (5,939 – 11,467)	1.58	121 (203 – 379)	2.11	332 (221 – 486)	2.07
08	April	362	2,345 (1,655 – 3,115)	0.98	2,983 (4,762 – 9,803)	1.28	35 (58 – 109)	0.60	100 (65 – 150)	0.62
09	May	246	1,113 (695 – 1,620)	0.47	1,734 (2,652 – 5,982)	0.74	4 (5 – 23)	0.07	29 (11 – 85)	0.18
10	June	433	1,861 (1,249 – 2,588)	0.78	3,435 (5,494 – 11,284)	1.47	5 (7 – 25)	0.09	25 (11 – 69)	0.16
11	July	1,570	9,033 (6,275 – 12,657)	3.79	9,765 (15,300 – 31,036)	4.18	408 (666 – 1,296)	7.08	1,185 (748 – 1,870)	7.40
12	August	2,759	15,629 (11,027 – 20,469)	6.55	16,996 (27,493 – 52,815)	7.28	710 (1,240 – 2,185)	12.31	1,943 (1,419 – 2,654)	12.14

Density Estimate Plots

4.3.6 Modelled kittiwake density estimates show strong seasonal variation across the Survey Area. Densities were spatially clustered in late-autumn and early-winter months (October to November), with birds mainly to the north and south of the PFI (Figure A4-3). Densities decrease through late winter to spring (December to April), becoming more widespread and reaching low levels within and around the PFI.

4.3.7 Although peak densities occurred in winter months (November), distinct hotspots developed north-east and south-west of the PFI boundary throughout the summer months (Figure A4-3).



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Kittiwake (km^2)

Job No: 281240

Datum: WGS84 **Projection:** UTM30N

Date: February 2026

Scale: 1:750,000 @A3

Status: Issue

Figure No: A4-3

4.4 Herring Gull

4.4.1 The following section presents model-based abundance estimates and density estimate plots derived from herring gull MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.4.2 Diagnostic plots indicated that the final model selected for herring gull MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.4.3 Herring gull were recorded in all surveys. The peak raw count within the Survey Area was recorded in August 2025, with 4,851 individuals. Raw counts for each month are presented in Table A4-7.

4.4.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-7 and Table A4-8. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 4,118 herring gull. Within the PFI alone, the peak unapportioned abundance was estimated at 1,634 herring gull for the same period. Across the Survey Area, the peak estimated abundance was 31,440 herring gull (Table A4-7).

4.4.5 It is noted that fishing activities were recorded within the Survey Area in November 2024, which might explain elevated abundance estimates in this month.

4.4.6 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 4,189 herring gull. Within the PFI alone, the peak apportioned abundance was estimated at 1,646 herring gull for the same period. Across the Survey Area, the peak estimated abundance was 33,085 herring gull in August 2025 (Table A4-8).

Table A4-7 The raw count of individual herring gull recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	1,190	6,629 (3,432 – 10,502)	2.84	8,052 (4,947 – 13,016)	3.45	88 (53 – 146)	1.53	252 (148 – 422)	1.57
02	October	1,166	5,836 (2,191 – 10,592)	2.50	7,898 (4,385 – 14,293)	3.38	87 (50 – 160)	1.50	247 (141 – 457)	1.54
03	November	1,539	7,568 (5,318 – 9,937)	3.24	7,897 (4,837 – 12,930)	3.38	504 (292 – 868)	8.75	1,319 (778 – 2,226)	8.24
04	December	604	3,049 (1,718 – 4,803)	1.31	4,074 (2,387 – 6,861)	1.75	45 (26 – 76)	0.78	127 (72 – 219)	0.80
05	January	671	3,066 (1,665 – 4,730)	1.31	4,537 (2,707 – 7,525)	1.94	50 (30 – 83)	0.86	142 (85 – 240)	0.89
06	February	1,416	4,662 (2,255 – 7,374)	2.00	9,560 (4,863 – 18,911)	4.10	105 (54 – 205)	1.82	299 (155 – 587)	1.87
07	March	416	1,450 (710 – 2,453)	0.62	2,816 (1,417 – 5,559)	1.21	31 (16 – 61)	0.54	88 (46 – 176)	0.55
08	April	754	4,220 (1,215 – 7,862)	1.81	4,210 (1,514 – 12,647)	1.80	15 (3 – 71)	0.25	45 (10 – 212)	0.28
09	May	982	6,021 (2,713 – 10,105)	2.58	8,164 (4,859 – 13,886)	3.50	90 (54 – 154)	1.55	255 (152 – 444)	1.60
10	June	1,116	3,665 (2,099 – 5,480)	1.57	11,019 (5,474 – 22,707)	4.72	1 (0 – 4)	0.01	3 (0 – 16)	0.02
11	July	3,656	17,108 (10,504 – 24,784)	7.33	23,487 (13,592 – 38,191)	10.07	248 (160 – 370)	4.30	715 (454 – 1,082)	4.46
12	August	4,851	24,710 (15,677 – 35,433)	10.59	31,440 (18,798 – 55,046)	13.47	1,634 (838 – 3,453)	28.35	4,118 (2,181 – 8,342)	25.71

Table A4-8 The raw count of individual herring gull recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

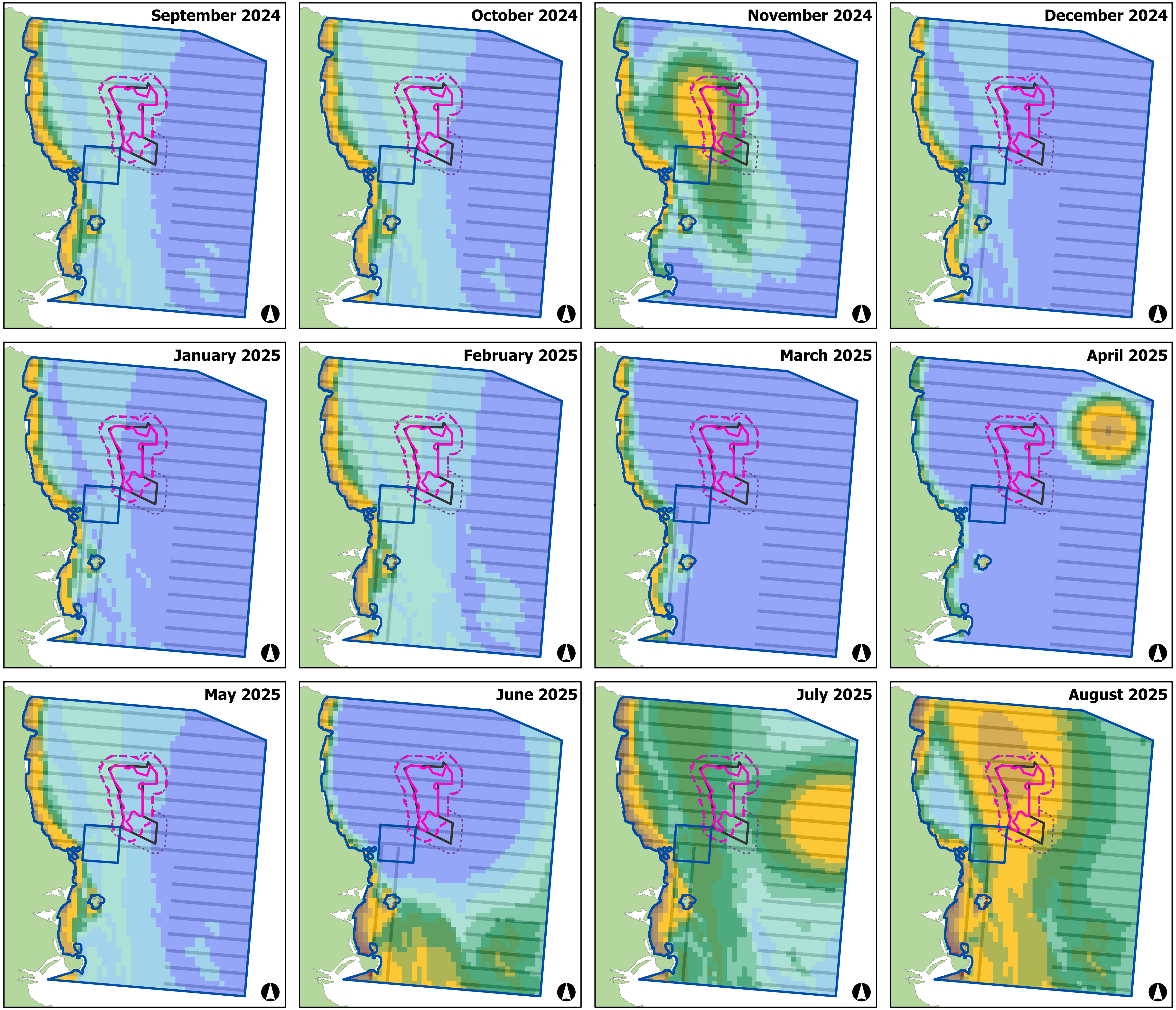
Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	1,190	7,759 (4,178 – 11,931)	3.25	9,182 (6,077 – 14,146)	3.94	88 (53 – 146)	1.53	270 (166 – 440)	1.69
02	October	1,166	6,076 (2,452 – 10,953)	2.55	8,138 (4,625 – 14,533)	3.49	87 (50 – 160)	1.50	271 (165 – 481)	1.69
03	November	1,539	8,220 (5,932 – 10,673)	3.45	8,549 (5,489 – 13,582)	3.66	563 (351 – 927)	9.77	1,406 (865 – 2,313)	8.78
04	December	604	3,189 (1,823 – 5,012)	1.34	4,214 (2,527 – 7,001)	1.81	45 (26 – 76)	0.78	127 (72 – 219)	0.80
05	January	671	3,156 (1,740 – 4,843)	1.32	4,627 (2,797 – 7,615)	1.98	50 (30 – 83)	0.86	144 (87 – 242)	0.90
06	February	1,416	4,848 (2,377 – 7,565)	2.03	9,746 (5,049 – 19,097)	4.18	105 (54 – 205)	1.82	299 (155 – 587)	1.87
07	March	416	1,476 (729 – 2,483)	0.62	2,842 (1,443 – 5,585)	1.22	33 (18 – 63)	0.57	89 (47 – 177)	0.56
08	April	754	4,412 (1,303 – 8,084)	1.85	4,402 (1,706 – 12,839)	1.89	15 (3 – 71)	0.25	45 (10 – 212)	0.28
09	May	982	6,083 (2,800 – 10,171)	2.55	8,226 (4,921 – 13,948)	3.53	90 (54 – 154)	1.55	255 (152 – 444)	1.60
10	June	1,116	3,714 (2,146 – 5,516)	1.56	11,068 (5,523 – 22,756)	4.74	1 (0 – 4)	0.01	3 (0 – 16)	0.02
11	July	3,656	17,458 (10,662 – 25,269)	7.32	23,837 (13,942 – 38,541)	10.22	248 (160 – 370)	4.30	719 (458 – 1,086)	4.49
12	August	4,851	26,355 (17,206 – 37,193)	11.05	33,085 (20,443 – 56,691)	14.18	1,646 (850 – 3,465)	28.56	4,189 (2,252 – 8,413)	26.16

Density Estimate Plot

4.4.7 Modelled herring gull density estimates show limited seasonal variation across the Survey Area. Densities were spatially homogenous in autumn, winter and spring months (September to March), with birds predominantly distributed to the west of the PFI in proximity to the coastline (Figure A4-4).

4.4.8 Densities became more spatially diffuse throughout the survey area within the summer months (June – August), with distinct hotspots north-east and south-west of the PFI boundary (Figure A4-4).

4.4.9 It is noted that spatial clustering observed in November 2024, April 2025 and July 2025 aligns with observed fishing vessel activities (Table A3-1) recorded within the survey report (APEM, 2024 & 2025).



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Herring Gull (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:750,000 @A3	**A4-4**
Status: Issue	

4.5 Great Black-backed Gull

4.5.1 The following section presents model-based abundance estimates and density estimate plots derived from great black-backed gull MRSea analysis. A detailed model summary, information regarding model setup and final model validation plots can be found in the appendices (A.2, A.3 & A.4).

4.5.2 Diagnostic plots indicated that the final model selected for great black-backed gull MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.5.3 Great black-backed gull were recorded in all surveys. The peak raw count within the Survey Area was recorded in April 2025, with 321 individuals. Raw counts for each month are presented in Table A4-9.

4.5.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-9 and Table A4-10. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in November 2024 at 331 great black-backed gull. Within the PFI alone, the peak unapportioned abundance was estimated at 163 great black-backed gull for the same period. Across the Survey Area, the peak estimated abundance was 1,622 great black-backed gull in April 2025 (Table A4-9).

4.5.5 It is noted that elevated abundance estimates within April 2025 may be a result of fishing activities that were taking place within the Survey Area.

4.5.6 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in November 2024 at 371 great black-backed gull. Within the PFI alone, the peak apportioned abundance was estimated at 198 great black-backed gull for the same period. Across the Survey Area, the peak estimated abundance was 1,804 great black-backed gull in April 2025 (Table A4-10).

Table A4-9 The raw count of individual great black-backed gull recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	104	435 (252 – 637)	0.19	588 (307 – 1,123)	0.25	18 (9 – 35)	0.32	44 (23 – 82)	0.28
02	October	136	444 (253 – 650)	0.19	765 (456 – 1,313)	0.33	24 (14 – 40)	0.42	58 (35 – 94)	0.36
03	November	177	875 (447 – 1,391)	0.37	904 (410 – 2,002)	0.39	163 (65 – 395)	2.84	331 (138 – 785)	2.07
04	December	95	402 (240 – 603)	0.17	536 (259 – 1,119)	0.23	17 (8 – 34)	0.29	40 (20 – 82)	0.25
05	January	68	287 (168 – 419)	0.12	382 (198 – 738)	0.16	12 (6 – 23)	0.21	29 (15 – 54)	0.18
06	February	195	856 (383 – 1,578)	0.37	1,096 (539 – 2,219)	0.47	34 (17 – 64)	0.60	83 (42 – 155)	0.52
07	March	161	745 (349 – 1,281)	0.32	875 (429 – 2,002)	0.37	20 (11 – 36)	0.34	48 (27 – 84)	0.30
08	April	321	1,966 (544 – 3,838)	0.84	1,622 (552 – 12,448)	0.69	13 (2 – 174)	0.23	40 (8 – 426)	0.25
09	May	37	233 (85 – 405)	0.10	297 (140 – 647)	0.13	9 (4 – 19)	0.16	22 (11 – 45)	0.14
10	June	110	478 (245 – 771)	0.20	624 (278 – 1,432)	0.27	1 (0 – 5)	0.02	5 (2 – 15)	0.03
11	July	151	695 (454 – 1,004)	0.30	851 (460 – 1,575)	0.36	27 (14 – 47)	0.46	64 (35 – 113)	0.40
12	August	265	1,243 (808 – 1,729)	0.53	1,492 (918 – 2,496)	0.64	47 (28 – 79)	0.81	113 (69 – 186)	0.70

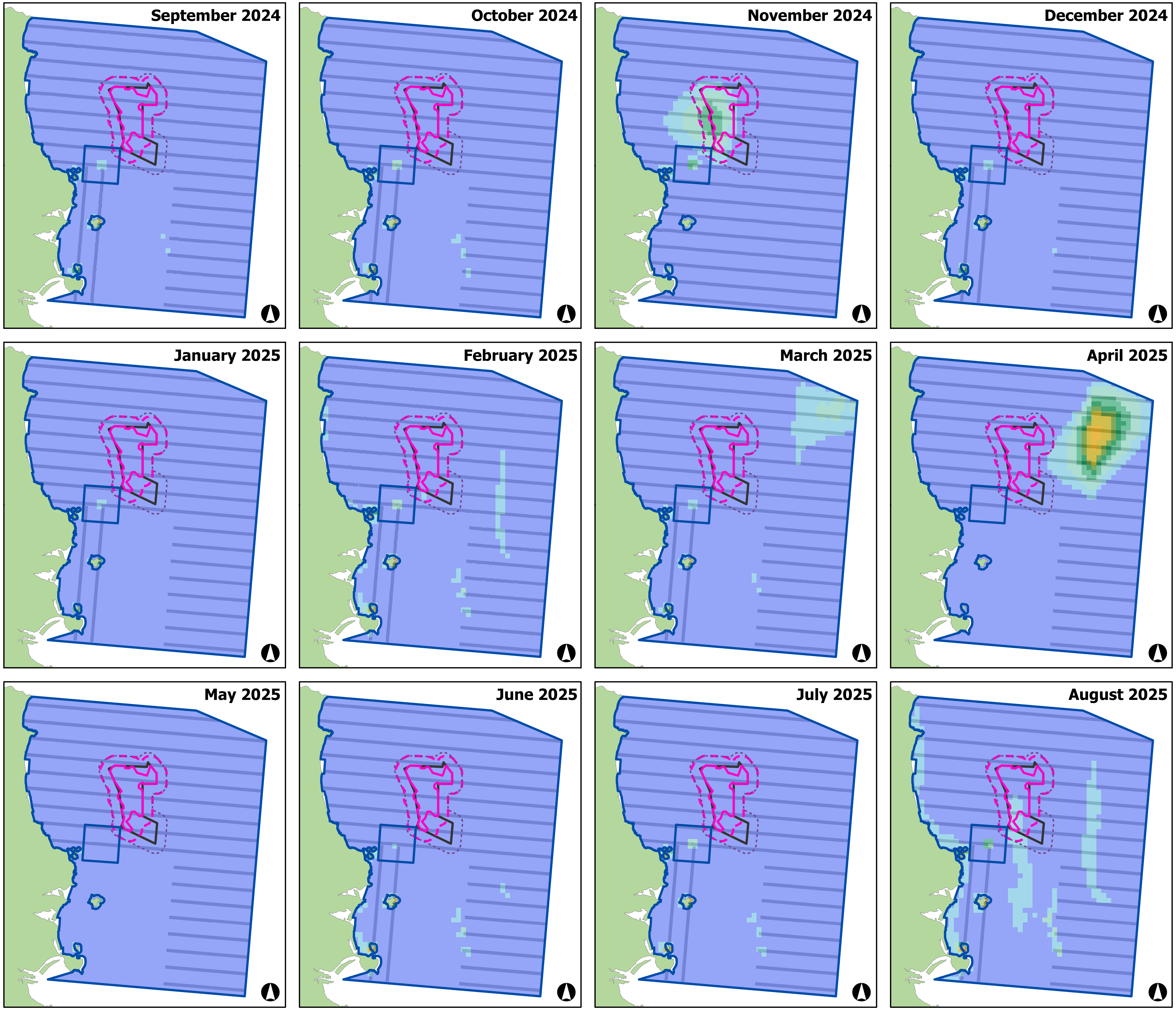
Table A4-10 The raw count of individual great black-backed gull recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	104	508 (289 – 780)	0.21	661 (380 – 1,196)	0.28	18 (9 – 35)	0.32	48 (27 – 86)	0.30
02	October	136	471 (274 – 685)	0.20	792 (483 – 1,340)	0.34	24 (14 – 40)	0.42	59 (36 – 95)	0.37
03	November	177	964 (495 – 1,508)	0.40	993 (499 – 2,091)	0.43	198 (100 – 430)	3.44	371 (178 – 825)	2.32
04	December	95	413 (246 – 624)	0.17	547 (270 – 1,130)	0.23	17 (8 – 34)	0.29	40 (20 – 82)	0.25
05	January	68	293 (173 – 427)	0.12	388 (204 – 744)	0.17	12 (6 – 23)	0.21	30 (16 – 55)	0.19
06	February	195	912 (412 – 1,674)	0.38	1,152 (595 – 2,275)	0.49	34 (17 – 64)	0.60	83 (42 – 155)	0.52
07	March	161	758 (357 – 1,308)	0.32	888 (442 – 2,015)	0.38	23 (14 – 39)	0.39	51 (30 – 87)	0.32
08	April	321	2,148 (607 – 4,043)	0.90	1,804 (734 – 12,630)	0.77	13 (2 – 174)	0.23	40 (8 – 426)	0.25
09	May	37	240 (92 – 414)	0.10	304 (147 – 654)	0.13	9 (4 – 19)	0.16	22 (11 – 45)	0.14
10	June	110	502 (260 – 803)	0.21	648 (302 – 1,456)	0.28	1 (0 – 5)	0.02	5 (2 – 15)	0.03
11	July	151	716 (469 – 1,045)	0.30	872 (481 – 1,596)	0.37	27 (14 – 47)	0.46	65 (36 – 114)	0.41
12	August	265	1,331 (862 – 1,868)	0.56	1,580 (1,006 – 2,584)	0.68	48 (29 – 80)	0.83	119 (75 – 192)	0.74

Density Estimate Plots

4.5.7 Modelled great black-backed gull density estimates show limited seasonal variation across the Survey Area. Densities were spatially clustered in November and April, with birds mainly to the west and north-east of the PFI (Figure A4-5).

4.5.8 Clustering aligns with fishing vessel activity (Table A3-1) that was recorded in similar areas within the Survey Area within both November 2024 and April 2025 survey reports (APEM, 2024 & 2025).



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Great Black-backed Gull (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:750,000 @A3	**A4-5**
Status: Issue	

4.6 Gannet

4.6.1 The following section presents model-based abundance estimates and density estimate plots derived from gannet MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.6.2 Diagnostic plots indicated that the final model selected for gannet MRSea analysis provided an acceptable fit, with the residuals showing no systematic spatial or temporal patterns and no evidence of undue influence from individual observations. Remaining autocorrelation was minimal, and model assumptions appeared to be met. On this basis, the resulting abundance and density estimates are considered robust and suitable for use in subsequent analyses.

Abundance Estimates

4.6.3 Gannet were recorded in all surveys. The raw count for December 2024 and January 2025 were too low to run model-based estimates (see Section 3.4). The peak raw count within the Survey Area was recorded in July 2025, with 910 individuals. Raw counts for each month are presented in Table A4-11.

4.6.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-11 and Table A4-12. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in July 2025 at 178 gannet. Within the PFI alone, the peak unapportioned abundance was estimated at 62 gannet for the same period. Across the Survey Area, the peak estimated abundance was 5,139 gannet in July 2025 (Table A4-11).

4.6.5 It is noted that elevated abundance estimates within April 2025 may be a result of fishing activities that were taking place within the Survey Area.

4.6.6 The peak apportioned abundance in thePFI plus 2 km buffer was estimated in July 2025 at 178 gannet. Within the PFI alone, the peak apportioned abundance was estimated at 62 gannet for the same period. Across the Survey Area, the peak estimated abundance was 5,139 gannet in July 2025 (Table A4-12).

Table A4-11 The raw count of individual gannet recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)
01	September	571	2,693 (1,610 – 4,037)	1.15	3,265 (1,804 – 6,000)	1.40	40 (22 – 76)	0.70	115 (64 – 212)	0.72
02	October	214	1,047 (493 – 1,907)	0.45	1,210 (549 – 2,779)	0.52	15 (6 – 35)	0.25	42 (18 – 101)	0.26
03	November	36	180 (91 – 281)	0.08	189 (81 – 439)	0.08	2 (1 – 5)	0.04	7 (3 – 16)	0.04
04	December	3	18 (3 – 36)	0.01	-	-	-	-	-	-
05	January	3	18 (3 – 48)	0.01	-	-	-	-	-	-
06	February	102	594 (185 – 1,124)	0.25	528 (280 – 996)	0.23	1 (0 – 3)	0.02	4 (2 – 9)	0.02
07	March	172	667 (259 – 1,299)	0.29	977 (384 – 2,699)	0.42	0 (0 – 1)	0.00	1 (0 – 3)	0.01
08	April	665	4,311 (1,519 – 8,271)	1.85	3,396 (1,701 – 7,137)	1.46	24 (16 – 39)	0.42	91 (58 – 148)	0.57
09	May	133	866 (438 – 1,379)	0.37	1,009 (520 – 1,938)	0.43	12 (6 – 24)	0.21	35 (18 – 67)	0.22
10	June	458	2,127 (1,375 – 2,970)	0.91	2,718 (1,525 – 4,949)	1.16	24 (12 – 47)	0.42	71 (36 – 140)	0.44
11	July	910	4,434 (2,756 – 6,420)	1.90	5,139 (2,950 – 9,134)	2.20	62 (34 – 120)	1.08	178 (96 – 338)	1.11
12	August	877	4,261 (2,908 – 5,767)	1.83	4,956 (2,821 – 8,666)	2.12	60 (36 – 99)	1.04	172 (102 – 283)	1.07

Table A4-12 The raw count of individual gannet recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

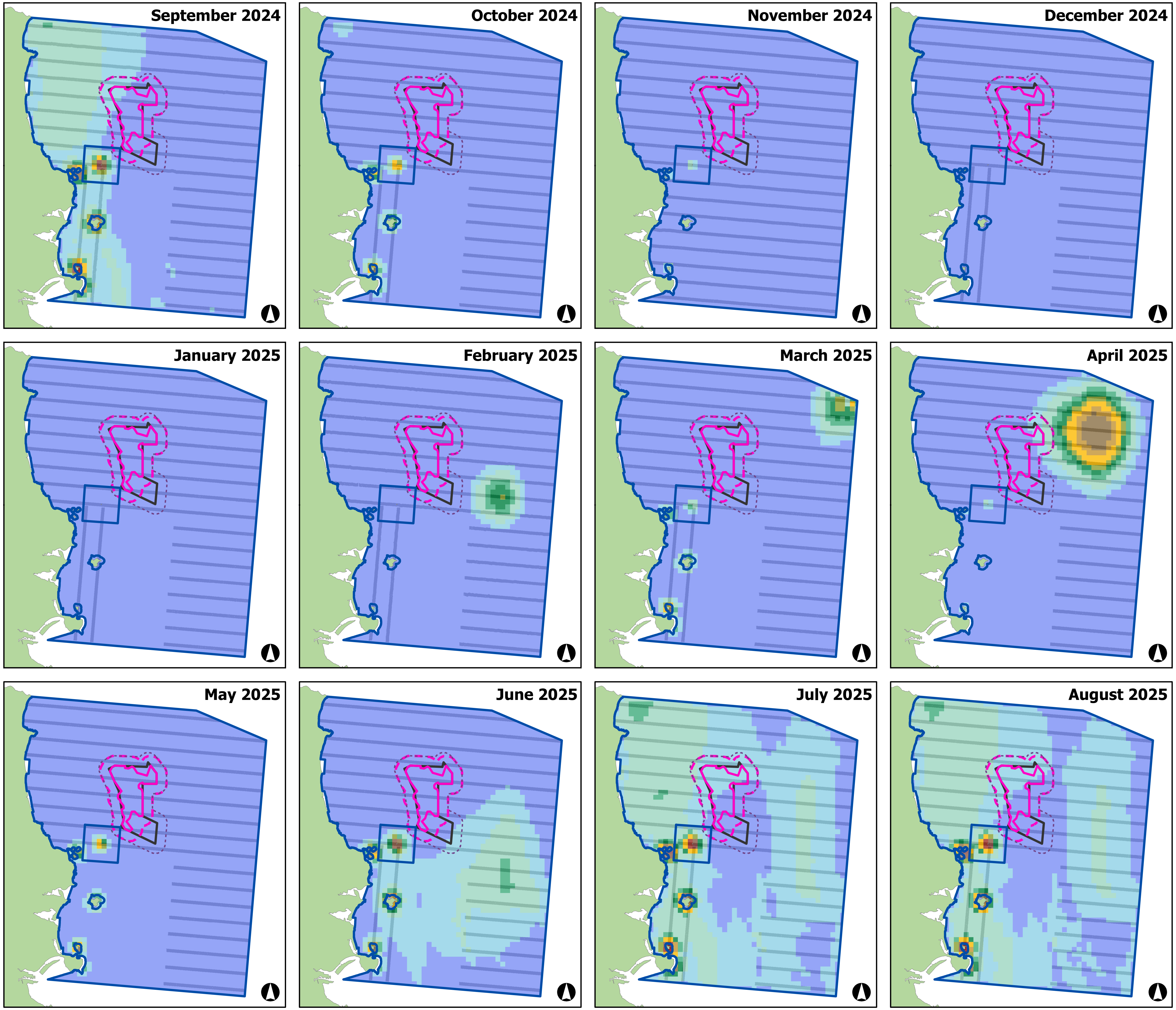
Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)
01	September	571	2,693 (1,610 – 4,037)	1.13	3,265 (1,804 – 6,000)	1.40	40 (22 – 76)	0.70	115 (64 – 212)	0.72
02	October	214	1,047 (493 – 1,907)	0.44	1,210 (549 – 2,779)	0.52	15 (6 – 35)	0.25	42 (18 – 101)	0.26
03	November	36	180 (91 – 281)	0.08	189 (81 – 439)	0.08	2 (1 – 5)	0.04	7 (3 – 16)	0.04
04	December	3	18 (3 – 36)	0.01	-	-	-	-	-	-
05	January	3	18 (3 – 48)	0.01	-	-	-	-	-	-
06	February	102	594 (185 – 1,124)	0.25	528 (280 – 996)	0.23	1 (0 – 3)	0.02	4 (2 – 9)	0.02
07	March	172	667 (259 – 1,299)	0.28	977 (384 – 2,699)	0.42	0 (0 – 1)	0.00	1 (0 – 3)	0.01
08	April	665	4,311 (1,519 – 8,271)	1.81	3,396 (1,701 – 7,137)	1.46	24 (16 – 39)	0.42	91 (58 – 148)	0.57
09	May	133	866 (438 – 1,379)	0.36	1,009 (520 – 1,938)	0.43	12 (6 – 24)	0.21	35 (18 – 67)	0.22
10	June	458	2,127 (1,375 – 2,970)	0.89	2,718 (1,525 – 4,949)	1.16	24 (12 – 47)	0.42	71 (36 – 140)	0.44
11	July	910	4,434 (2,756 – 6,420)	1.86	5,139 (2,950 – 9,134)	2.20	62 (34 – 120)	1.08	178 (96 – 338)	1.11
12	August	877	4,261 (2,908 – 5,767)	1.79	4,956 (2,821 – 8,666)	2.12	60 (36 – 99)	1.04	172 (102 – 283)	1.07

Density Estimate Plots

4.6.7 The modelled gannet density estimates show consistently low densities across the Survey Area for most of the year, with broad areas of uniform low density and only small, isolated, patches of elevated site usage (Figure A4-6).

4.6.8 From February, a localised area of increased density develops to the east of the PFI, strengthening into a clear hotspot during March and April, although densities elsewhere in the region remain low to moderate. Within the summer months (June – August), densities are slightly increased but are spread across the Survey Area. Overall, the plot indicates a brief period of elevated and spatially focused gannet activity in late spring and early summer, with low use of the Survey Area throughout the rest of the year.

4.6.9 Clustering aligns with fishing vessel activity (Table A3-1) that was recorded in similar areas within the Survey Area in the April 2025 survey report (APEM, 2024 & 2025).



▭ Array Area
⬚ 2km Buffer from Array Area
▭ Projected Footprint of Infrastructure
⬚ 2km Buffer from Projected Footprint of Infrastructure
▭ North-West Irish Sea SPA
═ Survey Transect Layout

Gannet Density (birds per km^2)

🟦	0 - 1
🟦	1 - 2
🟩	2 - 4
🟩	4 - 6
🟩	6 - 8
🟩	8 - 10
🟨	10 - 15
🟨	15 - 20
🟫	20 - 30
🟫	30 - 40
🟫	40 - 50
🟥	50 - 75
🟥	75 - 102.5

Kilometres
0 5 10 15 20 25

NISA
North Irish Sea Array

ARUP GoBe

Project
North Irish Sea Array Offshore Wind Farm

Figure Title
Model-based Density Estimates of Gannet (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:750,000 @A3	**A4-6**
Status: Issue	

4.7 Common Tern

4.7.1 The following section presents model-based abundance estimates and density estimate plots derived from common tern MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.7.2 It should be noted that the limited number of surveys with common tern observations does not support robust model-based inference, and diagnostic plots (A.4) indicate poor model fit, suggesting the model does not adequately represent the species' distribution. Consequently, design-based abundance and density estimates should be considered the more reliable and robust basis for further assessment.

Abundance Estimates

4.7.3 Common tern were only recorded in the September 2024 and May to August 2025 surveys. The raw count for September 2024 were too low to run model-based estimates (see Section 3.4). The peak raw count within the Survey Area was recorded in August 2025, with 349 individuals. Raw counts for each month are presented in Table A4-13.

4.7.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-13 and Table A4-14. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 148 common tern. Within the PFI alone, the peak unapportioned abundance was estimated at 59 common tern for the same period. Across the Survey Area, the peak estimated abundance was 2,036 common tern in August 2025 (Table A4-13).

4.7.5 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in July 2025 at 275 common tern. Within the PFI alone, the peak apportioned abundance was estimated at 96 common tern in both July and August 2025. Across the Survey Area, the peak estimated abundance was 3,973 common tern in August 2025 (Table A4-14).

Table A4-13 The raw count of individual common tern recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	3	18 (3 – 48)	0.01	-	-	-	-	-	-
02	October	0	0 (0 – 0)	-	-	-	-	-	-	-
03	November	0	0 (0 – 0)	-	-	-	-	-	-	-
04	December	0	0 (0 – 0)	-	-	-	-	-	-	-
05	January	0	0 (0 – 0)	-	-	-	-	-	-	-
06	February	0	0 (0 – 0)	-	-	-	-	-	-	-
07	March	0	0 (0 – 0)	-	-	-	-	-	-	-
08	April	0	0 (0 – 0)	-	-	-	-	-	-	-
09	May	17	97 (33 – 170)	0.04	113 (27 – 538)	0.05	0 (0 – 2)	0.01	1 (0 – 7)	0.01
10	June	34	132 (22 – 287)	0.06	209 (69 – 669)	0.09	0 (0 – 0)	0.00	0 (0 – 2)	0.00
11	July	93	517 (251 – 795)	0.22	505 (226 – 1,284)	0.22	38 (22 – 65)	0.66	97 (57 – 164)	0.61
12	August	349	1,997 (760 – 3,853)	0.86	2,036 (825 – 5,764)	0.87	59 (30 – 112)	1.02	148 (76 – 288)	0.92

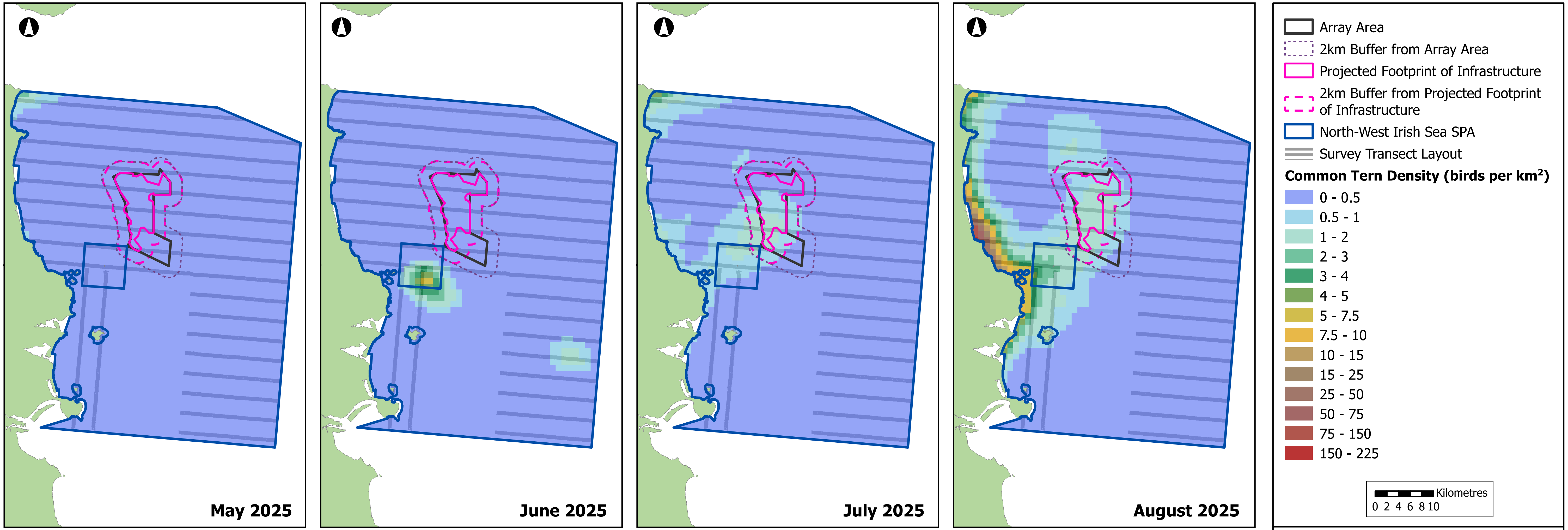
Table A4-14 The raw count of individual common tern recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	3	37 (6 – 74)	0.02	-	-	-	-	-	-
02	October	0	-	-	-	-	-	-	-	-
03	November	0	-	-	-	-	-	-	-	-
04	December	0	-	-	-	-	-	-	-	-
05	January	0	-	-	-	-	-	-	-	-
06	February	0	-	-	-	-	-	-	-	-
07	March	0	-	-	-	-	-	-	-	-
08	April	0	-	-	-	-	-	-	-	-
09	May	17	352 (124 – 670)	0.15	368 (282 – 793)	0.16	0 (0 – 2)	0.01	1 (0 – 7)	0.01
10	June	34	529 (22 – 1,270)	0.22	606 (466 – 1,066)	0.26	0 (0 – 0)	0.00	0 (0 – 2)	0.00
11	July	93	1,857 (1,090 – 2,899)	0.78	1,845 (1,566 – 2,624)	0.79	96 (80 – 123)	1.66	275 (235 – 342)	1.72
12	August	349	3,934 (1,727 – 6,990)	1.65	3,973 (2,762 – 7,701)	1.70	96 (67 – 149)	1.66	231 (159 – 371)	1.44

Density Estimate Plots

4.7.6 The modelled common tern density estimates show very low and spatially limited use of the Survey Area across all months assessed (Figure A4-7). In May, densities are uniformly low, with only a small, localised area of slightly elevated density near the southern boundary of the PFI. A similar pattern continues in June, with a small, isolated hotspot developing close to the inshore edge of the Survey Area but remaining highly limited in extent. By July, densities remain low across almost the entire region, with no meaningful areas of concentrated activity. In August, a narrow band of increased density appears along the inshore boundary outside the PFI footprint, but this remains localised and does not extend into the PFI.

4.7.7 Overall, there is consistently low common tern presence throughout the survey period, with only short-lived, small-scale areas of higher predicted density near the coast.



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Common Tern (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: December 2025	**Figure No:**
Scale: 1:675,000 @A3	**A4-7**
Status: Issue	

4.8 Roseate Tern

4.8.1 The following section presents model-based abundance estimates and density estimate plots derived from roseate tern MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.8.2 It should be noted that the limited number of surveys with roseate tern observations does not support robust model-based inference, and diagnostic plots (A.4) indicate poor model fit, suggesting the model does not adequately represent the species' distribution. Consequently, design-based abundance and density estimates should be considered the more reliable and robust basis for further assessment.

Abundance Estimates

4.8.3 Roseate tern were only recorded in the May to August 2025 surveys. The peak raw count within the Survey Area was recorded in August 2025, with 540 individuals. Raw counts for each month are presented in Table A4-15.

4.8.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-15 and Table A4-16. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 443 Roseate tern. Within the PFI alone, the peak unapportioned abundance was estimated at 125 Roseate tern for the same period. Across the Survey Area, the peak estimated abundance was 2,671 Roseate tern in August 2025 (Table A4-15).

4.8.5 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in August 2025 at 535 Roseate tern. Within the PFI alone, the peak apportioned abundance was estimated at 161 Roseate tern for the same period. Across the Survey Area, the peak estimated abundance was 4,042 Roseate tern in August 2025 (Table A4-16).

Table A4-15 The raw count of individual roseate tern recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Unapportioned)		Model-based Abundance Estimates - Survey Area (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	0	0 (0 – 0)	-	-	-	-	-	-	-
02	October	0	0 (0 – 0)	-	-	-	-	-	-	-
03	November	0	0 (0 – 0)	-	-	-	-	-	-	-
04	December	0	0 (0 – 0)	-	-	-	-	-	-	-
05	January	0	0 (0 – 0)	-	-	-	-	-	-	-
06	February	0	0 (0 – 0)	-	-	-	-	-	-	-
07	March	0	0 (0 – 0)	-	-	-	-	-	-	-
08	April	0	0 (0 – 0)	-	-	-	-	-	-	-
09	May	209	1,309 (497 – 2,451)	0.56	2,140 (1,003 – 4,723)	0.92	7 (3 – 19)	0.13	31 (13 – 77)	0.20
10	June	201	1,086 (221 – 2,426)	0.47	1,409 (614 – 3,282)	0.60	9 (4 – 21)	0.15	40 (18 – 91)	0.25
11	July	265	1,383 (550 – 2,487)	0.59	1,284 (739 – 2,472)	0.55	47 (27 – 85)	0.82	195 (108 – 362)	1.22
12	August	540	2,738 (610 – 5,761)	1.17	2,671 (1,780 – 4,379)	1.14	125 (82 – 199)	2.17	443 (309 – 658)	2.77

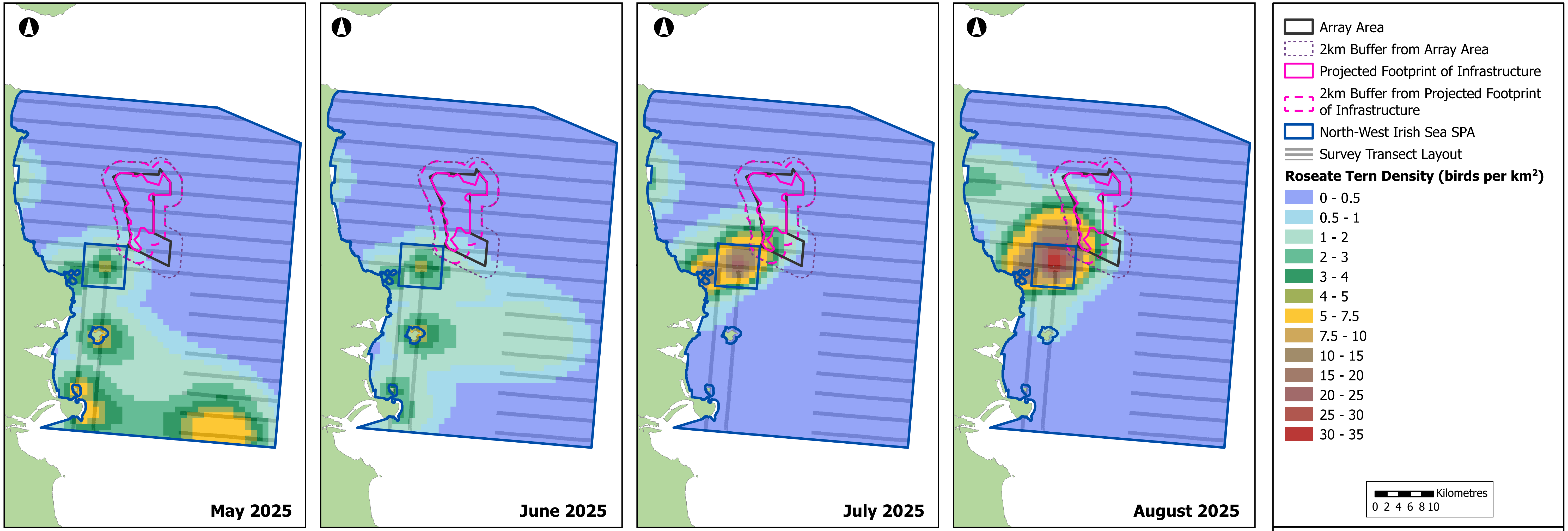
Table A4-16 The raw count of individual roseate tern recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned)		Model-based Abundance Estimates - Survey Area (Apportioned)		Model-based Abundance Estimates - PFI (Apportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	0	-	-	-	-	-	-	-	-
02	October	0	-	-	-	-	-	-	-	-
03	November	0	-	-	-	-	-	-	-	-
04	December	0	-	-	-	-	-	-	-	-
05	January	0	-	-	-	-	-	-	-	-
06	February	0	-	-	-	-	-	-	-	-
07	March	0	-	-	-	-	-	-	-	-
08	April	0	-	-	-	-	-	-	-	-
09	May	209	1,815 (806 – 3,079)	0.76	2,646 (1,509 – 5,229)	1.13	7 (3 – 19)	0.13	31 (13 – 77)	0.20
10	June	201	1,227 (309 – 2,549)	0.51	1,550 (755 – 3,423)	0.66	9 (4 – 21)	0.15	40 (18 – 91)	0.25
11	July	265	2,221 (958 – 3,678)	0.93	2,122 (1,577 – 3,310)	0.91	70 (50 – 108)	1.22	250 (163 – 417)	1.56
12	August	540	4,109 (1,349 – 7,631)	1.72	4,042 (3,151 – 5,750)	1.73	161 (118 – 235)	2.79	535 (401 – 750)	3.34

Density Estimate Plots

4.8.6 The modelled roseate tern density estimates shows low, and spatially limited, densities in May and June, during which only small patches of elevated spatial clustering occurs near inshore areas south of the PFI (Figure A4-8). Densities increase substantially by July, when a defined hotspot develops immediately south of the PFI footprint, indicating a period of concentrated local use. This intensifies further in August, with a strong and spatially extensive hotspot centred along the inshore boundary and extending into the buffer area, while densities across the wider region remain generally low to moderate.

4.8.7 Overall, the plots indicate that roseate tern activity within the Survey Area is focused in mid to late summer, with pronounced localised peaks near the southern boundary of the PFI and minimal use during early summer.



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Roseate Tern (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: December 2025	**Figure No:**
Scale: 1:675,000 @A3	**A4-8**
Status: Issue	

4.9 Red-throated Diver

4.9.1 The following section presents model-based abundance estimates and density estimate plots derived from red-throated diver MRSea analysis. Details regarding model summary, setup and final validation plots can be found in the appendices (A.2, A.3 & A.4).

4.9.2 It should be noted in the case of red-throated diver, model performance is constrained by the species' tendency to occur close to the coast, meaning it is absent from much of the wider Survey Area and therefore poorly represented in the model. Consequently, diagnostic plots (A.4) indicate poor model fit, suggesting it does not adequately represent the species' distribution. It is recommended that design-based abundance and density estimates should be considered the more reliable and robust basis for further assessment.

Abundance Estimates

4.9.3 Red-throated diver were recorded in all surveys apart from the June to August 2025 surveys. The peak raw count within the Survey Area was recorded in November 2024, with 233 individuals. Raw counts for each month are presented in Table A4-17.

4.9.4 Unapportioned and apportioned model-based population estimates are presented in Table A4-17 and Table A4-18. The peak unapportioned abundance in the PFI plus 2 km buffer was estimated in March 2025 at 8 red-throated diver. Within the PFI alone, the peak unapportioned abundance was estimated at 1 red-throated diver for the same period. Across the Survey Area, the peak estimated abundance was 1,125 red-throated diver in November 2024 (Table A4-17).

4.9.5 The peak apportioned abundance in the PFI plus 2 km buffer was estimated in March 2025 at 8 red-throated diver. Within the PFI alone, the peak apportioned abundance was estimated at 1 red-throated diver for the same period. Across the Survey Area, the peak estimated abundance was 1,481 red-throated diver in November 2024 (Table A4-18).

Table A4-17 The raw count of individual red-throated diver recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Estimates – Survey Area (Unapportioned)		Model-based Estimates - Survey Area (Unapportioned)		Model-based Estimates - PFI (Unapportioned)		Model-based Estimates - PFI plus 2km buffer (Unapportioned)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	102	610 (186 – 1,130)	0.26	527 (323 – 915)	0.23	0 (0 – 0)	0.00	0 (0 – 0)	0.00
02	October	188	1,136 (397 – 2,015)	0.49	990 (457 – 2,172)	0.42	0 (0 – 0)	0.00	0 (0 – 2)	0.00
03	November	233	1,148 (628 – 1,737)	0.49	1,125 (606 – 2,087)	0.48	0 (0 – 1)	0.00	1 (0 – 5)	0.01
04	December	135	813 (408 – 1,333)	0.35	700 (403 – 1,210)	0.30	0 (0 – 0)	0.00	1 (0 – 3)	0.00
05	January	150	903 (449 – 1,401)	0.39	808 (534 – 1,274)	0.35	0 (0 – 0)	0.00	0 (0 – 1)	0.00
06	February	121	714 (389 – 1,070)	0.31	640 (423 – 1,002)	0.27	0 (0 – 0)	0.00	0 (0 – 1)	0.00
07	March	131	799 (355 – 1,341)	0.34	658 (448 – 1,064)	0.28	1 (0 – 8)	0.02	8 (3 – 35)	0.05
08	April	133	840 (310 – 1,615)	0.36	899 (534 – 1,526)	0.39	0 (0 – 0)	0.00	0 (0 – 1)	0.00
09	May	19	126 (26 – 255)	0.05	198 (80 – 495)	0.08	0 (0 – 0)	0.00	0 (0 – 1)	0.00
10	June	0	-	-	-	-	-	-	-	-
11	July	0	-	-	-	-	-	-	-	-
12	August	0	-	-	-	-	-	-	-	-

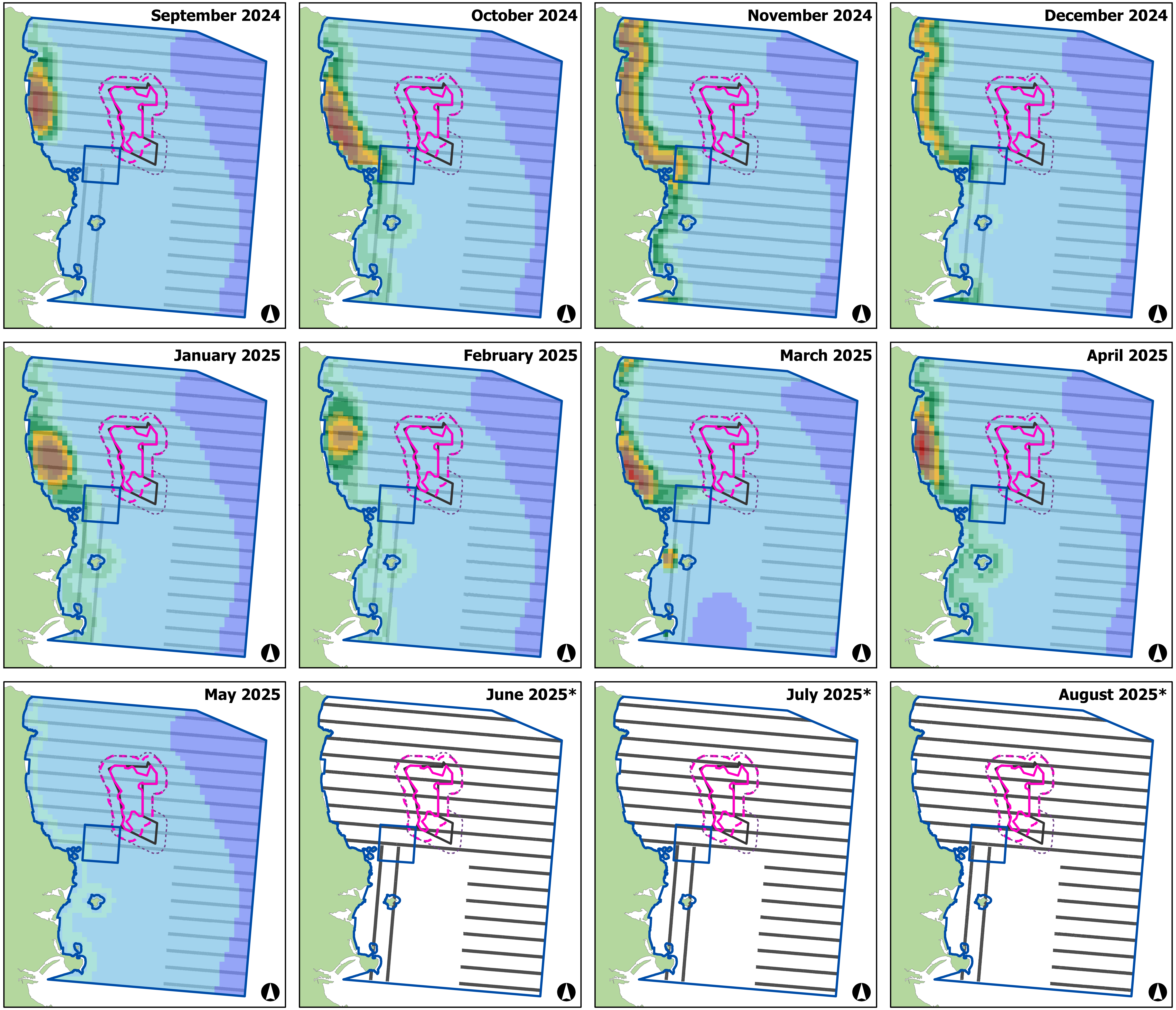
Table A4-18 The raw count of individual red-throated diver recorded within each survey, alongside apportioned and corrected for availability bias design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the Survey Area, PFI and PFI plus a 2km buffer. Raw count data in red indicates that the numbers were too small to successfully run the model.

Survey No.	Survey Month	Raw Count (Survey Area)	Design-based Abundance Estimates – Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - Survey Area (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI (Apportioned and Corrected for Availability Bias)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned and Corrected for Availability Bias)	
			Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	September	102	610 (186 – 1,130)	0.26	527 (323 – 915)	0.23	0 (0 – 0)	0.00	0 (0 – 0)	0.00
02	October	188	1,136 (397 – 2,015)	0.48	990 (457 – 2,172)	0.42	0 (0 – 0)	0.00	0 (0 – 2)	0.00
03	November	233	1,512 (832 – 2,272)	0.63	1,481 (804 – 2,737)	0.63	0 (0 – 0)	0.00	2 (3 – 19)	0.01
04	December	135	1,162 (596 – 1,867)	0.49	1,005 (595 – 1,711)	0.43	0 (0 – 0)	0.00	1 (2 – 7)	0.01
05	January	150	1,377 (734 – 2,101)	0.58	1,245 (863 – 1,894)	0.53	0 (0 – 0)	0.00	1 (1 – 1)	0.00
06	February	121	783 (459 – 1,152)	0.33	709 (492 – 1,071)	0.30	0 (0 – 0)	0.00	0 (0 – 1)	0.00
07	March	131	799 (355 – 1,341)	0.33	658 (448 – 1,064)	0.28	1 (0 – 9)	0.02	8 (3 – 43)	0.05
08	April	133	846 (311 – 1,622)	0.35	905 (540 – 1,532)	0.39	0 (0 – 0)	0.00	0 (0 – 2)	0.00
09	May	19	133 (27 – 266)	0.06	205 (87 – 502)	0.09	0 (0 – 0)	0.00	0 (0 – 1)	0.00
10	June	0	-	-	-	-	-	-	-	-
11	July	0	-	-	-	-	-	-	-	-
12	August	0	-	-	-	-	-	-	-	-

Density Estimate Plots

4.9.6 The modelled red-throated diver density estimates show strong seasonality, with the highest densities occurring during autumn and spring (Figure A4-9). From September through January, a pronounced inshore band of elevated density extends along the coastline, with the highest values concentrated west of the PFI and tapering southwards. During this period, modelled densities within the PFI footprint remain generally low, with most activity confined to near-coastal waters outside of the development area.

4.9.7 From February onward, overall densities decline markedly, with only specific high clustering observed in March and April. Very low densities persist in May and beyond that limited raw observations mean that no MRSea analysis was carried out across summer months.



*No MRSea analysis carried out for this month

NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Estimates of Red-throated Diver (km^2)

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:750,000 @A3	**A4-9**
Status: Issue	

5 Summary

5.1 Guillemot

5.1.1 Guillemot abundance estimates peaked across the survey area in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.1.2 Guillemot density estimates showed strong seasonality, with low winter use and pronounced hotspots forming south and northeast of the PFI during early summer in line with post-breeding season dispersal. Highest estimated densities of guillemot were predominantly outside of the PFI, and associated 2km buffer, occurring within the wider Survey Area.

5.1.3 Low counts in April 2025, potentially due to the implementation of survey Plan C, means that important areas may be under-represented in the modelled estimates. This limitation should be considered when interpreting modelled estimates.

5.1.4 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.2 Razorbill

5.2.1 Razorbill abundance estimates peaked in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.2.2 Razorbill density estimates were low through autumn and winter, peaking in early summer when a clear hotspot developed south of the PFI in line with expected post-breeding season dispersal. Highest estimated densities of razorbill occurred outside of the PFI, and associated 2km buffer, occurring within the wider Survey Area.

5.2.3 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.3 Kittiwake

5.3.1 Kittiwake abundance estimates peaked in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.3.2 Kittiwake densities varied seasonally, with highest values in winter and smaller summer hotspots occurring northeast and southwest of the PFI within the wider Survey Area.

5.3.3 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.4 Herring gull

5.4.1 Herring gull abundance estimates peaked in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.4.2 Herring gull densities peaked in winter, with additional localised summer hotspots forming on the northeast and southwest edges PFI boundary. Higher densities were predominantly located close to the coastline and across the wider Survey Area.

5.4.3 Elevated abundance estimates and spatial clustering appears to match fishing vessel activity occurring within the corresponding months.

5.4.4 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.5 Great black-backed gull

5.5.1 Great black-backed gull abundance estimates peaked in April 2025. However, peak abundance estimates for the PFI and PFI plus 2km buffer was noted in November 2024.

5.5.2 Great black-backed gull densities showed limited variation, with only localised clustering observed in early winter and late spring. While some of this clustering was observed to be within the PFI, highest densities were located within the wider Survey Area.

5.5.3 Elevated abundance estimates and spatial clustering appears to match fishing vessel activity occurring within the corresponding months.

5.5.4 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.6 Gannet

5.6.1 Gannet abundance estimates peaked in July 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.6.2 Gannet use of the Survey Area was generally low year-round, with a brief period of increased activity east of the PFI boundary and associated 2km buffer in late spring.

5.6.3 Elevated abundance estimates and spatial clustering appears to match fishing vessel activity occurring within the corresponding months.

5.6.4 It was determined that model-fit was acceptable resulting in realistic spatial distribution and robust abundance estimates. It is therefore recommended that MRSea abundance and density estimates are appropriate for use in associated assessments.

5.7 Common tern

5.7.1 Common tern abundance estimates peaked in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.7.2 Common tern densities remained consistently low across all months, with only small, short-lived coastal hotspots away from the PFI and associated 2km buffer.

5.7.3 It was determined that model-fit was not acceptable resulting in unrealistic spatial distribution and unstable abundance estimates. It is therefore recommended that design-based abundance and density estimates are appropriate for use in associated assessments.

5.8 Roseate tern

5.8.1 Roseate tern abundance estimates peaked in August 2025, with similar peaks noted for the PFI and PFI plus 2km buffer within the same period.

5.8.2 Roseate tern densities were low in early summer, rising sharply to form a focused hotspot south of the PFI, in July and August, that was predominantly within the wider Survey Area.

5.8.3 It was determined that model-fit was not acceptable resulting in unrealistic spatial distribution and unstable abundance estimates. It is therefore recommended that design-based abundance and density estimates are appropriate for use in associated assessments.

5.9 Red-throated Diver

5.9.1 Red-throated diver abundance estimates peaked in November 2024. However, peak abundance estimates for the PFI and PFI plus 2km buffer was noted in March 2025.

5.9.2 Red-throated diver densities peaked along the inshore coastline in autumn and winter, declining to very low levels from late spring onwards. Peak densities occurred in proximity to the coastline and not within the PFI or associated 2km buffer.

5.9.3 It was determined that model-fit was not acceptable resulting in unrealistic spatial distribution and unstable abundance estimates. It is therefore recommended that design-based abundance and density estimates are appropriate for use in associated assessments.

6 References

APEM (2024 & 2025) Ornithological and Marine Mammal Baseline Characterisation Survey Results: September 2024 – August 2025. APEM Scientific Monthly Report P00016983. NISA.

ArcGIS (2022) European Countries Data. https://www.arcgis.com/home/item.html?id=1c7a24b6b54e4877906213b4f97e12e0 (Last Accessed November 2025).

Dunn, R. E., Duckworth, J., O'Brien, S., Furness, R. W., Buckingham, L., Daunt, F., Bogdanova, M., & Green, J. A. (2024). Temporal and spatial variability in availability bias has consequences for marine bird abundance estimates during the non-breeding season. Ecological Solutions and Evidence, 5(4), e12373. https://doi.org/10.1002/2688-8319.12373

Fox, J., & Weisberg, S. (2015). Using car Functions in Other Functions.

GEBCO Compilation Group (2024). GEBCO 2024 Grid. https://doi.org/10.5285/1c44ce99-0a0d-5f4f-e063-7086abc0ea0f (Last Accessed November 2025).

JNCC (2024). Boundaries of Special Protection Areas (SPAs) of Great Britain (including offshore areas). Boundaries of Special Protection Areas (SPAs) of Great Britain (including offshore areas) | JNCC Resource Hub (Last Accessed November 2025).

Langton, R., Boulcott, P. and Wright P.J. (2021). A verified distribution model for the lesser sandeel *Ammodytes marinus*. Marine Ecology Progress Series. https://doi.org/10.3354/meps13693

National Parks & Wildlife Service (2024). Designated site data. https://www.npws.ie/maps-and-data/designated-site-data (Last Accessed November 2025).

NatureScot (2023). Guidance Note 2: Guidance to support Offshore Wind Application: Advice for Marine Ornithology Baseline Characterisation Surveys and Reporting. https://www.nature.scot/doc/guidance-note-2-guidance-support-offshore-wind-applications-advice-marine-ornithology-baseline#3.2+Baseline+characterisation (Last Accessed November 2025).

NPWS (2023) Conservation Objectives: North-west Irish Sea SPA 004236. Version 1. National Parks and Wildlife Service, Department of Housing, Local Government and Heritage. Available at: CO004236.pdf

Onoufriou, J., Russell, D. J. F., Thompson, D., Moss, S. E., & Hastie, G. D. (2021). Quantifying the effects of tidal turbine array operations on the distribution of marine mammals: Implications for collision risk. Renewable Energy, 180, 157–165. https://doi.org/10.1016/j.renene.2021.08.052

Scott-Hayward, L., Mackenzie, M. L., Donovan, C. R., Walker, C. G., & Ashe, E. (2014). Complex region spatial smoother (CReSS). Journal of Computational and Graphical Statistics, 23(2), 340–360. https://doi.org/10.1080/10618600.2012.762920

Scott-Hayward, L., Oedekoven, C., Mackenzie, M., & Rexstad, E. (2013). Statistical Modelling of bird and cetacean distributions in offshore renewables development areas.

Thaxter, C. B., Wanless, S., Daunt, F., Harris, M. P., Benvenuti, S., Watanuki, Y., Grémillet, D., & Hamer, K. C. (2010). Influence of wing loading on the trade-off between pursuit-diving and flight in common guillemots and razorbills. The Journal of Experimental Biology, 213, 1018–1025. https://doi.org/10.1242/jeb.037390

A Appendices

A.1 MRSea outputs (Original dataset)

Guillemot

Abundance estimates

A.1.1 Modelled outputs, alongside raw observations and associated DBE, from the original 29-month dataset have been clipped to the PFI and PFI plus 2km buffer to provide updated abundance and density estimates (Table A6-1 and Table A6-2). For further detail on the model setup and validation can be found in Appendix A15.2 MRSea Modelling for Offshore Ornithology.

Table A6-1 The raw count of individual guillemots recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer.

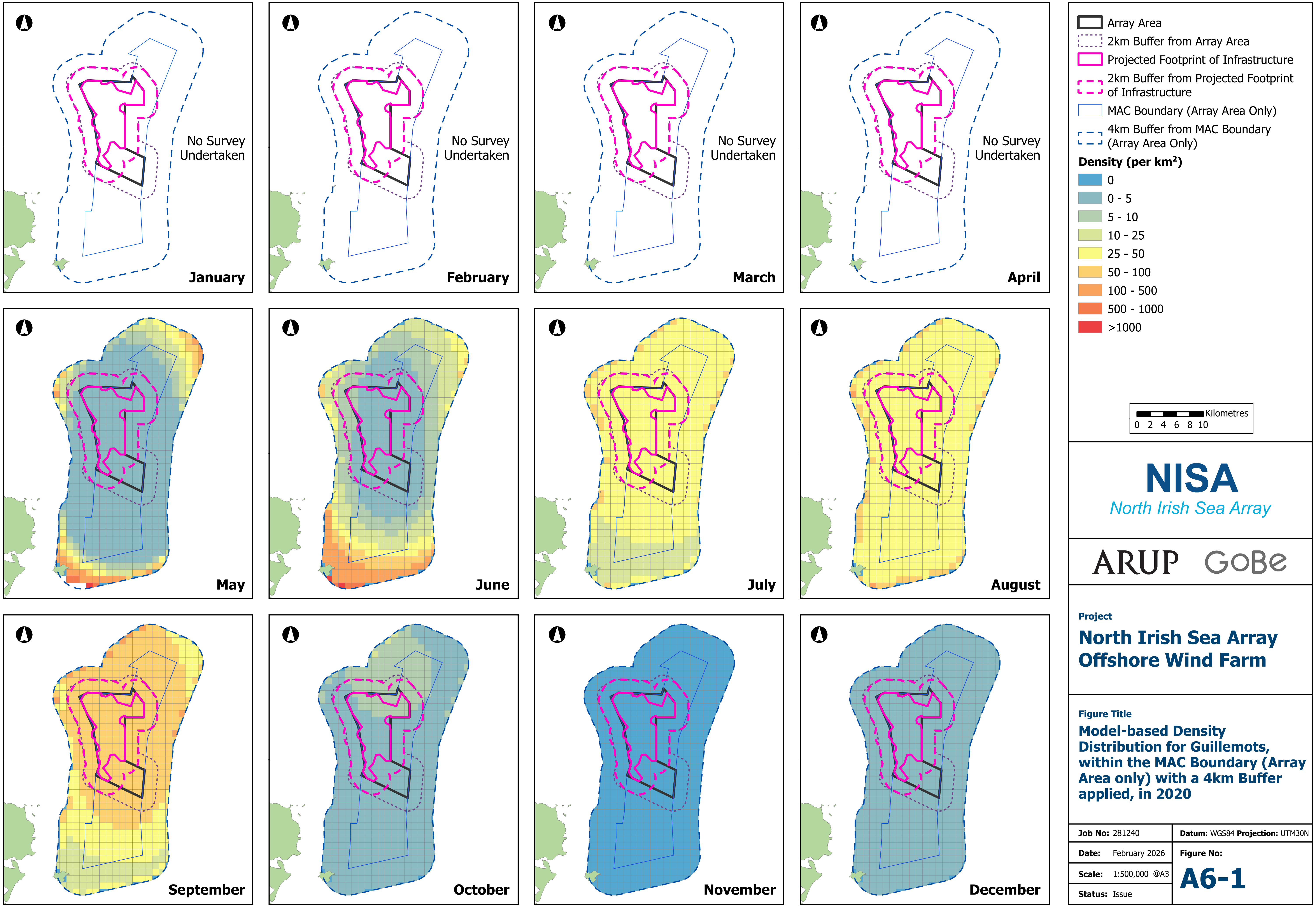
Survey No.	Survey Month	Raw Count (PFI)	Raw Count (PFI plus a 2km buffer)	Design-based Abundance Estimates – PFI (Unapportioned)		Design-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
				Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	May, 2020	48	122	195 (105 — 295)	3.38	493 (344 — 659)	3.08	19 (3 — 421)	0.34	114 (47 — 1,376)	0.71
02	June, 2020	36	133	116 (49 — 193)	2.01	442 (296 — 609)	2.76	200 (132 — 321)	3.48	659 (428 — 1,104)	4.12
03	July, 2020	607	1,666	2,040 (823 —	35.38	5,705 (3,615 —	35.62	2,274 (1,856 —	39.45	6,100 (4,986 — 7,597)	38.09
04	August, 2020	399	1,126	1,334 (701 —	23.14	3,891 (2,335 —	24.30	2,018 (1,390 —	35.01	5,593 (3,823 — 8,176)	34.92
05	September,	881	2,768	2,893 (2,275 —	50.18	9,253 (7,571 —	57.78	5,031 (4,043 —	87.26	13,137 (10,450 — 16,448)	82.03
06	October,	71	185	238 (71 — 545)	4.13	618 (219 —	3.86	240 (136 — 432)	4.17	615 (356 — 1,091)	3.84
07	November,	0	0	-	-	-	-	-	-	-	-
08	December,	13	29	42 (23 — 62)	0.73	95 (49 — 138)	0.59	39 (26 — 57)	0.67	104 (68 — 159)	0.65
09	February,	0	0	-	-	-	-	-	-	-	-
10	March, 2021	318	982	1,057 (455 —	18.33	3,320 (2,060 —	20.73	1,830 (1,454 —	31.74	4,888 (3,824 — 6,223)	30.52
11	April, 2021	39	168	128 (59 — 203)	2.22	562 (439 — 704)	3.51	408 (337 — 499)	7.07	1,098 (900 — 1,355)	6.85
12	May, 2021	86	197	287 (192 — 380)	4.98	663 (515 — 805)	4.14	93 (32 — 432)	1.61	336 (141 — 1,455)	2.10
13	June, 2021	52	135	171 (52 — 315)	2.97	451 (163 — 769)	2.82	92 (32 — 262)	1.59	301 (111 — 868)	1.88
14	July, 2021	84	220	281 (84 — 693)	4.87	753 (268 —	4.70	160 (51 — 547)	2.77	520 (185 — 1,684)	3.24
15	August, 2021	1,285	3,621	4,195 (2,205 —	72.76	11,985 (7,682 —	74.83	7,112 (5,804 —	123.36	19,639 (15,874 — 24,503)	122.62
16	September,	2,000	3,796	6,671 (2,689 —	115.71	12,836 (8,202 —	80.15	5,337 (4,118 —	92.57	15,124 (11,700 — 19,935)	94.43
17	October,	205	576	676 (235 —	11.73	1,936 (920 —	12.09	838 (634 —	14.54	2,263 (1,709 — 3,021)	14.13
18	November,	256	560	870 (407 —	15.09	1,968 (1,144 —	12.29	937 (751 —	16.25	2,525 (2,023 — 3,135)	15.77
19	December,	83	204	273 (115 — 438)	4.74	735 (386 —	4.59	359 (288 — 451)	6.24	960 (740 — 1,263)	6.00
20	January, 2022	243	535	792 (444 —	13.74	1,798 (1,052 —	11.23	285 (163 — 511)	4.94	919 (533 — 1,635)	5.74
21	February,	174	461	576 (357 — 859)	9.99	1,563 (1,149 —	9.76	681 (488 — 962)	11.82	1,962 (1,398 — 2,782)	12.25
22	March, 2022	52	166	164 (52 — 375)	2.84	550 (322 — 834)	3.43	203 (118 — 336)	3.52	623 (362 — 1,049)	3.89
23	April, 2022	119	243	388 (214 — 606)	6.73	814 (466 —	5.08	460 (325 — 653)	7.98	1,280 (870 — 1,878)	7.99
24	May, 2022	65	197	212 (99 — 360)	3.68	667 (470 — 845)	4.16	86 (23 — 525)	1.50	318 (113 — 1,738)	1.98
25	June, 2022	37	135	122 (66 — 185)	2.12	451 (261 — 653)	2.82	59 (19 — 242)	1.02	214 (79 — 856)	1.34
26	July, 2022	869	2,465	2,808 (914 —	48.71	8,184 (5,581 —	51.10	3,022 (2,220 —	52.42	8,238 (5,959 — 11,470)	51.44
27	August, 2022	1,113	3,118	3,613 (2,316 —	62.67	10,455 (8,284 —	65.28	5,230 (4,656 —	90.72	14,673 (12,798 — 16,864)	91.62
28	September,	1,571	4,000	5,106 (3,181 —	88.57	13,524 (9,713 —	84.44	5,524 (4,541 —	95.82	14,972 (12,181 — 18,433)	93.48
29	October,	587	1,556	1,906 (1,303 —	33.06	5,247 (3,901 —	32.76	2,627 (2,207 —	45.57	6,975 (5,810 — 8,378)	43.55

Table A6-2 The raw count of individual guillemots recorded within each survey, alongside apportioned, and corrected for availability bias, design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (PFI)	Raw Count (PFI plus 2km buffer)	Design-based Abundance Estimates – PFI (Apportioned and corrected)		Design-based Abundance Estimates – PFI plus 2km buffer (Apportioned and corrected)		Model-based Abundance Estimates - PFI (Apportioned and corrected)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned and corrected)	
				Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	May, 2020	48	122	287 (171 — 421)	4.98	720 (515 — 939)	4.50	70 (49 — 597)	1.21	248 (160 — 1,902)	1.55
02	June, 2020	36	133	193 (91 — 301)	3.35	710 (512 — 945)	4.43	313 (223 — 470)	5.42	1,027 (724 — 1,610)	6.41
03	July, 2020	607	1,666	2,493 (1,107 —	43.24	7,078 (4,704 — 9,628)	44.19	2,762 (2,283 —	47.92	7,530 (6,253 — 9,247)	47.02
04	August, 2020	399	1,126	1,549 (826 — 2,284)	26.87	4,401 (2,677 — 6,080)	27.48	2,306 (1,612 —	40.00	6,282 (4,325 — 9,138)	39.23
05	September,	881	2,768	3,640 (2,672 —	63.14	11,839 (9,393 —	73.91	6,017 (4,919 —	104.37	16,159 (13,171 —	100.90
06	October, 2020	71	185	2,276 (788 — 4,308)	39.48	4,363 (2,361 — 6,624)	27.24	2,278 (2,156 —	39.52	4,359 (4,055 — 4,919)	27.22
07	November,	0	0	616 (293 — 969)	10.68	1,188 (691 — 1,740)	7.42	-	-	-	-
08	December,	13	29	504 (223 — 901)	8.74	1,038 (543 — 1,657)	6.48	500 (481 — 526)	8.67	1,053 (1,002 — 1,129)	6.58
09	February, 2021	0	0	149 (58 — 233)	2.58	472 (303 — 640)	2.95	-	-	-	-
10	March, 2021	318	982	1,674 (776 — 2,675)	29.04	5,257 (3,454 — 7,163)	32.82	2,767 (2,236 —	47.99	7,474 (5,969 — 9,361)	46.67
11	April, 2021	39	168	242 (131 — 348)	4.2	974 (787 — 1,179)	6.08	608 (516 — 728)	10.55	1,678 (1,419 — 2,015)	10.48
12	May, 2021	86	197	393 (266 — 515)	6.82	913 (722 — 1,109)	5.70	144 (64 — 588)	2.50	504 (249 — 1,972)	3.15
13	June, 2021	52	135	243 (69 — 430)	4.21	621 (245 — 1,042)	3.88	143 (64 — 366)	2.47	434 (185 — 1,178)	2.71
14	July, 2021	84	220	342 (84 — 813)	5.93	898 (338 — 1,770)	5.61	202 (78 — 647)	3.51	630 (247 — 1,965)	3.93
15	August, 2021	1,285	3,621	4,989 (2,732 —	86.54	14,156 (9,494 —	88.38	8,214 (6,768 —	142.48	22,619 (18,458 —	141.23
16	September,	2,000	3,796	7,941 (3,512 —	137.74	15,378 (10,161 —	96.01	6,457 (5,101 —	112.00	17,922 (14,115 —	111.91
17	October, 2021	205	576	1,031 (447 — 1,683)	17.88	2,870 (1,563 — 4,384)	17.92	1,222 (982 — 1,552)	21.19	3,255 (2,604 — 4,145)	20.32
18	November,	256	560	1,691 (919 — 2,537)	29.33	4,113 (2,494 — 5,744)	25.68	1,772 (1,545 —	30.74	4,795 (4,180 — 5,541)	29.94
19	December,	83	204	642 (308 — 993)	11.14	1,671 (924 — 2,536)	10.43	764 (663 — 893)	13.25	1,987 (1,679 — 2,412)	12.41
20	January, 2022	243	535	1,361 (859 — 1,999)	23.61	3,123 (2,028 — 4,421)	19.50	639 (465 — 962)	11.09	1,870 (1,321 — 2,891)	11.67
21	February, 2022	174	461	1,223 (823 — 1,703)	21.21	3,067 (2,374 — 3,780)	19.15	1,373 (1,098 —	23.82	3,636 (2,831 — 4,806)	22.70
22	March, 2022	52	166	279 (94 — 585)	4.84	1,026 (672 — 1,466)	6.41	334 (213 — 522)	5.79	1,129 (759 — 1,730)	7.05
23	April, 2022	119	243	553 (315 — 833)	9.59	1,163 (696 — 1,673)	7.26	650 (473 — 903)	11.28	1,780 (1,243 — 2,564)	11.12
24	May, 2022	65	197	277 (131 — 471)	4.8	878 (613 — 1,109)	5.48	121 (37 — 696)	2.10	461 (192 — 2,322)	2.88
25	June, 2022	37	135	154 (81 — 235)	2.67	574 (336 — 840)	3.58	77 (25 — 318)	1.34	285 (108 — 1,126)	1.78
26	July, 2022	869	2,465	3,235 (1,062 —	56.11	9,508 (6,506 — 12,317)	59.36	3,480 (2,560 —	60.35	9,570 (6,957 — 13,276)	59.75
27	August, 2022	1,113	3,118	4,006 (2,565 —	69.49	11,595 (9,199 —	72.39	5,794 (5,159 —	100.50	16,258 (14,186 —	101.52
28	September,	1,571	4,000	5,918 (3,770 —	102.65	15,506 (11,280 —	96.81	6,383 (5,289 —	110.71	17,116 (14,012 —	106.87
29	October, 2022	587	1,556	3,150 (2,222 —	54.64	8,287 (6,202 — 10,191)	51.74	3,997 (3,503 —	69.34	10,318 (8,948 — 11,966)	64.42

Density Estimate Plots

A.1.2 The figures below illustrate the modelled density estimates incorporating the PFI boundary and associated 2km buffer (Figure A6-1, Figure A6-2 and Figure A6-3).



NISA

North Irish Sea Array

ARUP GoBe

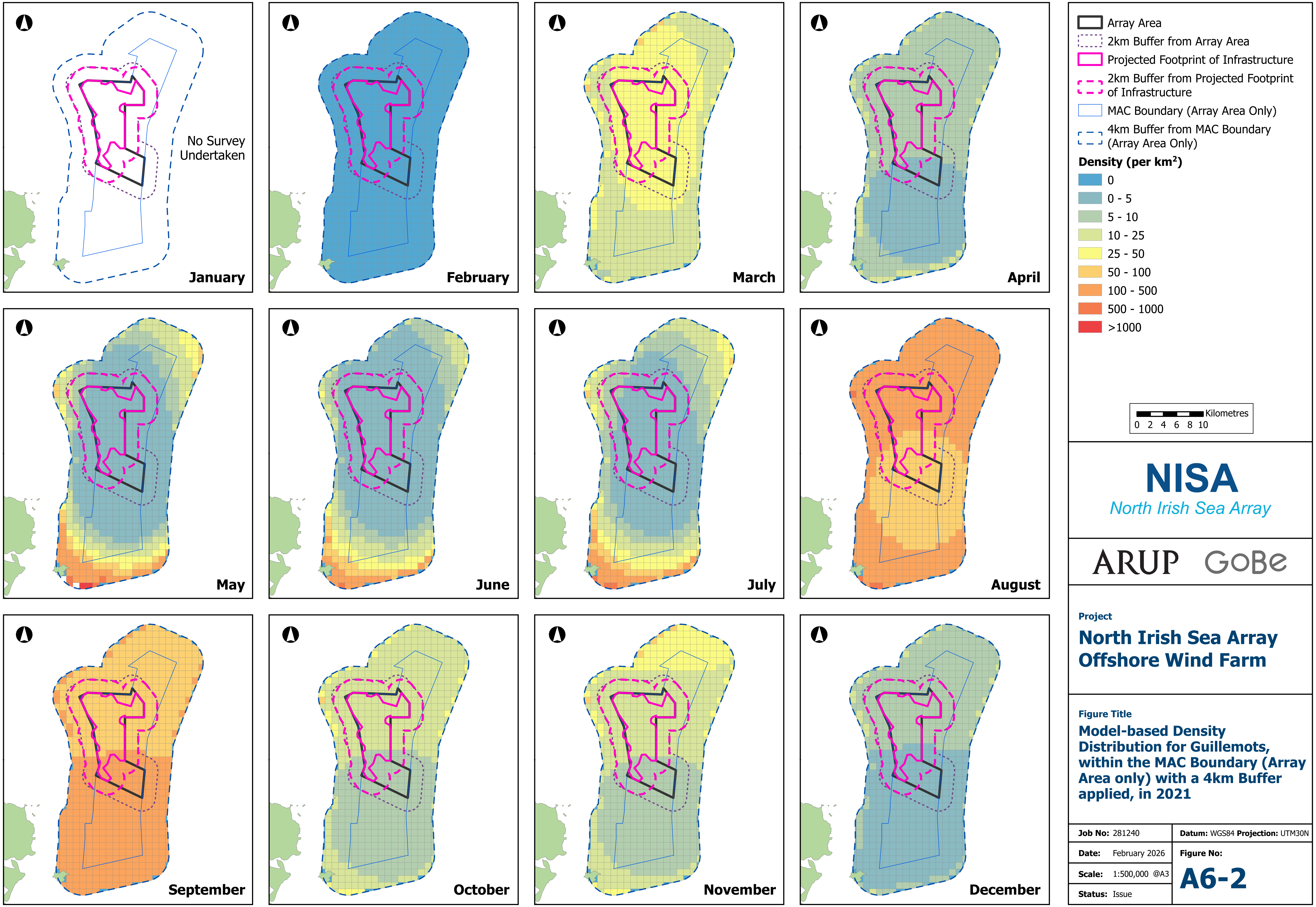
Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Distribution for Guillemots, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2020

Job No: 281240
Datum: WGS84 **Projection:** UTM30N
Date: February 2026
Figure No: A6-1
Scale: 1:500,000 @A3
Status: Issue



NISA

North Irish Sea Array

ARUP GoBe

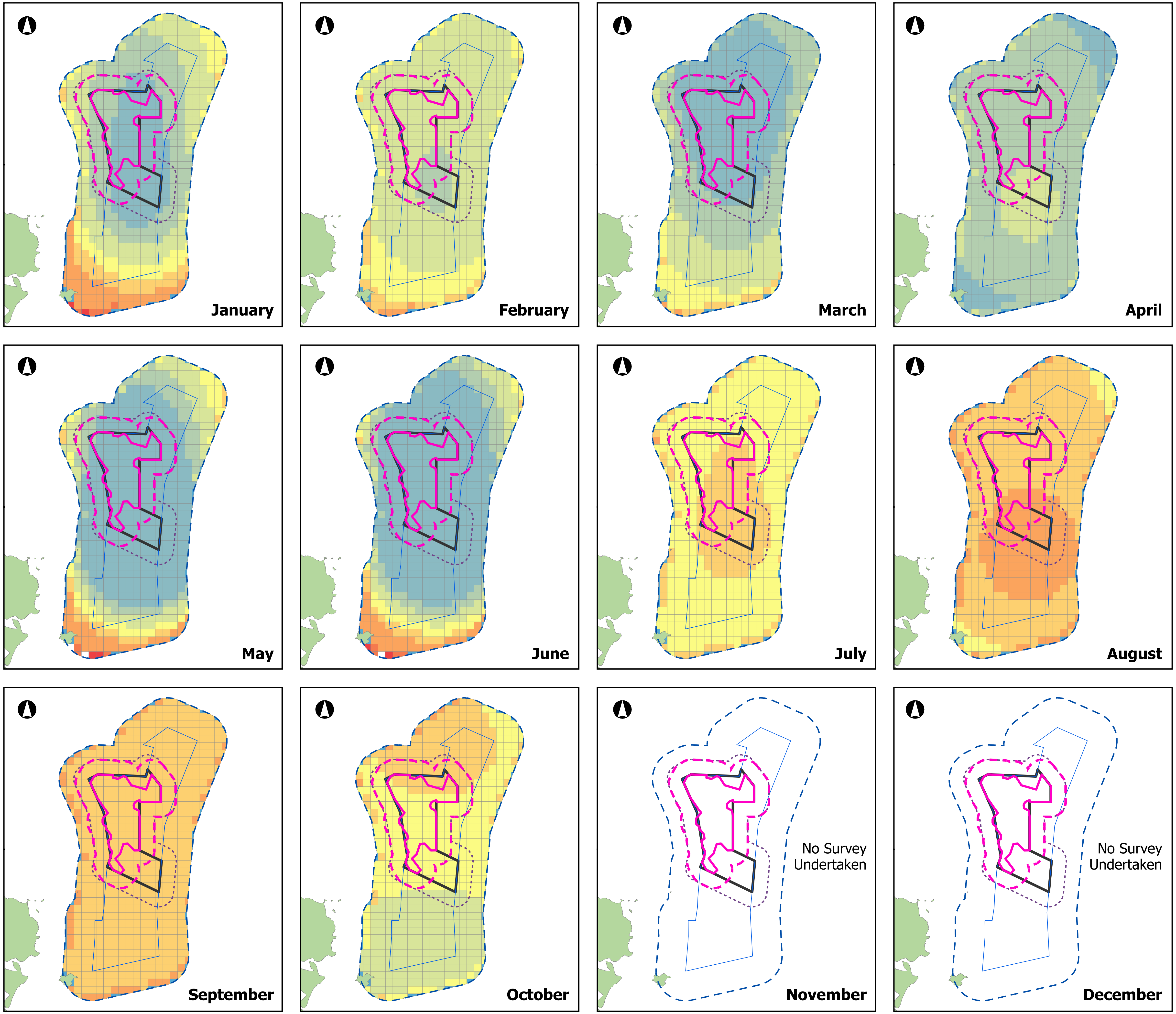
Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Distribution for Guillemots, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2021

Job No: 281240	**Datum:** WGS84 **Projection:** UTM30N
Date: February 2026	**Figure No:**
Scale: 1:500,000 @A3	**A6-2**
Status: Issue	



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Distribution for Guillemots, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2022

Job No: 281240

Datum: WGS84 **Projection:** UTM30N

Date: February 2026

Figure No: A6-3

Scale: 1:500,000 @A3

Status: Issue

Razorbill

Abundance estimates

A.1.3 Updated model-based population and density estimates (unapportioned and apportioned), alongside raw observations and associated DBE, for the original 29-month dataset, are provided below (Table A6-3 and Table A6-4).

Table A6-3 The raw count of individual razorbills recorded within each survey, alongside unapportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer.

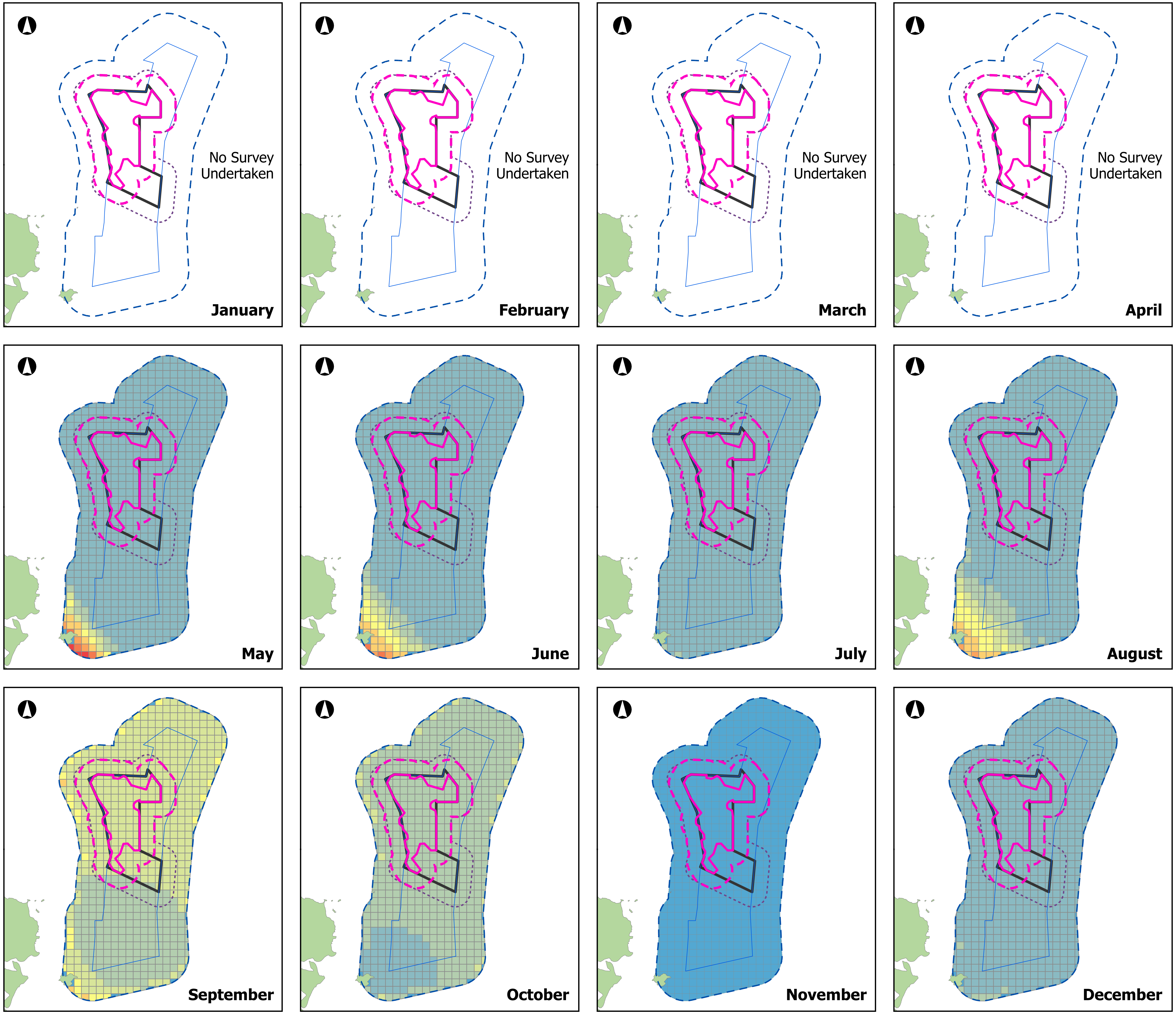
Survey No.	Survey Month	Raw Count (PFI)	Raw Count (PFI plus 2km buffer)	Design-based Abundance Estimates – PFI (Unapportioned)		Design-based Abundance Estimates – PFI plus 2km buffer (Unapportioned)		Model-based Abundance Estimates - PFI (Unapportioned)		Model-based Abundance Estimates - PFI plus 2km buffer (Unapportioned)	
				Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)	Population Estimate	Density Estimate (birds/km²)
01	May, 2020	2	12	8 (2 — 24)	0.14	48 (24 — 85)	0.30	25 (16 — 39)	0.43	63 (42 — 100)	0.40
02	June, 2020	4	7	13 (4 — 36)	0.23	24 (7 — 43)	0.15	20 (13 — 29)	0.34	54 (36 — 78)	0.33
03	July, 2020	5	27	17 (5 — 37)	0.29	93 (44 — 142)	0.58	35 (23 — 49)	0.60	94 (64 — 133)	0.59
04	August, 2020	5	9	16 (5 — 44)	0.28	32 (9 — 76)	0.20	31 (10 — 89)	0.54	86 (29 — 247)	0.54
05	September,	49	179	163 (49 — 358)	2.83	603 (205 — 1,058)	3.77	471 (170 — 1,281)	8.17	1,243 (466 — 3,305)	7.76
06	October, 2020	22	115	74 (22 — 149)	1.28	398 (137 — 688)	2.49	184 (129 — 266)	3.20	500 (345 — 740)	3.12
07	November,	0	0	-	-	-	-	-	-	-	-
08	December,	9	25	29 (9 — 65)	0.50	83 (30 — 148)	0.52	25 (13 — 48)	0.43	68 (36 — 132)	0.43
09	February, 2021	0	0	-	-	-	-	-	-	-	-
10	March, 2021	19	48	64 (33 — 93)	1.11	163 (108 — 217)	1.02	77 (61 — 96)	1.33	212 (167 — 268)	1.32
11	April, 2021	2	4	7 (2 — 20)	0.12	13 (4 — 27)	0.08	15 (9 — 25)	0.26	39 (24 — 65)	0.25
12	May, 2021	1	4	3 (1 — 10)	0.05	13 (4 — 30)	0.08	5 (2 — 14)	0.09	14 (5 — 37)	0.08
13	June, 2021	0	0	0 (0 — 0)	0.00	0 (0 — 0)	0.00	3 (1 — 14)	0.06	9 (2 — 38)	0.06
14	July, 2021	0	2	0 (0 — 0)	0.00	7 (2 — 20)	0.04	6 (3 — 11)	0.10	15 (8 — 30)	0.09
15	August, 2021	55	141	179 (65 — 295)	3.10	467 (213 — 714)	2.92	245 (171 — 363)	4.26	775 (542 — 1,140)	4.84
16	September,	30	63	99 (30 — 194)	1.72	210 (63 — 380)	1.31	137 (91 — 203)	2.37	408 (278 — 594)	2.55
17	October, 2021	27	78	90 (33 — 152)	1.56	260 (104 — 458)	1.62	190 (128 — 279)	3.30	490 (328 — 724)	3.06
18	November,	133	282	456 (133 — 827)	7.91	995 (391 — 1,566)	6.21	853 (539 — 1,369)	14.80	2,189 (1,359 — 3,598)	13.67
19	December,	74	197	248 (89 — 448)	4.30	715 (408 — 1,123)	4.46	828 (483 — 1,443)	14.36	2,192 (1,284 — 3,807)	13.69
20	January, 2022	5	14	17 (5 — 33)	0.29	47 (17 — 77)	0.29	11 (6 — 20)	0.20	33 (19 — 57)	0.21
21	February, 2022	23	72	76 (33 — 119)	1.32	242 (95 — 441)	1.51	116 (81 — 166)	2.01	300 (207 — 436)	1.88
22	March, 2022	2	5	7 (2 — 13)	0.12	16 (5 — 33)	0.10	18 (10 — 36)	0.31	48 (25 — 100)	0.30
23	April, 2022	3	15	10 (3 — 23)	0.17	51 (20 — 84)	0.32	36 (23 — 52)	0.62	106 (67 — 159)	0.66
24	May, 2022	1	6	3 (1 — 10)	0.05	20 (7 — 34)	0.12	18 (11 — 32)	0.32	47 (27 — 82)	0.29
25	June, 2022	0	0	0 (0 — 0)	0.00	0 (0 — 0)	0.00	4 (1 — 17)	0.07	10 (2 — 44)	0.06
26	July, 2022	9	29	29 (23 — 36)	0.50	95 (76 — 116)	0.59	44 (28 — 69)	0.77	124 (78 — 192)	0.77
27	August, 2022	1	3	3 (1 — 10)	0.05	10 (3 — 20)	0.06	1 (0 — 5)	0.02	4 (1 — 16)	0.03
28	September,	27	90	91 (30 — 161)	1.58	305 (90 — 626)	1.90	62 (23 — 187)	1.08	194 (78 — 547)	1.21
29	October, 2022	292	768	974 (377 — 1,732)	16.89	2,615 (1,293 — 4,108)	16.33	690 (452 — 1,072)	11.98	1,856 (1,212 — 2,887)	11.59

Table A6-4 The raw count of individual razorbills recorded within each survey, alongside apportioned design-based abundance estimates (CIs provided in brackets), model-based abundance estimates (CIs provided in brackets) and density estimates within the PFI and PFI plus a 2km buffer.

Survey No.	Survey Month	Raw Count (PFI)	Raw Count (PFI plus 2km buffer)	Design-based Abundance Estimates – PFI (Apportioned and corrected)		Design-based Abundance Estimates – PFI plus 2km buffer (Apportioned and corrected)		Model-based Abundance Estimates - PFI (Apportioned and corrected)		Model-based Abundance Estimates - PFI plus 2km buffer (Apportioned and corrected)	
				Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)	Population Estimate	Density Estimate (birds/km^2)
01	May, 2020	2	12	12 (2 — 35)	0.21	69 (34 — 119)	0.43	38 (28 — 54)	0.66	130 (104 — 172)	2.25
02	June, 2020	4	7	21 (4 — 58)	0.36	38 (12 — 72)	0.24	38 (31 — 49)	0.66	90 (70 — 119)	1.57
03	July, 2020	5	27	19 (5 — 43)	0.33	110 (51 — 167)	0.69	60 (47 — 77)	1.04	218 (182 — 263)	3.77
04	August, 2020	5	9	18 (5 — 49)	0.31	35 (9 — 83)	0.22	53 (30 — 119)	0.92	134 (69 — 316)	2.32
05	September,	49	179	207 (49 — 455)	3.59	778 (276 — 1,366)	4.86	759 (398 — 1,727)	13.16	2,208 (1,278 —	38.31
06	October, 2020	22	115	608 (94 — 1,830)	10.55	1,647 (687 — 3,425)	10.28	323 (254 — 425)	5.60	1,123 (929 — 1,423)	19.48
07	November, 2020	0	0	11 (4 — 19)	0.19	214 (84 — 364)	1.34	-	-	-	-
08	December, 2020	9	25	258 (40 — 542)	4.48	684 (296 — 1,162)	4.27	69 (54 — 99)	1.19	194 (153 — 276)	3.36
09	February, 2021	0	0	7 (3 — 11)	0.12	23 (15 — 32)	0.14	-	-	-	-
10	March, 2021	19	48	88 (45 — 129)	1.53	225 (147 — 299)	1.4	180 (160 — 205)	3.13	480 (422 — 552)	8.33
11	April, 2021	2	4	12 (2 — 35)	0.21	22 (5 — 45)	0.14	26 (19 — 37)	0.45	61 (43 — 91)	1.06
12	May, 2021	1	4	5 (1 — 19)	0.09	17 (4 — 38)	0.11	9 (6 — 20)	0.16	31 (21 — 58)	0.54
13	June, 2021	0	0	0 (0 — 0)	0.00	0 (0 — 0)	0	4 (1 — 16)	0.07	10 (2 — 44)	0.18
14	July, 2021	0	2	0 (0 — 0)	0.00	8 (2 — 24)	0.05	6 (3 — 13)	0.11	26 (17 — 43)	0.44
15	August, 2021	55	141	208 (75 — 343)	3.61	540 (243 — 822)	3.37	479 (395 — 612)	8.31	1,403 (1,139 —	24.33
16	September,	30	63	120 (30 — 241)	2.08	255 (72 — 468)	1.59	282 (227 — 362)	4.89	739 (584 — 962)	12.82
17	October, 2021	27	78	133 (47 — 235)	2.31	377 (152 — 659)	2.35	350 (273 — 462)	6.08	937 (735 — 1,230)	16.26
18	November, 2021	133	282	781 (228 — 1,363)	13.55	1,908 (761 — 2,987)	11.91	1,676 (1,274 —	29.07	4,076 (3,013 —	70.70
19	December, 2021	74	197	470 (196 — 797)	8.15	1,340 (810 — 2,010)	8.37	1,377 (936 — 2,165)	23.89	3,722 (2,559 —	64.55
20	January, 2022	5	14	25 (5 — 49)	0.43	72 (26 — 120)	0.45	36 (30 — 47)	0.63	103 (84 — 133)	1.78
21	February, 2022	23	72	140 (55 — 240)	2.43	412 (165 — 738)	2.57	245 (200 — 310)	4.26	694 (575 — 868)	12.04
22	March, 2022	2	5	11 (2 — 25)	0.19	27 (5 — 55)	0.17	32 (22 — 55)	0.56	82 (53 — 149)	1.43
23	April, 2022	3	15	13 (3 — 31)	0.23	66 (27 — 106)	0.41	53 (38 — 72)	0.92	183 (137 — 244)	3.17
24	May, 2022	1	6	4 (1 — 12)	0.07	26 (9 — 42)	0.16	25 (16 — 40)	0.43	78 (55 — 119)	1.36
25	June, 2022	0	0	0 (0 — 0)	0.00	0 (0 — 0)	0.00	4 (1 — 20)	0.08	11 (2 — 51)	0.19
26	July, 2022	9	29	32 (25 — 39)	0.56	105 (84 — 128)	0.66	85 (66 — 114)	1.48	255 (202 — 334)	4.43
27	August, 2022	1	3	4 (1 — 11)	0.07	11 (4 — 22)	0.07	5 (4 — 9)	0.08	16 (13 — 30)	0.28
28	September,	27	90	107 (33 — 191)	1.86	353 (90 — 719)	2.20	184 (137 — 333)	3.18	597 (458 — 1,018)	10.36
29	October, 2022	292	768	1,539 (638 —	26.69	3,955 (1,997 — 5,990)	24.69	2,081 (1,782 —	36.09	5,589 (4,783 —	96.94

Density Estimate Plots

A.1.4 The figures below illustrate the modelled density estimates incorporating the PFI boundary and associated 2km buffer (Figure A6-4, Figure A6-5 and Figure A6-6).



▭ Array Area

⬚ 2km Buffer from Array Area

▭ Projected Footprint of Infrastructure

⬚ 2km Buffer from Projected Footprint of Infrastructure

▭ MAC Boundary (Array Area Only)

⬚ 4km Buffer from MAC Boundary (Array Area Only)

Density (per km^2)

🟦	0
🟦	0 - 2
🟩	2 - 5
🟩	5 - 10
🟨	10 - 20
🟨	20 - 50
🟧	50 - 100
🟧	100 - 200
🟥	>200

Kilometres
0 2 4 6 8 10

NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Distribution for Razorbills, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2020

Job No: 281240

Datum: WGS84 **Projection:** UTM30N

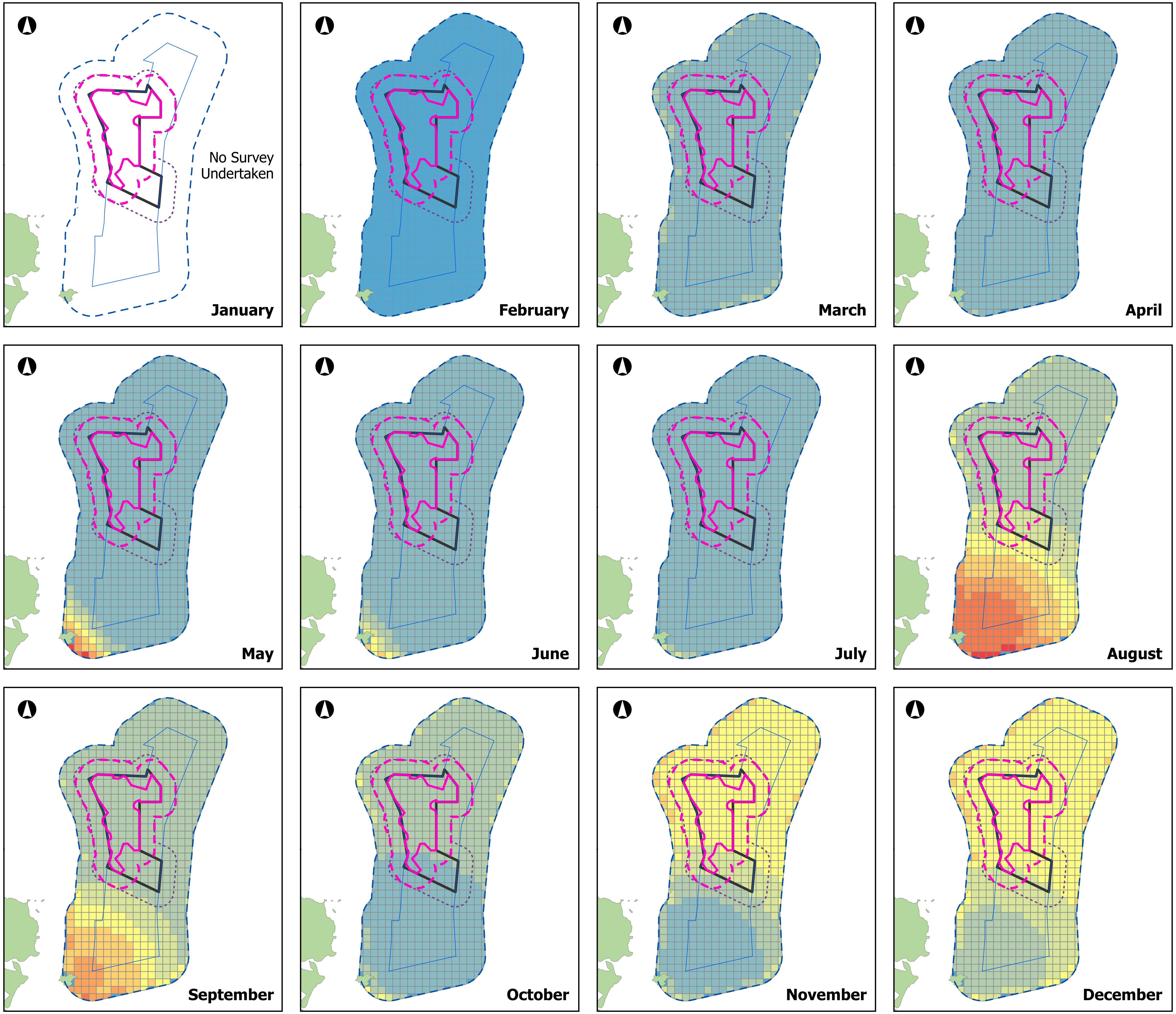
Date: February 2026

Figure No:

Scale: 1:500,000 @A3

A6-4

Status: Issue



NISA

North Irish Sea Array

ARUP GoBe

Project

North Irish Sea Array Offshore Wind Farm

Figure Title

Model-based Density Distribution for Razorbills, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2021

Job No: 281240

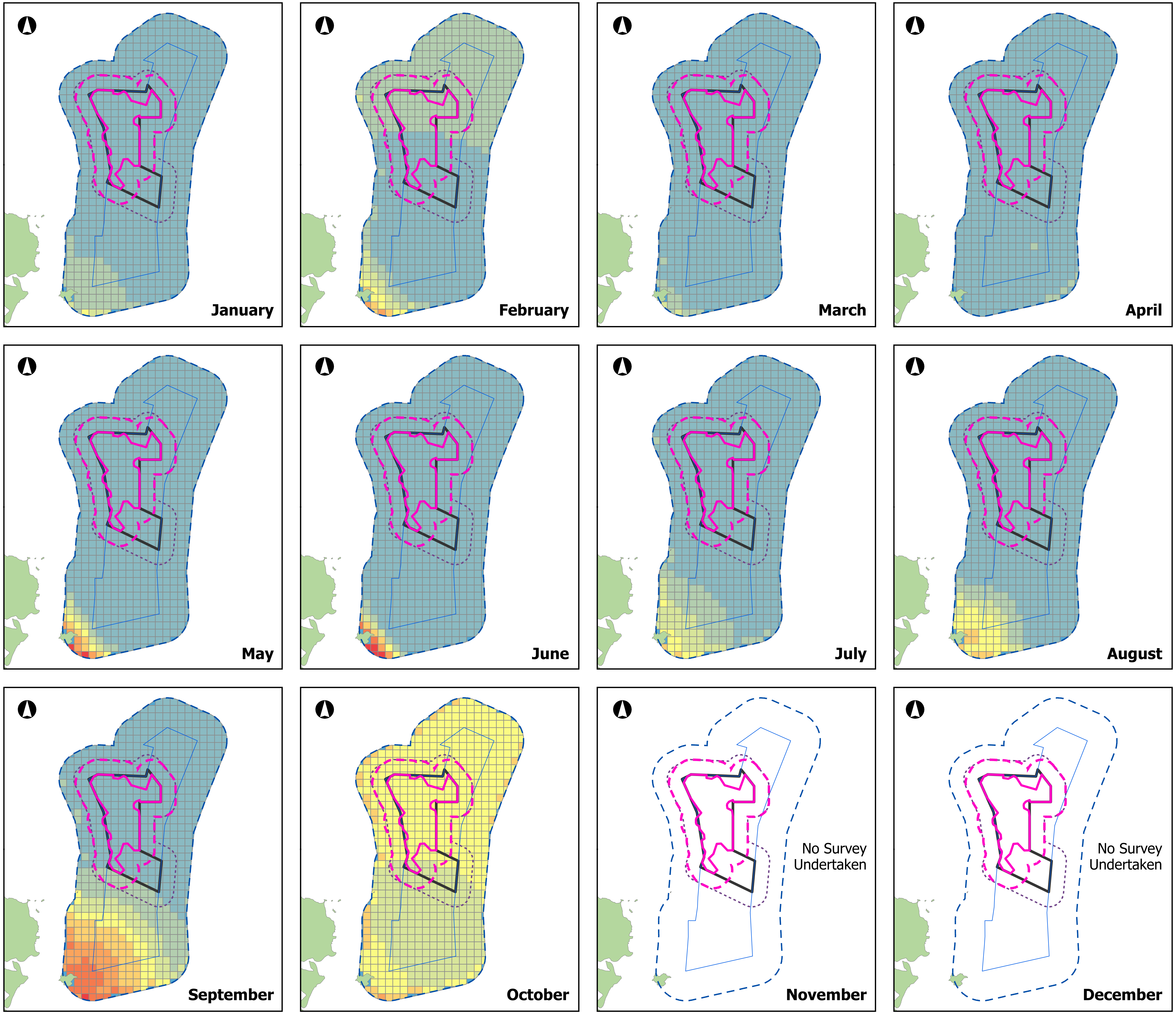
Datum: WGS84 **Projection:** UTM30N

Date: February 2026

Figure No: **A6-5**

Scale: 1:500,000 @A3

Status: Issue



▭ Array Area
⬚ 2km Buffer from Array Area
▭ Projected Footprint of Infrastructure
⬚ 2km Buffer from Projected Footprint of Infrastructure
▭ MAC Boundary (Array Area Only)
⬚ 4km Buffer from MAC Boundary (Array Area Only)

Density (per km^2)

🟦 0
🟦 0 - 2
🟩 2 - 5
🟩 5 - 10
🟨 10 - 20
🟨 20 - 50
🟧 50 - 100
🟧 100 - 200
🟥 >200

Kilometres
0 2 4 6 8 10

NISA
North Irish Sea Array

ARUP GoBe

Project
North Irish Sea Array Offshore Wind Farm

Figure Title
Model-based Density Distribution for Razorbills, within the MAC Boundary (Array Area only) with a 4km Buffer applied, in 2022

Job No: 281240
Datum: WGS84 **Projection:** UTM30N
Date: February 2026
Figure No:
Scale: 1:500,000 @A3
A6-6
Status: Issue

A.2 Final Model Summary

A.2.1 Final MRSea models (detailed) that were selected for each species are provided below (Table A6-5).

Table A6-5 Final MRSea model selections for each species.

Species	Final Model
Guillemot	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, Boundary.knots = splineParams[[3]]$bd) + bs(slope, knots = splineParams[[4]]$knots, degree = splineParams[[4]]$degree, Boundary.knots = splineParams[[4]]$bd) + bs(prey_density, knots = splineParams[[5]]$knots, degree = splineParams[[5]]$degree, Boundary.knots = splineParams[[5]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Razorbill	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, Boundary.knots = splineParams[[3]]$bd) + bs(prey_density, knots = splineParams[[4]]$knots, degree = splineParams[[4]]$degree, Boundary.knots = splineParams[[4]]$bd) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Kittiwake	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Herring gull	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Great black-backed gull	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, Boundary.knots = splineParams[[3]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Gannet	response (count) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, Boundary.knots = splineParams[[3]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Common tern	response (count) ~ Survey + depth + bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, Boundary.knots = splineParams[[3]]$bd) + dist_coast + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Roseate tern	response (count) ~ Survey + dist_colony + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)
Red-throated diver	response (count) ~ Survey + bs(dist_coast, knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

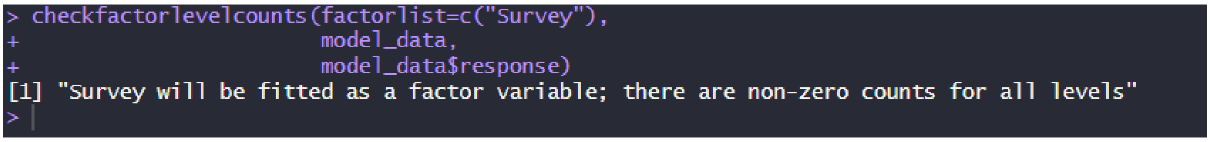
A.3 Model Setup

A.3.1 Code excerpts and plots are provided to illustrate the model setup process for each species. These include outputs for collinearity assessment, auto-correlation function (ACF), factor level checks, runs test results, and cross-validation error estimates. Additional information is also provided with regard to model setup and summaries for the initial GLM, 1D SALSA, and 2D SALSA models.

Guillemot

Factor Level Check

A.3.2 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-7).

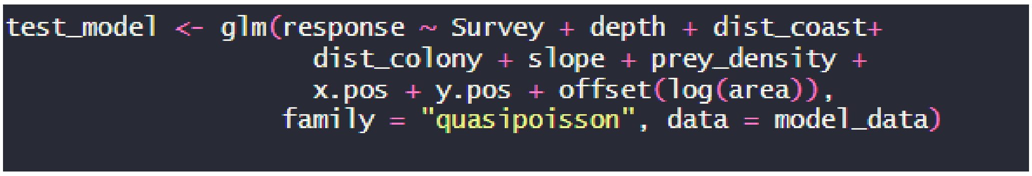


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                        model_data,
+                        model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-7 Code excerpt detailing factor level check for categorical variables considered within guillemot MRSea modelling.

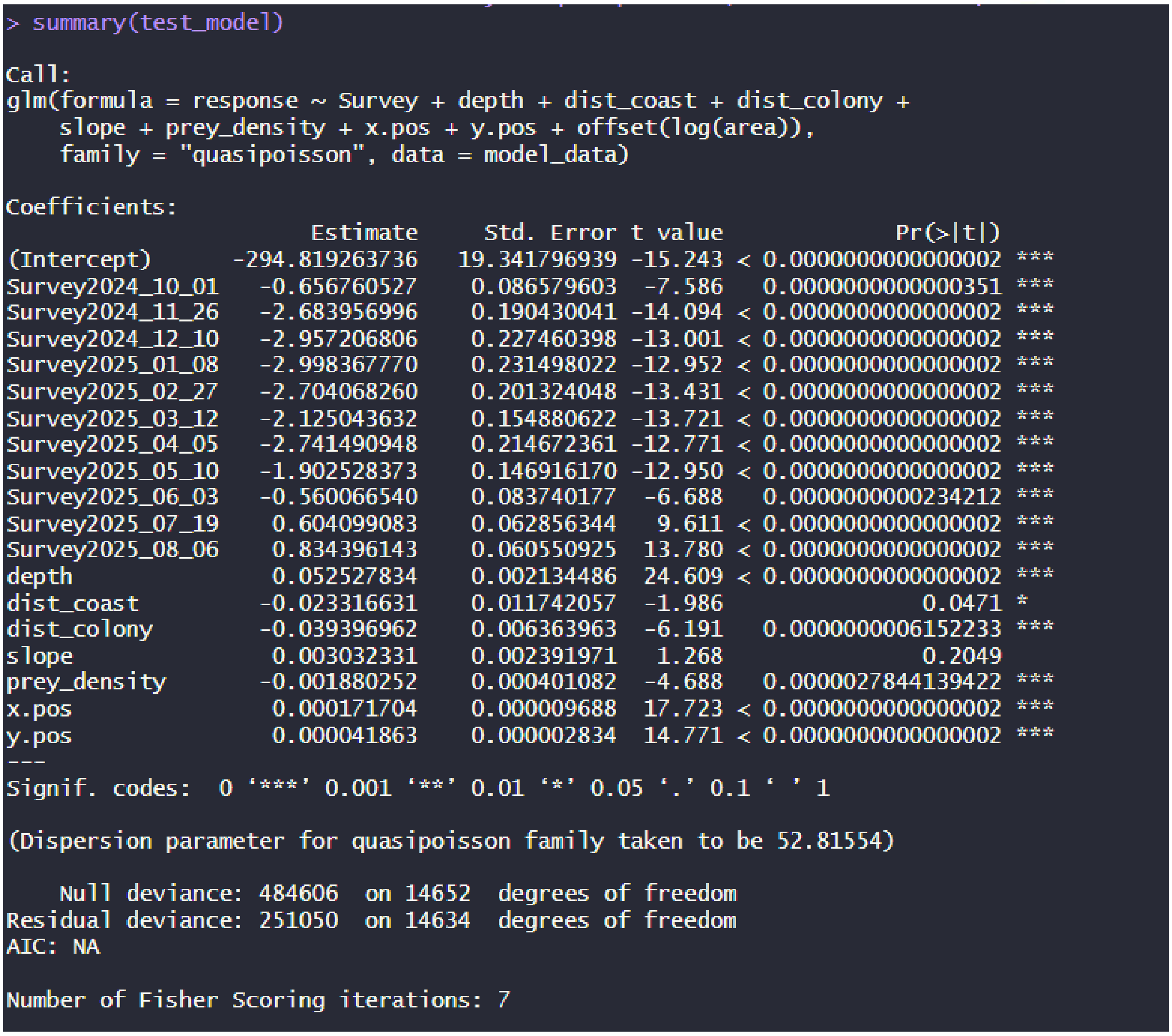
Initial GLM

A.3.3 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-8 and Figure A6-9).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-8 Code excerpt detailing the setup of the initial GLM considered within guillemot MRSea analysis.

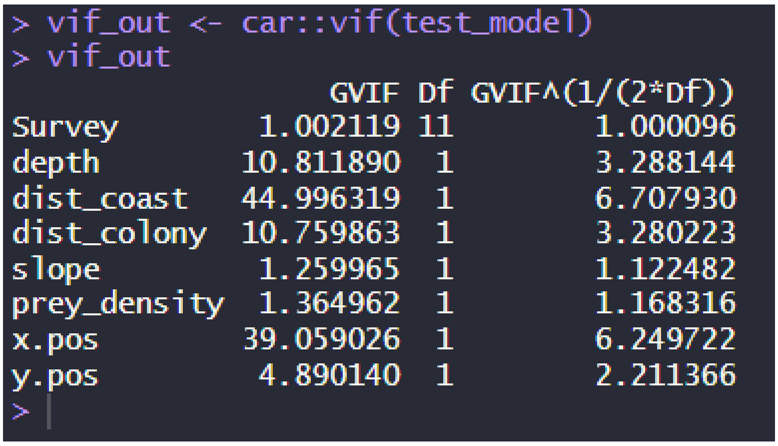


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_coast + dist_colony + 
    slope + prey_density + x.pos + y.pos + offset(log(area)), 
    family = "quasipoisson", data = model_data)

Coefficients:
                       Estimate     Std. Error t value             Pr(>|t|)    
(Intercept)      -294.819263736   19.341796939 -15.243 < 0.0000000000000002 ***
Survey2024_10_01   -0.656760527    0.086579603  -7.586   0.0000000000000351 ***
Survey2024_11_26   -2.683956996    0.190430041 -14.094 < 0.0000000000000002 ***
Survey2024_12_10   -2.957206806    0.227460398 -13.001 < 0.0000000000000002 ***
Survey2025_01_08   -2.998367770    0.231498022 -12.952 < 0.0000000000000002 ***
Survey2025_02_27   -2.704068260    0.201324048 -13.431 < 0.0000000000000002 ***
Survey2025_03_12   -2.125043632    0.154880622 -13.721 < 0.0000000000000002 ***
Survey2025_04_05   -2.741490948    0.214672361 -12.771 < 0.0000000000000002 ***
Survey2025_05_10   -1.902528373    0.146916170 -12.950 < 0.0000000000000002 ***
Survey2025_06_03   -0.560066540    0.083740177  -6.688   0.0000000000234212 ***
Survey2025_07_19    0.604099083    0.062856344   9.611 < 0.0000000000000002 ***
Survey2025_08_06    0.834396143    0.060550925  13.780 < 0.0000000000000002 ***
depth               0.052527834    0.002134486  24.609 < 0.0000000000000002 ***
dist_coast         -0.023316631    0.011742057  -1.986               0.0471 *  
dist_colony        -0.039396962    0.006363963  -6.191   0.0000000006152233 ***
slope               0.003032331    0.002391971   1.268               0.2049    
prey_density       -0.001880252    0.000401082  -4.688   0.0000027844139422 ***
x.pos               0.000171704    0.000009688  17.723 < 0.0000000000000002 ***
y.pos               0.000041863    0.000002834  14.771 < 0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 52.81554)

    Null deviance: 484606  on 14652  degrees of freedom
Residual deviance: 251050  on 14634  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 7
```

Figure A6-9 Code excerpt detailing the summary of the initial GLM considered within guillemot MRSea analysis.

Collinearity Check

A.3.4 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-10).

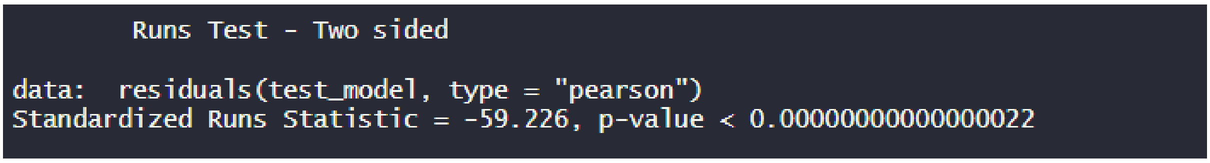


```
> vif_out <- car::vif(test_model)
> vif_out
                   GVIF Df GVIF^(1/(2*Df))
Survey         1.002119 11        1.000096
depth         10.811890  1        3.288144
dist_coast    44.996319  1        6.707930
dist_colony   10.759863  1        3.280223
slope          1.259965  1        1.122482
prey_density   1.364962  1        1.168316
x.pos         39.059026  1        6.249722
y.pos          4.890140  1        2.211366
> 
```

Figure A6-10 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within guillemot MRSea analysis.

Runs Test

A.3.5 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-11).



```
        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -59.226, p-value < 0.00000000000000022
```

Figure A6-11 Code excerpt detailing runs test carried out on the initial GLM considered within guillemot MRSea analysis.

A.3.6 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-12).

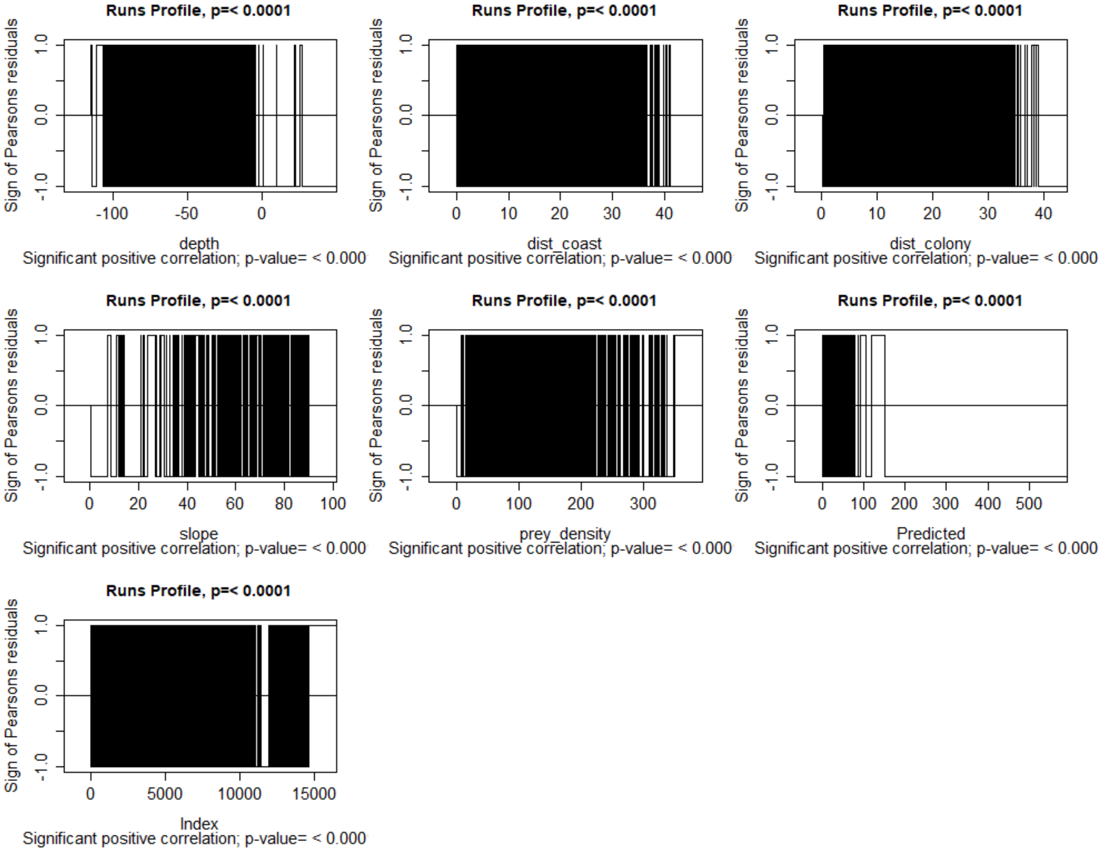


Figure A6-12 Runs profiles for explanatory variables considered within the initial GLM considered for guillemot MRSea analysis.

Autocorrelation

A.3.7 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-13).

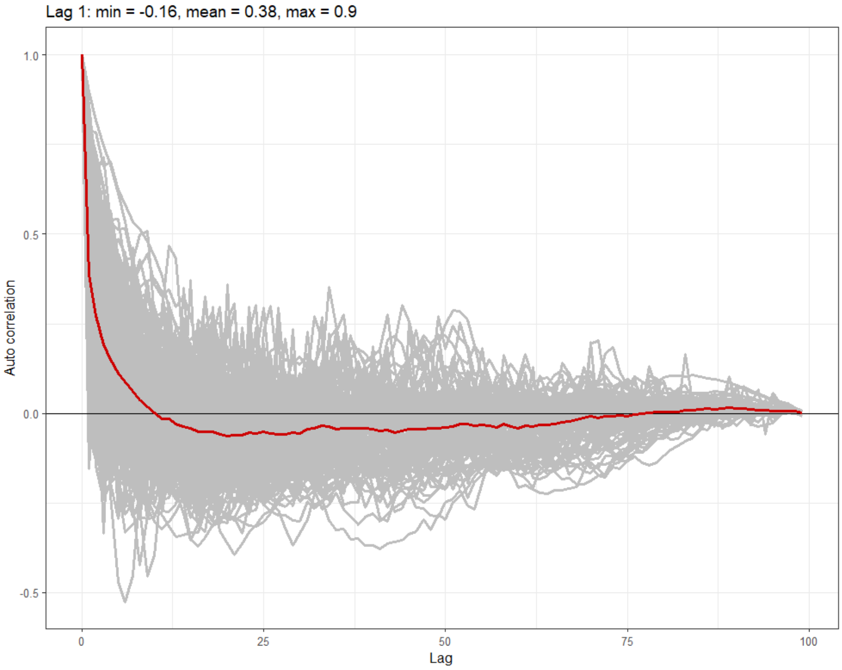


Figure A6-13 ACF plot for the initial GLM considered for guillemot MRSea analysis.

Cumulative Residuals

A.3.8 Cumulative residuals for explanatory variables were plotted (Figure A6-14, Figure A6-15, Figure A6-16 & Figure A6-17). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

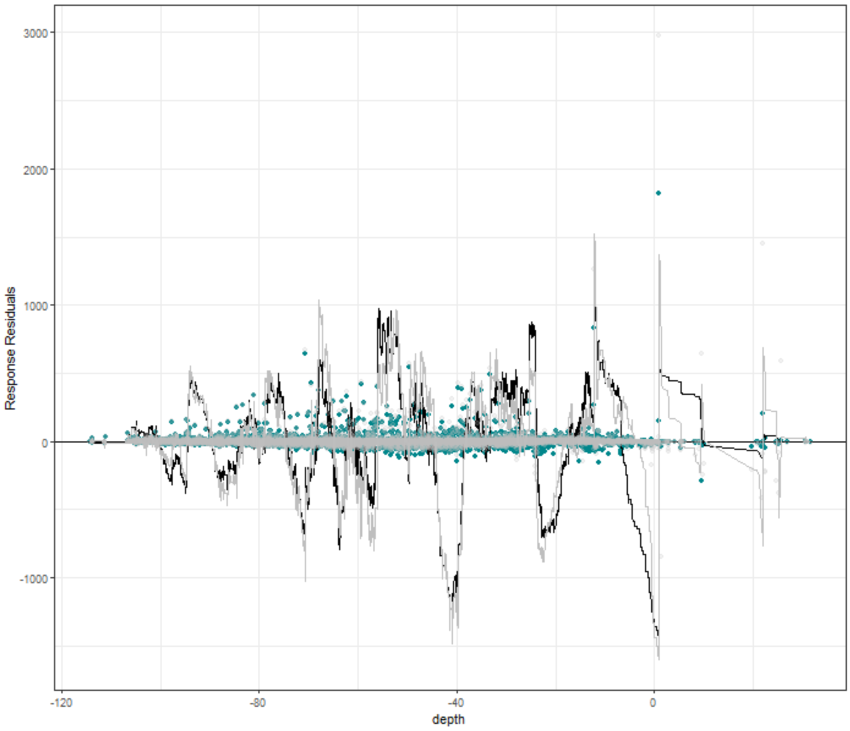


Figure A6-14 Cumulative residuals of depth plotted for the initial GLM considered for guillemot MRSea analysis.

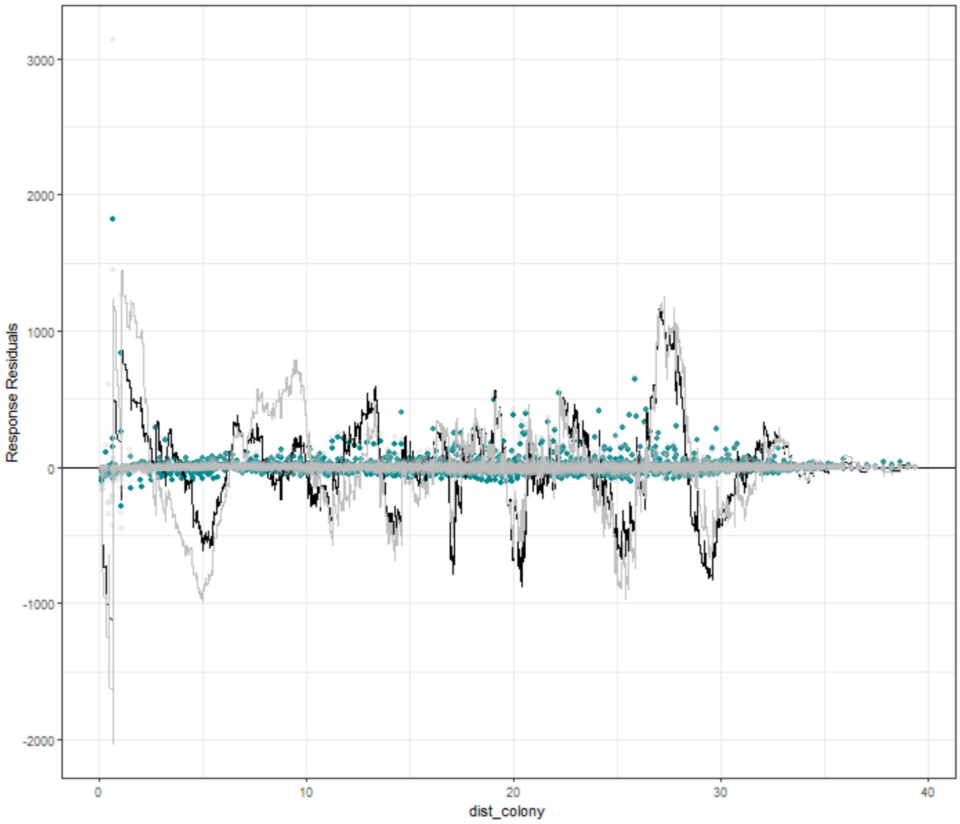


Figure A6-15 Cumulative residuals of distance to colony plotted for the initial GLM considered for guillemot MRSea analysis.

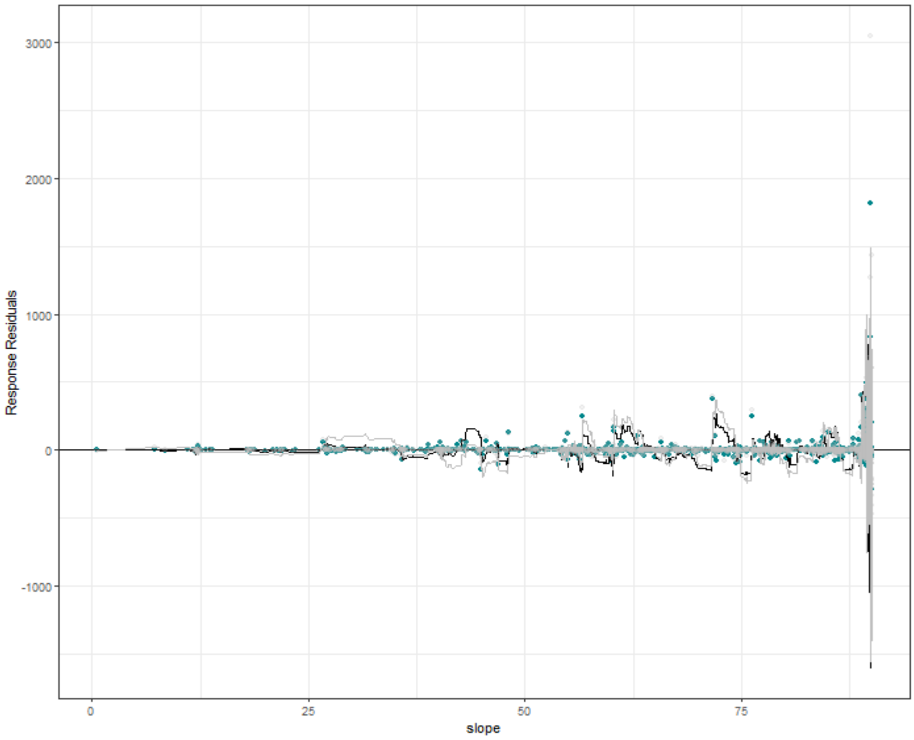


Figure A6-16 Cumulative residuals of seabed slope plotted for the initial GLM considered for guillemot MRSea analysis.

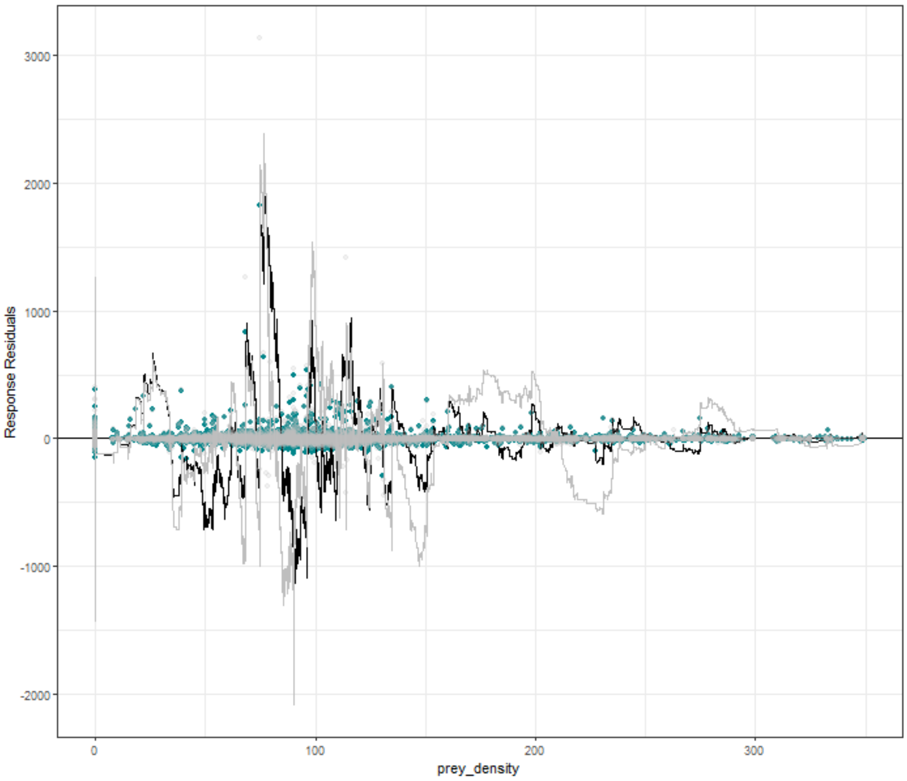
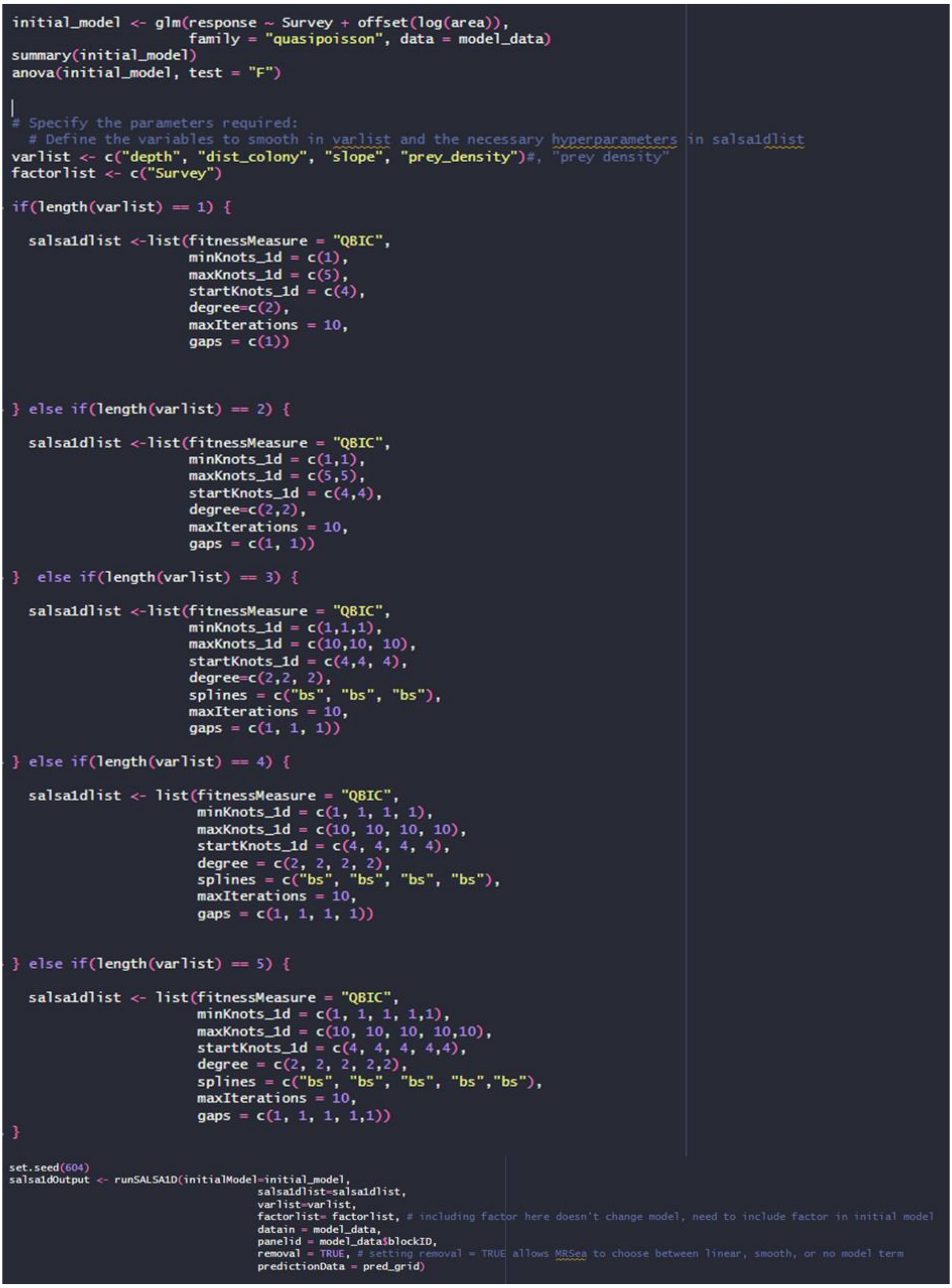


Figure A6-17 Cumulative residuals of prey density plotted for the initial GLM considered for guillemot MRSea analysis.

SALSA 1D

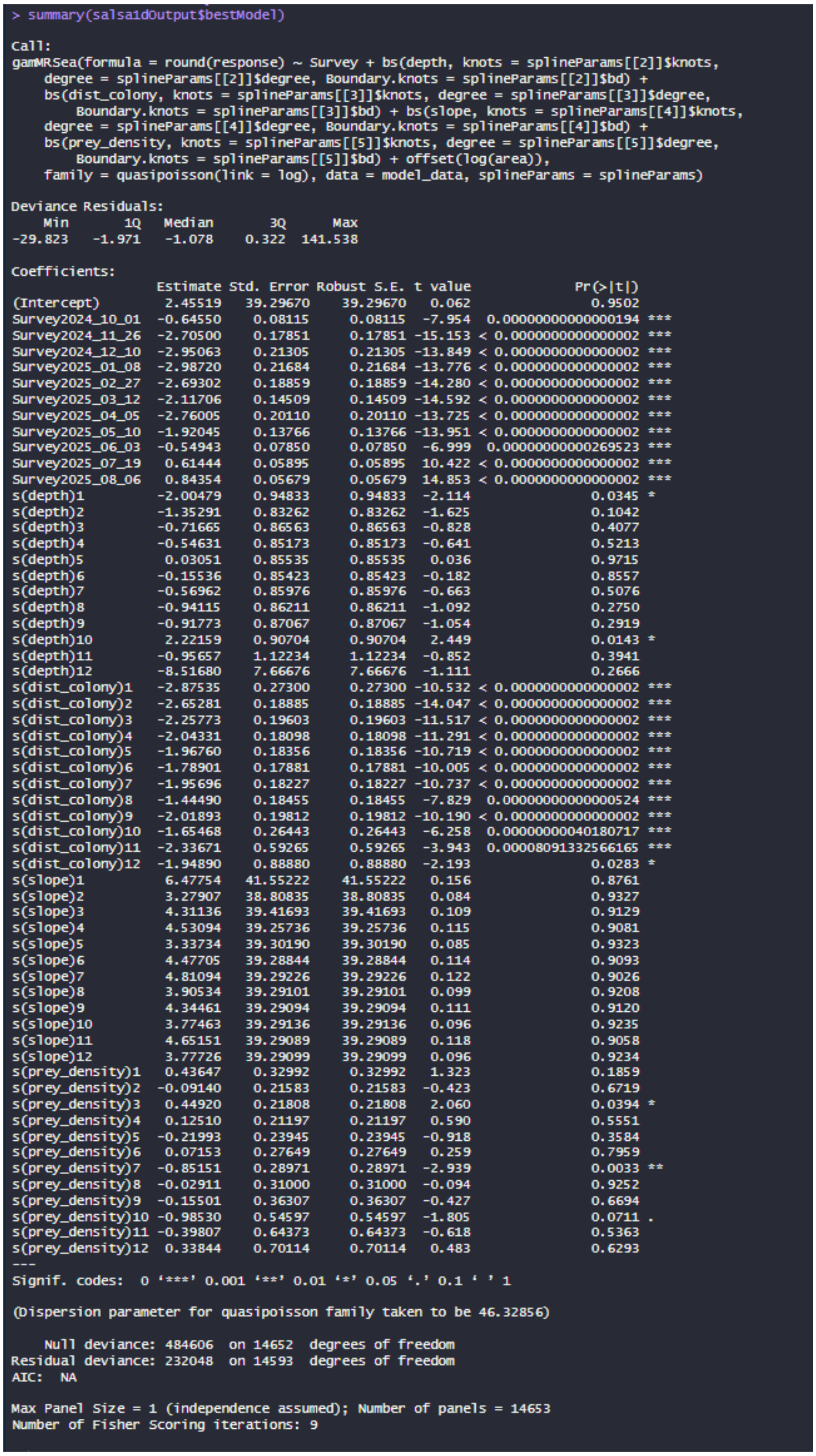
A.3.9 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.10 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-18 & Figure A6-19).



```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")

|
# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth", "dist_colony", "slope", "prey_density")#, "prey density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))

} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1,1),
                      maxKnots_1d = c(10, 10, 10, 10,10),
                      startKnots_1d = c(4, 4, 4, 4,4),
                      degree = c(2, 2, 2, 2,2),
                      splines = c("bs", "bs", "bs", "bs","bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1,1))
}

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                            salsa1dlist=salsa1dlist,
                            varlist=varlist,
                            factorlist= factorlist, # including factor here doesn't change model, need to include factor in initial model
                            datain = model_data,
                            panelid = model_data$blockID,
                            removal = TRUE, # setting removal = TRUE allows MRSea to choose between linear, smooth, or no model term
                            predictionData = pred_grid)
```

Figure A6-18 Code excerpt detailing the setup process for the SALSA 1D model considered for guillemot MRSea analysis.



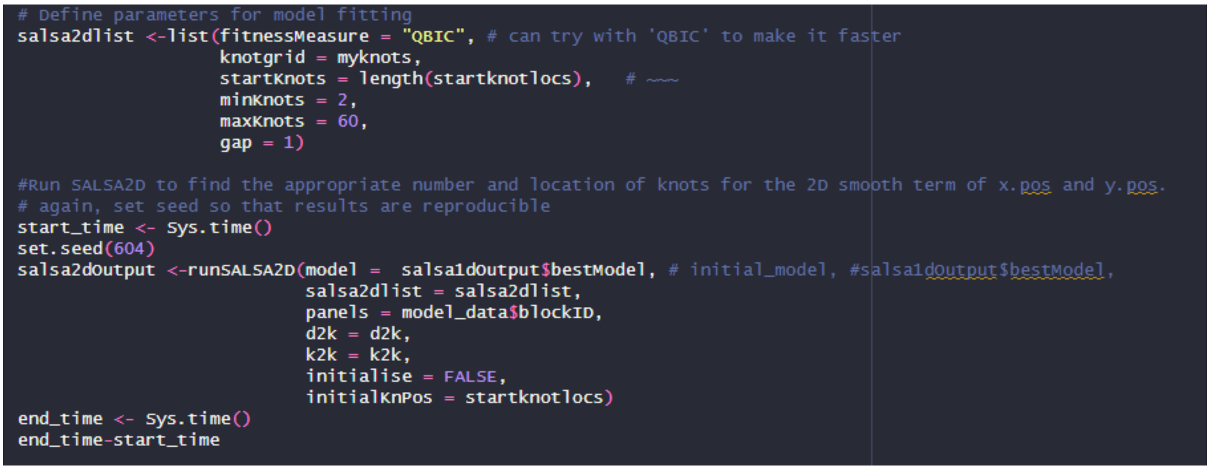
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, 
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + 
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, 
        Boundary.knots = splineParams[[3]]$bd) + bs(slope, knots = splineParams[[4]]$knots, 
    degree = splineParams[[4]]$degree, Boundary.knots = splineParams[[4]]$bd) + 
    bs(prey_density, knots = splineParams[[5]]$knots, degree = splineParams[[5]]$degree, 
        Boundary.knots = splineParams[[5]]$bd) + offset(log(area)), 
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-29.823   -1.971   -1.078    0.322  141.538  

Coefficients:
                  Estimate Std. Error Robust S.E. t value             Pr(>|t|)    
(Intercept)        2.45519   39.29670    39.29670   0.062               0.9502    
Survey2024_10_01  -0.64550    0.08115     0.08115  -7.954   0.0000000000000194 ***
Survey2024_11_26  -2.70500    0.17851     0.17851 -15.153 < 0.0000000000000002 ***
Survey2024_12_10  -2.95063    0.21305     0.21305 -13.849 < 0.0000000000000002 ***
Survey2025_01_08  -2.98720    0.21684     0.21684 -13.776 < 0.0000000000000002 ***
Survey2025_02_27  -2.69302    0.18859     0.18859 -14.280 < 0.0000000000000002 ***
Survey2025_03_12  -2.11706    0.14509     0.14509 -14.592 < 0.0000000000000002 ***
Survey2025_04_05  -2.76005    0.20110     0.20110 -13.725 < 0.0000000000000002 ***
Survey2025_05_10  -1.92045    0.13766     0.13766 -13.951 < 0.0000000000000002 ***
Survey2025_06_03  -0.54943    0.07850     0.07850  -6.999   0.0000000000269523 ***
Survey2025_07_19   0.61444    0.05895     0.05895  10.422 < 0.0000000000000002 ***
Survey2025_08_06   0.84354    0.05679     0.05679  14.853 < 0.0000000000000002 ***
s(depth)1         -2.00479    0.94833     0.94833  -2.114               0.0345 *  
s(depth)2         -1.35291    0.83262     0.83262  -1.625               0.1042    
s(depth)3         -0.71665    0.86563     0.86563  -0.828               0.4077    
s(depth)4         -0.54631    0.85173     0.85173  -0.641               0.5213    
s(depth)5          0.03051    0.85535     0.85535   0.036               0.9715    
s(depth)6         -0.15536    0.85423     0.85423  -0.182               0.8557    
s(depth)7         -0.56962    0.85976     0.85976  -0.663               0.5076    
s(depth)8         -0.94115    0.86211     0.86211  -1.092               0.2750    
s(depth)9         -0.91773    0.87067     0.87067  -1.054               0.2919    
s(depth)10         2.22159    0.90704     0.90704   2.449               0.0143 *  
s(depth)11        -0.95657    1.12234     1.12234  -0.852               0.3941    
s(depth)12        -8.51680    7.66676     7.66676  -1.111               0.2666    
s(dist_colony)1   -2.87535    0.27300     0.27300 -10.532 < 0.0000000000000002 ***
s(dist_colony)2   -2.65281    0.18885     0.18885 -14.047 < 0.0000000000000002 ***
s(dist_colony)3   -2.25773    0.19603     0.19603 -11.517 < 0.0000000000000002 ***
s(dist_colony)4   -2.04331    0.18098     0.18098 -11.291 < 0.0000000000000002 ***
s(dist_colony)5   -1.96760    0.18356     0.18356 -10.719 < 0.0000000000000002 ***
s(dist_colony)6   -1.78901    0.17881     0.17881 -10.005 < 0.0000000000000002 ***
s(dist_colony)7   -1.95696    0.18227     0.18227 -10.737 < 0.0000000000000002 ***
s(dist_colony)8   -1.44490    0.18455     0.18455  -7.829   0.0000000000000524 ***
s(dist_colony)9   -2.01893    0.19812     0.19812 -10.190 < 0.0000000000000002 ***
s(dist_colony)10  -1.65468    0.26443     0.26443  -6.258   0.0000000004018071 ***
s(dist_colony)11  -2.33671    0.59265     0.59265  -3.943   0.0000809133256616 ***
s(dist_colony)12  -1.94890    0.88880     0.88880  -2.193               0.0283 *  
s(slope)1          6.47754   41.55222    41.55222   0.156               0.8761    
s(slope)2          3.27907   38.80835    38.80835   0.084               0.9327    
s(slope)3          4.31136   39.41693    39.41693   0.109               0.9129    
s(slope)4          4.53094   39.25736    39.25736   0.115               0.9081    
s(slope)5          3.33734   39.30190    39.30190   0.085               0.9323    
s(slope)6          4.47705   39.28844    39.28844   0.114               0.9093    
s(slope)7          4.81094   39.29226    39.29226   0.122               0.9026    
s(slope)8          3.90534   39.29101    39.29101   0.099               0.9208    
s(slope)9          4.34461   39.29094    39.29094   0.111               0.9120    
s(slope)10         3.77463   39.29136    39.29136   0.096               0.9235    
s(slope)11         4.65151   39.29089    39.29089   0.118               0.9058    
s(slope)12         3.77726   39.29099    39.29099   0.096               0.9234    
s(prey_density)1   0.43647    0.32992     0.32992   1.323               0.1859    
s(prey_density)2  -0.09140    0.21583     0.21583  -0.423               0.6719    
s(prey_density)3   0.44920    0.21808     0.21808   2.060               0.0394 *  
s(prey_density)4   0.12510    0.21197     0.21197   0.590               0.5551    
s(prey_density)5  -0.21993    0.23945     0.23945  -0.918               0.3584    
s(prey_density)6   0.07153    0.27649     0.27649   0.259               0.7959    
s(prey_density)7  -0.85151    0.28971     0.28971  -2.939               0.0033 ** 
s(prey_density)8  -0.02911    0.31000     0.31000  -0.094               0.9252    
s(prey_density)9  -0.15501    0.36307     0.36307  -0.427               0.6694    
s(prey_density)10 -0.98530    0.54597     0.54597  -1.805               0.0711 .  
s(prey_density)11 -0.39807    0.64373     0.64373  -0.618               0.5363    
s(prey_density)12  0.33844    0.70114     0.70114   0.483               0.6293    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 46.32856)

    Null deviance: 484606  on 14652  degrees of freedom
Residual deviance: 232048  on 14593  degrees of freedom
AIC:  NA

Max Panel Size = 1 (independence assumed); Number of panels = 14653
Number of Fisher Scoring iterations: 9
```

Figure A6-19 Code excerpt detailing the summary for the SALSA 1D model for guillemot MRSea analysis.

SALSA 2D

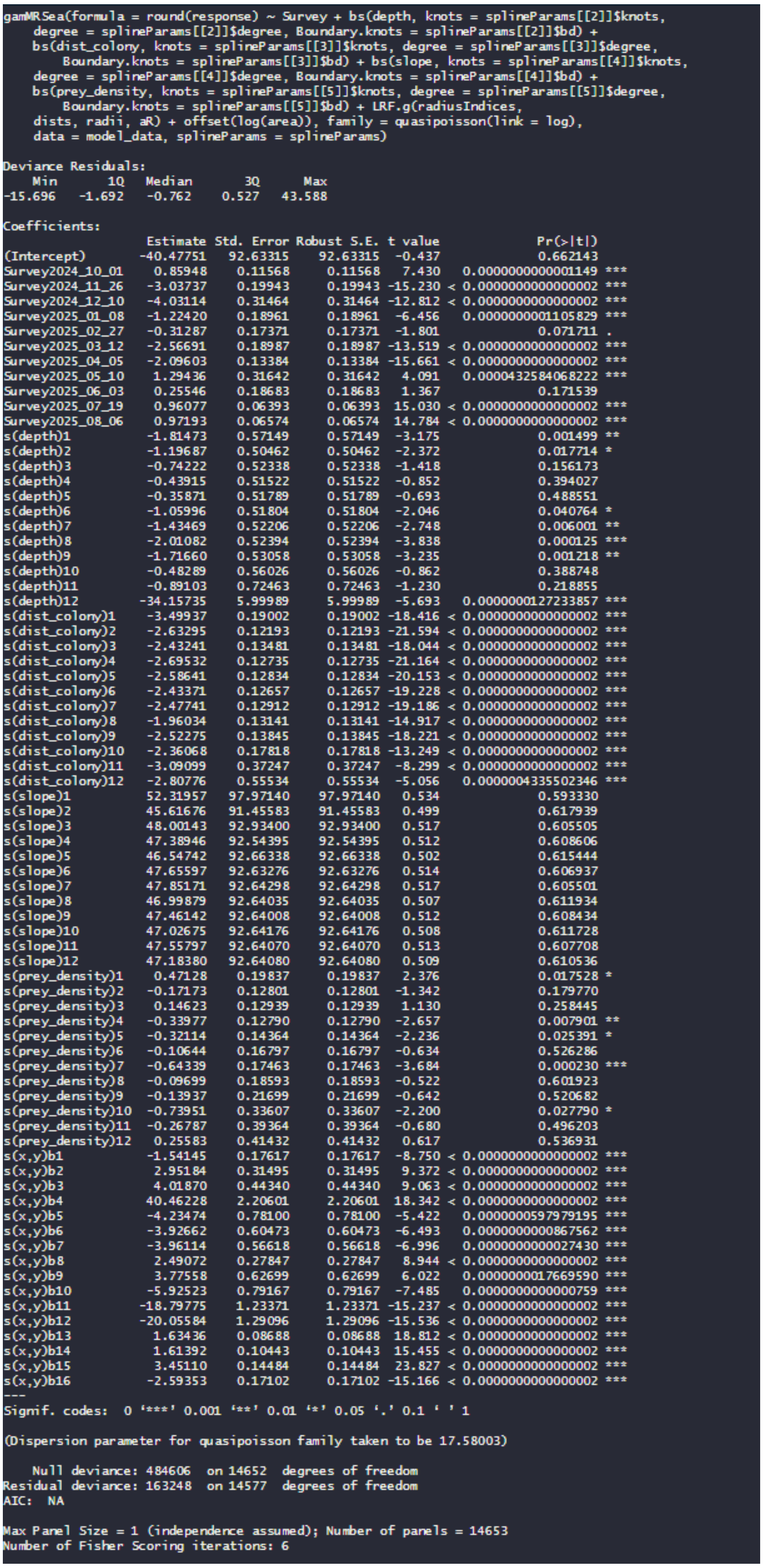
A.3.11 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.12 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-20 & Figure A6-21).



```r
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC' to make it faster
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),   # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1)

#Run SALSA2D to find the appropriate number and location of knots for the 2D smooth term of x.pos and y.pos.
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2doutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # initial_model, #salsa1dOutput$bestModel,
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-20 Code excerpt for SALSA 2D model setup for guillemot MRSea analysis.

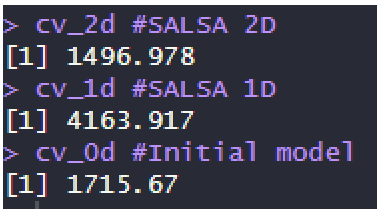


```
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, 
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + 
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, 
        Boundary.knots = splineParams[[3]]$bd) + bs(slope, knots = splineParams[[4]]$knots, 
    degree = splineParams[[4]]$degree, Boundary.knots = splineParams[[4]]$bd) + 
    bs(prey_density, knots = splineParams[[5]]$knots, degree = splineParams[[5]]$degree, 
        Boundary.knots = splineParams[[5]]$bd) + LRF.g(radiusIndices, 
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), 
    data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-15.696   -1.692   -0.762    0.527   43.588  

Coefficients:
                   Estimate Std. Error Robust S.E. t value             Pr(>|t|)    
(Intercept)       -40.47751   92.63315    92.63315  -0.437             0.662143    
Survey2024_10_01    0.85948    0.11568     0.11568   7.430   0.0000000000001149 ***
Survey2024_11_26   -3.03737    0.19943     0.19943 -15.230 < 0.0000000000000002 ***
Survey2024_12_10   -4.03114    0.31464     0.31464 -12.812 < 0.0000000000000002 ***
Survey2025_01_08   -1.22420    0.18961     0.18961  -6.456   0.0000000001105829 ***
Survey2025_02_27   -0.31287    0.17371     0.17371  -1.801             0.071711 .  
Survey2025_03_12   -2.56691    0.18987     0.18987 -13.519 < 0.0000000000000002 ***
Survey2025_04_05   -2.09603    0.13384     0.13384 -15.661 < 0.0000000000000002 ***
Survey2025_05_10    1.29436    0.31642     0.31642   4.091   0.0000432584068222 ***
Survey2025_06_03    0.25546    0.18683     0.18683   1.367             0.171539    
Survey2025_07_19    0.96077    0.06393     0.06393  15.030 < 0.0000000000000002 ***
Survey2025_08_06    0.97193    0.06574     0.06574  14.784 < 0.0000000000000002 ***
s(depth)1          -1.81473    0.57149     0.57149  -3.175             0.001499 ** 
s(depth)2          -1.19687    0.50462     0.50462  -2.372             0.017714 *  
s(depth)3          -0.74222    0.52338     0.52338  -1.418             0.156173    
s(depth)4          -0.43915    0.51522     0.51522  -0.852             0.394027    
s(depth)5          -0.35871    0.51789     0.51789  -0.693             0.488551    
s(depth)6          -1.05996    0.51804     0.51804  -2.046             0.040764 *  
s(depth)7          -1.43469    0.52206     0.52206  -2.748             0.006001 ** 
s(depth)8          -2.01082    0.52394     0.52394  -3.838             0.000125 ***
s(depth)9          -1.71660    0.53058     0.53058  -3.235             0.001218 ** 
s(depth)10         -0.48289    0.56026     0.56026  -0.862             0.388748    
s(depth)11         -0.89103    0.72463     0.72463  -1.230             0.218855    
s(depth)12        -34.15735    5.99989     5.99989  -5.693   0.0000000127233857 ***
s(dist_colony)1    -3.49937    0.19002     0.19002 -18.416 < 0.0000000000000002 ***
s(dist_colony)2    -2.63295    0.12193     0.12193 -21.594 < 0.0000000000000002 ***
s(dist_colony)3    -2.43241    0.13481     0.13481 -18.044 < 0.0000000000000002 ***
s(dist_colony)4    -2.69532    0.12735     0.12735 -21.164 < 0.0000000000000002 ***
s(dist_colony)5    -2.58641    0.12834     0.12834 -20.153 < 0.0000000000000002 ***
s(dist_colony)6    -2.43371    0.12657     0.12657 -19.228 < 0.0000000000000002 ***
s(dist_colony)7    -2.47741    0.12912     0.12912 -19.186 < 0.0000000000000002 ***
s(dist_colony)8    -1.96034    0.13141     0.13141 -14.917 < 0.0000000000000002 ***
s(dist_colony)9    -2.52275    0.13845     0.13845 -18.221 < 0.0000000000000002 ***
s(dist_colony)10   -2.36068    0.17818     0.17818 -13.249 < 0.0000000000000002 ***
s(dist_colony)11   -3.09099    0.37247     0.37247  -8.299 < 0.0000000000000002 ***
s(dist_colony)12   -2.80776    0.55534     0.55534  -5.056   0.0000004335502346 ***
s(slope)1          52.31957   97.97140    97.97140   0.534             0.593330    
s(slope)2          45.61676   91.45583    91.45583   0.499             0.617939    
s(slope)3          48.00143   92.93400    92.93400   0.517             0.605505    
s(slope)4          47.38946   92.54395    92.54395   0.512             0.608606    
s(slope)5          46.54742   92.66338    92.66338   0.502             0.615444    
s(slope)6          47.65597   92.63276    92.63276   0.514             0.606937    
s(slope)7          47.85171   92.64298    92.64298   0.517             0.605501    
s(slope)8          46.99879   92.64035    92.64035   0.507             0.611934    
s(slope)9          47.46142   92.64008    92.64008   0.512             0.608434    
s(slope)10         47.02675   92.64176    92.64176   0.508             0.611728    
s(slope)11         47.55797   92.64070    92.64070   0.513             0.607708    
s(slope)12         47.18380   92.64080    92.64080   0.509             0.610536    
s(prey_density)1    0.47128    0.19837     0.19837   2.376             0.017528 *  
s(prey_density)2   -0.17173    0.12801     0.12801  -1.342             0.179770    
s(prey_density)3    0.14623    0.12939     0.12939   1.130             0.258445    
s(prey_density)4   -0.33977    0.12790     0.12790  -2.657             0.007901 ** 
s(prey_density)5   -0.32114    0.14364     0.14364  -2.236             0.025391 *  
s(prey_density)6   -0.10644    0.16797     0.16797  -0.634             0.526286    
s(prey_density)7   -0.64339    0.17463     0.17463  -3.684             0.000230 ***
s(prey_density)8   -0.09699    0.18593     0.18593  -0.522             0.601923    
s(prey_density)9   -0.13937    0.21699     0.21699  -0.642             0.520682    
s(prey_density)10  -0.73951    0.33607     0.33607  -2.200             0.027790 *  
s(prey_density)11  -0.26787    0.39364     0.39364  -0.680             0.496203    
s(prey_density)12   0.25583    0.41432     0.41432   0.617             0.536931    
s(x,y)b1           -1.54145    0.17617     0.17617  -8.750 < 0.0000000000000002 ***
s(x,y)b2            2.95184    0.31495     0.31495   9.372 < 0.0000000000000002 ***
s(x,y)b3            4.01870    0.44340     0.44340   9.063 < 0.0000000000000002 ***
s(x,y)b4           40.46228    2.20601     2.20601  18.342 < 0.0000000000000002 ***
s(x,y)b5           -4.23474    0.78100     0.78100  -5.422   0.0000000597979195 ***
s(x,y)b6           -3.92662    0.60473     0.60473  -6.493   0.0000000000867562 ***
s(x,y)b7           -3.96114    0.56618     0.56618  -6.996   0.0000000000027430 ***
s(x,y)b8            2.49072    0.27847     0.27847   8.944 < 0.0000000000000002 ***
s(x,y)b9            3.77558    0.62699     0.62699   6.022   0.0000000017669590 ***
s(x,y)b10          -5.92523    0.79167     0.79167  -7.485   0.0000000000000759 ***
s(x,y)b11         -18.79775    1.23371     1.23371 -15.237 < 0.0000000000000002 ***
s(x,y)b12         -20.05584    1.29096     1.29096 -15.536 < 0.0000000000000002 ***
s(x,y)b13           1.63436    0.08688     0.08688  18.812 < 0.0000000000000002 ***
s(x,y)b14           1.61392    0.10443     0.10443  15.455 < 0.0000000000000002 ***
s(x,y)b15           3.45110    0.14484     0.14484  23.827 < 0.0000000000000002 ***
s(x,y)b16          -2.59353    0.17102     0.17102 -15.166 < 0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 17.58003)

    Null deviance: 484606  on 14652  degrees of freedom
Residual deviance: 163248  on 14577  degrees of freedom
AIC:  NA

Max Panel Size = 1 (independence assumed); Number of panels = 14653
Number of Fisher Scoring iterations: 6
```

Figure A6-21 Code excerpt detailing SALSA 2D summary for guillemot MRSea analysis.

Cross-Validation

A.3.13 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. The 2D model exhibited the lowest CV score comparative to the 1D and GAM model, indicating a better fit. As a result the 2D model was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-22).



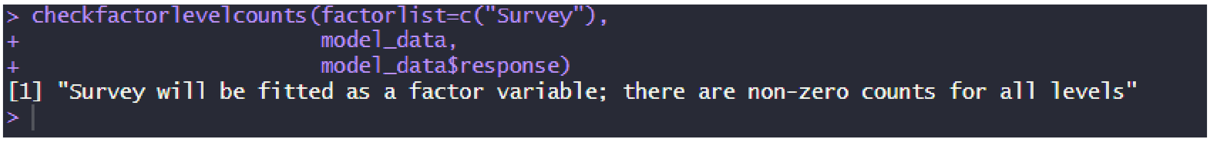
```
> cv_2d #SALSA 2D
[1] 1496.978
> cv_1d #SALSA 1D
[1] 4163.917
> cv_0d #Initial model
[1] 1715.67
```

Figure A6-22 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for guillemot MRSea analysis.

Razorbill

Factor Level Check

A.3.14 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-23).

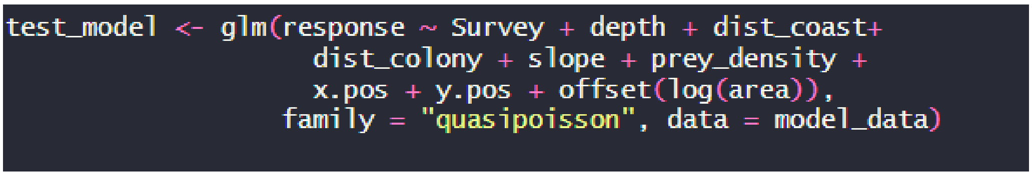


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                          model_data,
+                          model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-23 Code excerpt detailing factor level check for categorical variables considered within razorbill MRSea modelling.

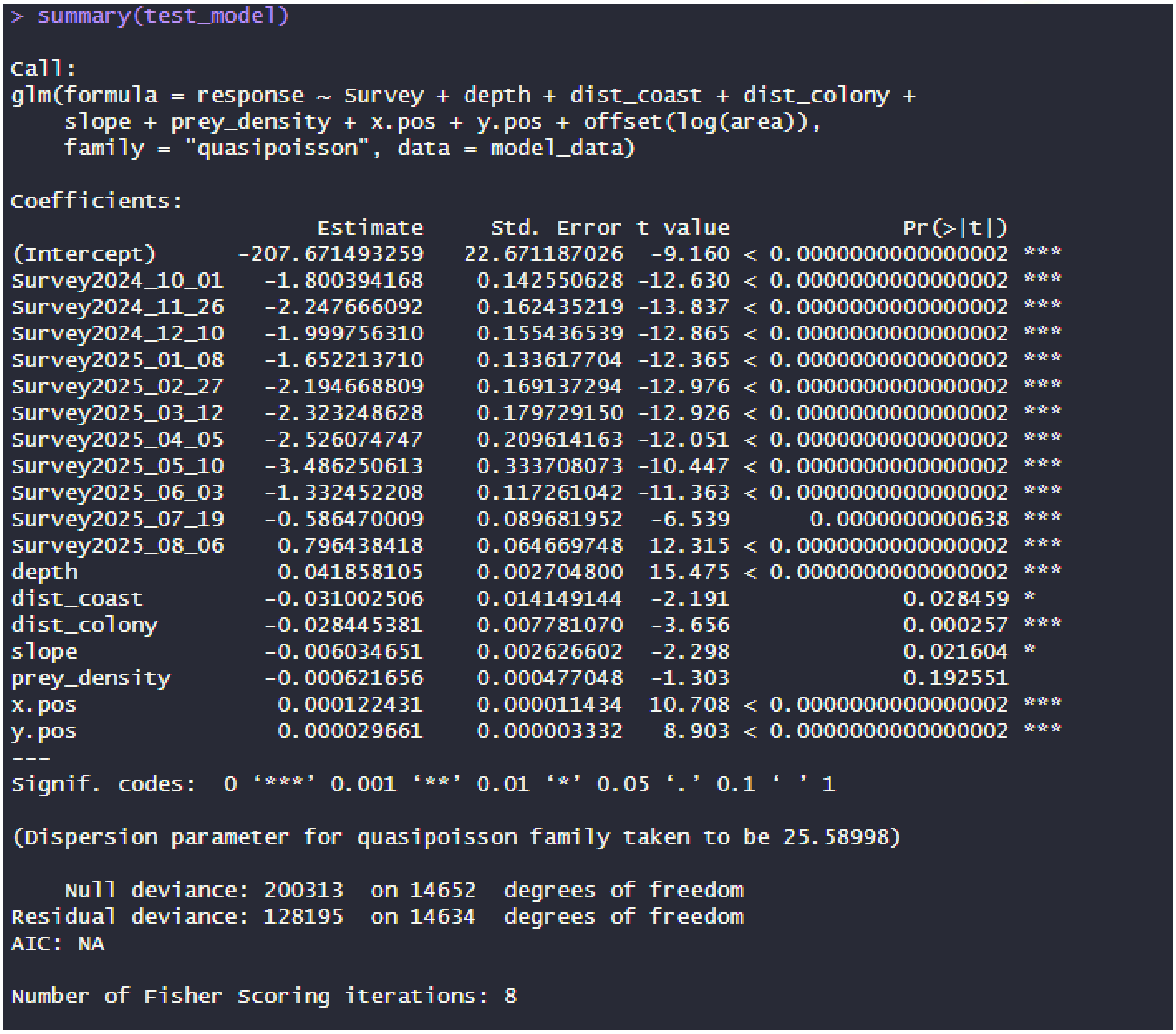
Initial GLM

A.3.15 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-24 & Figure A6-25).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-24 Code excerpt detailing the setup of the initial GLM considered within razorbill MRSea analysis.



```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_coast + dist_colony +
    slope + prey_density + x.pos + y.pos + offset(log(area)),
    family = "quasipoisson", data = model_data)

Coefficients:
                      Estimate     Std. Error t value             Pr(>|t|)    
(Intercept)      -207.671493259   22.671187026  -9.160 < 0.0000000000000002 ***
Survey2024_10_01   -1.800394168    0.142550628 -12.630 < 0.0000000000000002 ***
Survey2024_11_26   -2.247666092    0.162435219 -13.837 < 0.0000000000000002 ***
Survey2024_12_10   -1.999756310    0.155436539 -12.865 < 0.0000000000000002 ***
Survey2025_01_08   -1.652213710    0.133617704 -12.365 < 0.0000000000000002 ***
Survey2025_02_27   -2.194668809    0.169137294 -12.976 < 0.0000000000000002 ***
Survey2025_03_12   -2.323248628    0.179729150 -12.926 < 0.0000000000000002 ***
Survey2025_04_05   -2.526074747    0.209614163 -12.051 < 0.0000000000000002 ***
Survey2025_05_10   -3.486250613    0.333708073 -10.447 < 0.0000000000000002 ***
Survey2025_06_03   -1.332452208    0.117261042 -11.363 < 0.0000000000000002 ***
Survey2025_07_19   -0.586470009    0.089681952  -6.539    0.0000000000638 ***
Survey2025_08_06    0.796438418    0.064669748  12.315 < 0.0000000000000002 ***
depth               0.041858105    0.002704800  15.475 < 0.0000000000000002 ***
dist_coast         -0.031002506    0.014149144  -2.191             0.028459 *  
dist_colony        -0.028445381    0.007781070  -3.656             0.000257 ***
slope              -0.006034651    0.002626602  -2.298             0.021604 *  
prey_density       -0.000621656    0.000477048  -1.303             0.192551    
x.pos               0.000122431    0.000011434  10.708 < 0.0000000000000002 ***
y.pos               0.000029661    0.000003332   8.903 < 0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 25.58998)

    Null deviance: 200313  on 14652  degrees of freedom
Residual deviance: 128195  on 14634  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 8
```

Figure A6-25 Code excerpt detailing the summary of the initial GLM considered within razorbill MRSea analysis.

Collinearity Check

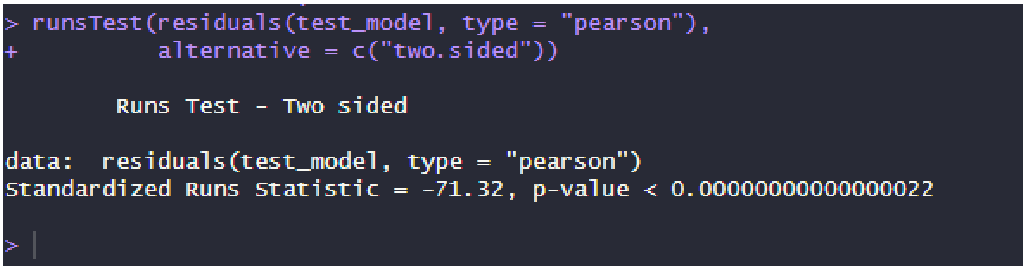
A.3.16 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-26).

```
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
survey        1.003264 11        1.000148
depth        10.854975  1        3.294689
dist_coast   41.234933  1        6.421443
dist_colony  10.065570  1        3.172628
slope         1.271693  1        1.127693
prey_density  1.398971  1        1.182781
x.pos        36.069306  1        6.005773
y.pos         4.474766  1        2.115364
```

Figure A6-26 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within razorbill MRSea analysis.

Runs Test

A.3.17 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-27).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -71.32, p-value < 0.00000000000000022

> |
```

Figure A6-27 Code excerpt detailing runs test carried out on the initial GLM considered within razorbill MRSea analysis.

A.3.18 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-28).

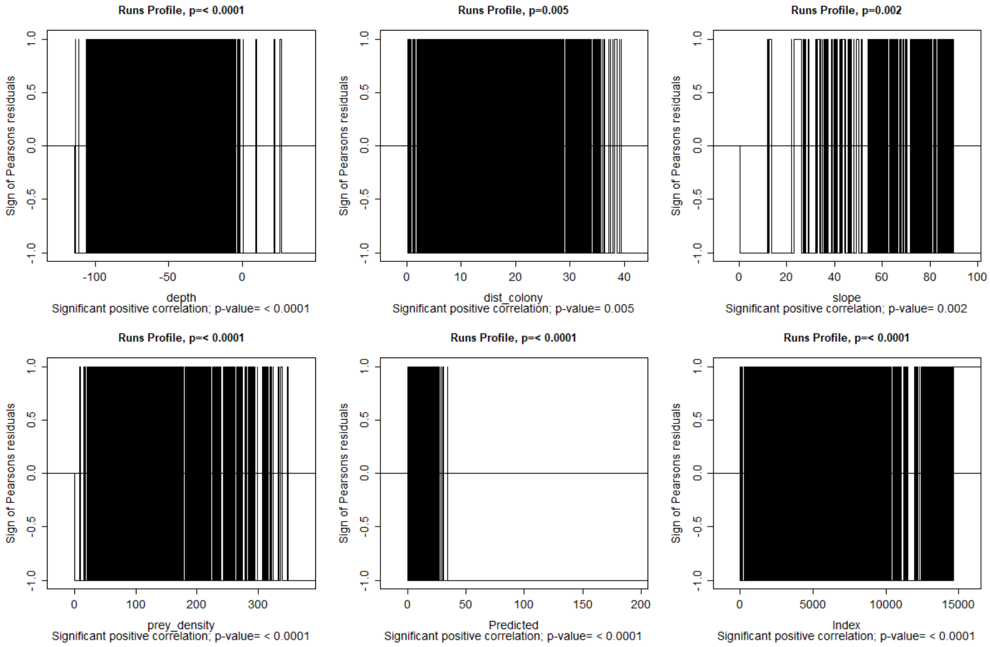


Figure A6-28 Runs profiles for explanatory variables considered within the initial GLM considered for razorbill MRSea analysis.

Autocorrelation

A.3.19 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-29).

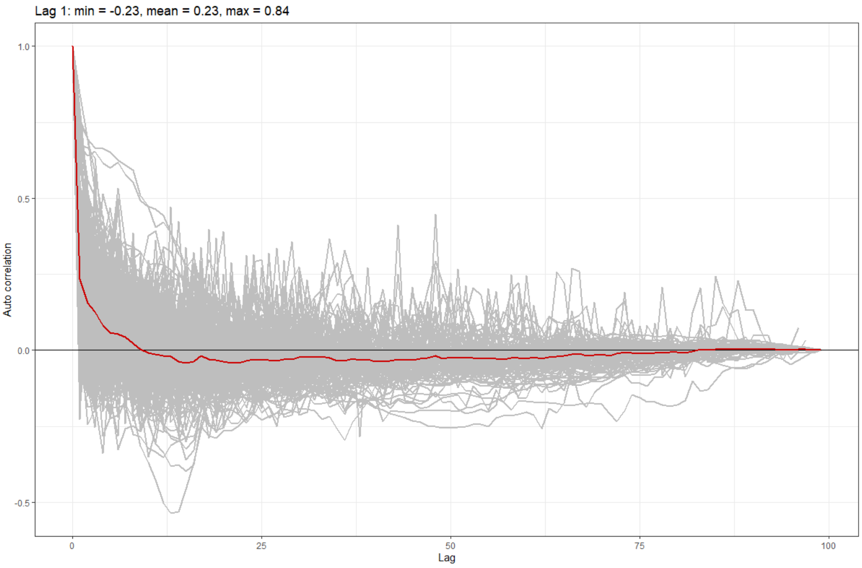


Figure A6-29 ACF plot for the initial GLM considered for razorbill MRSea analysis.

Cumulative Residuals

A.3.20 Cumulative residuals for explanatory variables were plotted (Figure A6-30, Figure A6-31, Figure A6-32 & Figure A6-33). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

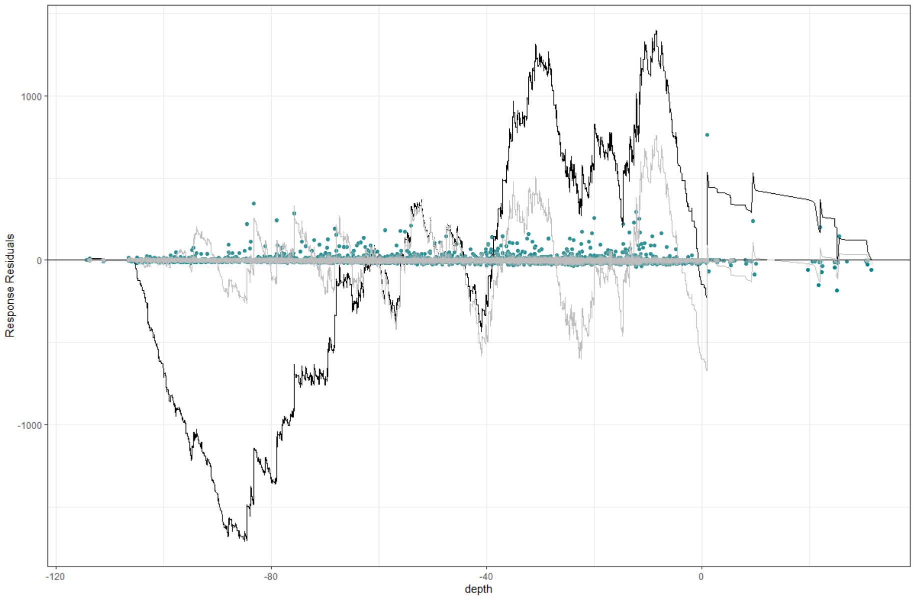


Figure A6-30 Cumulative residuals of depth plotted for the initial GLM considered for razorbill MRSea analysis.

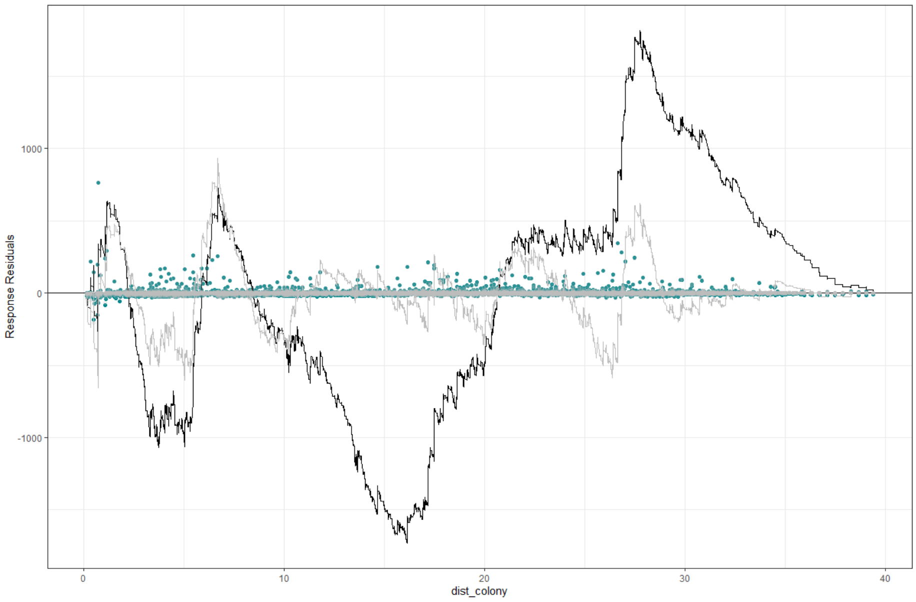


Figure A6-31 Cumulative residuals of distance to colony plotted for the initial GLM considered for razorbill MRSea analysis.

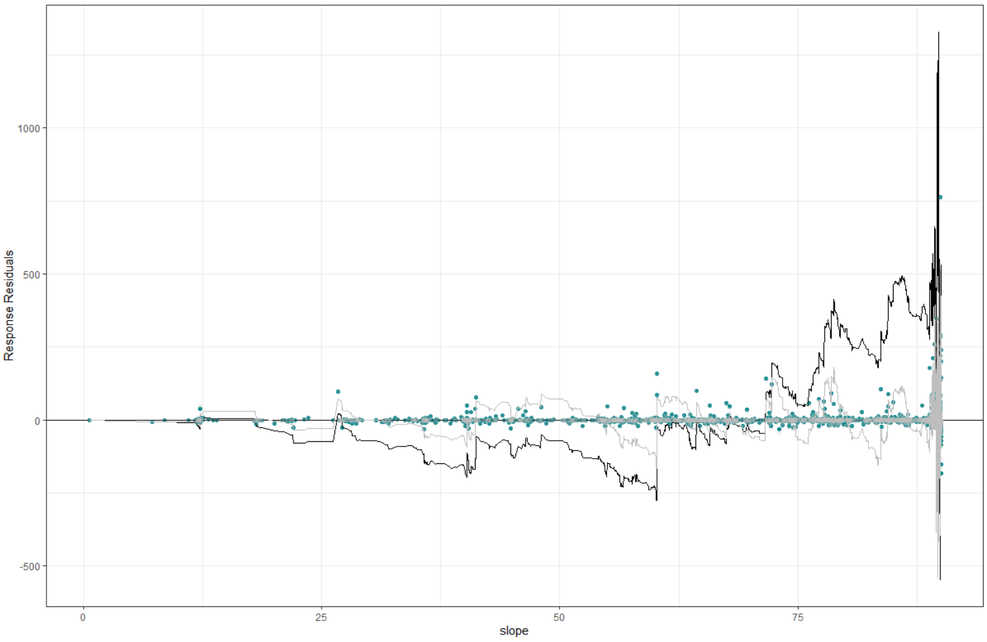


Figure A6-32 Cumulative residuals of seabed slope plotted for the initial GLM considered for razorbill MRSea analysis.

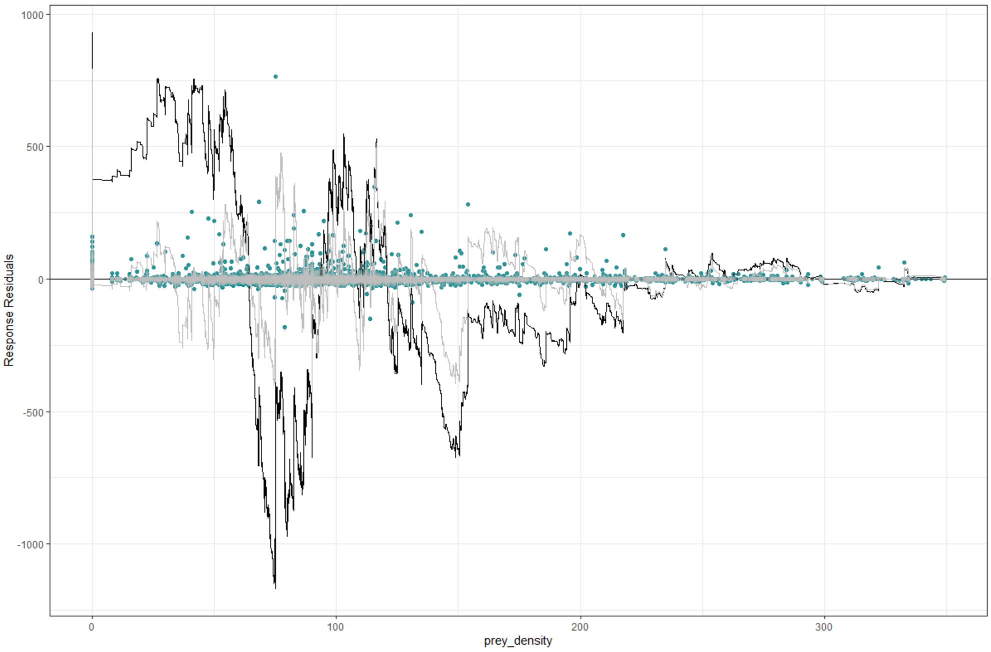
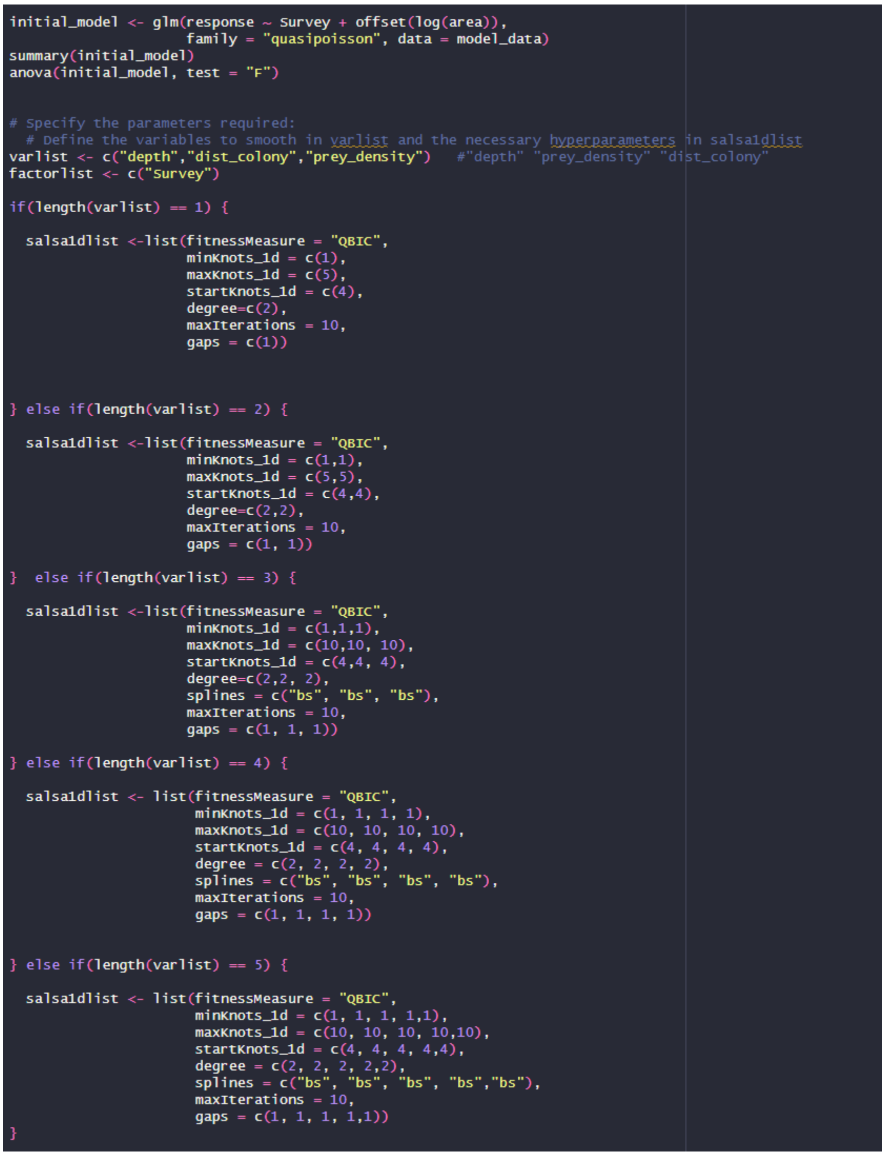


Figure A6-33 Cumulative residuals of prey density plotted for the initial GLM considered for razorbill MRSea analysis.

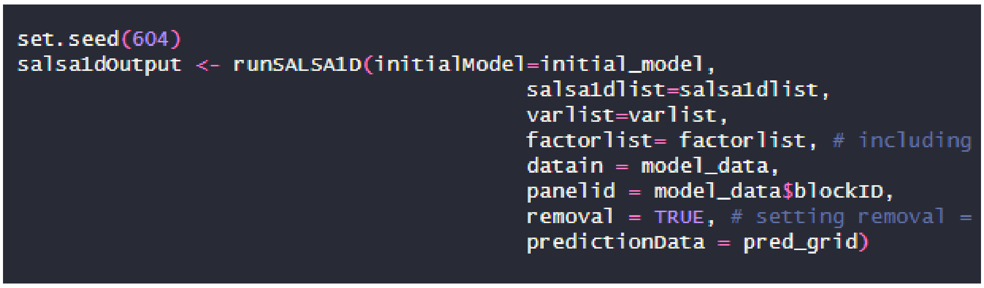
SALSA 1D

A.3.21 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.22 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-34 & Figure A6-35).

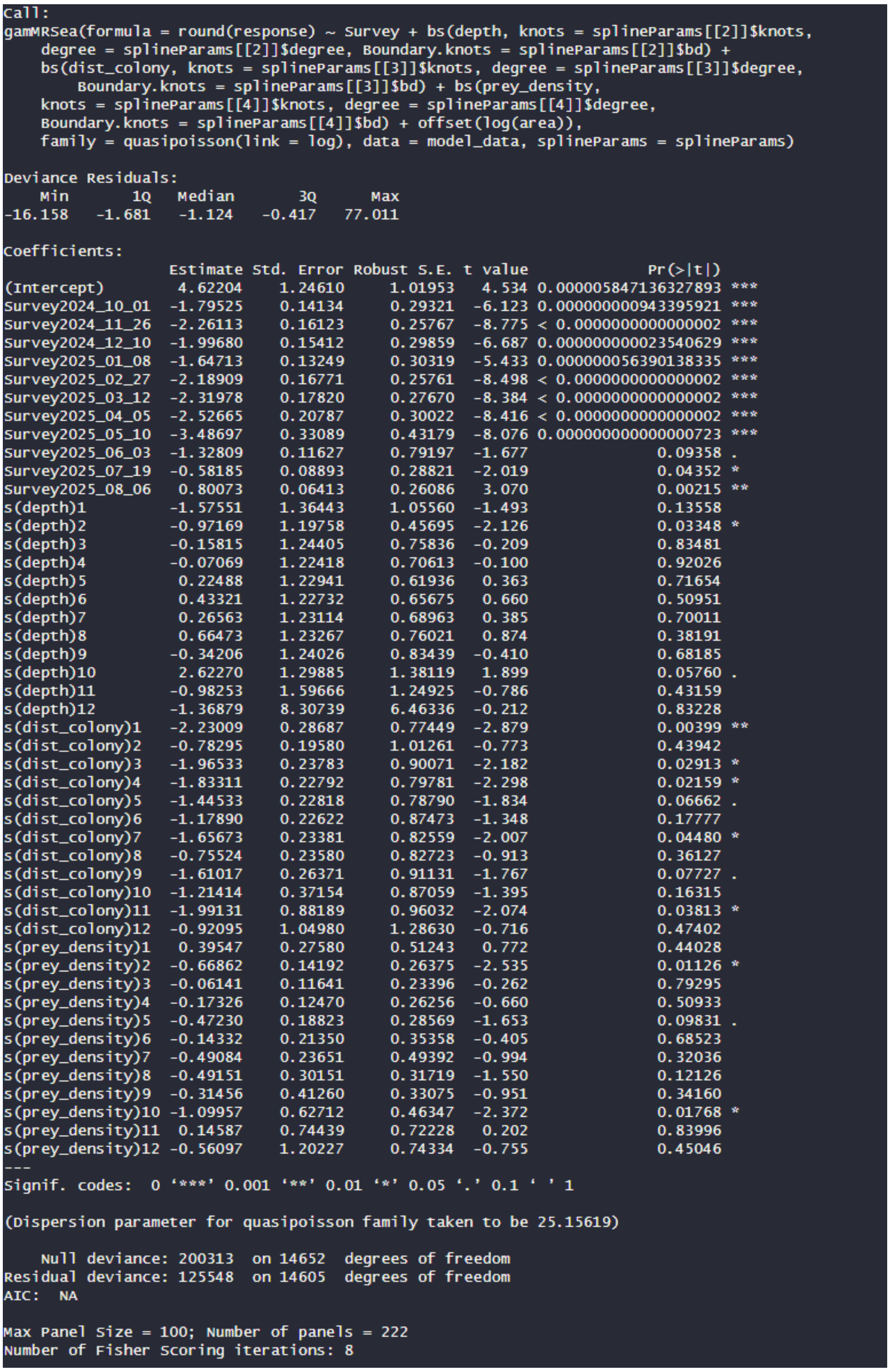


```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth","dist_colony","prey_density")   #"depth" "prey_density" "dist_colony"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     maxIterations = 10,
                     gaps = c(1))



} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1,1),
                      maxKnots_1d = c(10, 10, 10, 10,10),
                      startKnots_1d = c(4, 4, 4, 4,4),
                      degree = c(2, 2, 2, 2,2),
                      splines = c("bs", "bs", "bs", "bs","bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1,1))
}
```



```
set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                            salsa1dlist=salsa1dlist,
                            varlist=varlist,
                            factorlist= factorlist, # including
                            datain = model_data,
                            panelid = model_data$blockID,
                            removal = TRUE, # setting removal =
                            predictionData = pred_grid)
```

Figure A6-34 Code excerpts detailing the setup process for the SALSA 1D model considered for razorbill MRSea analysis.



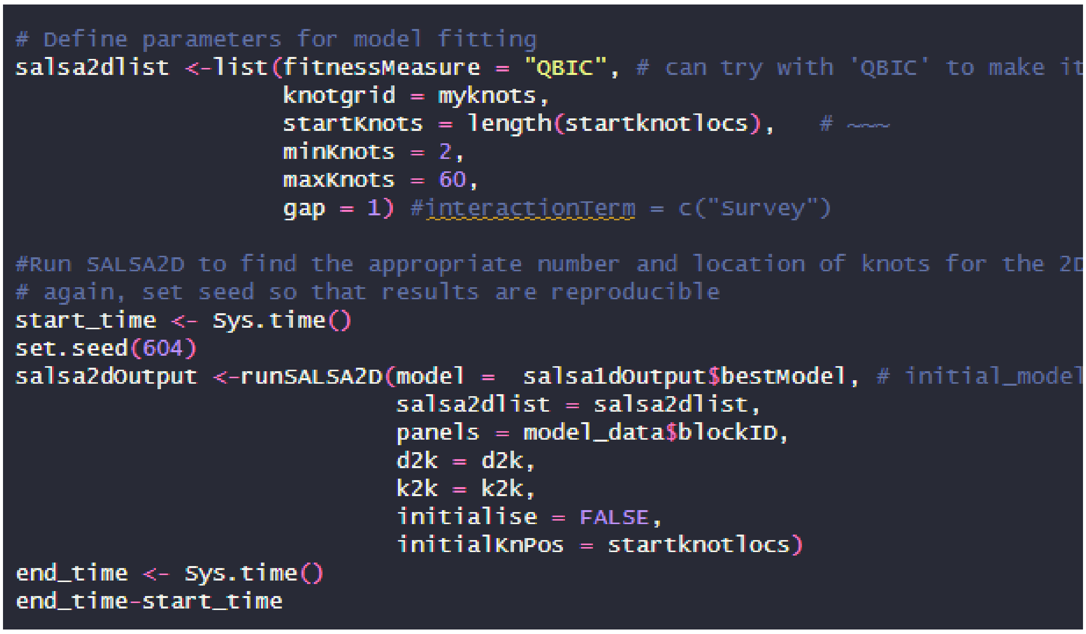
```
Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, 
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + 
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, 
        Boundary.knots = splineParams[[3]]$bd) + bs(prey_density, 
    knots = splineParams[[4]]$knots, degree = splineParams[[4]]$degree, 
    Boundary.knots = splineParams[[4]]$bd) + offset(log(area)), 
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-16.158   -1.681   -1.124   -0.417   77.011  

Coefficients:
                  Estimate Std. Error Robust S.E. t value             Pr(>|t|)    
(Intercept)        4.62204    1.24610     1.01953   4.534 0.000005847136327893 ***
Survey2024_10_01  -1.79525    0.14134     0.29321  -6.123 0.000000000943395921 ***
Survey2024_11_26  -2.26113    0.16123     0.25767  -8.775 < 0.0000000000000002 ***
Survey2024_12_10  -1.99680    0.15412     0.29859  -6.687 0.000000000023540629 ***
Survey2025_01_08  -1.64713    0.13249     0.30319  -5.433 0.000000056390138335 ***
Survey2025_02_27  -2.18909    0.16771     0.25761  -8.498 < 0.0000000000000002 ***
Survey2025_03_12  -2.31978    0.17820     0.27670  -8.384 < 0.0000000000000002 ***
Survey2025_04_05  -2.52665    0.20787     0.30022  -8.416 < 0.0000000000000002 ***
Survey2025_05_10  -3.48697    0.33089     0.43179  -8.076 0.000000000000000723 ***
Survey2025_06_03  -1.32809    0.11627     0.79197  -1.677              0.09358 .  
Survey2025_07_19  -0.58185    0.08893     0.28821  -2.019              0.04352 *  
Survey2025_08_06   0.80073    0.06413     0.26086   3.070              0.00215 ** 
s(depth)1         -1.57551    1.36443     1.05560  -1.493              0.13558    
s(depth)2         -0.97169    1.19758     0.45695  -2.126              0.03348 *  
s(depth)3         -0.15815    1.24405     0.75836  -0.209              0.83481    
s(depth)4         -0.07069    1.22418     0.70613  -0.100              0.92026    
s(depth)5          0.22488    1.22941     0.61936   0.363              0.71654    
s(depth)6          0.43321    1.22732     0.65675   0.660              0.50951    
s(depth)7          0.26563    1.23114     0.68963   0.385              0.70011    
s(depth)8          0.66473    1.23267     0.76021   0.874              0.38191    
s(depth)9         -0.34206    1.24026     0.83439  -0.410              0.68185    
s(depth)10         2.62270    1.29885     1.38119   1.899              0.05760 .  
s(depth)11        -0.98253    1.59666     1.24925  -0.786              0.43159    
s(depth)12        -1.36879    8.30739     6.46336  -0.212              0.83228    
s(dist_colony)1   -2.23009    0.28687     0.77449  -2.879              0.00399 ** 
s(dist_colony)2   -0.78295    0.19580     1.01261  -0.773              0.43942    
s(dist_colony)3   -1.96533    0.23783     0.90071  -2.182              0.02913 *  
s(dist_colony)4   -1.83311    0.22792     0.79781  -2.298              0.02159 *  
s(dist_colony)5   -1.44533    0.22818     0.78790  -1.834              0.06662 .  
s(dist_colony)6   -1.17890    0.22622     0.87473  -1.348              0.17777    
s(dist_colony)7   -1.65673    0.23381     0.82559  -2.007              0.04480 *  
s(dist_colony)8   -0.75524    0.23580     0.82723  -0.913              0.36127    
s(dist_colony)9   -1.61017    0.26371     0.91131  -1.767              0.07727 .  
s(dist_colony)10  -1.21414    0.37154     0.87059  -1.395              0.16315    
s(dist_colony)11  -1.99131    0.88189     0.96032  -2.074              0.03813 *  
s(dist_colony)12  -0.92095    1.04980     1.28630  -0.716              0.47402    
s(prey_density)1   0.39547    0.27580     0.51243   0.772              0.44028    
s(prey_density)2  -0.66862    0.14192     0.26375  -2.535              0.01126 *  
s(prey_density)3  -0.06141    0.11641     0.23396  -0.262              0.79295    
s(prey_density)4  -0.17326    0.12470     0.26256  -0.660              0.50933    
s(prey_density)5  -0.47230    0.18823     0.28569  -1.653              0.09831 .  
s(prey_density)6  -0.14332    0.21350     0.35358  -0.405              0.68523    
s(prey_density)7  -0.49084    0.23651     0.49392  -0.994              0.32036    
s(prey_density)8  -0.49151    0.30151     0.31719  -1.550              0.12126    
s(prey_density)9  -0.31456    0.41260     0.33075  -0.951              0.34160    
s(prey_density)10 -1.09957    0.62712     0.46347  -2.372              0.01768 *  
s(prey_density)11  0.14587    0.74439     0.72228   0.202              0.83996    
s(prey_density)12 -0.56097    1.20227     0.74334  -0.755              0.45046    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 25.15619)

    Null deviance: 200313  on 14652  degrees of freedom
Residual deviance: 125548  on 14605  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 8
```

Figure A6-35 Code excerpt detailing the summary for the SALSA 1D model for razorbill MRSea analysis.

SALSA 2D

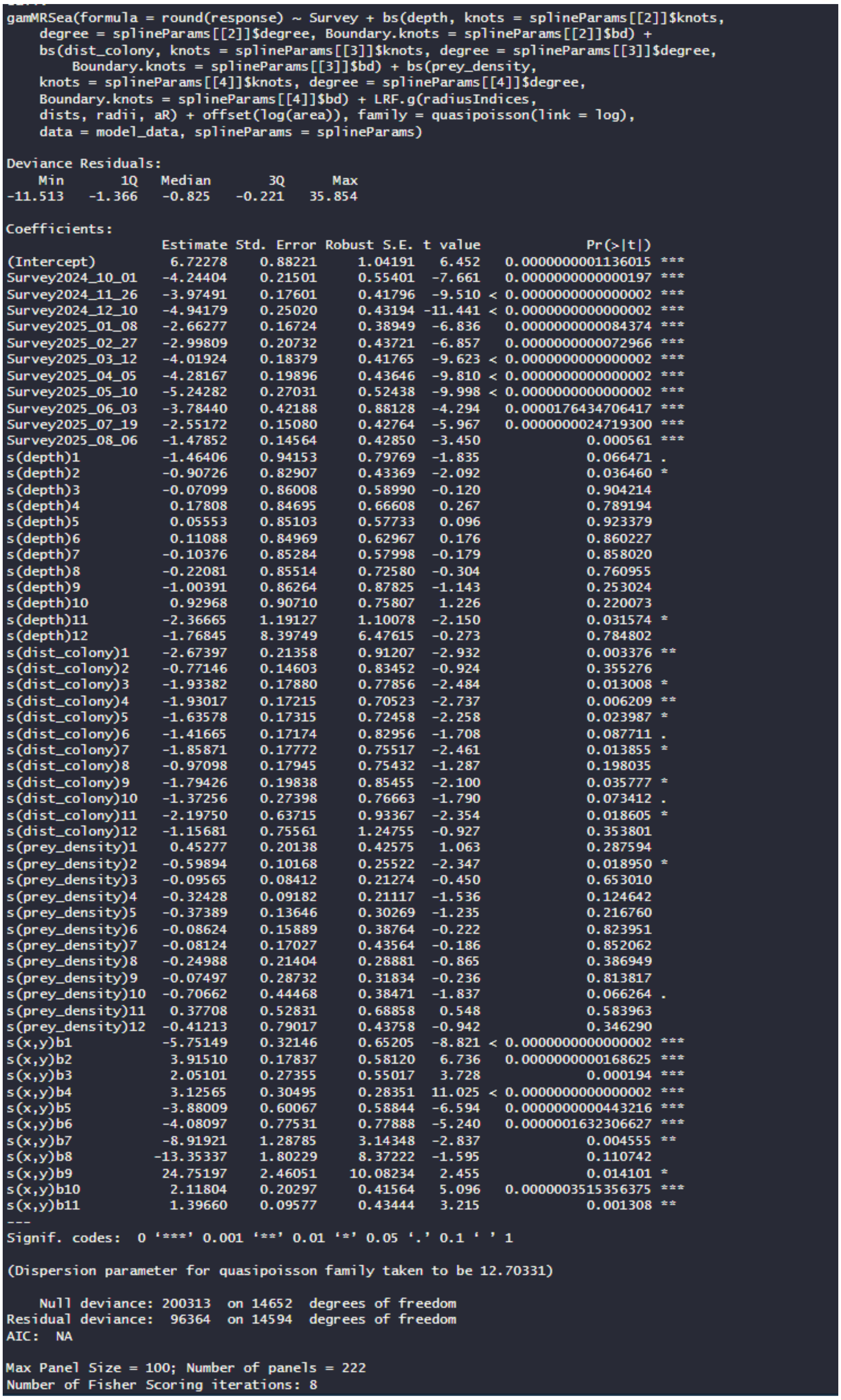
A.3.23 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.24 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-36 & Figure A6-37).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC' to make it
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),    # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots for the 2D
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2doutput <-runSALSA2D(model =  salsa1doutput$bestModel, # initial_model
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-36 Code excerpt for SALSA 2D model setup for razorbill MRSea analysis.

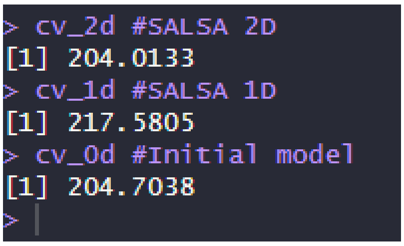


```
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree,
        Boundary.knots = splineParams[[3]]$bd) + bs(prey_density,
    knots = splineParams[[4]]$knots, degree = splineParams[[4]]$degree,
    Boundary.knots = splineParams[[4]]$bd) + LRF.g(radiusIndices,
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log),
    data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-11.513   -1.366   -0.825   -0.221   35.854

Coefficients:
                   Estimate Std. Error Robust S.E. t value             Pr(>|t|)
(Intercept)         6.72278    0.88221     1.04191   6.452   0.0000000001136015 ***
Survey2024_10_01   -4.24404    0.21501     0.55401  -7.661   0.0000000000000197 ***
Survey2024_11_26   -3.97491    0.17601     0.41796  -9.510 < 0.0000000000000002 ***
Survey2024_12_10   -4.94179    0.25020     0.43194 -11.441 < 0.0000000000000002 ***
Survey2025_01_08   -2.66277    0.16724     0.38949  -6.836   0.0000000000084374 ***
Survey2025_02_27   -2.99809    0.20732     0.43721  -6.857   0.0000000000072966 ***
Survey2025_03_12   -4.01924    0.18379     0.41765  -9.623 < 0.0000000000000002 ***
Survey2025_04_05   -4.28167    0.19896     0.43646  -9.810 < 0.0000000000000002 ***
Survey2025_05_10   -5.24282    0.27031     0.52438  -9.998 < 0.0000000000000002 ***
Survey2025_06_03   -3.78440    0.42188     0.88128  -4.294   0.0000176434706417 ***
Survey2025_07_19   -2.55172    0.15080     0.42764  -5.967   0.0000000024719300 ***
Survey2025_08_06   -1.47852    0.14564     0.42850  -3.450             0.000561 ***
s(depth)1          -1.46406    0.94153     0.79769  -1.835             0.066471 .
s(depth)2          -0.90726    0.82907     0.43369  -2.092             0.036460 *
s(depth)3          -0.07099    0.86008     0.58990  -0.120             0.904214
s(depth)4           0.17808    0.84695     0.66608   0.267             0.789194
s(depth)5           0.05553    0.85103     0.57733   0.096             0.923379
s(depth)6           0.11088    0.84969     0.62967   0.176             0.860227
s(depth)7          -0.10376    0.85284     0.57998  -0.179             0.858020
s(depth)8          -0.22081    0.85514     0.72580  -0.304             0.760955
s(depth)9          -1.00391    0.86264     0.87825  -1.143             0.253024
s(depth)10          0.92968    0.90710     0.75807   1.226             0.220073
s(depth)11         -2.36665    1.19127     1.10078  -2.150             0.031574 *
s(depth)12         -1.76845    8.39749     6.47615  -0.273             0.784802
s(dist_colony)1    -2.67397    0.21358     0.91207  -2.932             0.003376 **
s(dist_colony)2    -0.77146    0.14603     0.83452  -0.924             0.355276
s(dist_colony)3    -1.93382    0.17880     0.77856  -2.484             0.013008 *
s(dist_colony)4    -1.93017    0.17215     0.70523  -2.737             0.006209 **
s(dist_colony)5    -1.63578    0.17315     0.72458  -2.258             0.023987 *
s(dist_colony)6    -1.41665    0.17174     0.82956  -1.708             0.087711 .
s(dist_colony)7    -1.85871    0.17772     0.75517  -2.461             0.013855 *
s(dist_colony)8    -0.97098    0.17945     0.75432  -1.287             0.198035
s(dist_colony)9    -1.79426    0.19838     0.85455  -2.100             0.035777 *
s(dist_colony)10   -1.37256    0.27398     0.76663  -1.790             0.073412 .
s(dist_colony)11   -2.19750    0.63715     0.93367  -2.354             0.018605 *
s(dist_colony)12   -1.15681    0.75561     1.24755  -0.927             0.353801
s(prey_density)1    0.45277    0.20138     0.42575   1.063             0.287594
s(prey_density)2   -0.59894    0.10168     0.25522  -2.347             0.018950 *
s(prey_density)3   -0.09565    0.08412     0.21274  -0.450             0.653010
s(prey_density)4   -0.32428    0.09182     0.21117  -1.536             0.124642
s(prey_density)5   -0.37389    0.13646     0.30269  -1.235             0.216760
s(prey_density)6   -0.08624    0.15889     0.38764  -0.222             0.823951
s(prey_density)7   -0.08124    0.17027     0.43564  -0.186             0.852062
s(prey_density)8   -0.24988    0.21404     0.28881  -0.865             0.386949
s(prey_density)9   -0.07497    0.28732     0.31834  -0.236             0.813817
s(prey_density)10  -0.70662    0.44468     0.38471  -1.837             0.066264 .
s(prey_density)11   0.37708    0.52831     0.68858   0.548             0.583963
s(prey_density)12  -0.41213    0.79017     0.43758  -0.942             0.346290
s(x,y)b1           -5.75149    0.32146     0.65205  -8.821 < 0.0000000000000002 ***
s(x,y)b2            3.91510    0.17837     0.58120   6.736   0.0000000000168625 ***
s(x,y)b3            2.05101    0.27355     0.55017   3.728             0.000194 ***
s(x,y)b4            3.12565    0.30495     0.28351  11.025 < 0.0000000000000002 ***
s(x,y)b5           -3.88009    0.60067     0.58844  -6.594   0.0000000000443216 ***
s(x,y)b6           -4.08097    0.77531     0.77888  -5.240   0.0000001632306627 ***
s(x,y)b7           -8.91921    1.28785     3.14348  -2.837             0.004555 **
s(x,y)b8          -13.35337    1.80229     8.37222  -1.595             0.110742
s(x,y)b9           24.75197    2.46051    10.08234   2.455             0.014101 *
s(x,y)b10           2.11804    0.20297     0.41564   5.096   0.0000003515356375 ***
s(x,y)b11           1.39660    0.09577     0.43444   3.215             0.001308 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 12.70331)

    Null deviance: 200313  on 14652  degrees of freedom
Residual deviance:  96364  on 14594  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 8
```

Figure A6-37 Code excerpt detailing SALSA 2D summary for razorbill MRSea analysis.

Cross-Validation

A.3.25 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a slightly higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-38).



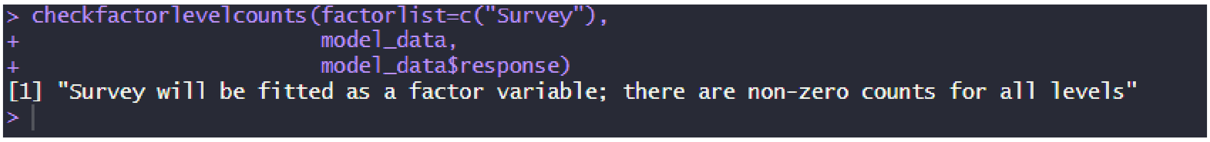
```
> cv_2d #SALSA 2D
[1] 204.0133
> cv_1d #SALSA 1D
[1] 217.5805
> cv_0d #Initial model
[1] 204.7038
> |
```

Figure A6-38 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for razorbill MRSea analysis.

Kittiwake

Factor Level Check

A.3.26 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, "Survey" (a proxy for month), was checked and required no further action (Figure A6-39).

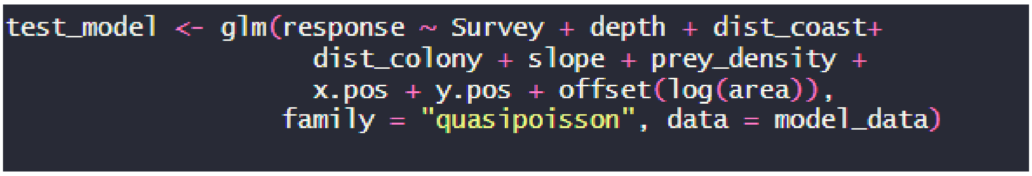


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                         model_data,
+                         model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-39 Code excerpt detailing factor level check for categorical variables considered within kittiwake MRSea modelling.

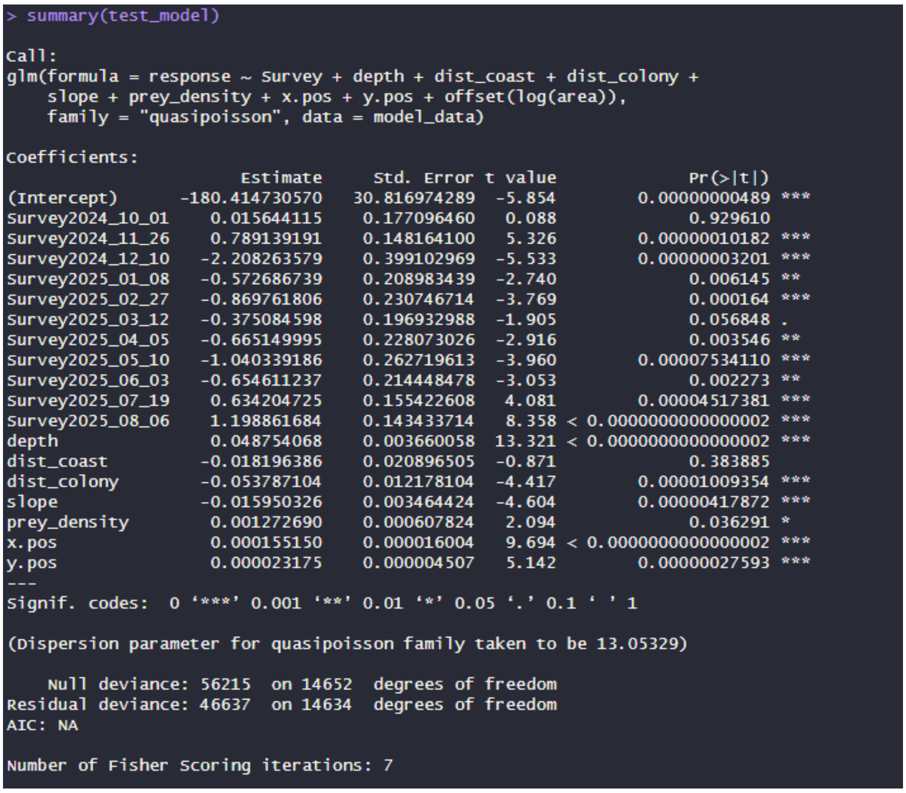
Initial GLM

A.3.27 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-40 & Figure A6-41).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-40 Code excerpt detailing the setup of the initial GLM considered within kittiwake MRSea analysis.

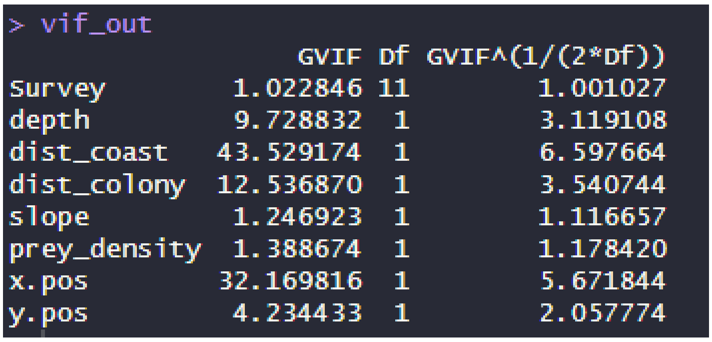


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_coast + dist_colony +
    slope + prey_density + x.pos + y.pos + offset(log(area)),
    family = "quasipoisson", data = model_data)

Coefficients:
                    Estimate    Std. Error t value            Pr(>|t|)
(Intercept)      -180.414730570  30.816974289  -5.854        0.00000000489 ***
Survey2024_10_01    0.015644115   0.177096460   0.088            0.929610
Survey2024_11_26    0.789139191   0.148164100   5.326        0.00000010182 ***
Survey2024_12_10   -2.208263579   0.399102969  -5.533        0.00000003201 ***
Survey2025_01_08   -0.572686739   0.208983439  -2.740            0.006145 **
Survey2025_02_27   -0.869761806   0.230746714  -3.769            0.000164 ***
Survey2025_03_12   -0.375084598   0.196932988  -1.905            0.056848 .
Survey2025_04_05   -0.665149995   0.228073026  -2.916            0.003546 **
Survey2025_05_10   -1.040339186   0.262719613  -3.960        0.00007534110 ***
Survey2025_06_03   -0.654611237   0.214448478  -3.053            0.002273 **
Survey2025_07_19    0.634204725   0.155422608   4.081        0.00004517381 ***
Survey2025_08_06    1.198861684   0.143433714   8.358 < 0.0000000000000002 ***
depth               0.048754068   0.003660058  13.321 < 0.0000000000000002 ***
dist_coast         -0.018196386   0.020896505  -0.871            0.383885
dist_colony        -0.053787104   0.012178104  -4.417        0.00001009354 ***
slope              -0.015950326   0.003464424  -4.604        0.00000417872 ***
prey_density        0.001272690   0.000607824   2.094            0.036291 *
x.pos               0.000155150   0.000016004   9.694 < 0.0000000000000002 ***
y.pos               0.000023175   0.000004507   5.142        0.00000027593 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 13.05329)

    Null deviance: 56215  on 14652  degrees of freedom
Residual deviance: 46637  on 14634  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 7
```

Figure A6-41 Code excerpt detailing the summary of the initial GLM considered within kittiwake MRSea analysis.

Collinearity Check

A.3.28 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-42).

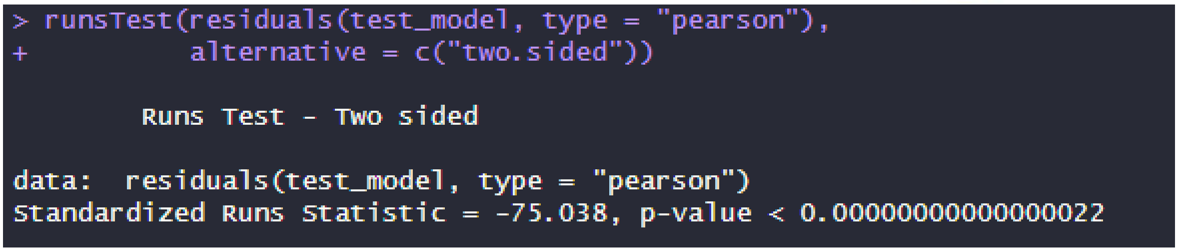


```
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
Survey        1.022846 11        1.001027
depth         9.728832  1        3.119108
dist_coast   43.529174  1        6.597664
dist_colony  12.536870  1        3.540744
slope         1.246923  1        1.116657
prey_density  1.388674  1        1.178420
x.pos        32.169816  1        5.671844
y.pos         4.234433  1        2.057774
```

Figure A6-42 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within kittiwake MRSea analysis.

Runs Test

A.3.29 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-43).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -75.038, p-value < 0.00000000000000022
```

Figure A6-43 Code excerpt detailing runs test carried out on the initial GLM considered within kittiwake MRSea analysis.

A.3.30 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-44).

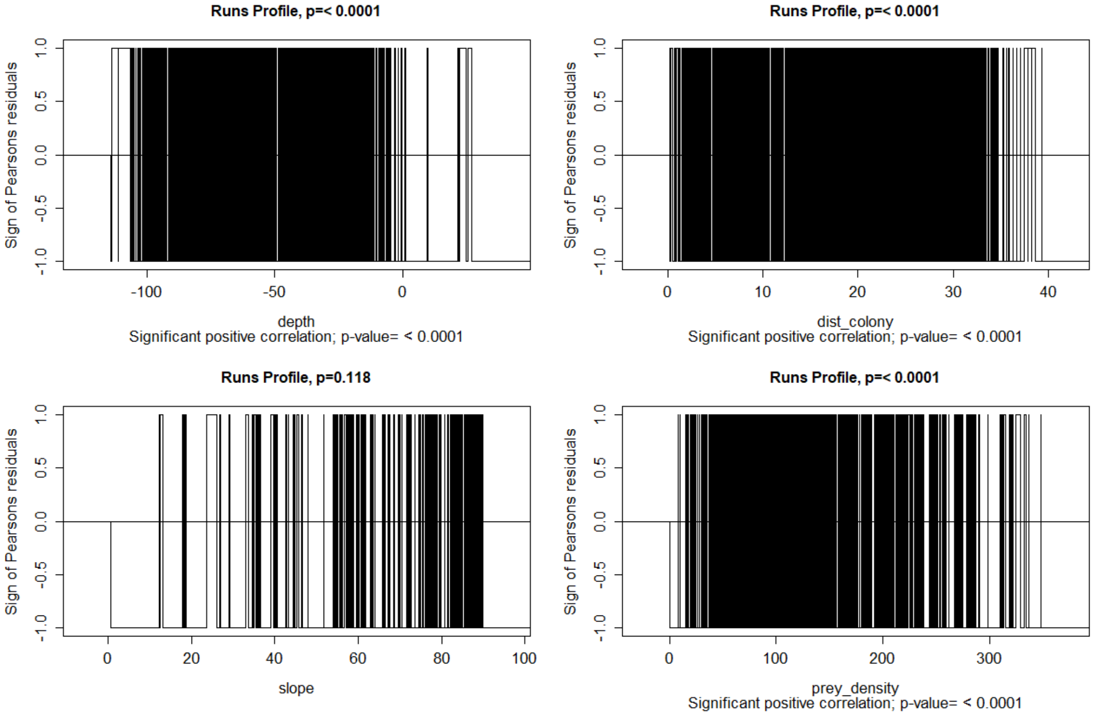


Figure A6-44 Runs profiles for explanatory variables considered within the initial GLM considered for kittiwake MRSea analysis.

Autocorrelation

A.3.31 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-45).

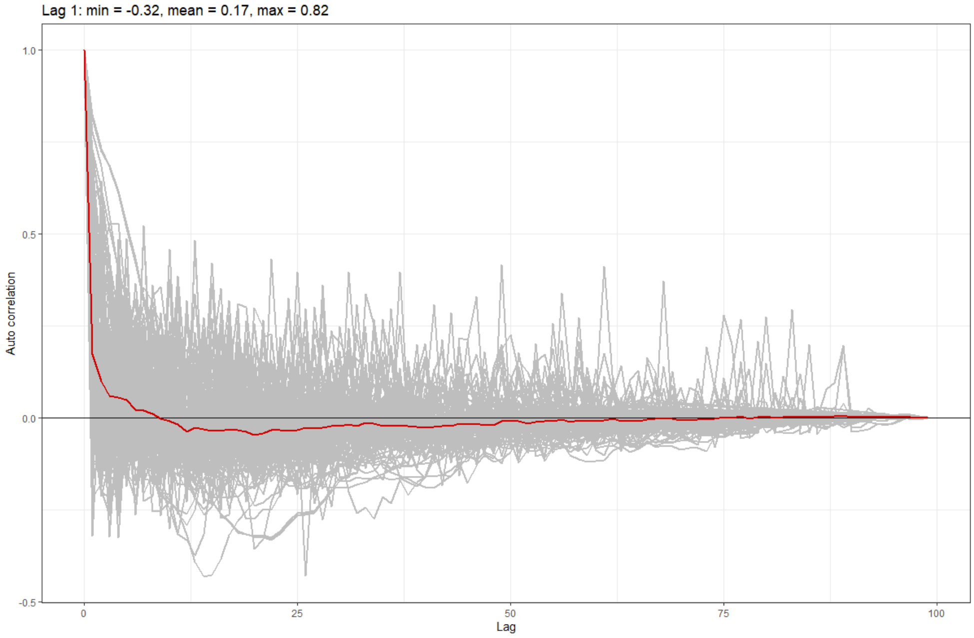


Figure A6-45 ACF plot for the initial GLM considered for kittiwake MRSea analysis.

Cumulative Residuals

A.3.32 Cumulative residuals for explanatory variables were plotted (Figure A6-46, Figure A6-47, Figure A6-48 & Figure A6-49). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

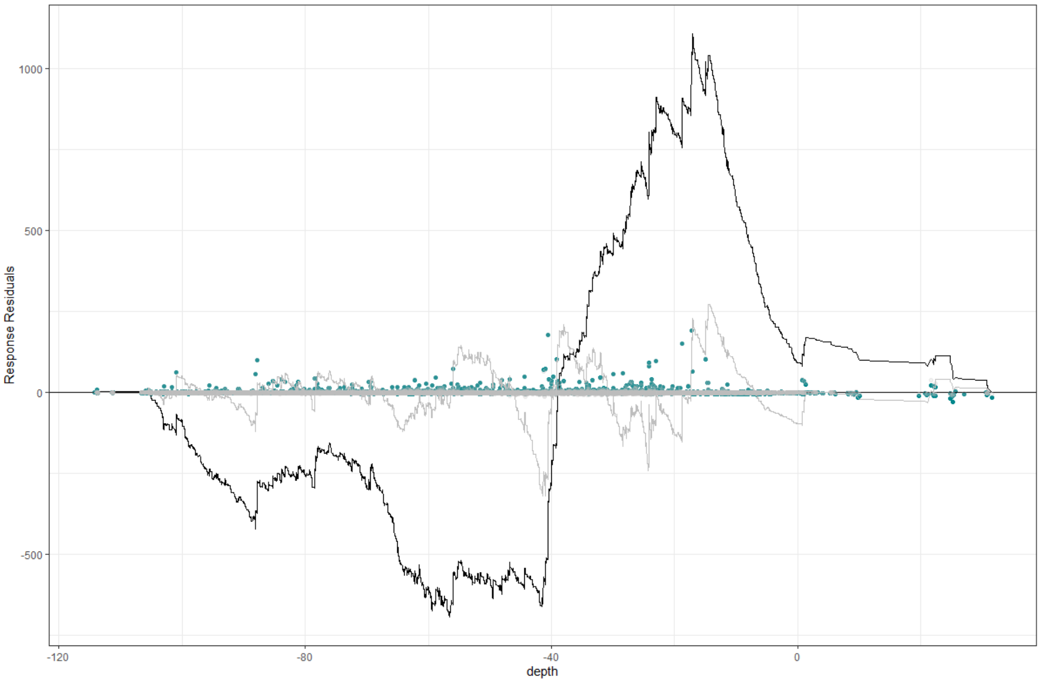


Figure A6-46 Cumulative residuals of depth plotted for the initial GLM considered for kittiwake MRSea analysis.

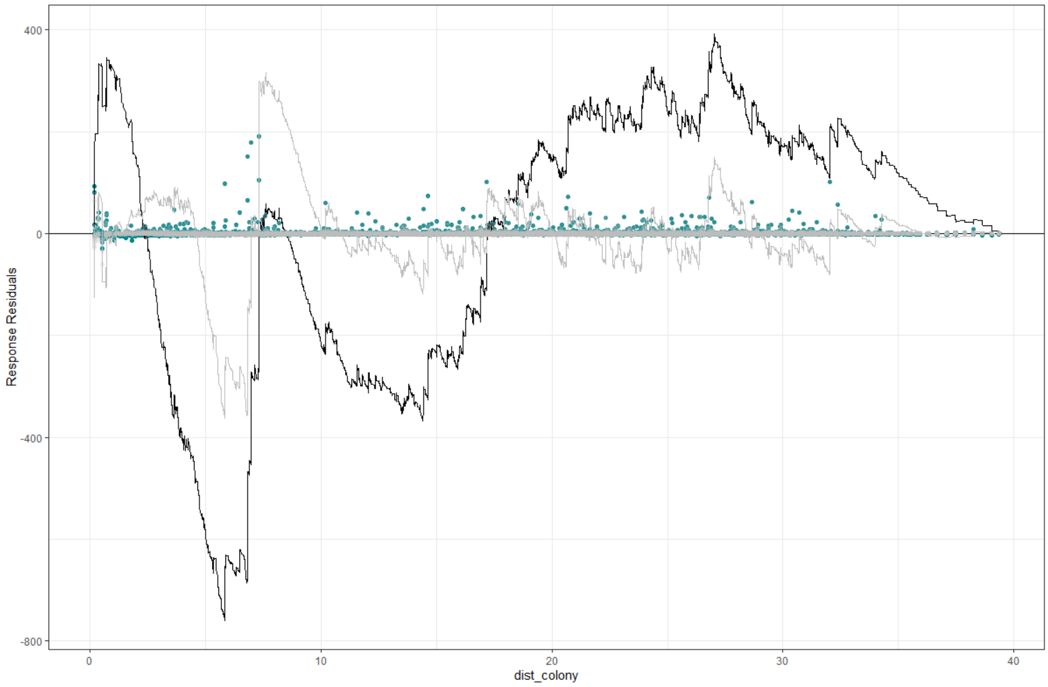


Figure A6-47 Cumulative residuals of distance to colony plotted for the initial GLM considered for kittiwake MRSea analysis.

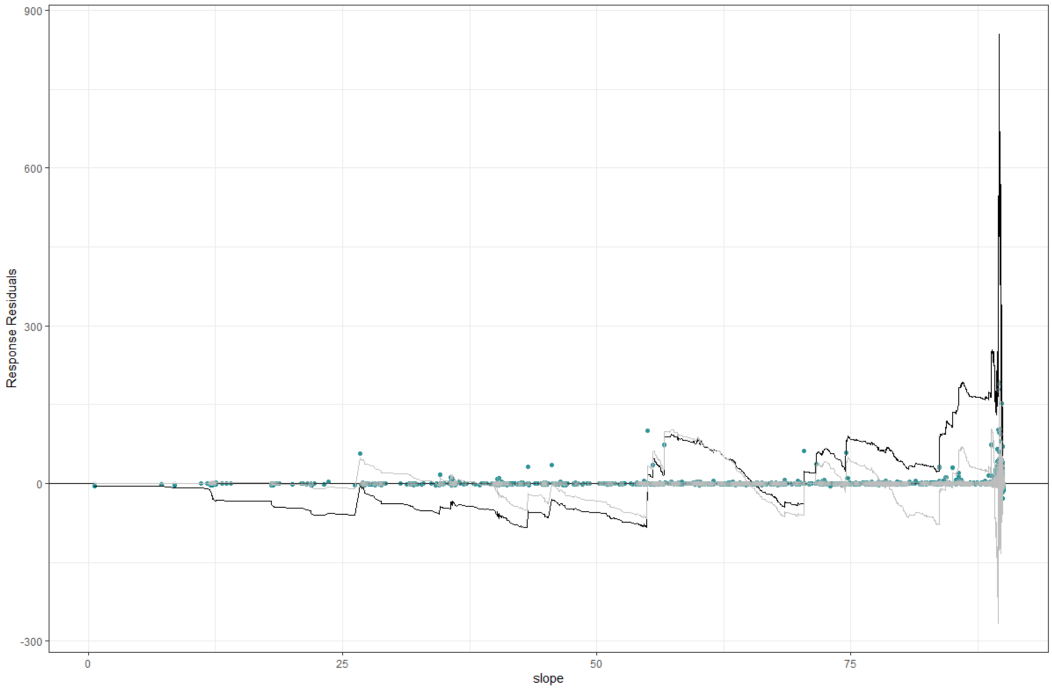


Figure A6-48 Cumulative residuals of seabed slope plotted for the initial GLM considered for kittiwake MRSea analysis.

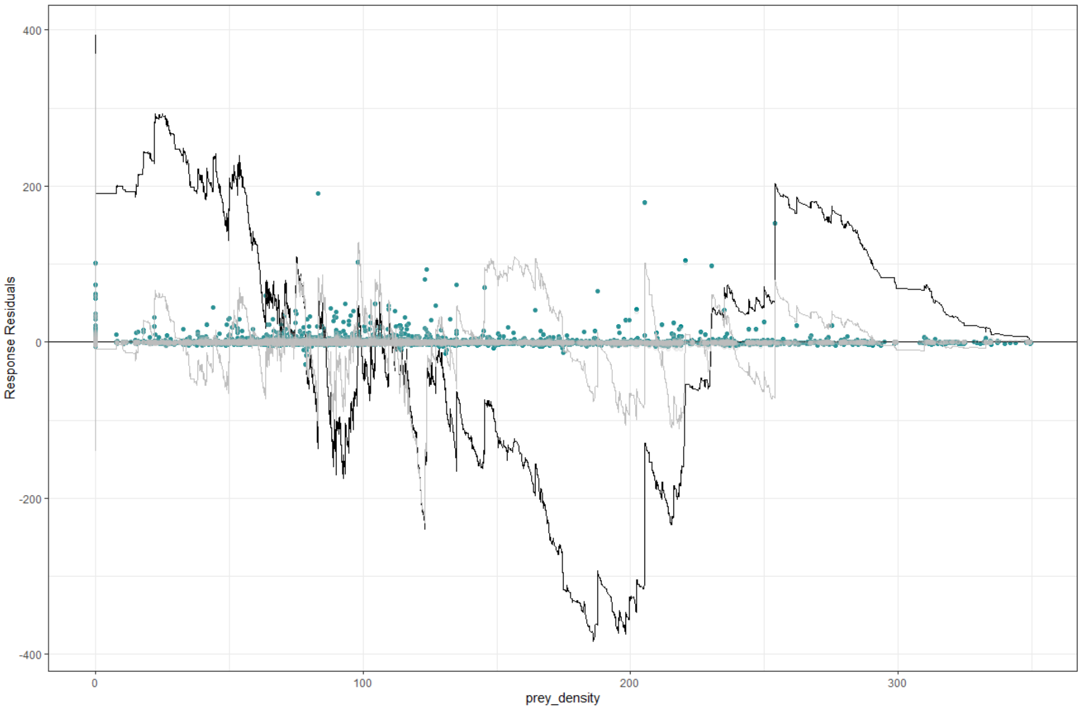


Figure A6-49 Cumulative residuals of prey density plotted for the initial GLM considered for kittiwake MRSea analysis.

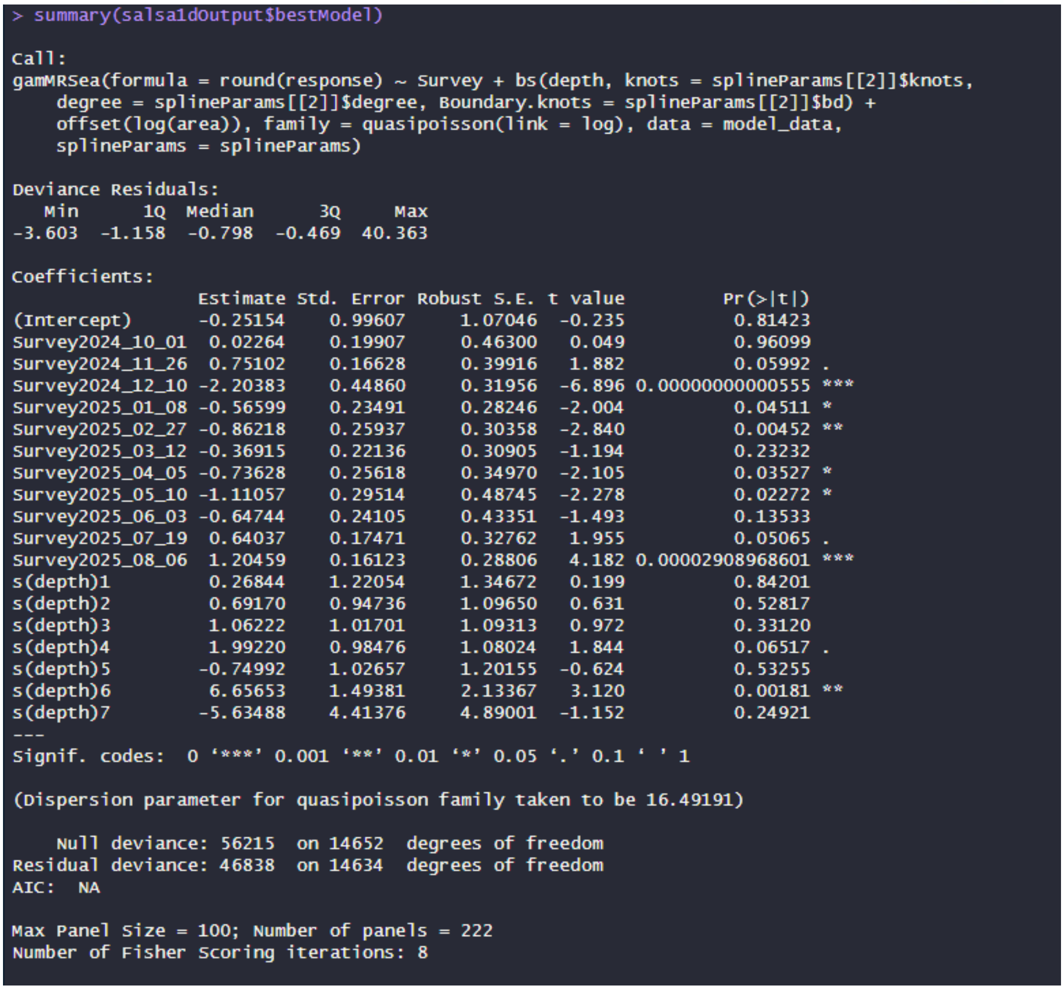
SALSA 1D

A.3.33 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.34 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). During trial runs all covariates, except depth, were removed from modelling and as a result were not carried forward for further analysis. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-50 & Figure A6-51).

```r
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     maxIterations = 10,
                     gaps = c(1))



} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d    = rep(1, 5),
                      maxKnots_1d    = rep(2, 5),
                      startKnots_1d  = rep(1, 5),
                      degree         = rep(2, 5),
                      splines        = rep("bs", 5),
                      maxIterations  = 10,
                      gaps           = rep(1, 5))


}
# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smooth.
  # SALSA selects the smoothness of each term (number and location of knots) and CV is used to choose
  # between the best smooth term, a linear term or no term at all. To not allow the removal process the
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                                        salsa1dlist=salsa1dlist,
                                        varlist=varlist,
                                        factorlist= factorlist, # including factor here doesn't change model, need to include factor in initial model
                                        datain = model_data,
                                        panelid = model_data$blockID,
                                        removal = TRUE, # setting removal = TRUE allows MRSea to choose between linear, smooth, or no model term
                                        predictionData = pred_grid)

# Warning messages:
```

Figure A6-50 Code excerpt detailing the setup process for the SALSA 1D model considered for kittiwake MRSea analysis.



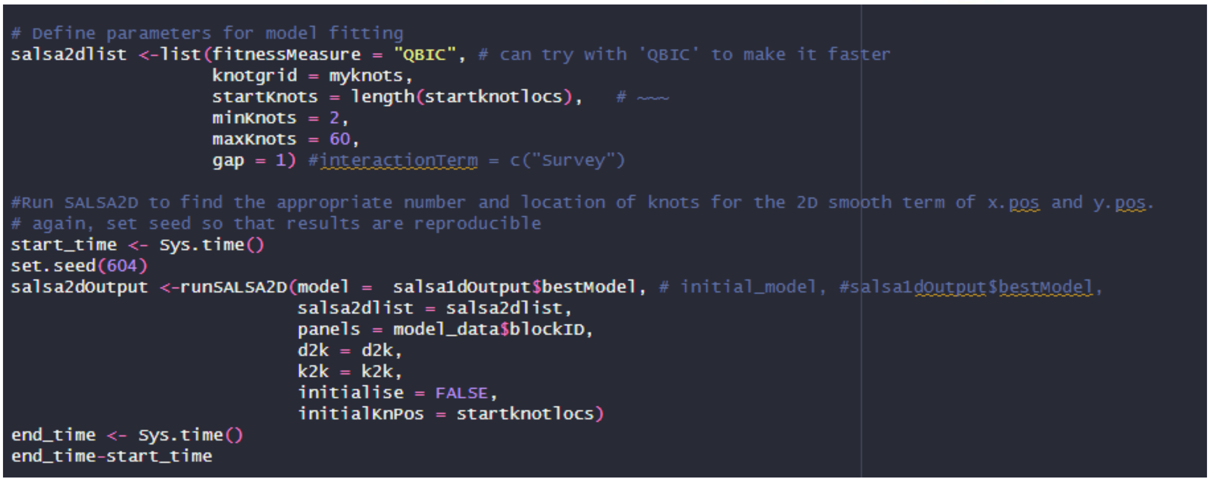
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    offset(log(area)), family = quasipoisson(link = log), data = model_data,
    splineParams = splineParams)

Deviance Residuals:
   Min       1Q  Median       3Q      Max
-3.603  -1.158  -0.798  -0.469  40.363

Coefficients:
                 Estimate Std. Error Robust S.E. t value         Pr(>|t|)
(Intercept)      -0.25154    0.99607     1.07046  -0.235          0.81423
Survey2024_10_01  0.02264    0.19907     0.46300   0.049          0.96099
Survey2024_11_26  0.75102    0.16628     0.39916   1.882          0.05992 .
Survey2024_12_10 -2.20383    0.44860     0.31956  -6.896 0.00000000000555 ***
Survey2025_01_08 -0.56599    0.23491     0.28246  -2.004          0.04511 *
Survey2025_02_27 -0.86218    0.25937     0.30358  -2.840          0.00452 **
Survey2025_03_12 -0.36915    0.22136     0.30905  -1.194          0.23232
Survey2025_04_05 -0.73628    0.25618     0.34970  -2.105          0.03527 *
Survey2025_05_10 -1.11057    0.29514     0.48745  -2.278          0.02272 *
Survey2025_06_03 -0.64744    0.24105     0.43351  -1.493          0.13533
Survey2025_07_19  0.64037    0.17471     0.32762   1.955          0.05065 .
Survey2025_08_06  1.20459    0.16123     0.28806   4.182 0.00002908968601 ***
s(depth)1         0.26844    1.22054     1.34672   0.199          0.84201
s(depth)2         0.69170    0.94736     1.09650   0.631          0.52817
s(depth)3         1.06222    1.01701     1.09313   0.972          0.33120
s(depth)4         1.99220    0.98476     1.08024   1.844          0.06517 .
s(depth)5        -0.74992    1.02657     1.20155  -0.624          0.53255
s(depth)6         6.65653    1.49381     2.13367   3.120          0.00181 **
s(depth)7        -5.63488    4.41376     4.89001  -1.152          0.24921
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 16.49191)

    Null deviance: 56215  on 14652  degrees of freedom
Residual deviance: 46838  on 14634  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 8
```

Figure A6-51 Code excerpt detailing the summary for the SALSA 1D model for kittiwake MRSea analysis.

SALSA 2D

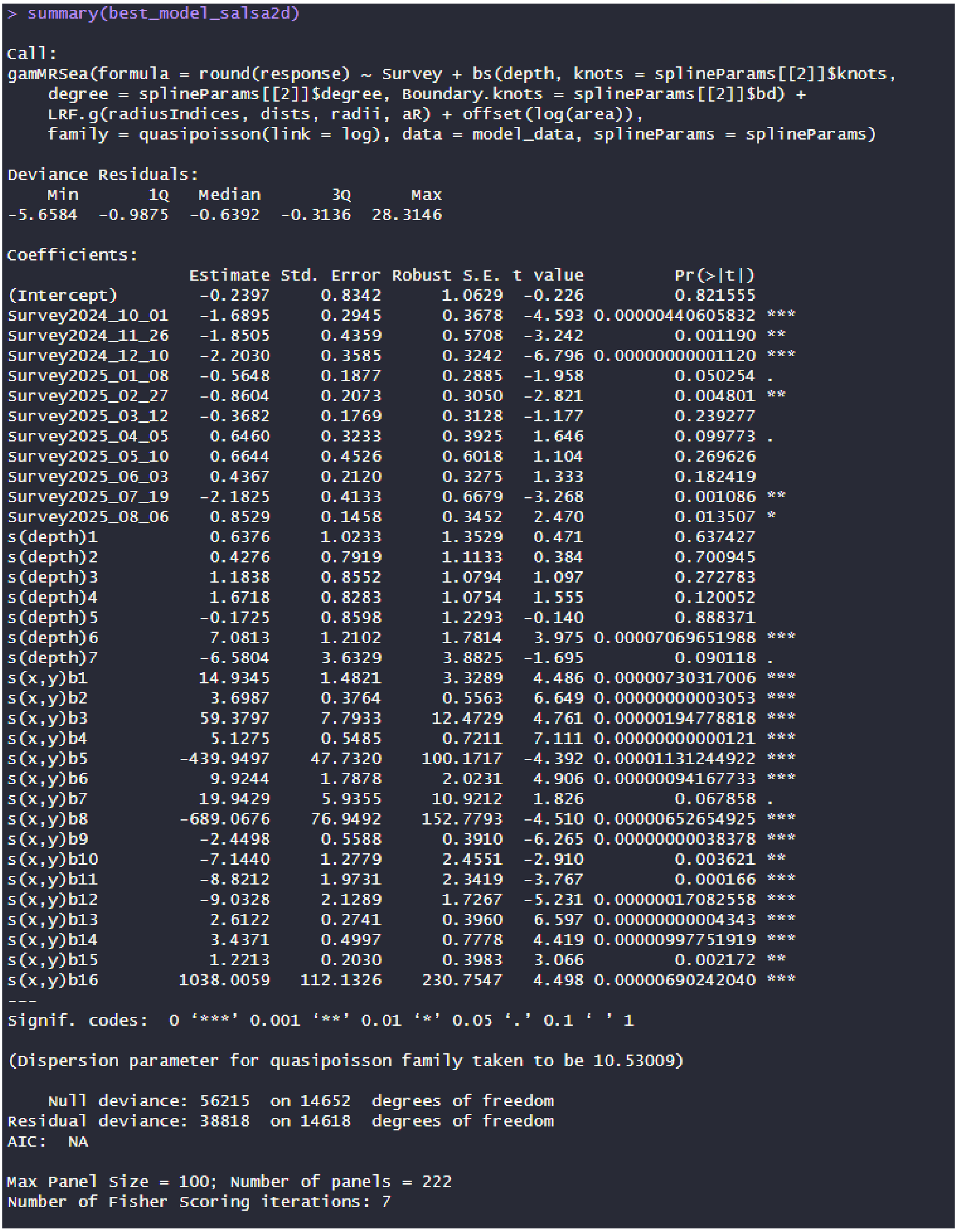
A.3.35 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.36 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-52 & Figure A6-53).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC' to make it faster
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),    # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots for the 2D smooth term of x.pos and y.pos.
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2dOutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # initial_model, #salsa1dOutput$bestModel,
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-52 Code excerpt for SALSA 2D model setup for kittiwake MRSea analysis.

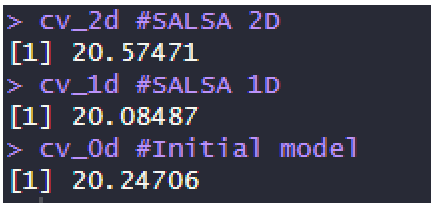


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)),
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.6584  -0.9875  -0.6392  -0.3136  28.3146

Coefficients:
                  Estimate Std. Error Robust S.E. t value         Pr(>|t|)
(Intercept)        -0.2397     0.8342      1.0629  -0.226         0.821555
Survey2024_10_01   -1.6895     0.2945      0.3678  -4.593 0.00000440605832 ***
Survey2024_11_26   -1.8505     0.4359      0.5708  -3.242         0.001190 **
Survey2024_12_10   -2.2030     0.3585      0.3242  -6.796 0.00000000001120 ***
Survey2025_01_08   -0.5648     0.1877      0.2885  -1.958         0.050254 .
Survey2025_02_27   -0.8604     0.2073      0.3050  -2.821         0.004801 **
Survey2025_03_12   -0.3682     0.1769      0.3128  -1.177         0.239277
Survey2025_04_05    0.6460     0.3233      0.3925   1.646         0.099773 .
Survey2025_05_10    0.6644     0.4526      0.6018   1.104         0.269626
Survey2025_06_03    0.4367     0.2120      0.3275   1.333         0.182419
Survey2025_07_19   -2.1825     0.4133      0.6679  -3.268         0.001086 **
Survey2025_08_06    0.8529     0.1458      0.3452   2.470         0.013507 *
s(depth)1           0.6376     1.0233      1.3529   0.471         0.637427
s(depth)2           0.4276     0.7919      1.1133   0.384         0.700945
s(depth)3           1.1838     0.8552      1.0794   1.097         0.272783
s(depth)4           1.6718     0.8283      1.0754   1.555         0.120052
s(depth)5          -0.1725     0.8598      1.2293  -0.140         0.888371
s(depth)6           7.0813     1.2102      1.7814   3.975 0.00007069651988 ***
s(depth)7          -6.5804     3.6329      3.8825  -1.695         0.090118 .
s(x,y)b1           14.9345     1.4821      3.3289   4.486 0.00000730317006 ***
s(x,y)b2            3.6987     0.3764      0.5563   6.649 0.00000000003053 ***
s(x,y)b3           59.3797     7.7933     12.4729   4.761 0.00000194778818 ***
s(x,y)b4            5.1275     0.5485      0.7211   7.111 0.00000000000121 ***
s(x,y)b5         -439.9497    47.7320    100.1717  -4.392 0.00001131244922 ***
s(x,y)b6            9.9244     1.7878      2.0231   4.906 0.00000094167733 ***
s(x,y)b7           19.9429     5.9355     10.9212   1.826         0.067858 .
s(x,y)b8         -689.0676    76.9492    152.7793  -4.510 0.00000652654925 ***
s(x,y)b9           -2.4498     0.5588      0.3910  -6.265 0.00000000038378 ***
s(x,y)b10          -7.1440     1.2779      2.4551  -2.910         0.003621 **
s(x,y)b11          -8.8212     1.9731      2.3419  -3.767         0.000166 ***
s(x,y)b12          -9.0328     2.1289      1.7267  -5.231 0.00000017082558 ***
s(x,y)b13           2.6122     0.2741      0.3960   6.597 0.00000000004343 ***
s(x,y)b14           3.4371     0.4997      0.7778   4.419 0.00000997751919 ***
s(x,y)b15           1.2213     0.2030      0.3983   3.066         0.002172 **
s(x,y)b16        1038.0059   112.1326    230.7547   4.498 0.00000690242040 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 10.53009)

    Null deviance: 56215  on 14652  degrees of freedom
Residual deviance: 38818  on 14618  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 7
```

Figure A6-53 Code excerpt detailing SALSA 2D summary for kittiwake MRSea analysis.

Cross-Validation

A.3.37 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-54).



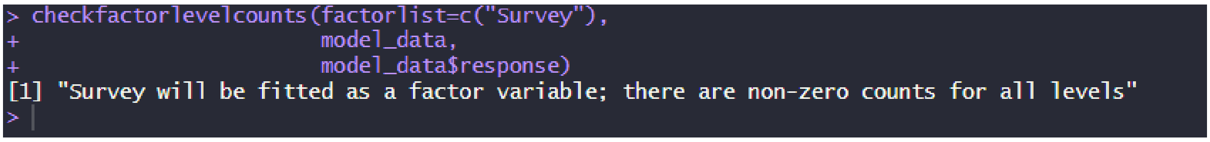
```
> cv_2d #SALSA 2D
[1] 20.57471
> cv_1d #SALSA 1D
[1] 20.08487
> cv_0d #Initial model
[1] 20.24706
```

Figure A6-54 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for kittiwake MRSea analysis.

Herring Gull

Factor Level Check

A.3.38 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, "Survey" (a proxy for month), was checked and required no further action (Figure A6-55).

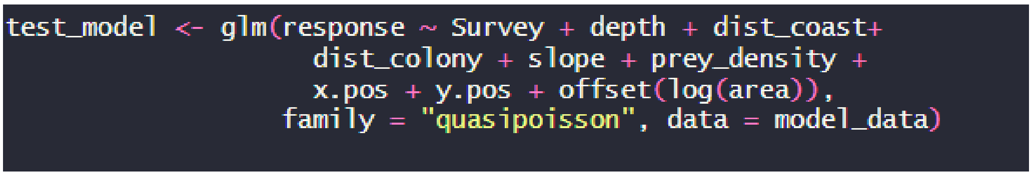


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                         model_data,
+                         model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-55 Code excerpt detailing factor level check for categorical variables considered within herring gull MRSea modelling.

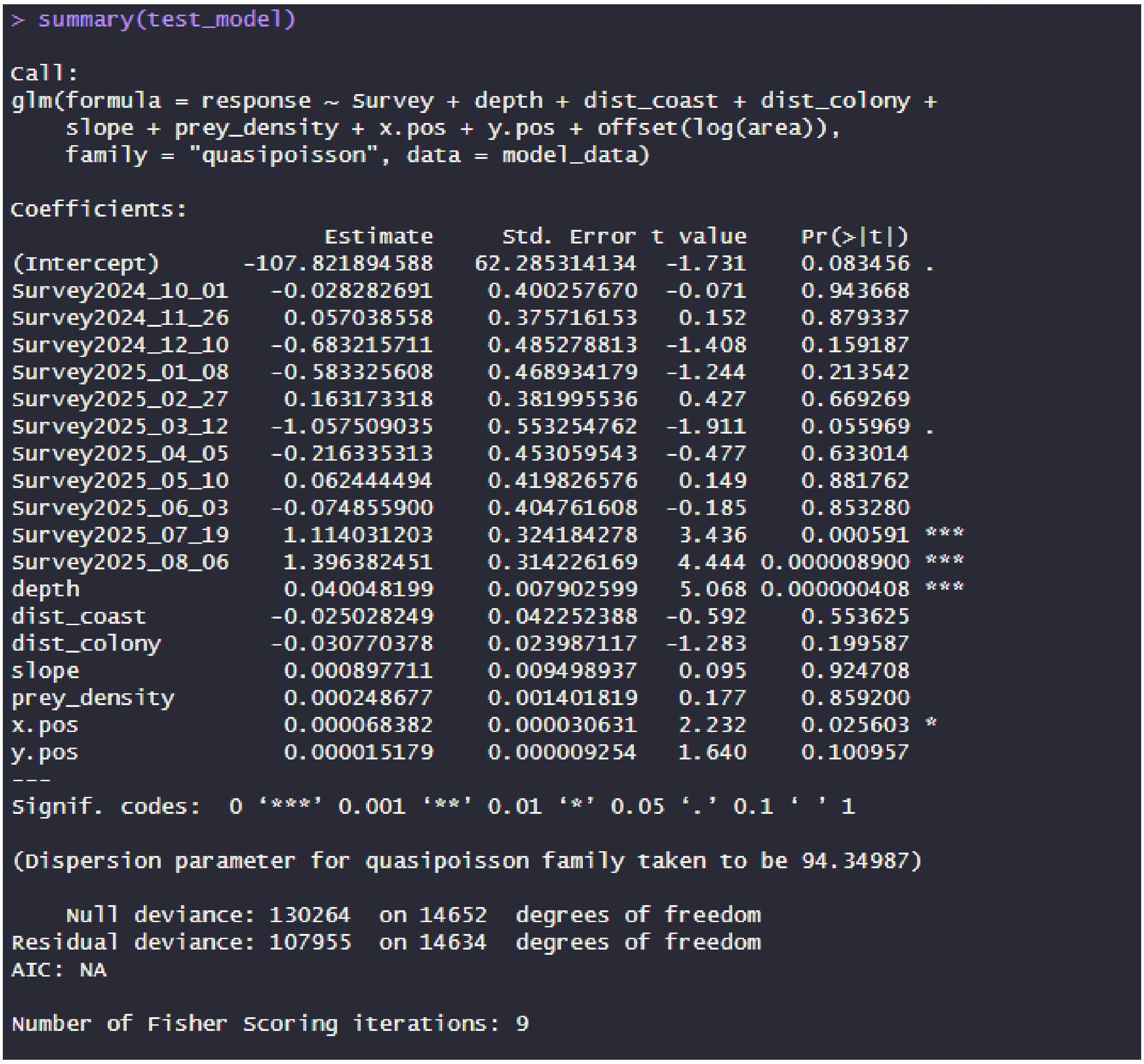
Initial GLM

A.3.39 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-56 and Figure A6-57).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                      dist_colony + slope + prey_density +
                      x.pos + y.pos + offset(log(area)),
                    family = "quasipoisson", data = model_data)
```

Figure A6-56 Code excerpt detailing the setup of the initial GLM considered within herring gull MRSea analysis.

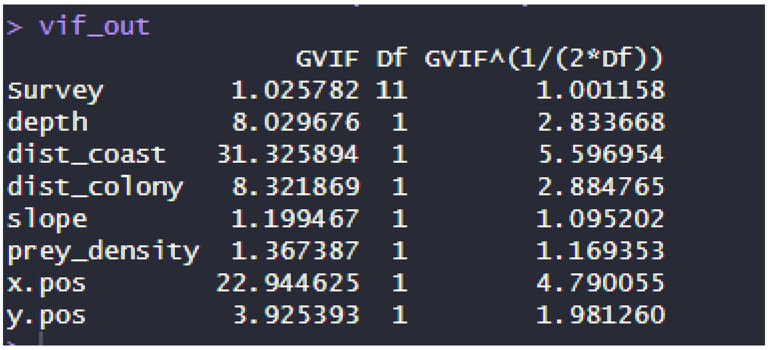


```
> summary(test_model)

Call:
glm(formula = response ~ survey + depth + dist_coast + dist_colony +
    slope + prey_density + x.pos + y.pos + offset(log(area)),
    family = "quasipoisson", data = model_data)

Coefficients:
                      Estimate     Std. Error t value    Pr(>|t|)
(Intercept)      -107.821894588  62.285314134  -1.731    0.083456 .
survey2024_10_01   -0.028282691   0.400257670  -0.071    0.943668
survey2024_11_26    0.057038558   0.375716153   0.152    0.879337
survey2024_12_10   -0.683215711   0.485278813  -1.408    0.159187
survey2025_01_08   -0.583325608   0.468934179  -1.244    0.213542
survey2025_02_27    0.163173318   0.381995536   0.427    0.669269
survey2025_03_12   -1.057509035   0.553254762  -1.911    0.055969 .
survey2025_04_05   -0.216335313   0.453059543  -0.477    0.633014
survey2025_05_10    0.062444494   0.419826576   0.149    0.881762
survey2025_06_03   -0.074855900   0.404761608  -0.185    0.853280
survey2025_07_19    1.114031203   0.324184278   3.436    0.000591 ***
survey2025_08_06    1.396382451   0.314226169   4.444 0.000008900 ***
depth               0.040048199   0.007902599   5.068 0.000000408 ***
dist_coast         -0.025028249   0.042252388  -0.592    0.553625
dist_colony        -0.030770378   0.023987117  -1.283    0.199587
slope               0.000897711   0.009498937   0.095    0.924708
prey_density        0.000248677   0.001401819   0.177    0.859200
x.pos               0.000068382   0.000030631   2.232    0.025603 *
y.pos               0.000015179   0.000009254   1.640    0.100957
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 94.34987)

    Null deviance: 130264  on 14652  degrees of freedom
Residual deviance: 107955  on 14634  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 9
```

Figure A6-57 Code excerpt detailing the summary of the initial GLM considered within herring gull MRSea analysis.

Collinearity Check

A.3.40 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-58).

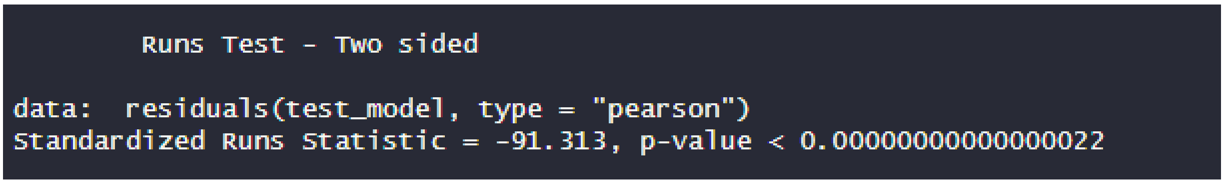


```
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
survey        1.025782 11        1.001158
depth         8.029676  1        2.833668
dist_coast   31.325894  1        5.596954
dist_colony   8.321869  1        2.884765
slope         1.199467  1        1.095202
prey_density  1.367387  1        1.169353
x.pos        22.944625  1        4.790055
y.pos         3.925393  1        1.981260
```

Figure A6-58 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within herring gull MRSea analysis.

Runs Test

A.3.41 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-59).



```
        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -91.313, p-value < 0.0000000000000022
```

Figure A6-59 Code excerpt detailing runs test carried out on the initial GLM considered within herring gull MRSea analysis.

A.3.42 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-60).

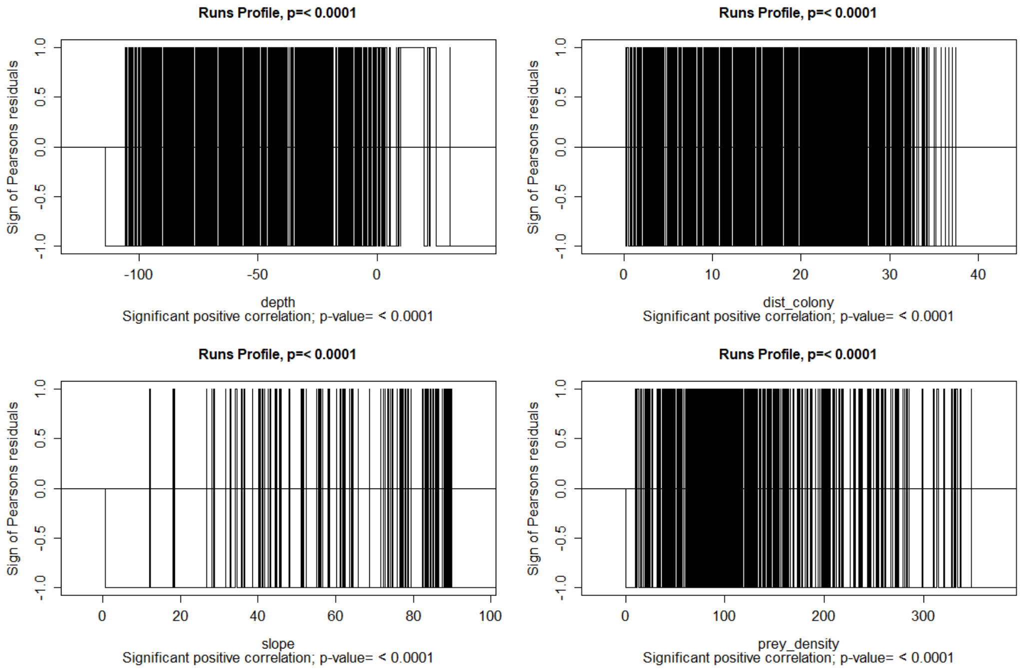


Figure A6-60 Runs profiles for explanatory variables considered within the initial GLM considered for herring gull MRSea analysis.

Autocorrelation

A.3.43 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-61).

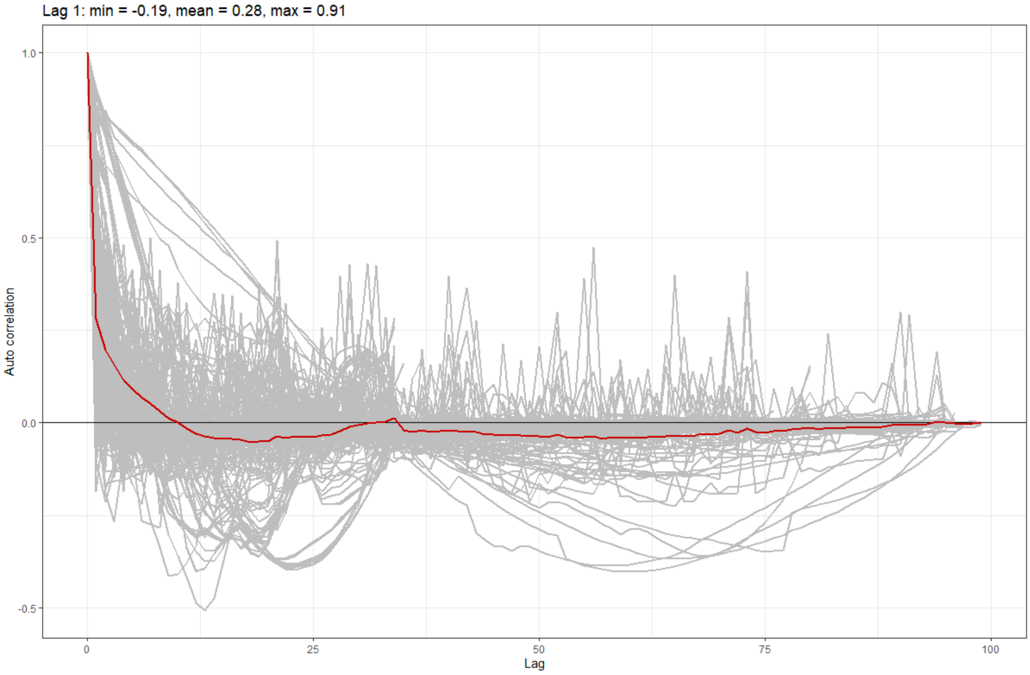


Figure A6-61 ACF plot for the initial GLM considered for herring gull MRSea analysis.

Cumulative Residuals

A.3.44 Cumulative residuals for explanatory variables were plotted (Figure A6-62, Figure A6-63, Figure A6-64 & Figure A6-65). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

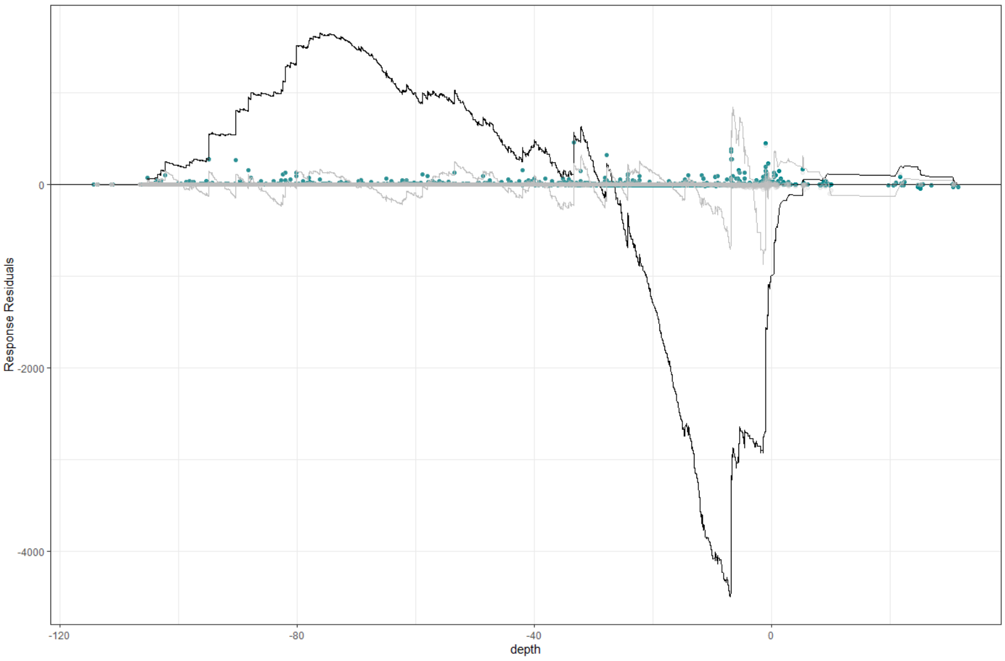


Figure A6-62 Cumulative residuals of depth plotted for the initial GLM considered for herring gull MRSea analysis.

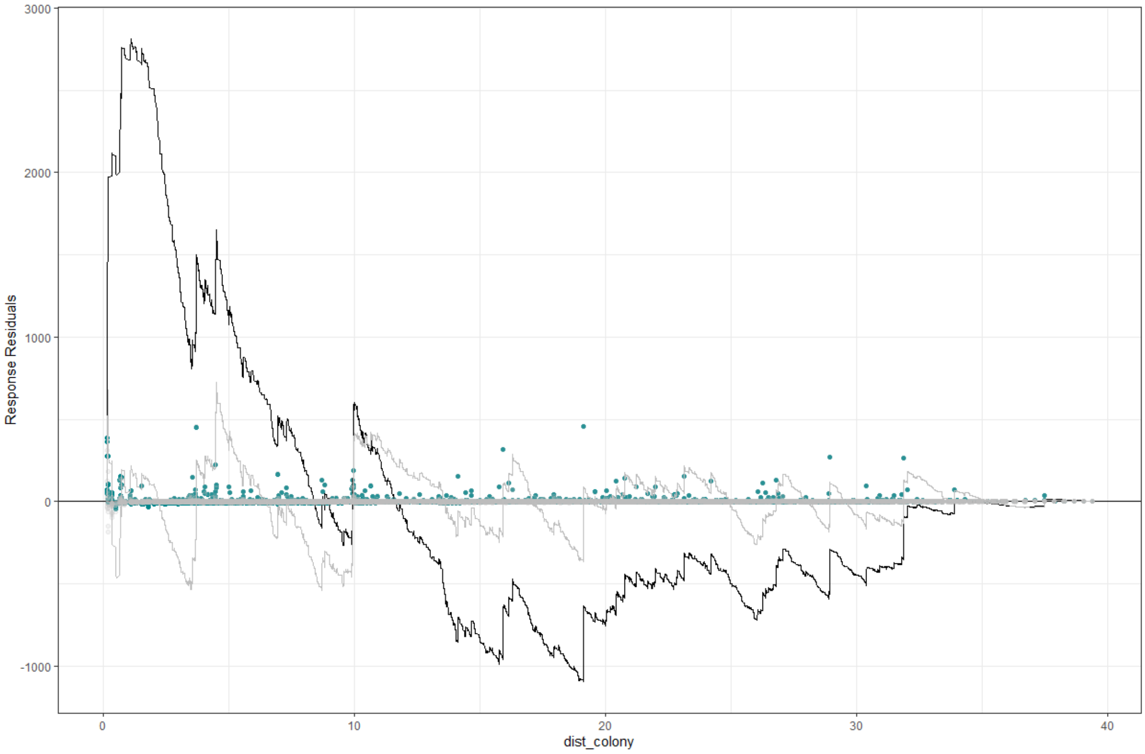


Figure A6-63 Cumulative residuals of distance to colony plotted for the initial GLM considered for herring gull MRSea analysis.

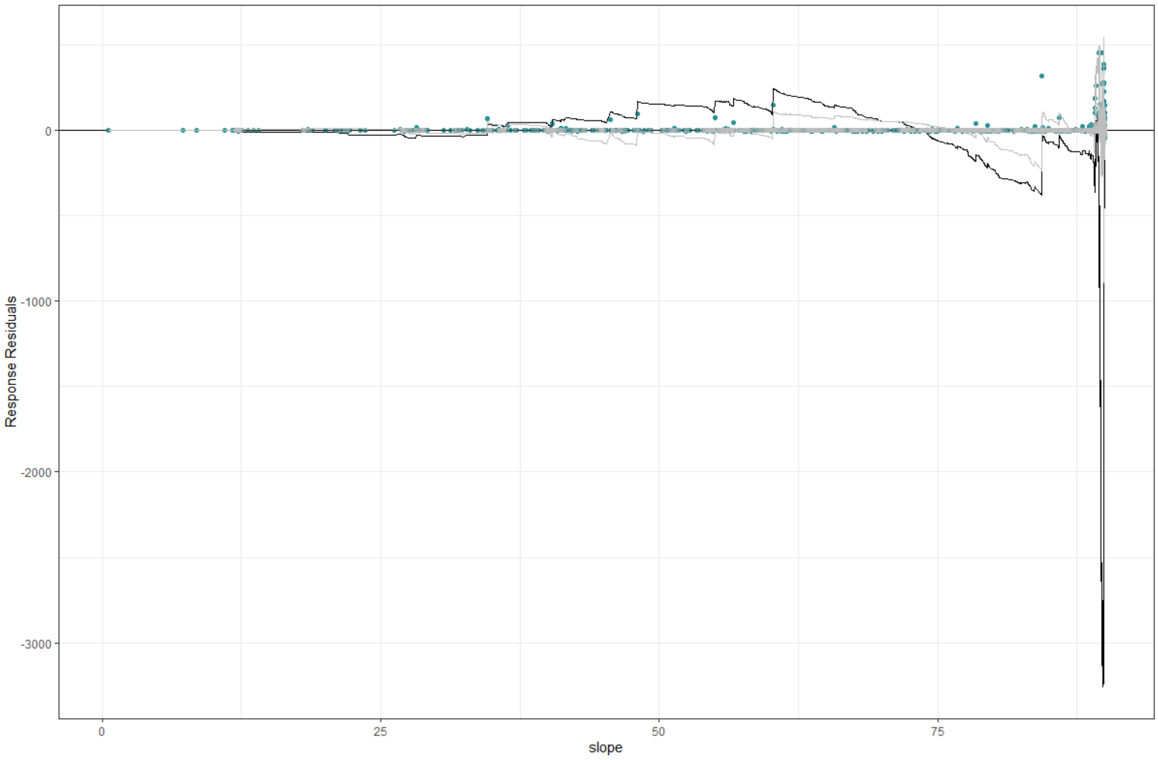


Figure A6-64 Cumulative residuals of seabed slope plotted for the initial GLM considered for herring gull MRSea analysis.

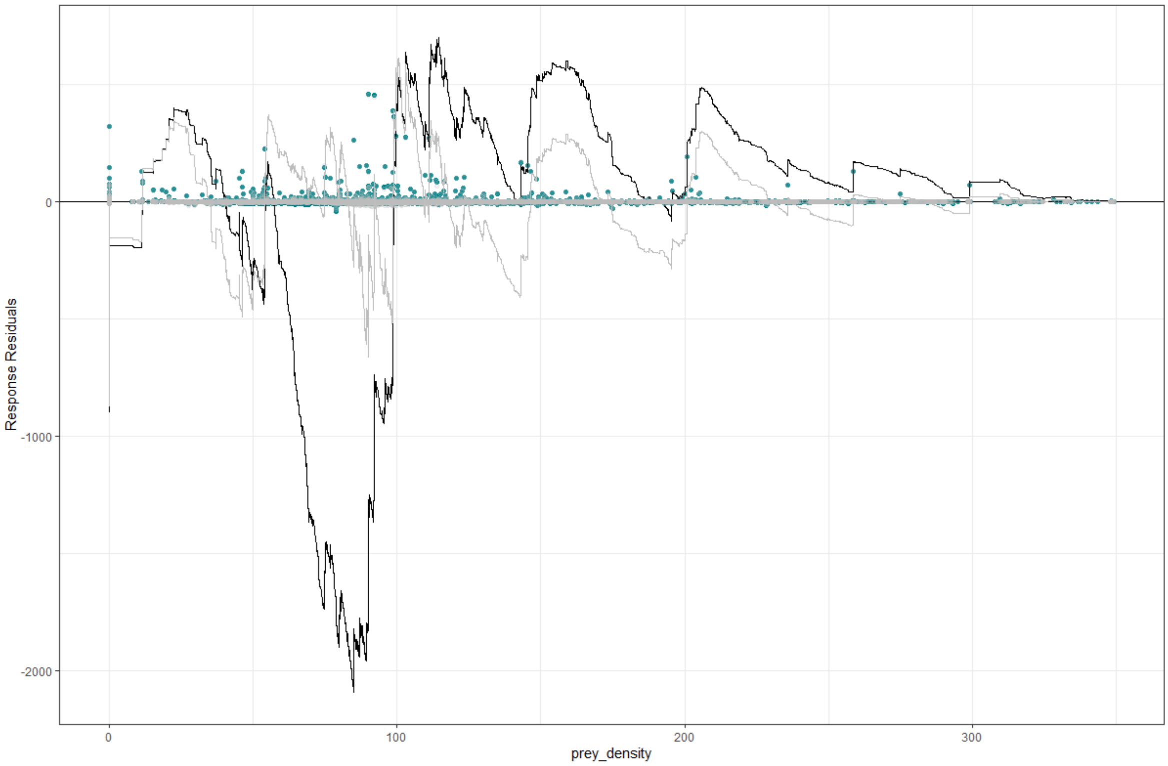
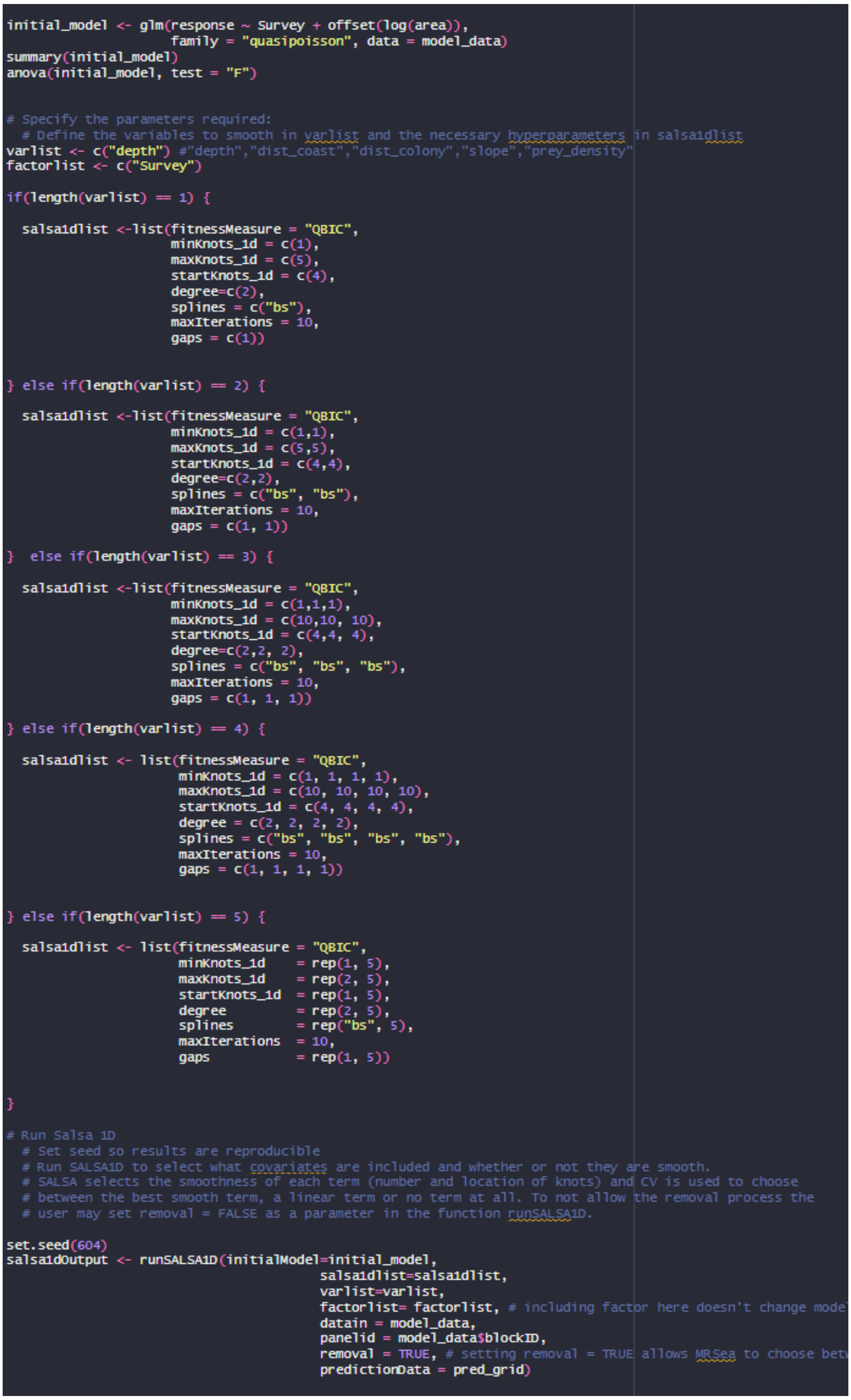


Figure A6-65 Cumulative residuals of prey density plotted for the initial GLM considered for herring gull MRSea analysis.

SALSA 1D

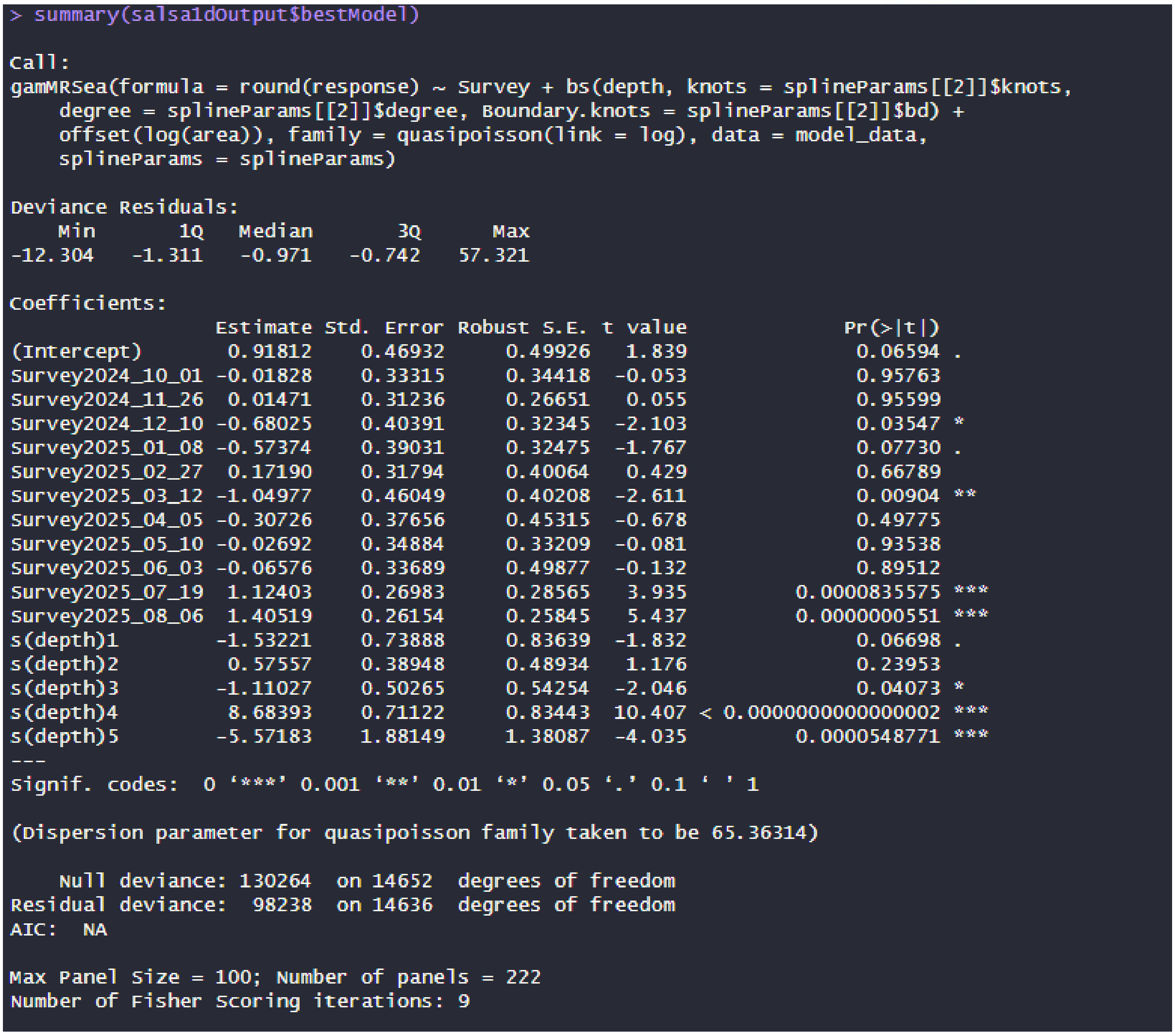
A.3.45 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.46 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). During trial runs all covariates, except depth, were removed from modelling and as a result were not carried forward for further analysis. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-66 & Figure A6-67).



```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d     = rep(1, 5),
                      maxKnots_1d     = rep(2, 5),
                      startKnots_1d   = rep(1, 5),
                      degree          = rep(2, 5),
                      splines         = rep("bs", 5),
                      maxIterations   = 10,
                      gaps            = rep(1, 5))


}

# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smooth.
  # SALSA selects the smoothness of each term (number and location of knots) and CV is used to choose
  # between the best smooth term, a linear term or no term at all. To not allow the removal process the
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                            salsa1dlist=salsa1dlist,
                            varlist=varlist,
                            factorlist= factorlist, # including factor here doesn't change mode
                            datain = model_data,
                            panelid = model_data$blockID,
                            removal = TRUE, # setting removal = TRUE allows MRSea to choose bet
                            predictionData = pred_grid)
```

Figure A6-66 Code excerpt detailing the setup process for the SALSA 1D model considered for herring gull MRSea analysis.



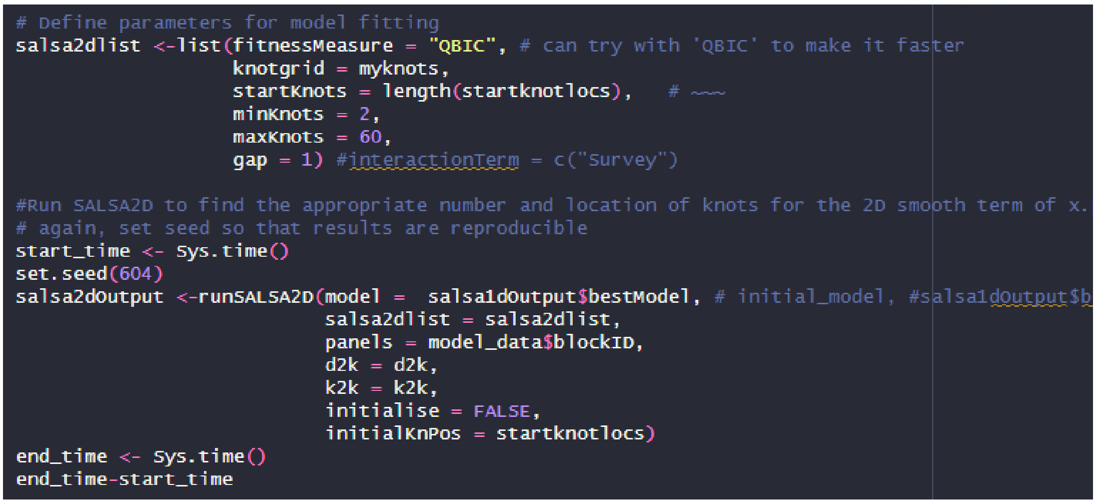
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    offset(log(area)), family = quasipoisson(link = log), data = model_data,
    splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-12.304   -1.311   -0.971   -0.742   57.321

Coefficients:
                 Estimate Std. Error Robust S.E. t value             Pr(>|t|)
(Intercept)       0.91812    0.46932     0.49926   1.839              0.06594 .
Survey2024_10_01 -0.01828    0.33315     0.34418  -0.053              0.95763
Survey2024_11_26  0.01471    0.31236     0.26651   0.055              0.95599
Survey2024_12_10 -0.68025    0.40391     0.32345  -2.103              0.03547 *
Survey2025_01_08 -0.57374    0.39031     0.32475  -1.767              0.07730 .
Survey2025_02_27  0.17190    0.31794     0.40064   0.429              0.66789
Survey2025_03_12 -1.04977    0.46049     0.40208  -2.611              0.00904 **
Survey2025_04_05 -0.30726    0.37656     0.45315  -0.678              0.49775
Survey2025_05_10 -0.02692    0.34884     0.33209  -0.081              0.93538
Survey2025_06_03 -0.06576    0.33689     0.49877  -0.132              0.89512
Survey2025_07_19  1.12403    0.26983     0.28565   3.935         0.0000835575 ***
Survey2025_08_06  1.40519    0.26154     0.25845   5.437         0.0000000551 ***
s(depth)1        -1.53221    0.73888     0.83639  -1.832              0.06698 .
s(depth)2         0.57557    0.38948     0.48934   1.176              0.23953
s(depth)3        -1.11027    0.50265     0.54254  -2.046              0.04073 *
s(depth)4         8.68393    0.71122     0.83443  10.407 < 0.0000000000000002 ***
s(depth)5        -5.57183    1.88149     1.38087  -4.035         0.0000548771 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 65.36314)

    Null deviance: 130264  on 14652  degrees of freedom
Residual deviance:  98238  on 14636  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 9
```

Figure A6-67 Code excerpt detailing the summary for the SALSA 1D model for herring gull MRSea analysis.

SALSA 2D

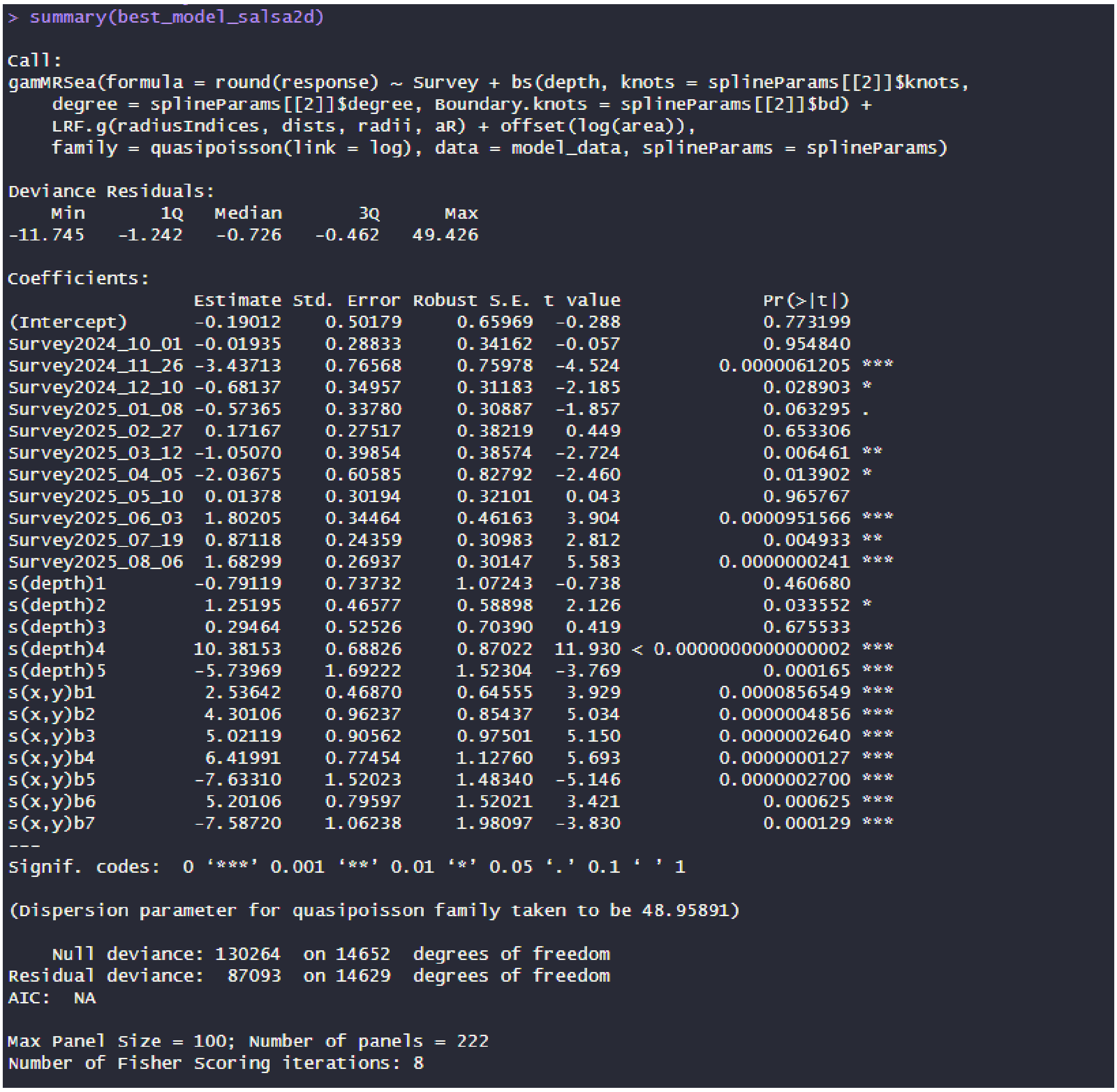
A.3.47 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.48 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-68 & Figure A6-69).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC' to make it faster
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),    # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots for the 2D smooth term of x.
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2dOutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # initial_model, #salsa1dOutput$b
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-68 Code excerpt for SALSA 2D model setup for herring gull MRSea analysis.

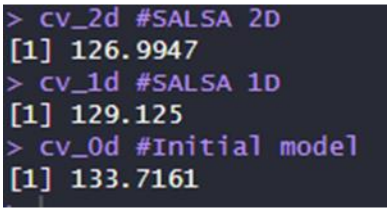


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, 
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + 
    LRF.g(radiusIndices, dists, radii, aR) + offset(log(area)), 
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-11.745   -1.242   -0.726   -0.462   49.426  

Coefficients:
                 Estimate Std. Error Robust S.E. t value             Pr(>|t|)    
(Intercept)      -0.19012    0.50179     0.65969  -0.288             0.773199    
Survey2024_10_01 -0.01935    0.28833     0.34162  -0.057             0.954840    
Survey2024_11_26 -3.43713    0.76568     0.75978  -4.524         0.0000061205 ***
Survey2024_12_10 -0.68137    0.34957     0.31183  -2.185             0.028903 *  
Survey2025_01_08 -0.57365    0.33780     0.30887  -1.857             0.063295 .  
Survey2025_02_27  0.17167    0.27517     0.38219   0.449             0.653306    
Survey2025_03_12 -1.05070    0.39854     0.38574  -2.724             0.006461 ** 
Survey2025_04_05 -2.03675    0.60585     0.82792  -2.460             0.013902 *  
Survey2025_05_10  0.01378    0.30194     0.32101   0.043             0.965767    
Survey2025_06_03  1.80205    0.34464     0.46163   3.904         0.0000951566 ***
Survey2025_07_19  0.87118    0.24359     0.30983   2.812             0.004933 ** 
Survey2025_08_06  1.68299    0.26937     0.30147   5.583         0.0000000241 ***
s(depth)1        -0.79119    0.73732     1.07243  -0.738             0.460680    
s(depth)2         1.25195    0.46577     0.58898   2.126             0.033552 *  
s(depth)3         0.29464    0.52526     0.70390   0.419             0.675533    
s(depth)4        10.38153    0.68826     0.87022  11.930 < 0.0000000000000002 ***
s(depth)5        -5.73969    1.69222     1.52304  -3.769             0.000165 ***
s(x,y)b1          2.53642    0.46870     0.64555   3.929         0.0000856549 ***
s(x,y)b2          4.30106    0.96237     0.85437   5.034         0.0000004856 ***
s(x,y)b3          5.02119    0.90562     0.97501   5.150         0.0000002640 ***
s(x,y)b4          6.41991    0.77454     1.12760   5.693         0.0000000127 ***
s(x,y)b5         -7.63310    1.52023     1.48340  -5.146         0.0000002700 ***
s(x,y)b6          5.20106    0.79597     1.52021   3.421             0.000625 ***
s(x,y)b7         -7.58720    1.06238     1.98097  -3.830             0.000129 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 48.95891)

    Null deviance: 130264  on 14652  degrees of freedom
Residual deviance:  87093  on 14629  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 8
```

Figure A6-69 Code excerpt detailing SALSA 2D summary for herring gull MRSea analysis.

Cross-Validation

A.3.49 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. The 2D model exhibited a lower CV score than the 1D model, indicating a better fit, it was therefore retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-70).



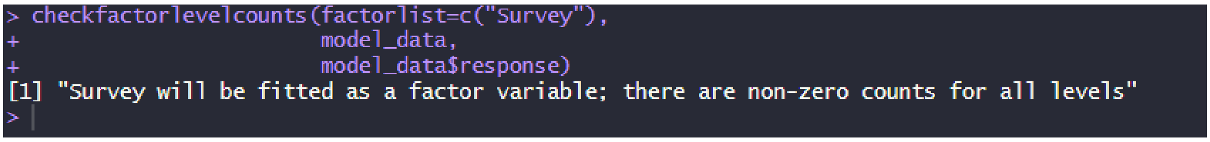
```
> cv_2d #SALSA 2D
[1] 126.9947
> cv_1d #SALSA 1D
[1] 129.125
> cv_0d #Initial model
[1] 133.7161
```

Figure A6-70 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for herring gull MRSea analysis.

Great Black-backed Gull

Factor Level Check

A.3.50 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-71).

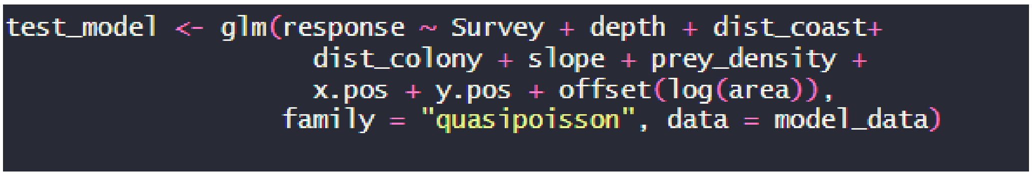


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                          model_data,
+                          model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> |
```

Figure A6-71 Code excerpt detailing factor level check for categorical variables considered within great black-backed gull MRSea analysis.

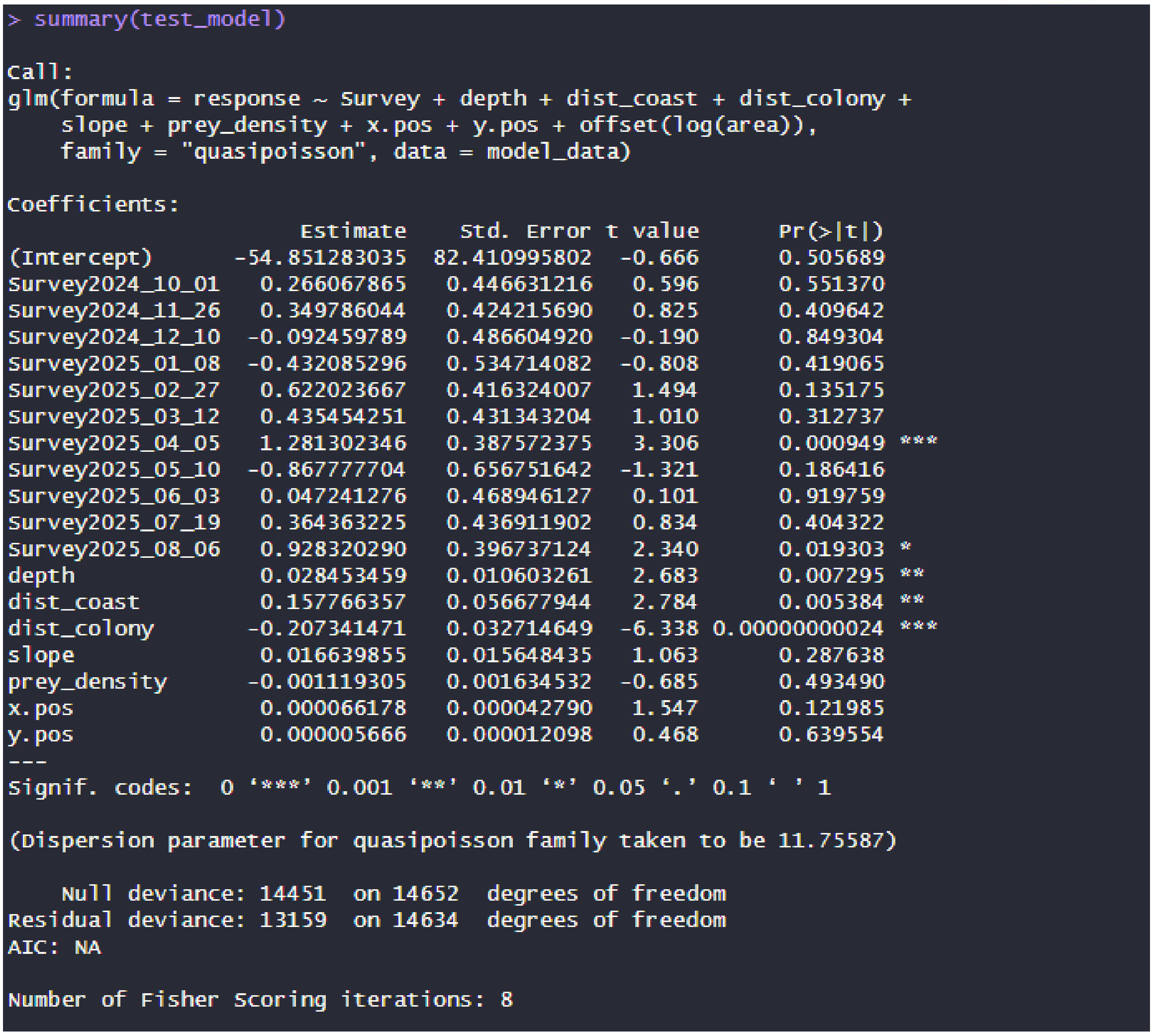
Initial GLM

A.3.51 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-72 and Figure A6-73).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-72 Code excerpt detailing the setup of the initial GLM considered within great black-backed gull MRSea analysis.

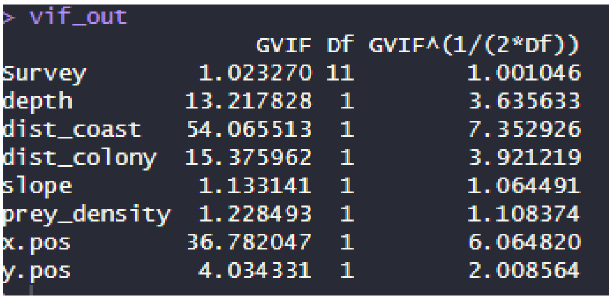


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_coast + dist_colony +
    slope + prey_density + x.pos + y.pos + offset(log(area)),
    family = "quasipoisson", data = model_data)

Coefficients:
                      Estimate    Std. Error t value      Pr(>|t|)
(Intercept)       -54.851283035  82.410995802  -0.666      0.505689
Survey2024_10_01    0.266067865   0.446631216   0.596      0.551370
Survey2024_11_26    0.349786044   0.424215690   0.825      0.409642
Survey2024_12_10   -0.092459789   0.486604920  -0.190      0.849304
Survey2025_01_08   -0.432085296   0.534714082  -0.808      0.419065
Survey2025_02_27    0.622023667   0.416324007   1.494      0.135175
Survey2025_03_12    0.435454251   0.431343204   1.010      0.312737
Survey2025_04_05    1.281302346   0.387572375   3.306      0.000949 ***
Survey2025_05_10   -0.867777704   0.656751642  -1.321      0.186416
Survey2025_06_03    0.047241276   0.468946127   0.101      0.919759
Survey2025_07_19    0.364363225   0.436911902   0.834      0.404322
Survey2025_08_06    0.928320290   0.396737124   2.340      0.019303 *
depth               0.028453459   0.010603261   2.683      0.007295 **
dist_coast          0.157766357   0.056677944   2.784      0.005384 **
dist_colony        -0.207341471   0.032714649  -6.338 0.00000000024 ***
slope               0.016639855   0.015648435   1.063      0.287638
prey_density       -0.001119305   0.001634532  -0.685      0.493490
x.pos               0.000066178   0.000042790   1.547      0.121985
y.pos               0.000005666   0.000012098   0.468      0.639554
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 11.75587)

    Null deviance: 14451  on 14652  degrees of freedom
Residual deviance: 13159  on 14634  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 8
```

Figure A6-73 Code excerpt detailing the summary of the initial GLM considered within great black-backed gull MRSea analysis.

Collinearity Check

A.3.52 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-74).

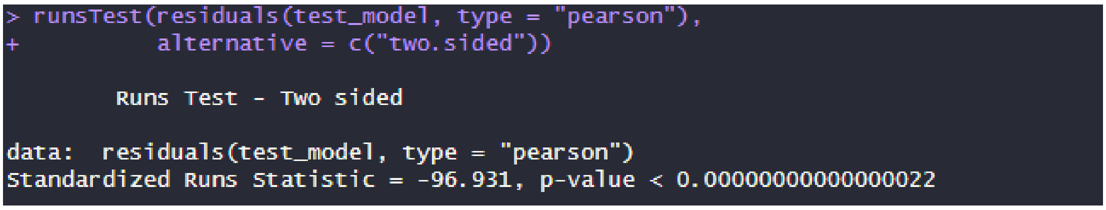


```
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
Survey        1.023270 11        1.001046
depth        13.217828  1        3.635633
dist_coast   54.065513  1        7.352926
dist_colony  15.375962  1        3.921219
slope         1.133141  1        1.064491
prey_density  1.228493  1        1.108374
x.pos        36.782047  1        6.064820
y.pos         4.034331  1        2.008564
```

Figure A6-74 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within great black-backed gull MRSea analysis.

Runs Test

A.3.53 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-75).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -96.931, p-value < 0.00000000000000022
```

Figure A6-75 Code excerpt detailing runs test carried out on the initial GLM considered within great black-backed gull MRSea analysis.

A.3.54 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-76).

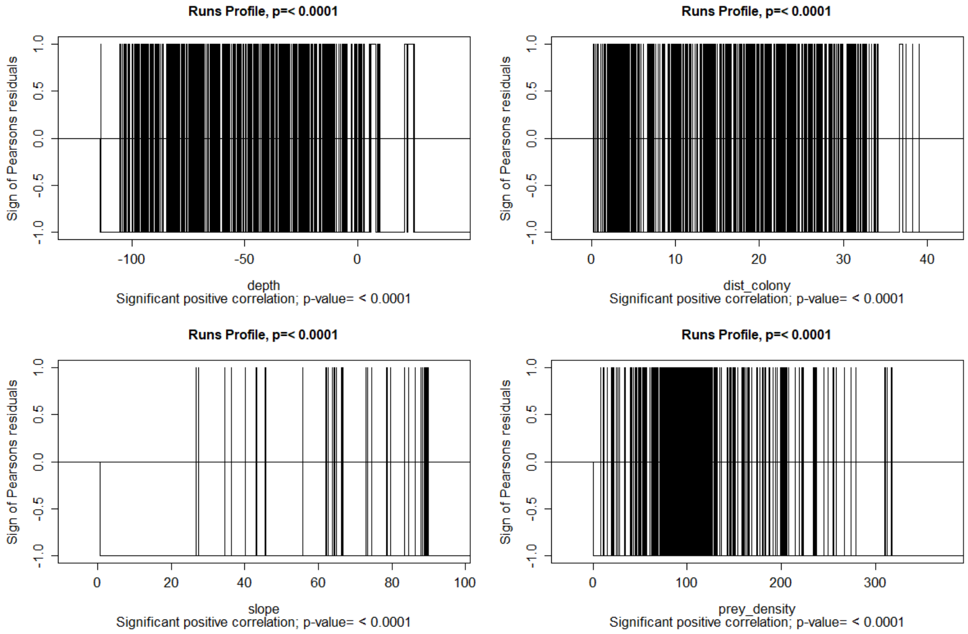


Figure A6-76 Runs profiles for explanatory variables considered within the initial GLM considered for great black-backed gull MRSea analysis.

Autocorrelation

A.3.55 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-77).

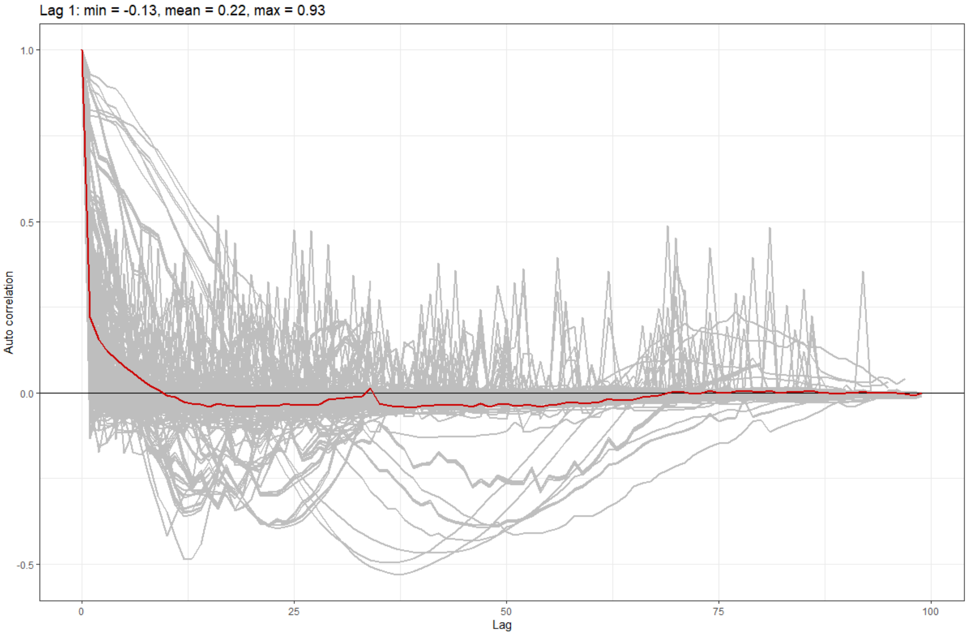


Figure A6-77 ACF plot for the initial GLM considered for great black-backed gull MRSea analysis.

Cumulative Residuals

A.3.56 Cumulative residuals for explanatory variables were plotted (Figure A6-78, Figure A6-79, Figure A6-80 & Figure A6-81). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

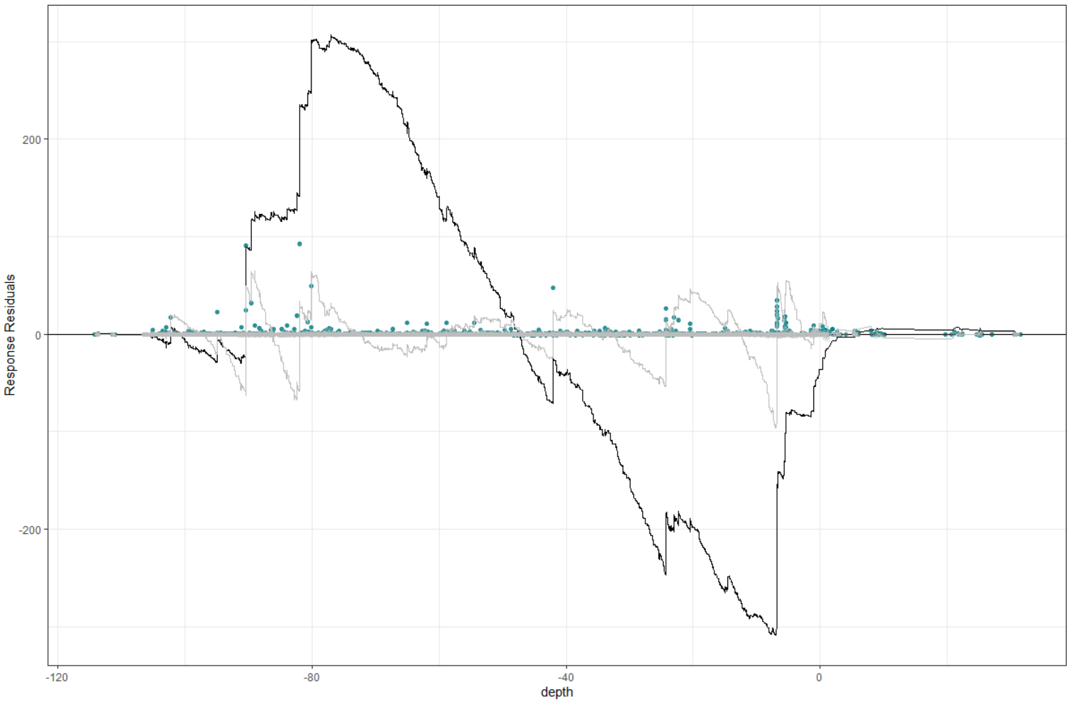


Figure A6-78 Cumulative residuals of depth plotted for the initial GLM considered for great black-backed gull MRSea analysis.

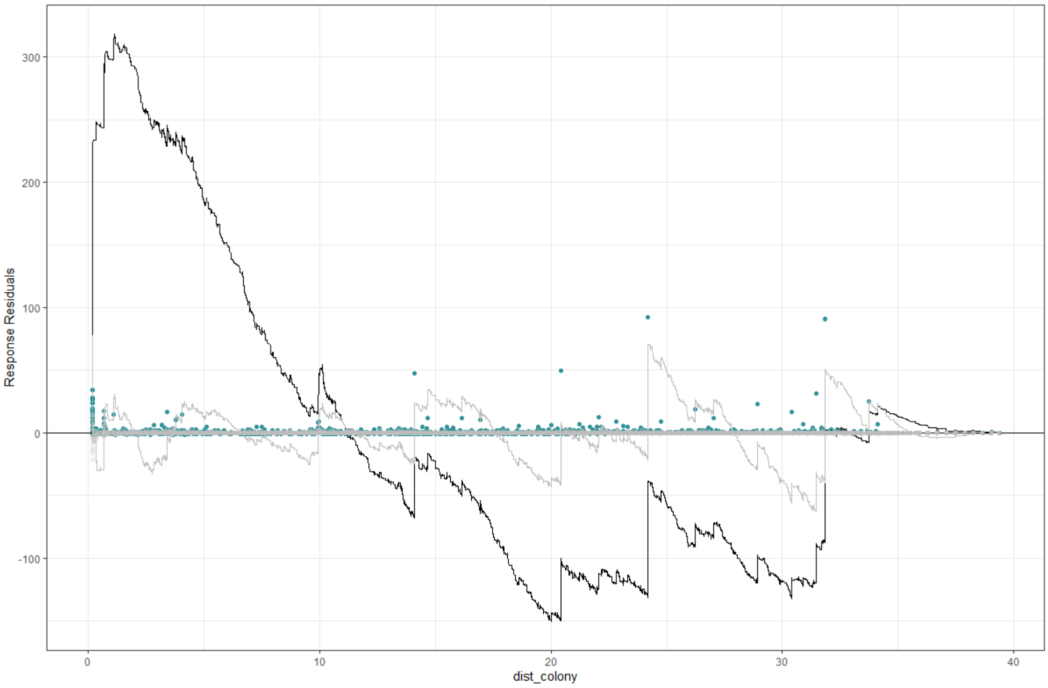


Figure A6-79 Cumulative residuals of distance to colony plotted for the initial GLM considered for great black-backed gull MRSea analysis.

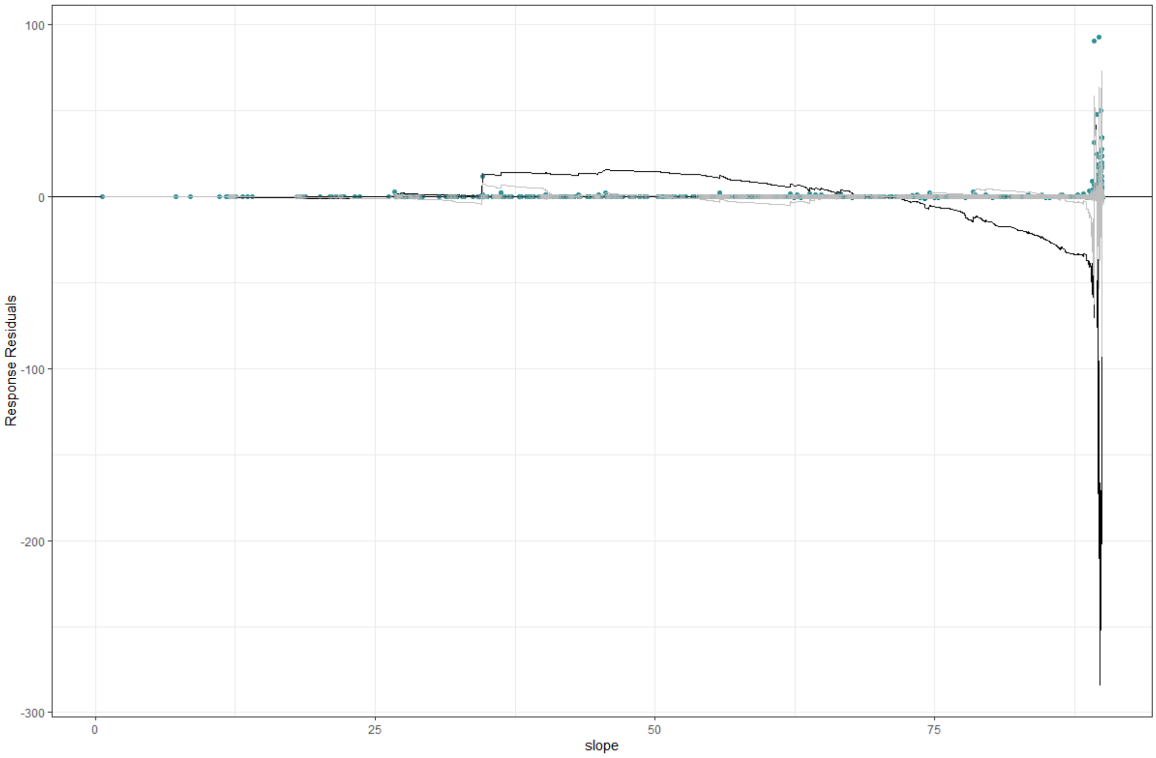


Figure A6-80 Cumulative residuals of seabed slope plotted for the initial GLM considered for great black-backed gull MRSea analysis.

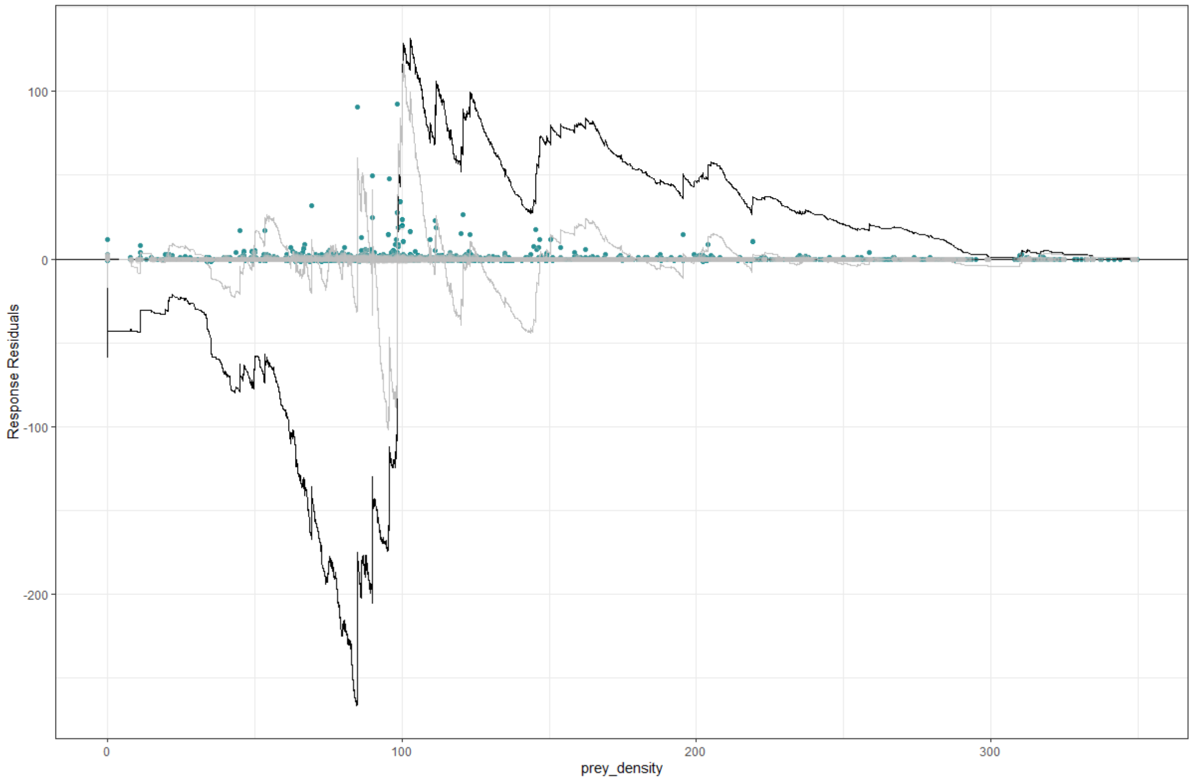
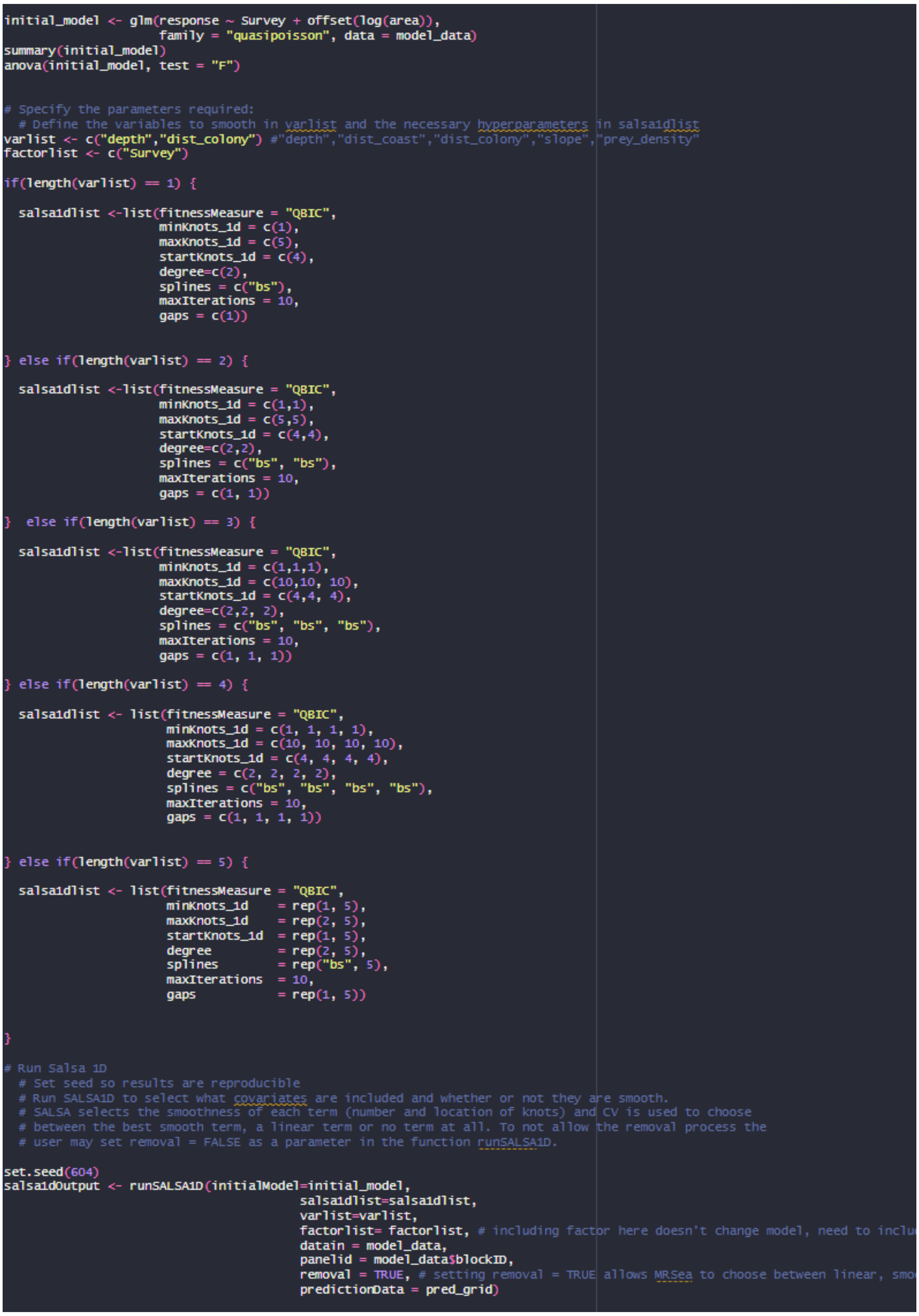


Figure A6-81 Cumulative residuals of prey density plotted for the initial GLM considered for great black-backed gull MRSea analysis.

SALSA 1D

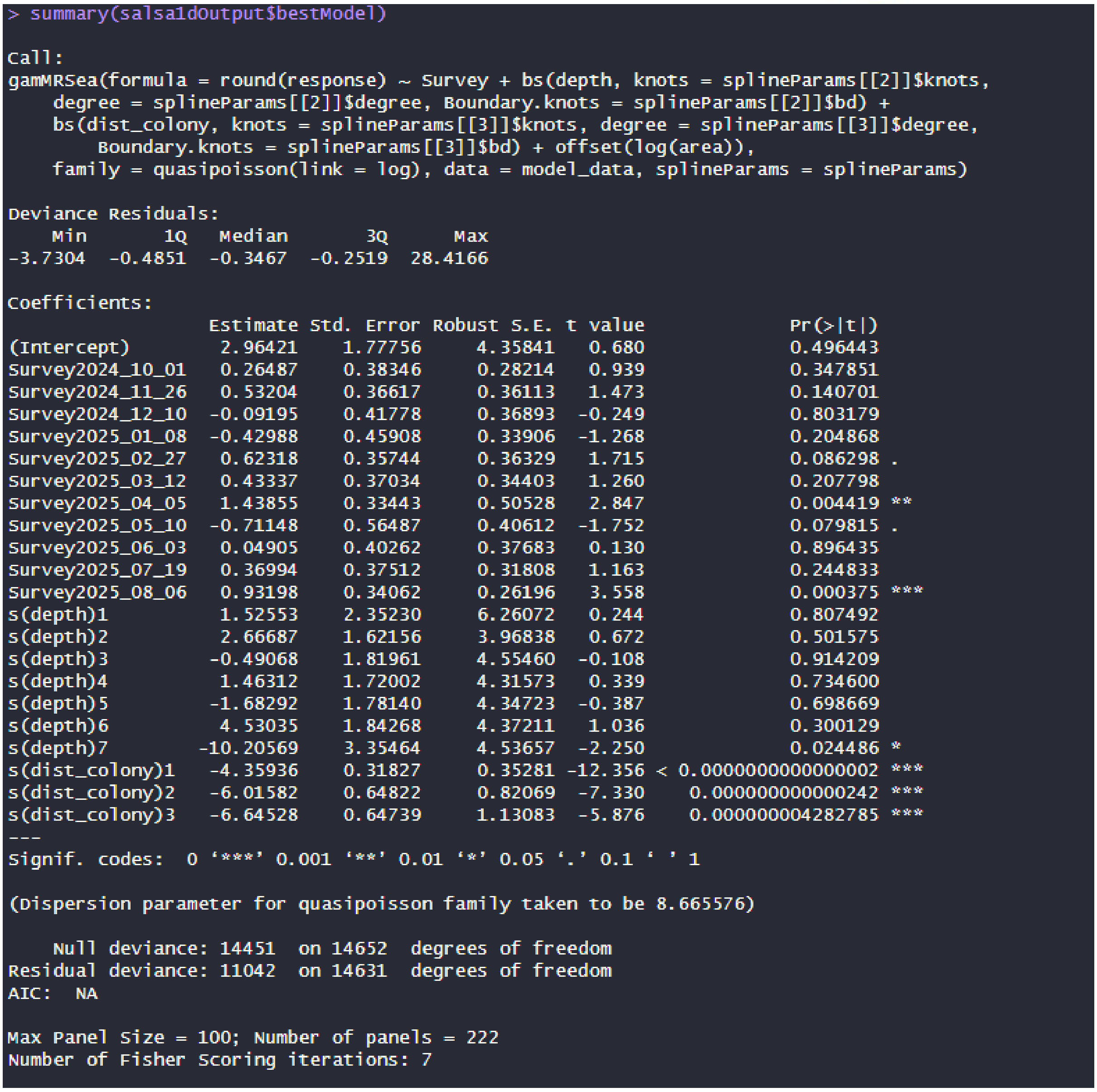
A.3.57 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.58 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). During trial runs all covariates, except depth and distance to coast, were removed from modelling and as a result were not carried forward for further analysis. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-82 & Figure A6-83).



```r
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth","dist_colony") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d    = rep(1, 5),
                      maxKnots_1d    = rep(2, 5),
                      startKnots_1d  = rep(1, 5),
                      degree         = rep(2, 5),
                      splines        = rep("bs", 5),
                      maxIterations  = 10,
                      gaps           = rep(1, 5))


}
# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smooth.
  # SALSA selects the smoothness of each term (number and location of knots) and CV is used to choose
  # between the best smooth term, a linear term or no term at all. To not allow the removal process the
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                            salsa1dlist=salsa1dlist,
                            varlist=varlist,
                            factorlist= factorlist, # including factor here doesn't change model, need to inclu
                            datain = model_data,
                            panelid = model_data$blockID,
                            removal = TRUE, # setting removal = TRUE allows MRSea to choose between linear, smo
                            predictionData = pred_grid)
```

Figure A6-82 Code excerpt detailing the setup process for the SALSA 1D model considered for great black-backed gull MRSea analysis.



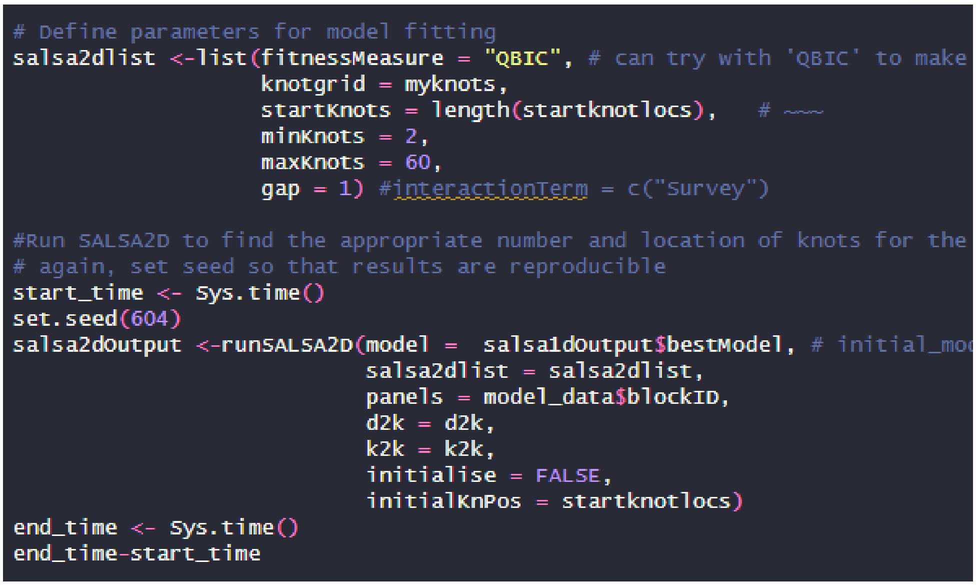
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots, 
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) + 
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, 
        Boundary.knots = splineParams[[3]]$bd) + offset(log(area)), 
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-3.7304  -0.4851  -0.3467  -0.2519  28.4166  

Coefficients:
                  Estimate Std. Error Robust S.E. t value             Pr(>|t|)    
(Intercept)        2.96421    1.77756     4.35841   0.680             0.496443    
Survey2024_10_01   0.26487    0.38346     0.28214   0.939             0.347851    
Survey2024_11_26   0.53204    0.36617     0.36113   1.473             0.140701    
Survey2024_12_10  -0.09195    0.41778     0.36893  -0.249             0.803179    
Survey2025_01_08  -0.42988    0.45908     0.33906  -1.268             0.204868    
Survey2025_02_27   0.62318    0.35744     0.36329   1.715             0.086298 .  
Survey2025_03_12   0.43337    0.37034     0.34403   1.260             0.207798    
Survey2025_04_05   1.43855    0.33443     0.50528   2.847             0.004419 ** 
Survey2025_05_10  -0.71148    0.56487     0.40612  -1.752             0.079815 .  
Survey2025_06_03   0.04905    0.40262     0.37683   0.130             0.896435    
Survey2025_07_19   0.36994    0.37512     0.31808   1.163             0.244833    
Survey2025_08_06   0.93198    0.34062     0.26196   3.558             0.000375 ***
s(depth)1          1.52553    2.35230     6.26072   0.244             0.807492    
s(depth)2          2.66687    1.62156     3.96838   0.672             0.501575    
s(depth)3         -0.49068    1.81961     4.55460  -0.108             0.914209    
s(depth)4          1.46312    1.72002     4.31573   0.339             0.734600    
s(depth)5         -1.68292    1.78140     4.34723  -0.387             0.698669    
s(depth)6          4.53035    1.84268     4.37211   1.036             0.300129    
s(depth)7        -10.20569    3.35464     4.53657  -2.250             0.024486 *  
s(dist_colony)1   -4.35936    0.31827     0.35281 -12.356 < 0.0000000000000002 ***
s(dist_colony)2   -6.01582    0.64822     0.82069  -7.330    0.000000000000242 ***
s(dist_colony)3   -6.64528    0.64739     1.13083  -5.876    0.000000004282785 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 8.665576)

    Null deviance: 14451  on 14652  degrees of freedom
Residual deviance: 11042  on 14631  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 7
```

Figure A6-83 Code excerpt detailing the summary for the SALSA 1D model for great black-backed gull MRSea analysis.

SALSA 2D

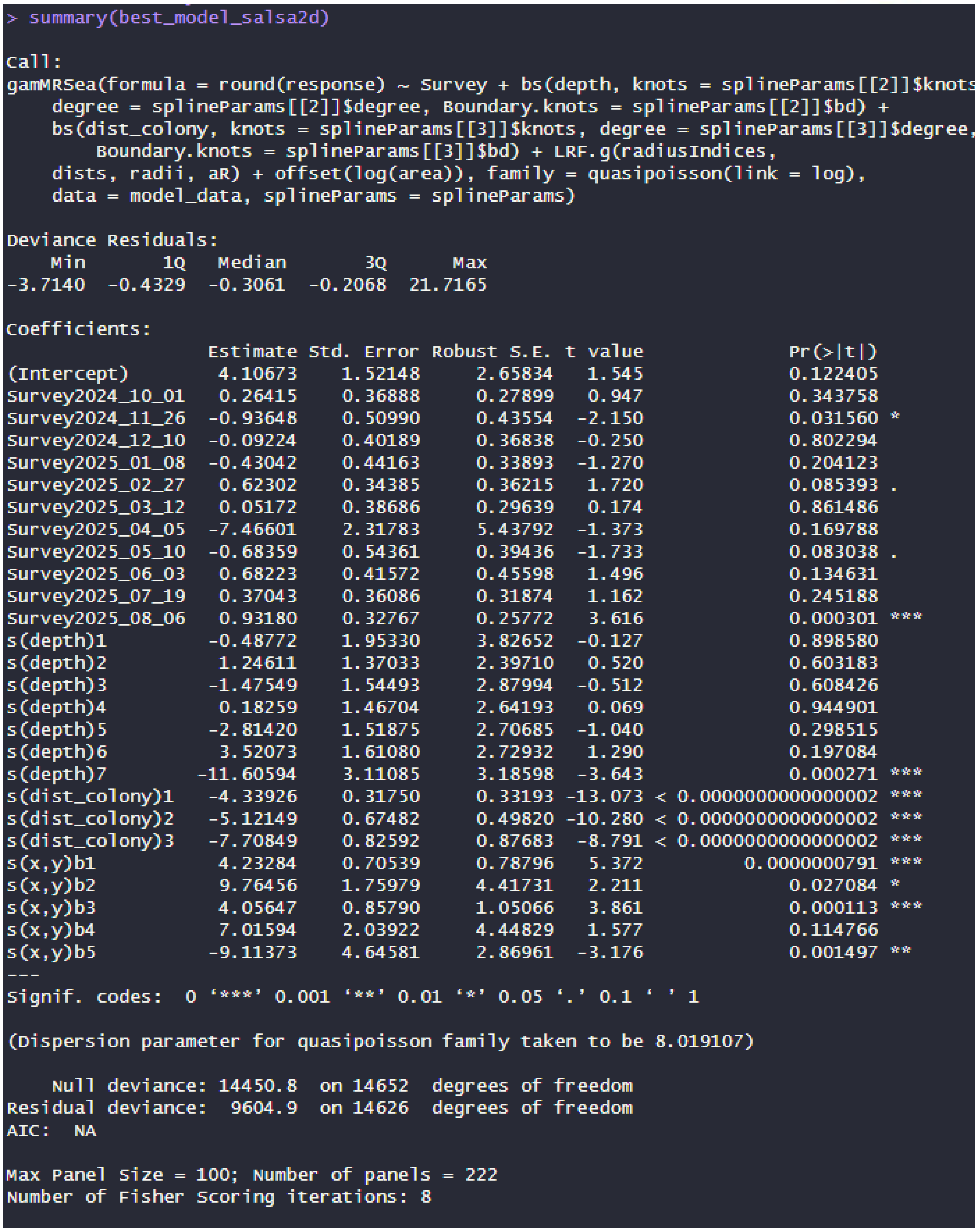
A.3.59 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.60 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-84 & Figure A6-85).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC' to make
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),   # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots for the
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2dOutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # initial_mo
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-84 Code excerpt for SALSA 2D model setup for great black-backed gull MRSea analysis.

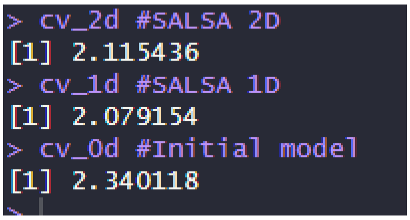


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree,
        Boundary.knots = splineParams[[3]]$bd) + LRF.g(radiusIndices,
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log),
    data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-3.7140  -0.4329  -0.3061  -0.2068  21.7165

Coefficients:
                 Estimate Std. Error Robust S.E. t value              Pr(>|t|)
(Intercept)       4.10673    1.52148     2.65834   1.545              0.122405
Survey2024_10_01  0.26415    0.36888     0.27899   0.947              0.343758
Survey2024_11_26 -0.93648    0.50990     0.43554  -2.150              0.031560 *
Survey2024_12_10 -0.09224    0.40189     0.36838  -0.250              0.802294
Survey2025_01_08 -0.43042    0.44163     0.33893  -1.270              0.204123
Survey2025_02_27  0.62302    0.34385     0.36215   1.720              0.085393 .
Survey2025_03_12  0.05172    0.38686     0.29639   0.174              0.861486
Survey2025_04_05 -7.46601    2.31783     5.43792  -1.373              0.169788
Survey2025_05_10 -0.68359    0.54361     0.39436  -1.733              0.083038 .
Survey2025_06_03  0.68223    0.41572     0.45598   1.496              0.134631
Survey2025_07_19  0.37043    0.36086     0.31874   1.162              0.245188
Survey2025_08_06  0.93180    0.32767     0.25772   3.616              0.000301 ***
s(depth)1        -0.48772    1.95330     3.82652  -0.127              0.898580
s(depth)2         1.24611    1.37033     2.39710   0.520              0.603183
s(depth)3        -1.47549    1.54493     2.87994  -0.512              0.608426
s(depth)4         0.18259    1.46704     2.64193   0.069              0.944901
s(depth)5        -2.81420    1.51875     2.70685  -1.040              0.298515
s(depth)6         3.52073    1.61080     2.72932   1.290              0.197084
s(depth)7       -11.60594    3.11085     3.18598  -3.643              0.000271 ***
s(dist_colony)1  -4.33926    0.31750     0.33193 -13.073 < 0.0000000000000002 ***
s(dist_colony)2  -5.12149    0.67482     0.49820 -10.280 < 0.0000000000000002 ***
s(dist_colony)3  -7.70849    0.82592     0.87683  -8.791 < 0.0000000000000002 ***
s(x,y)b1          4.23284    0.70539     0.78796   5.372          0.0000000791 ***
s(x,y)b2          9.76456    1.75979     4.41731   2.211              0.027084 *
s(x,y)b3          4.05647    0.85790     1.05066   3.861              0.000113 ***
s(x,y)b4          7.01594    2.03922     4.44829   1.577              0.114766
s(x,y)b5         -9.11373    4.64581     2.86961  -3.176              0.001497 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 8.019107)

    Null deviance: 14450.8  on 14652  degrees of freedom
Residual deviance:  9604.9  on 14626  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 222
Number of Fisher Scoring iterations: 8
```

Figure A6-85 Code excerpt detailing SALSA 2D summary for great black-backed gull MRSea analysis.

Cross-Validation

A.3.61 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a slightly higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-86).



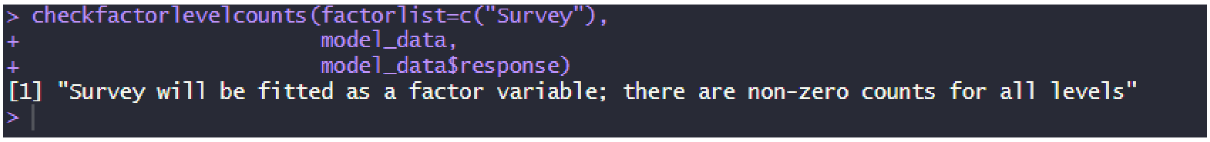
```
> cv_2d #SALSA 2D
[1] 2.115436
> cv_1d #SALSA 1D
[1] 2.079154
> cv_0d #Initial model
[1] 2.340118
```

Figure A6-86 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for great black-backed gull MRSea analysis.

Gannet

Factor Level Check

A.3.62 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-87).

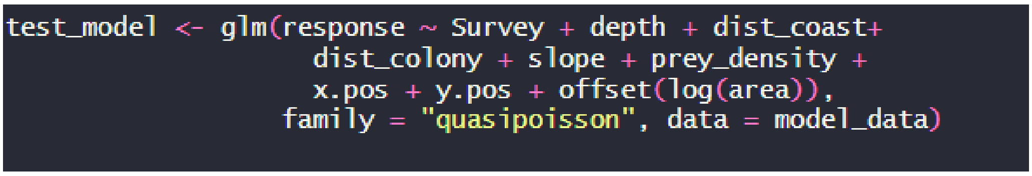


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                         model_data,
+                         model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-87 Code excerpt detailing factor level check for categorical variables considered within gannet MRSea modelling.

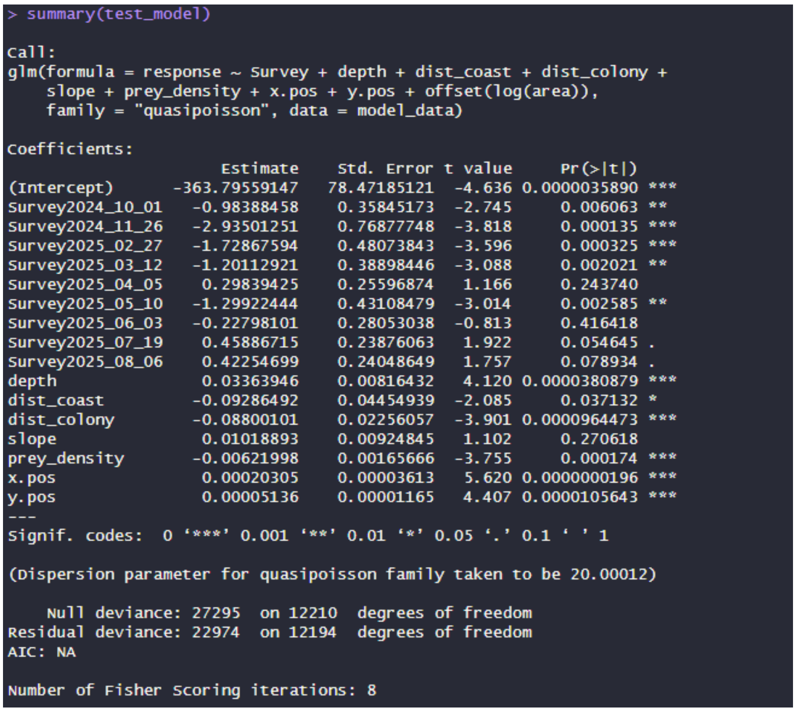
Initial GLM

A.3.63 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-88 and Figure A6-89).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-88 Code excerpt detailing the setup of the initial GLM considered within gannet MRSea analysis.

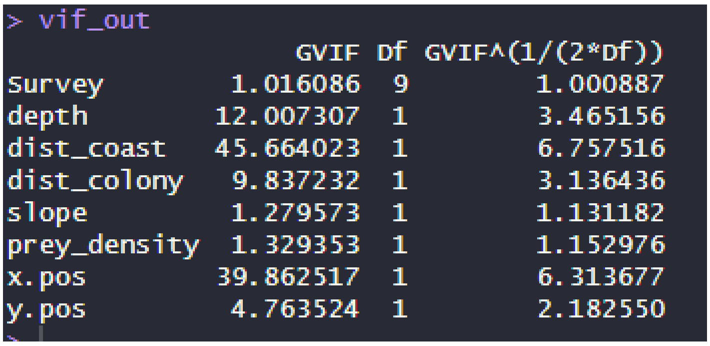


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_coast + dist_colony + 
    slope + prey_density + x.pos + y.pos + offset(log(area)), 
    family = "quasipoisson", data = model_data)

Coefficients:
                    Estimate    Std. Error t value     Pr(>|t|)    
(Intercept)      -363.79559147   78.47185121  -4.636 0.0000035890 ***
Survey2024_10_01   -0.98388458    0.35845173  -2.745     0.006063 ** 
Survey2024_11_26   -2.93501251    0.76877748  -3.818     0.000135 ***
Survey2025_02_27   -1.72867594    0.48073843  -3.596     0.000325 ***
Survey2025_03_12   -1.20112921    0.38898446  -3.088     0.002021 ** 
Survey2025_04_05    0.29839425    0.25596874   1.166     0.243740    
Survey2025_05_10   -1.29922444    0.43108479  -3.014     0.002585 ** 
Survey2025_06_03   -0.22798101    0.28053038  -0.813     0.416418    
Survey2025_07_19    0.45886715    0.23876063   1.922     0.054645 .  
Survey2025_08_06    0.42254699    0.24048649   1.757     0.078934 .  
depth               0.03363946    0.00816432   4.120 0.0000380879 ***
dist_coast         -0.09286492    0.04454939  -2.085     0.037132 *  
dist_colony        -0.08800101    0.02256057  -3.901 0.0000964473 ***
slope               0.01018893    0.00924845   1.102     0.270618    
prey_density       -0.00621998    0.00165666  -3.755     0.000174 ***
x.pos               0.00020305    0.00003613   5.620 0.0000000196 ***
y.pos               0.00005136    0.00001165   4.407 0.0000105643 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 20.00012)

    Null deviance: 27295  on 12210  degrees of freedom
Residual deviance: 22974  on 12194  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 8
```

Figure A6-89 Code excerpt detailing the summary of the initial GLM considered within gannet MRSea analysis.

Collinearity Check

A.3.64 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-90).

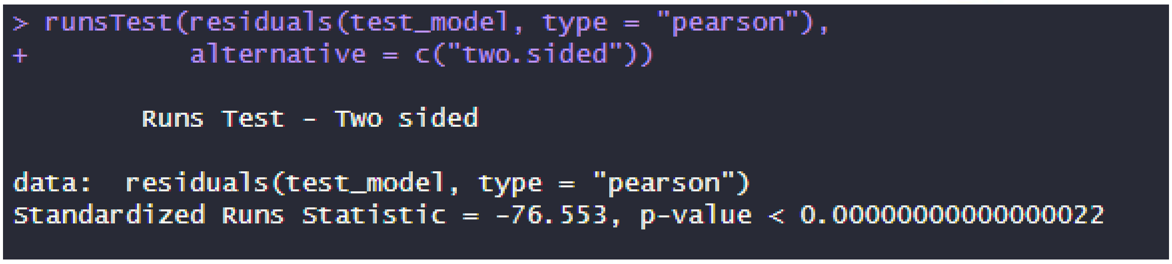


```
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
Survey        1.016086  9        1.000887
depth        12.007307  1        3.465156
dist_coast   45.664023  1        6.757516
dist_colony   9.837232  1        3.136436
slope         1.279573  1        1.131182
prey_density  1.329353  1        1.152976
x.pos        39.862517  1        6.313677
y.pos         4.763524  1        2.182550
```

Figure A6-90 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within gannet MRSea analysis.

Runs Test

A.3.65 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-91).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -76.553, p-value < 0.00000000000000022
```

Figure A6-91 Code excerpt detailing runs test carried out on the initial GLM considered within gannet MRSea analysis.

A.3.66 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-92).

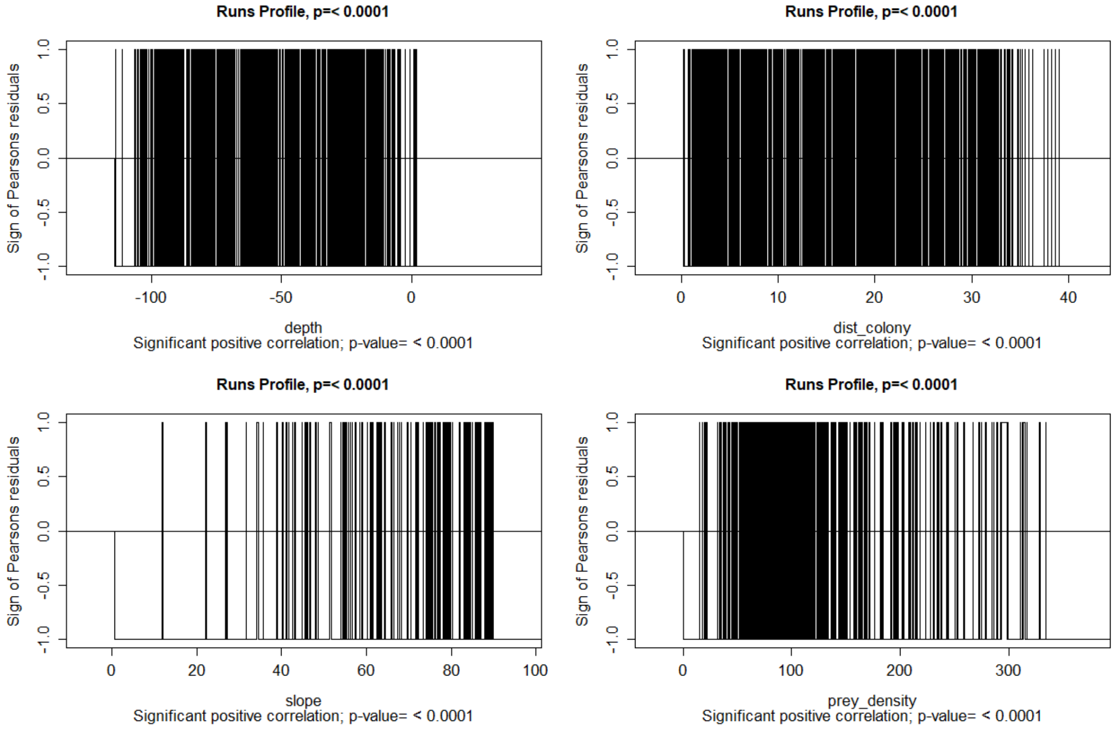


Figure A6-92 Runs profiles for explanatory variables considered within the initial GLM considered for gannet MRSea analysis.

Autocorrelation

A.3.67 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-93).

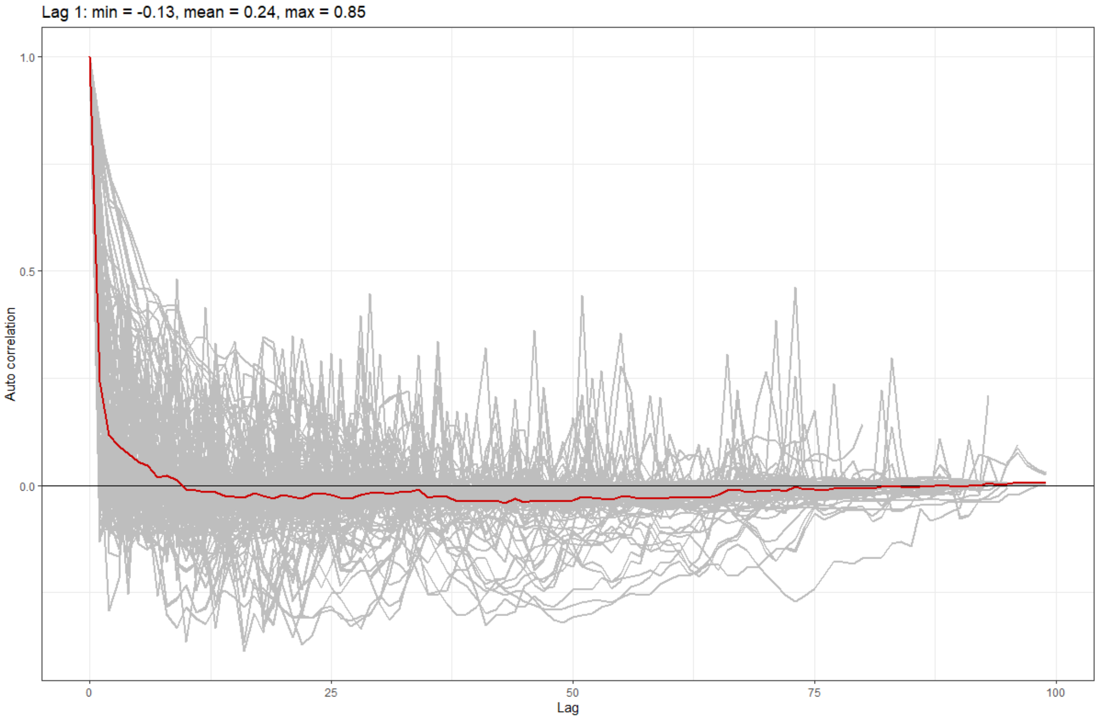


Figure A6-93 ACF plot for the initial GLM considered for gannet MRSea analysis.

Cumulative Residuals

A.3.68 Cumulative residuals for explanatory variables were plotted (Figure A6-94, Figure A6-95, Figure A6-96 & Figure A6-97). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

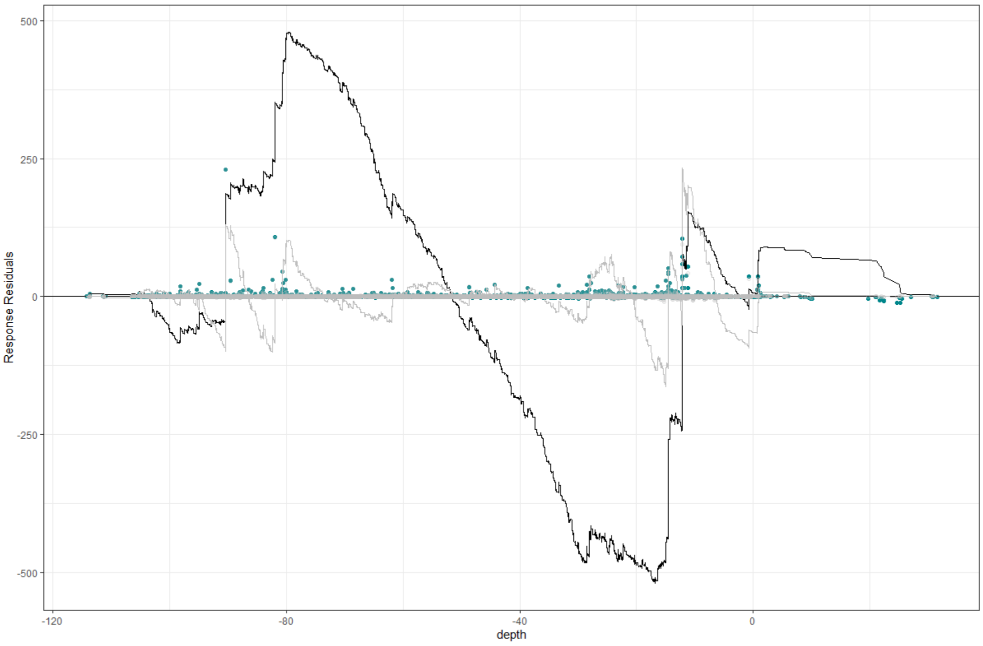


Figure A6-94 Cumulative residuals of depth plotted for the initial GLM considered for gannet MRSea analysis.

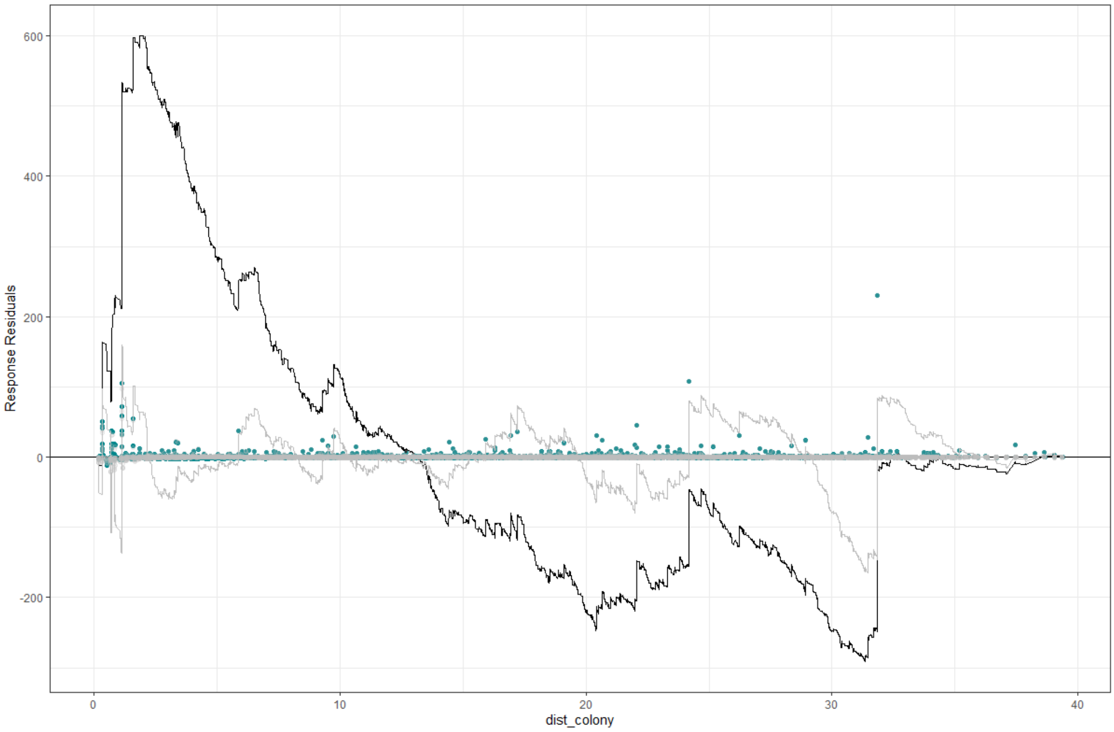


Figure A6-95 Cumulative residuals of distance to colony plotted for the initial GLM considered for gannet MRSea analysis.

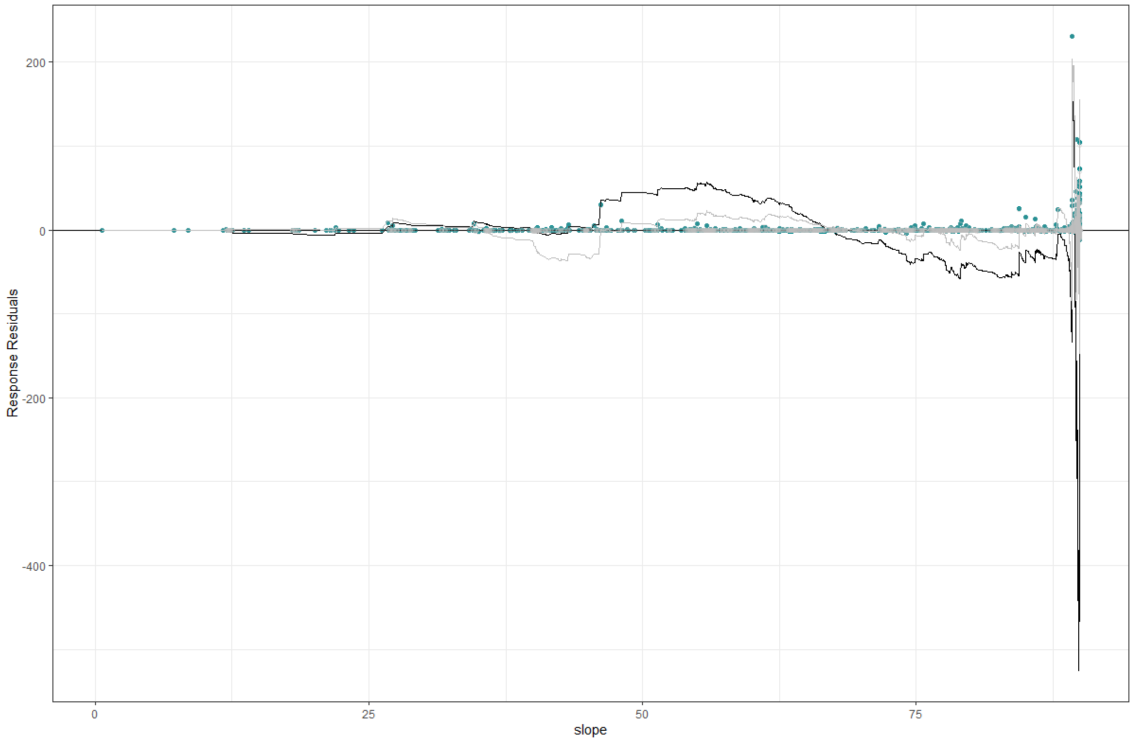


Figure A6-96 Cumulative residuals of seabed slope plotted for the initial GLM considered for gannet MRSea analysis.

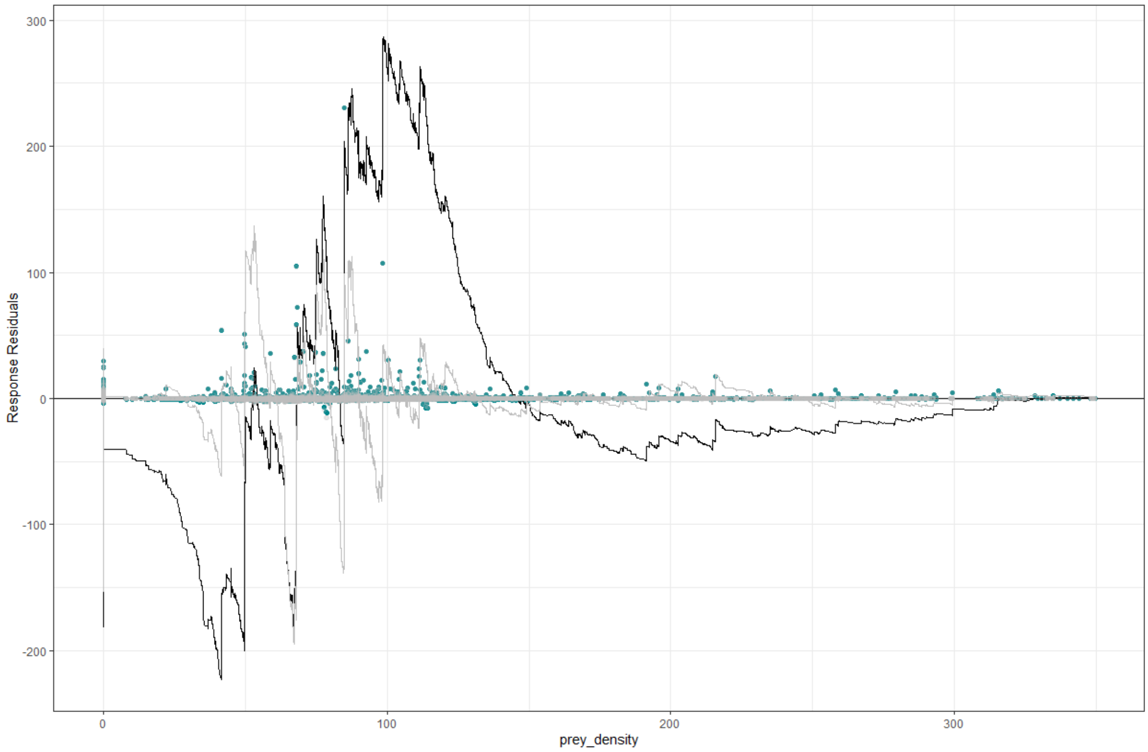
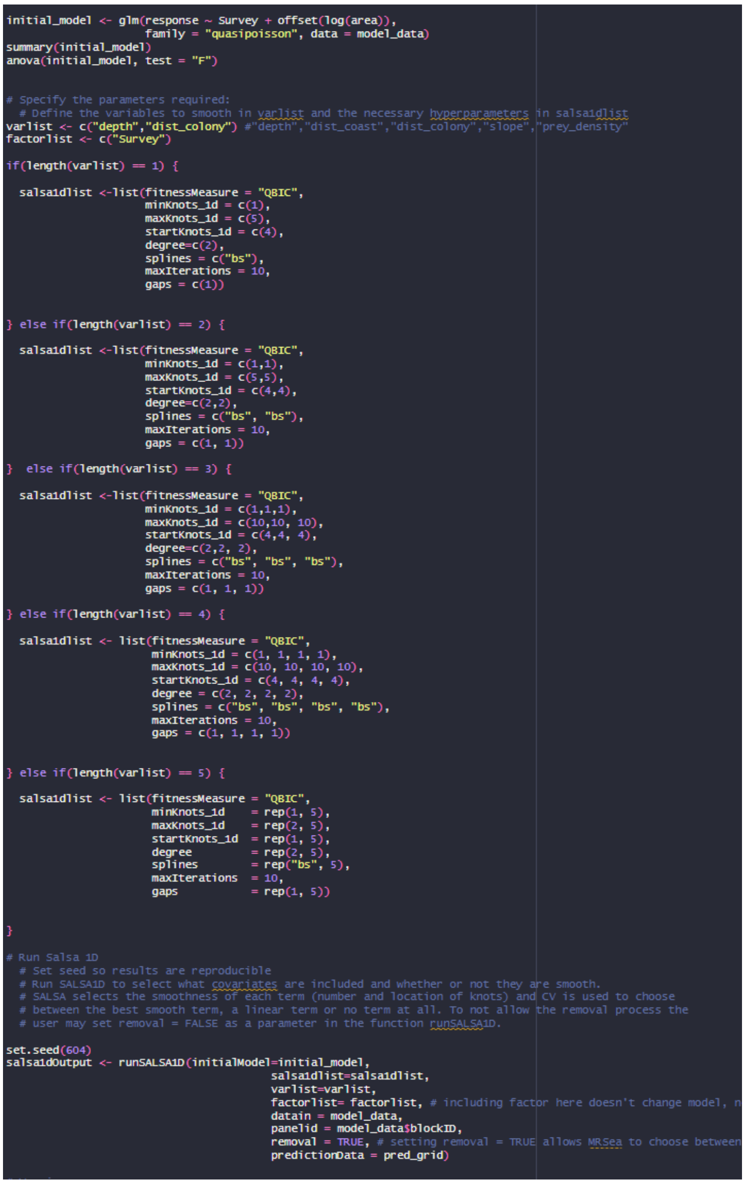


Figure A6-97 Cumulative residuals of prey density plotted for the initial GLM considered for gannet MRSea analysis.

SALSA 1D

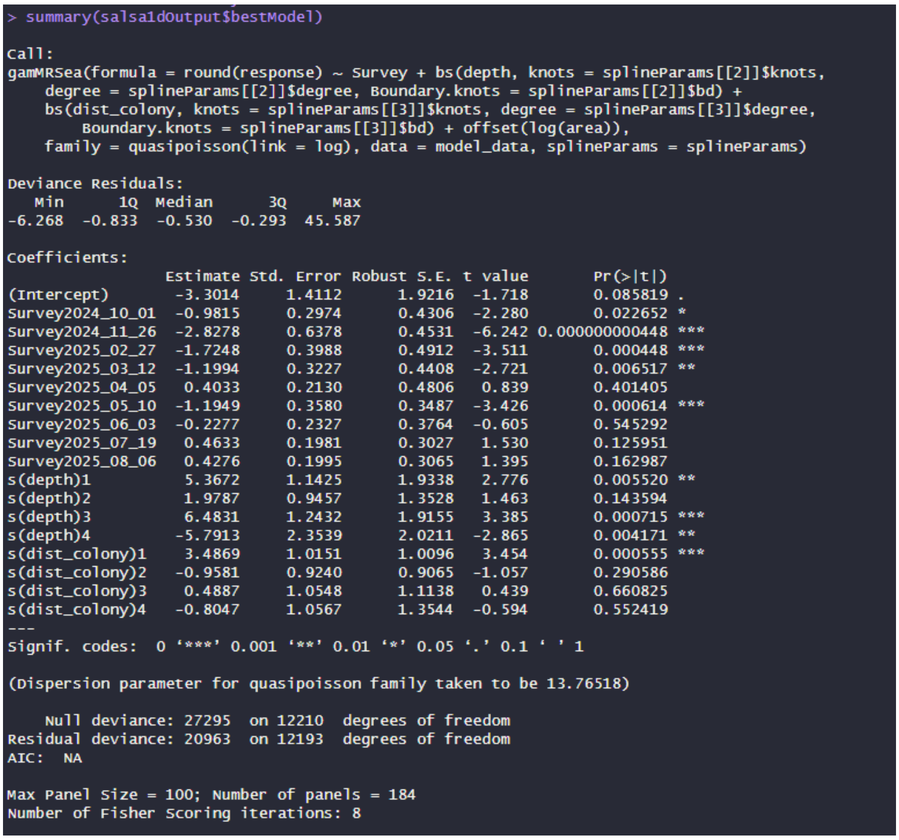
A.3.69 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.70 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). During trial runs all covariates, except depth and distance to colony, were removed from modelling and as a result were not carried forward for further analysis. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-98 & Figure A6-99).



```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dlist
varlist <- c("depth","dist_colony") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d    = rep(1, 5),
                      maxKnots_1d    = rep(2, 5),
                      startKnots_1d  = rep(1, 5),
                      degree         = rep(2, 5),
                      splines        = rep("bs", 5),
                      maxIterations  = 10,
                      gaps           = rep(1, 5))


}
# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smooth.
  # SALSA selects the smoothness of each term (number and location of knots) and CV is used to choose
  # between the best smooth term, a linear term or no term at all. To not allow the removal process the
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                                      salsa1dlist=salsa1dlist,
                                      varlist=varlist,
                                      factorlist= factorlist, # including factor here doesn't change model, n
                                      datain = model_data,
                                      panelid = model_data$blockID,
                                      removal = TRUE, # setting removal = TRUE allows MRSea to choose between
                                      predictionData = pred_grid)
```

Figure A6-98 Code excerpt detailing the setup process for the SALSA 1D model considered for gannet MRSea analysis.



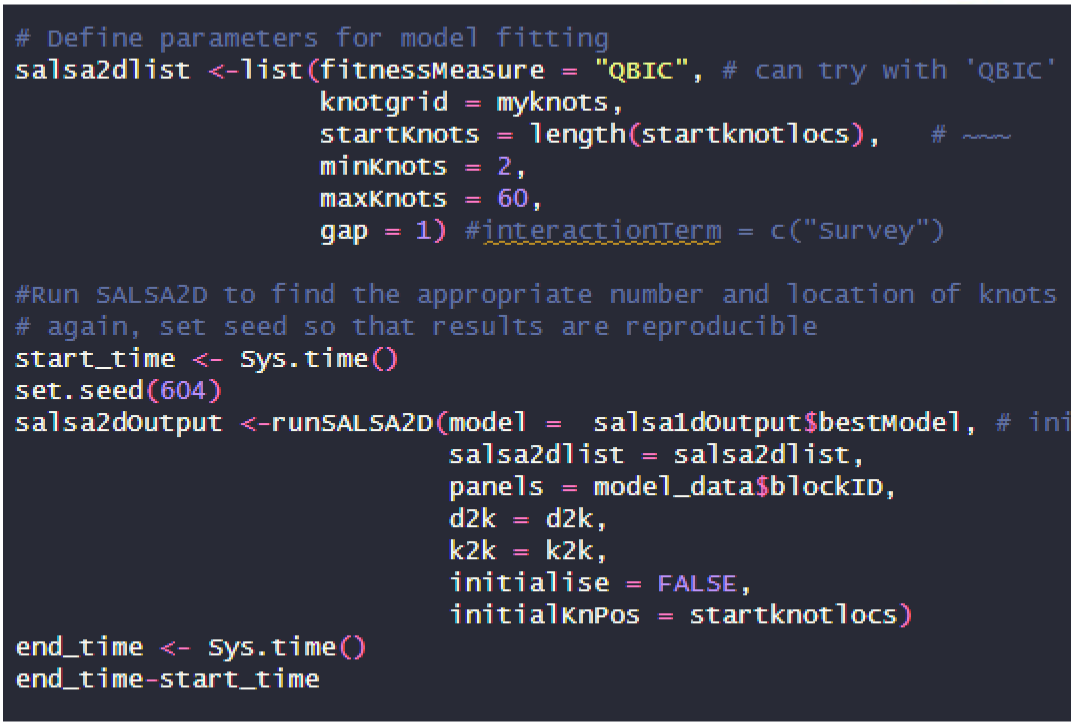
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree,
        Boundary.knots = splineParams[[3]]$bd) + offset(log(area)),
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals:
   Min       1Q  Median       3Q      Max
-6.268  -0.833  -0.530  -0.293  45.587

Coefficients:
                 Estimate Std. Error Robust S.E. t value        Pr(>|t|)
(Intercept)       -3.3014     1.4112      1.9216  -1.718        0.085819 .
Survey2024_10_01  -0.9815     0.2974      0.4306  -2.280        0.022652 *
Survey2024_11_26  -2.8278     0.6378      0.4531  -6.242 0.000000000448 ***
Survey2025_02_27  -1.7248     0.3988      0.4912  -3.511        0.000448 ***
Survey2025_03_12  -1.1994     0.3227      0.4408  -2.721        0.006517 **
Survey2025_04_05   0.4033     0.2130      0.4806   0.839        0.401405
Survey2025_05_10  -1.1949     0.3580      0.3487  -3.426        0.000614 ***
Survey2025_06_03  -0.2277     0.2327      0.3764  -0.605        0.545292
Survey2025_07_19   0.4633     0.1981      0.3027   1.530        0.125951
Survey2025_08_06   0.4276     0.1995      0.3065   1.395        0.162987
s(depth)1          5.3672     1.1425      1.9338   2.776        0.005520 **
s(depth)2          1.9787     0.9457      1.3528   1.463        0.143594
s(depth)3          6.4831     1.2432      1.9155   3.385        0.000715 ***
s(depth)4         -5.7913     2.3539      2.0211  -2.865        0.004171 **
s(dist_colony)1    3.4869     1.0151      1.0096   3.454        0.000555 ***
s(dist_colony)2   -0.9581     0.9240      0.9065  -1.057        0.290586
s(dist_colony)3    0.4887     1.0548      1.1138   0.439        0.660825
s(dist_colony)4   -0.8047     1.0567      1.3544  -0.594        0.552419
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 13.76518)

    Null deviance: 27295  on 12210  degrees of freedom
Residual deviance: 20963  on 12193  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 184
Number of Fisher Scoring iterations: 8
```

Figure A6-99 Code excerpt detailing the summary for the SALSA 1D model for gannet MRSea analysis.

SALSA 2D

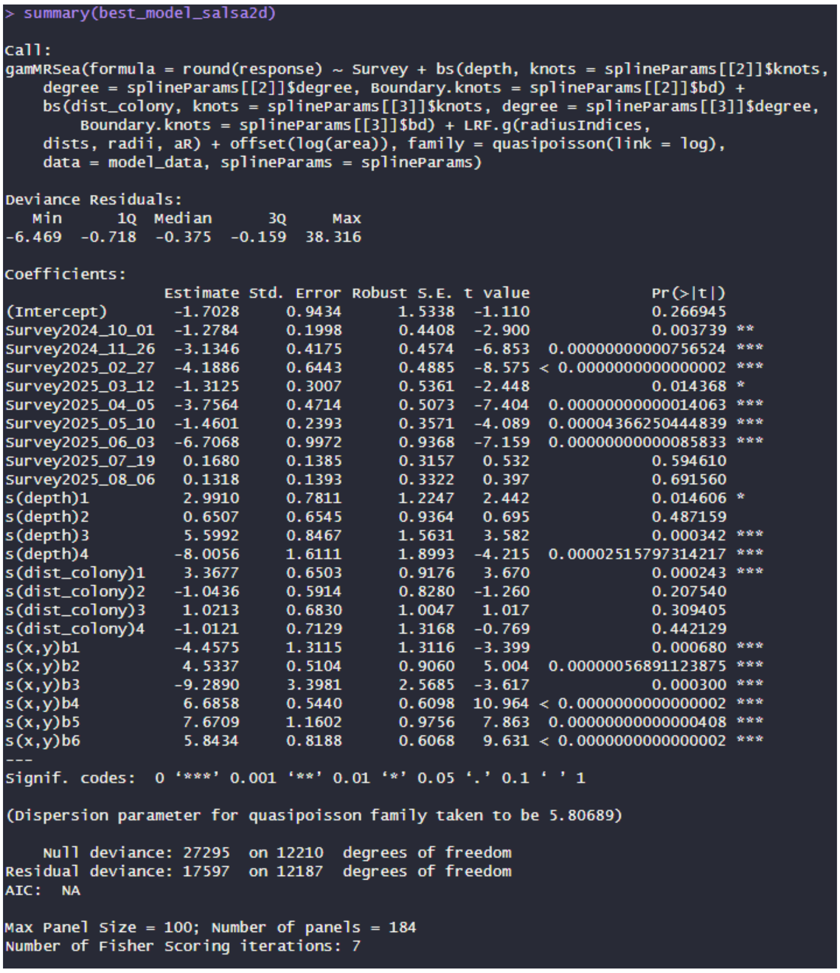
A.3.71 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.72 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-100 & Figure A6-101).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC'
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),    # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2dOutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # ini
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-100 Code excerpt for SALSA 2D model setup for gannet MRSea analysis.

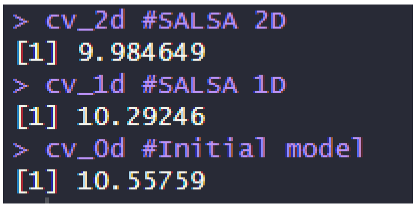


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(depth, knots = splineParams[[2]]$knots,
    degree = splineParams[[2]]$degree, Boundary.knots = splineParams[[2]]$bd) +
    bs(dist_colony, knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree,
        Boundary.knots = splineParams[[3]]$bd) + LRF.g(radiusIndices,
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log),
    data = model_data, splineParams = splineParams)

Deviance Residuals:
   Min       1Q   Median       3Q      Max
-6.469   -0.718   -0.375   -0.159   38.316

Coefficients:
                  Estimate Std. Error Robust S.E. t value              Pr(>|t|)
(Intercept)        -1.7028     0.9434      1.5338  -1.110              0.266945
Survey2024_10_01   -1.2784     0.1998      0.4408  -2.900              0.003739 **
Survey2024_11_26   -3.1346     0.4175      0.4574  -6.853   0.00000000000756524 ***
Survey2025_02_27   -4.1886     0.6443      0.4885  -8.575 < 0.0000000000000002 ***
Survey2025_03_12   -1.3125     0.3007      0.5361  -2.448              0.014368 *
Survey2025_04_05   -3.7564     0.4714      0.5073  -7.404   0.00000000000014063 ***
Survey2025_05_10   -1.4601     0.2393      0.3571  -4.089   0.00004366250444839 ***
Survey2025_06_03   -6.7068     0.9972      0.9368  -7.159   0.00000000000085833 ***
Survey2025_07_19    0.1680     0.1385      0.3157   0.532              0.594610
Survey2025_08_06    0.1318     0.1393      0.3322   0.397              0.691560
s(depth)1           2.9910     0.7811      1.2247   2.442              0.014606 *
s(depth)2           0.6507     0.6545      0.9364   0.695              0.487159
s(depth)3           5.5992     0.8467      1.5631   3.582              0.000342 ***
s(depth)4          -8.0056     1.6111      1.8993  -4.215   0.00002515797314217 ***
s(dist_colony)1     3.3677     0.6503      0.9176   3.670              0.000243 ***
s(dist_colony)2    -1.0436     0.5914      0.8280  -1.260              0.207540
s(dist_colony)3     1.0213     0.6830      1.0047   1.017              0.309405
s(dist_colony)4    -1.0121     0.7129      1.3168  -0.769              0.442129
s(x,y)b1           -4.4575     1.3115      1.3116  -3.399              0.000680 ***
s(x,y)b2            4.5337     0.5104      0.9060   5.004   0.00000056891123875 ***
s(x,y)b3           -9.2890     3.3981      2.5685  -3.617              0.000300 ***
s(x,y)b4            6.6858     0.5440      0.6098  10.964 < 0.0000000000000002 ***
s(x,y)b5            7.6709     1.1602      0.9756   7.863   0.00000000000000408 ***
s(x,y)b6            5.8434     0.8188      0.6068   9.631 < 0.0000000000000002 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 5.80689)

    Null deviance: 27295  on 12210  degrees of freedom
Residual deviance: 17597  on 12187  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 184
Number of Fisher Scoring iterations: 7
```

Figure A6-101 Code excerpt detailing SALSA 2D summary for gannet MRSea analysis.

Cross-Validation

A.3.73 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a slightly higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-102).



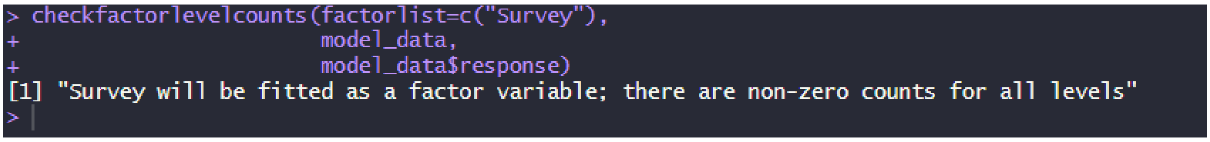
```
> cv_2d #SALSA 2D
[1] 9.984649
> cv_1d #SALSA 1D
[1] 10.29246
> cv_0d #Initial model
[1] 10.55759
```

Figure A6-102 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for gannet MRSea analysis.

Common Tern

Factor Level Check

A.3.74 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-103).

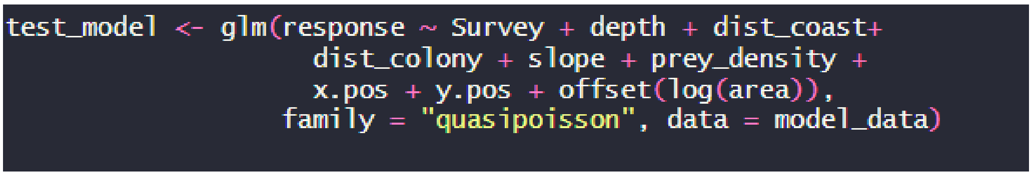


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                        model_data,
+                        model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> |
```

Figure A6-103 Code excerpt detailing factor level check for categorical variables considered within guillemot MRSea modelling.

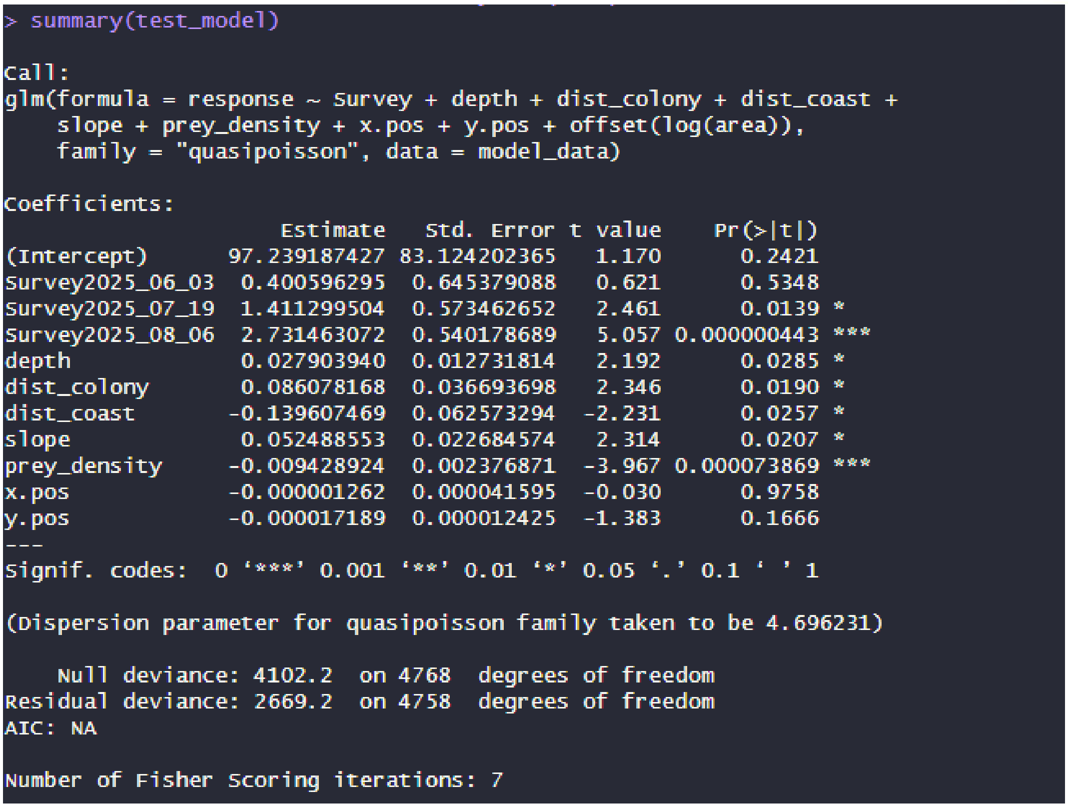
Initial GLM

A.3.75 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-104 and Figure A6-105).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-104 Code excerpt detailing the setup of the initial GLM considered within common tern MRSea analysis.

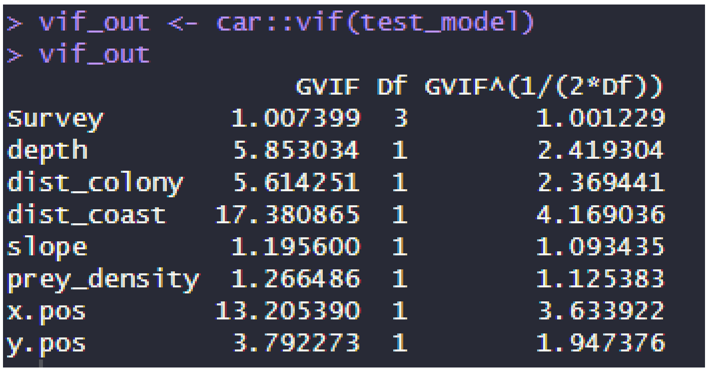


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_colony + dist_coast +
    slope + prey_density + x.pos + y.pos + offset(log(area)),
    family = "quasipoisson", data = model_data)

Coefficients:
                    Estimate   Std. Error t value    Pr(>|t|)
(Intercept)      97.239187427 83.124202365   1.170      0.2421
Survey2025_06_03  0.400596295  0.645379088   0.621      0.5348
Survey2025_07_19  1.411299504  0.573462652   2.461      0.0139 *
Survey2025_08_06  2.731463072  0.540178689   5.057 0.000000443 ***
depth             0.027903940  0.012731814   2.192      0.0285 *
dist_colony       0.086078168  0.036693698   2.346      0.0190 *
dist_coast       -0.139607469  0.062573294  -2.231      0.0257 *
slope             0.052488553  0.022684574   2.314      0.0207 *
prey_density     -0.009428924  0.002376871  -3.967 0.000073869 ***
x.pos            -0.000001262  0.000041595  -0.030      0.9758
y.pos            -0.000017189  0.000012425  -1.383      0.1666
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 4.696231)

    Null deviance: 4102.2  on 4768  degrees of freedom
Residual deviance: 2669.2  on 4758  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 7
```

Figure A6-105 Code excerpt detailing the summary of the initial GLM considered within common tern MRSea analysis.

Collinearity Check

A.3.76 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold (Figure A6-106).

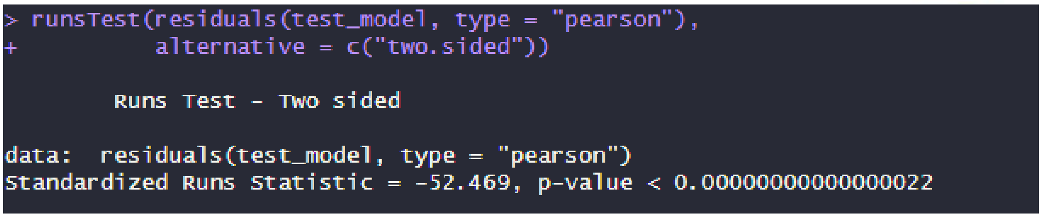


```
> vif_out <- car::vif(test_model)
> vif_out
                  GVIF Df GVIF^(1/(2*Df))
Survey        1.007399  3        1.001229
depth         5.853034  1        2.419304
dist_colony   5.614251  1        2.369441
dist_coast   17.380865  1        4.169036
slope         1.195600  1        1.093435
prey_density  1.266486  1        1.125383
x.pos        13.205390  1        3.633922
y.pos         3.792273  1        1.947376
```

Figure A6-106 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within common tern MRSea analysis.

Runs Test

A.3.77 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-107).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -52.469, p-value < 0.00000000000000022
```

Figure A6-107 Code excerpt detailing runs test carried out on the initial GLM considered within common tern MRSea analysis.

A.3.78 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-108).

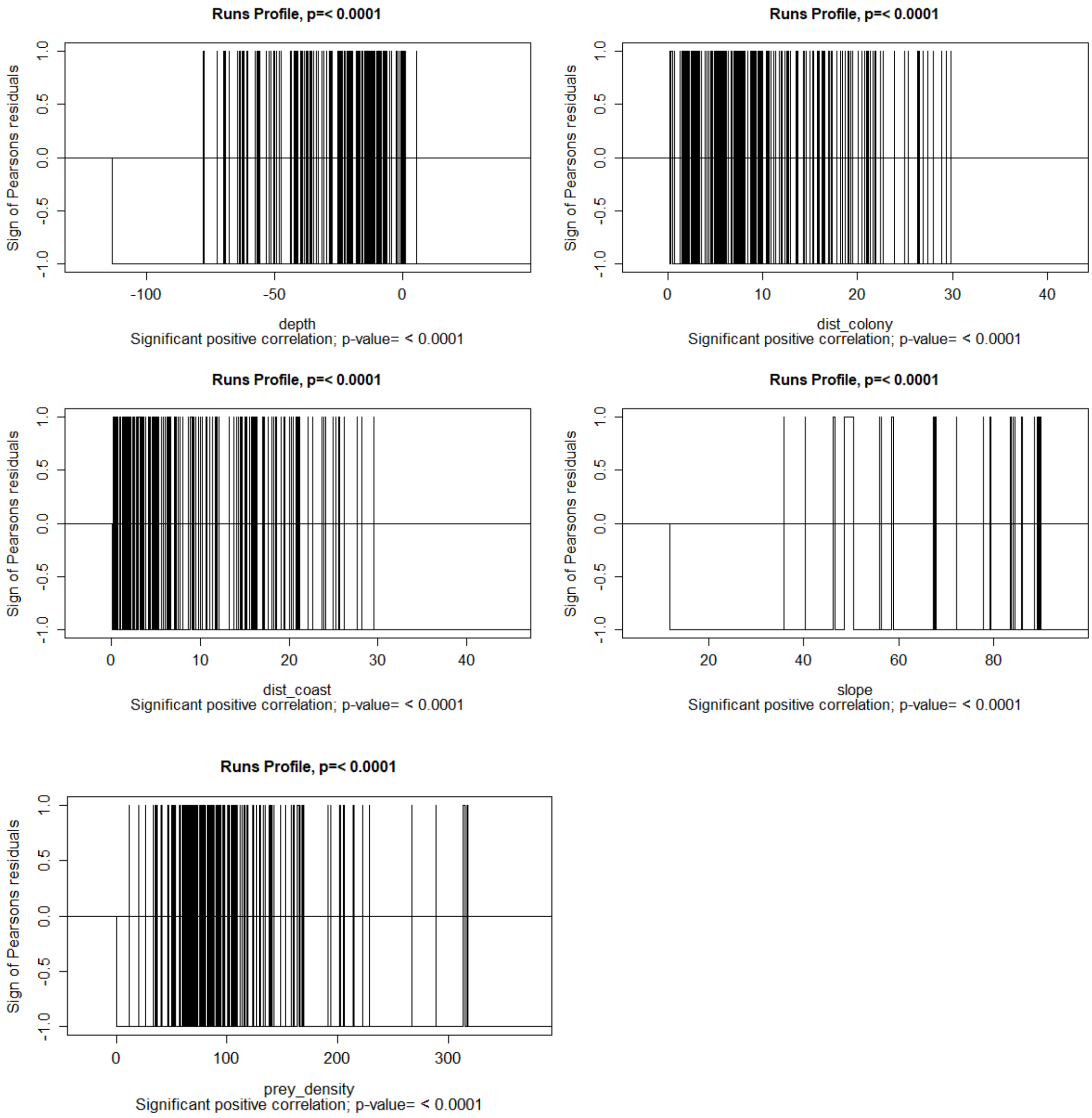


Figure A6-108 Runs profiles for explanatory variables considered within the initial GLM considered for common tern MRSea analysis.

Autocorrelation

A.3.79 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-109).

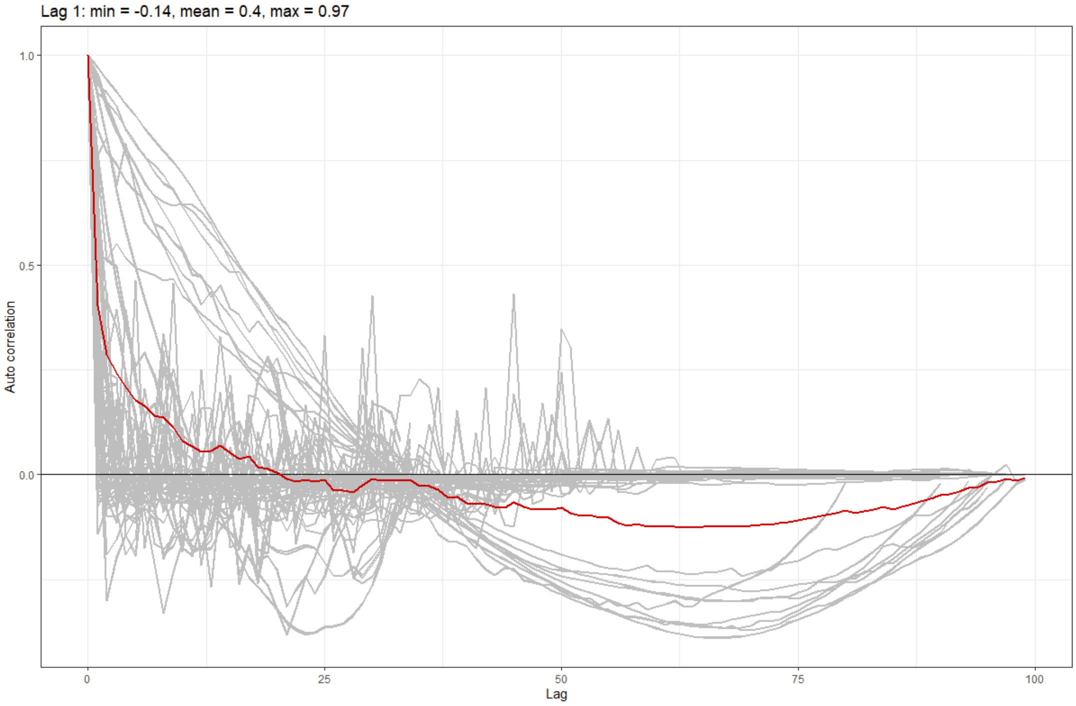


Figure A6-109 ACF plot for the initial GLM considered for common tern MRSea analysis.

Cumulative Residuals

A.3.80 Cumulative residuals for explanatory variables were plotted (Figure A6-110, Figure A6-111, Figure A6-112, Figure A6-113 & Figure A6-114). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

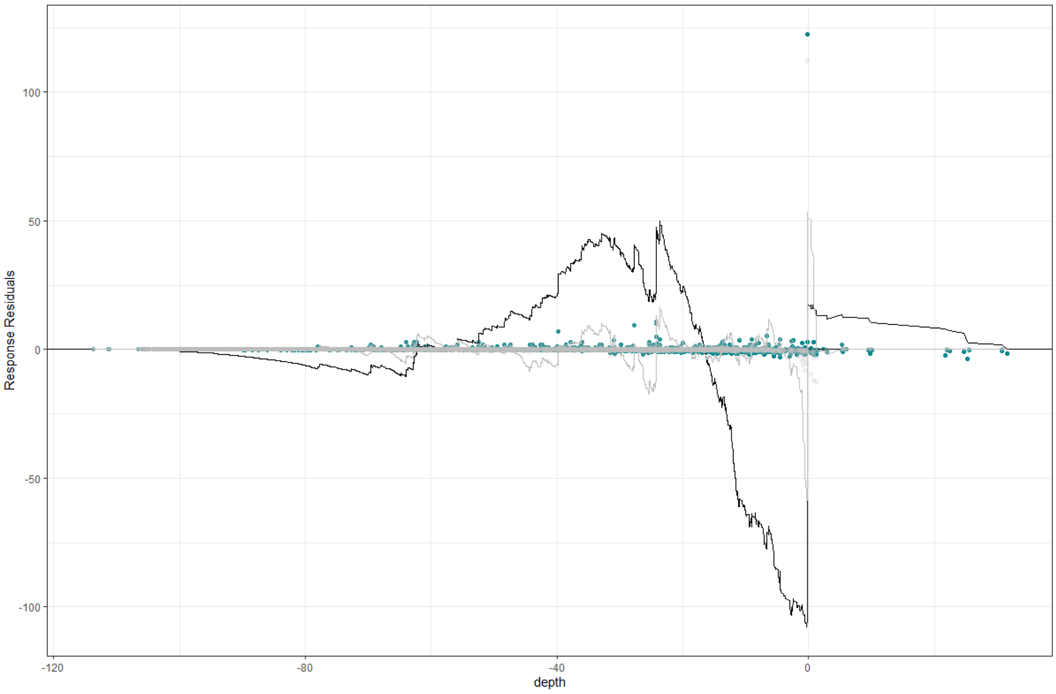


Figure A6-110 Cumulative residuals of depth plotted for the initial GLM considered for common tern MRSea analysis.

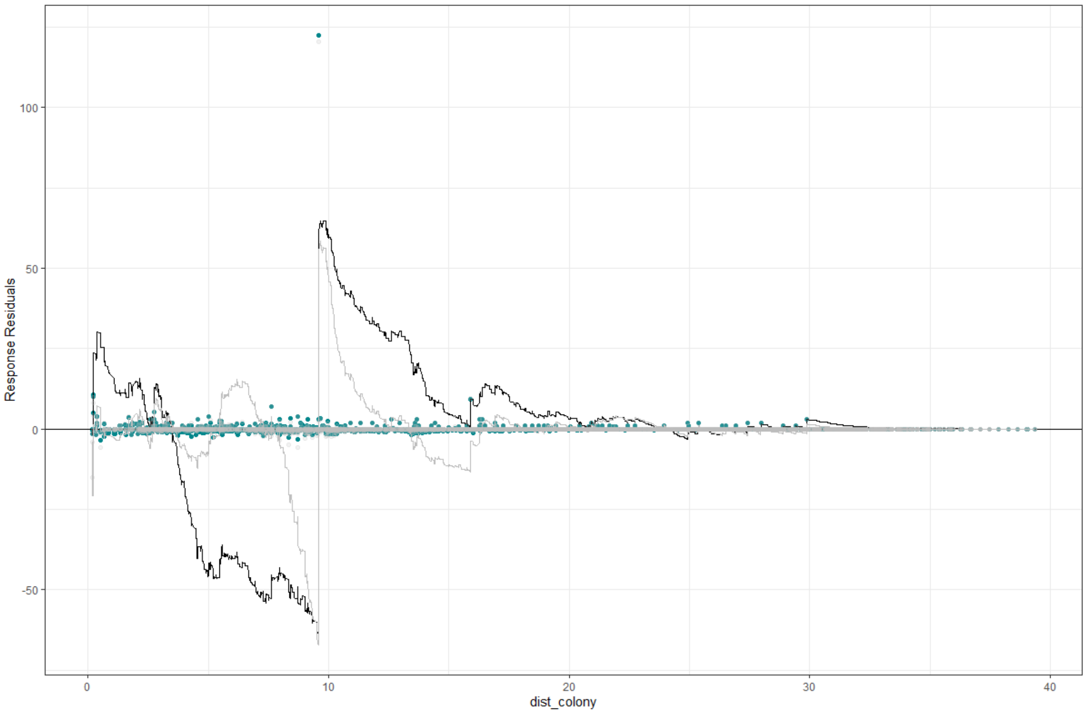


Figure A6-111 Cumulative residuals of distance to colony plotted for the initial GLM considered for common tern MRSea analysis.

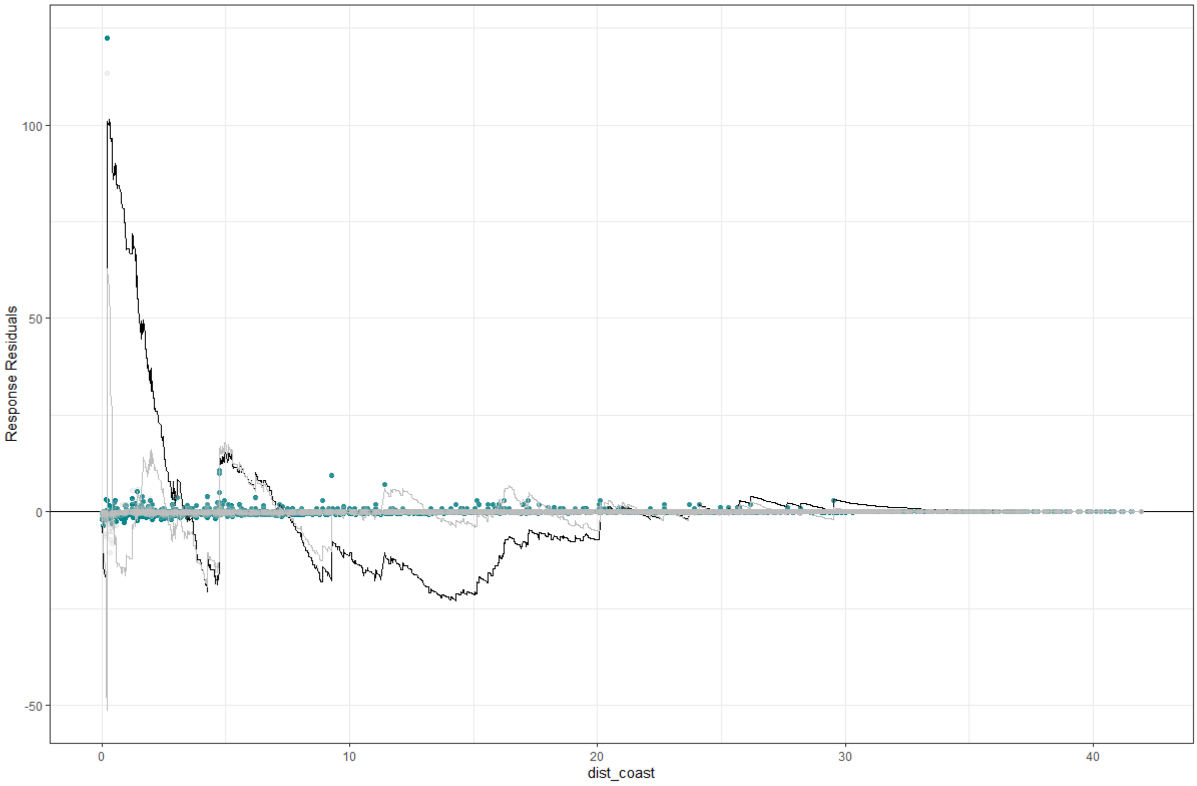


Figure A6-112 Cumulative residuals of distance to coast plotted for the initial GLM considered for common tern MRSea analysis.

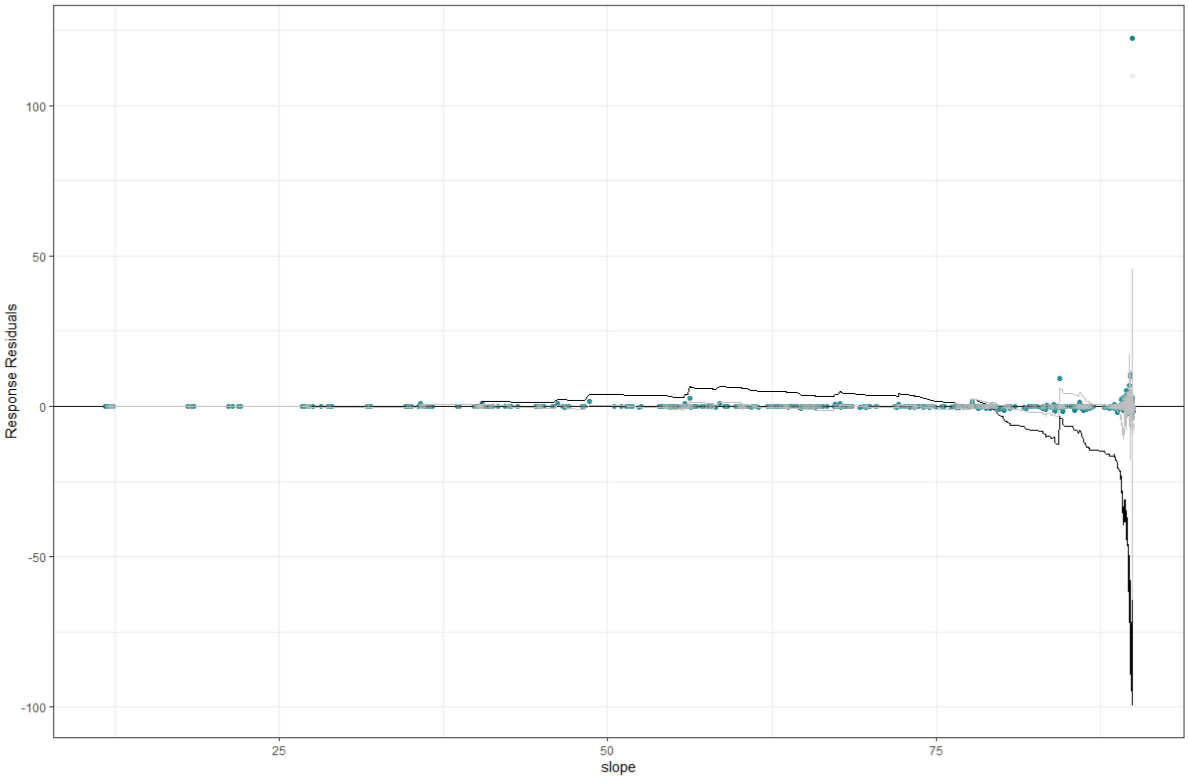


Figure A6-113 Cumulative residuals of seabed slope plotted for the initial GLM considered for common tern MRSea analysis.

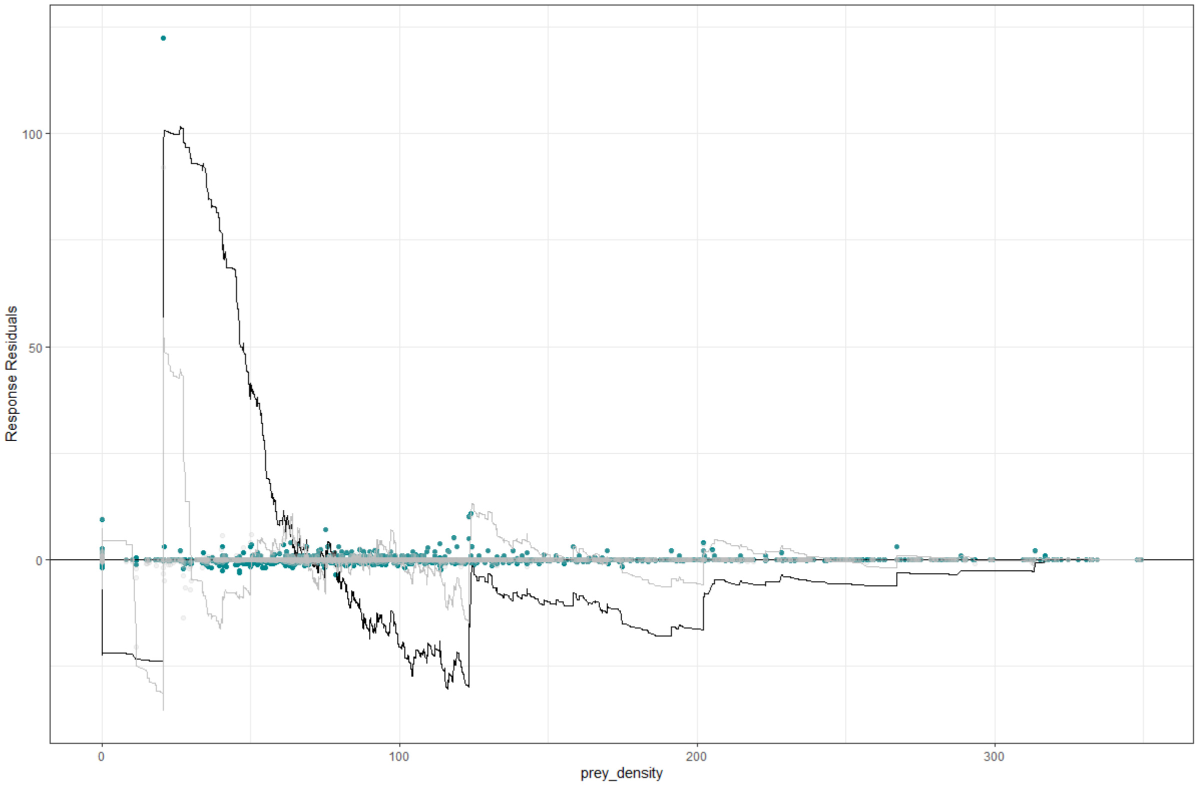
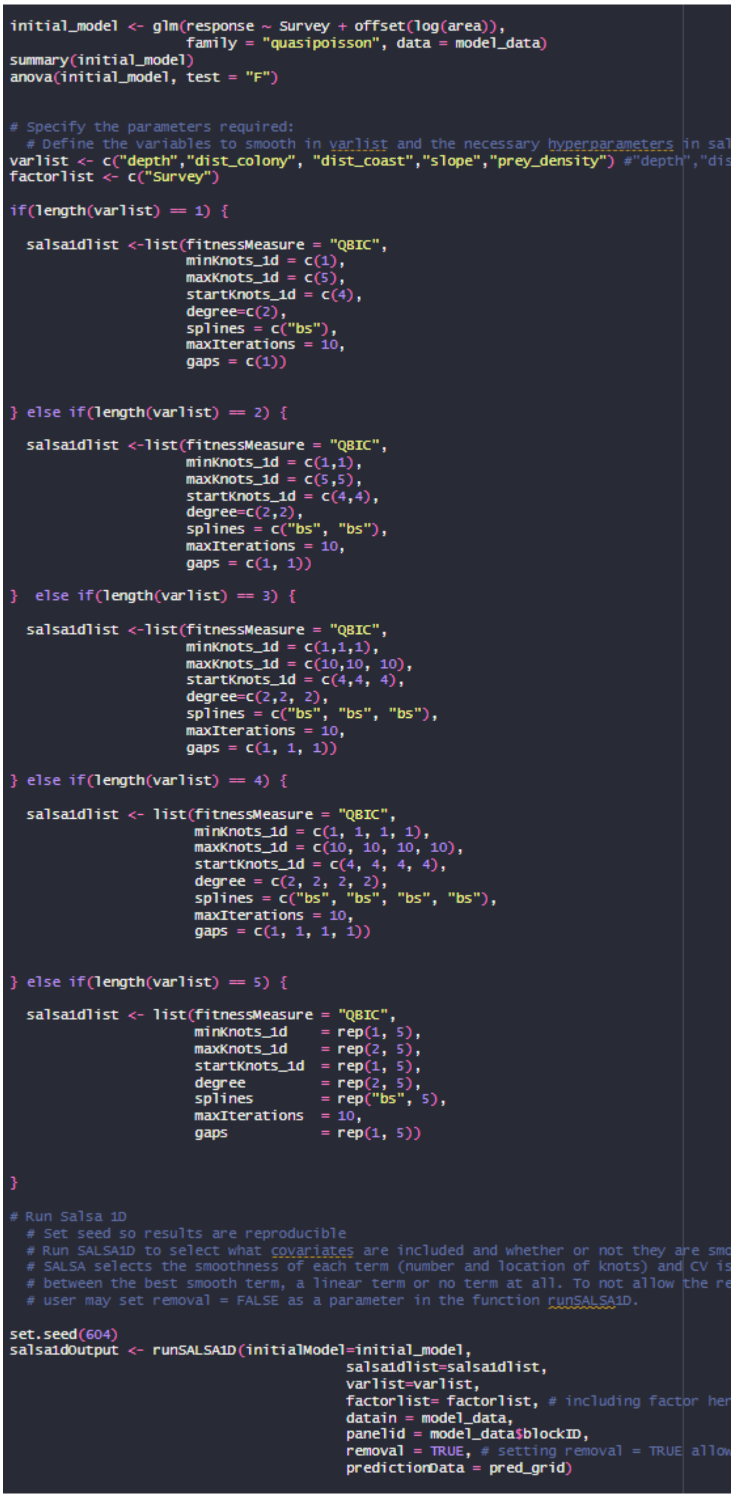


Figure A6-114 Cumulative residuals of prey density plotted for the initial GLM considered for common tern MRSea analysis.

SALSA 1D

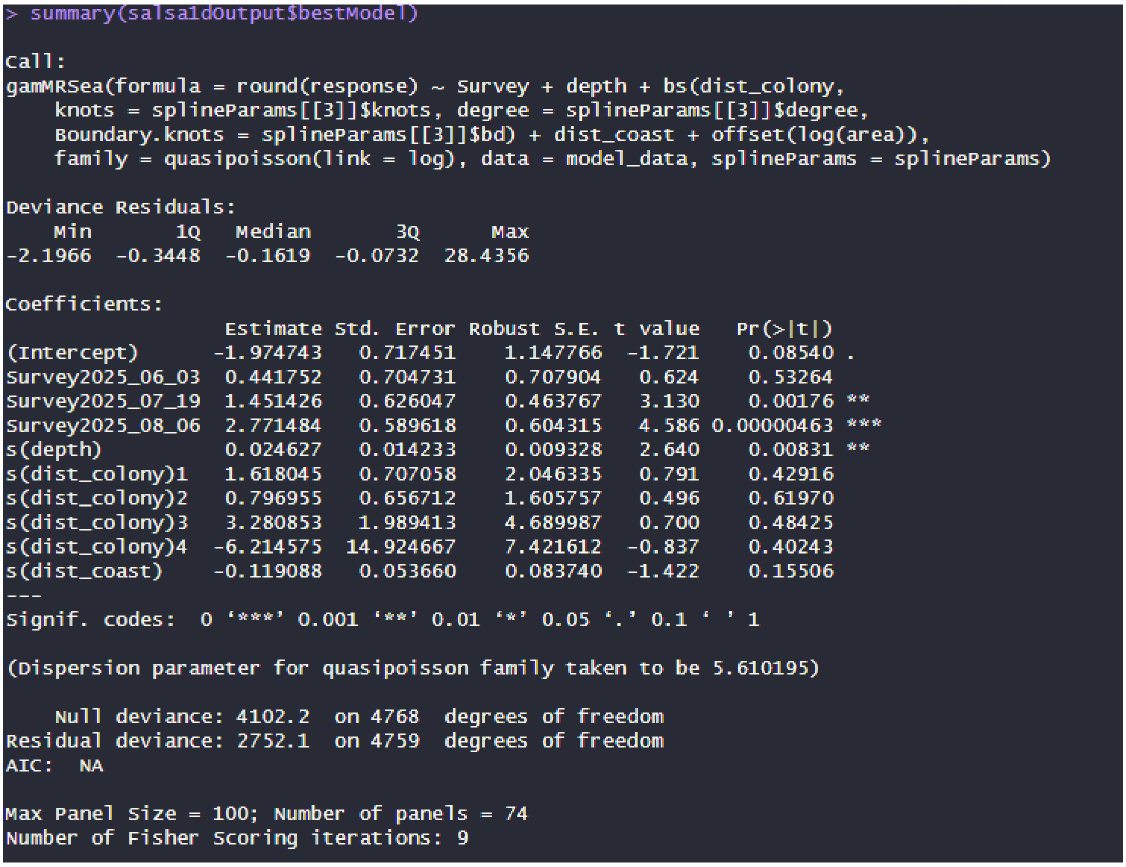
A.3.81 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.82 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). The following covariates: depth, slope and prey density were dropped from the model at this point. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the “best model.” Code excerpts for model setup and summary are provided below (Figure A6-115 & Figure A6-116).



```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in sal
varlist <- c("depth","dist_colony", "dist_coast","slope","prey_density") #"depth","dis
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d     = rep(1, 5),
                      maxKnots_1d     = rep(2, 5),
                      startKnots_1d   = rep(1, 5),
                      degree          = rep(2, 5),
                      splines         = rep("bs", 5),
                      maxIterations   = 10,
                      gaps            = rep(1, 5))


}

# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smo
  # SALSA selects the smoothness of each term (number and location of knots) and CV is
  # between the best smooth term, a linear term or no term at all. To not allow the re
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                            salsa1dlist=salsa1dlist,
                            varlist=varlist,
                            factorlist= factorlist, # including factor her
                            datain = model_data,
                            panelid = model_data$blockID,
                            removal = TRUE, # setting removal = TRUE allow
                            predictionData = pred_grid)
```

Figure A6-115 Code excerpt detailing the setup process for the SALSA 1D model considered for common tern MRSea analysis.



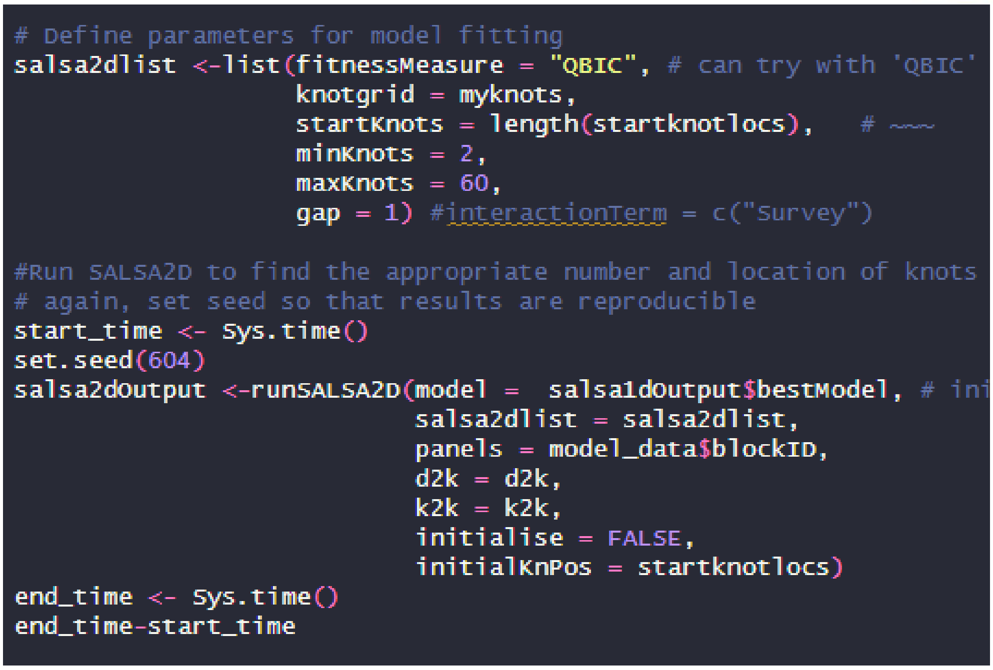
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + depth + bs(dist_colony,
    knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree,
    Boundary.knots = splineParams[[3]]$bd) + dist_coast + offset(log(area)),
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.1966  -0.3448  -0.1619  -0.0732  28.4356

Coefficients:
                  Estimate Std. Error Robust S.E. t value    Pr(>|t|)
(Intercept)      -1.974743   0.717451    1.147766  -1.721     0.08540 .
Survey2025_06_03  0.441752   0.704731    0.707904   0.624     0.53264
Survey2025_07_19  1.451426   0.626047    0.463767   3.130     0.00176 **
Survey2025_08_06  2.771484   0.589618    0.604315   4.586 0.00000463 ***
s(depth)          0.024627   0.014233    0.009328   2.640     0.00831 **
s(dist_colony)1   1.618045   0.707058    2.046335   0.791     0.42916
s(dist_colony)2   0.796955   0.656712    1.605757   0.496     0.61970
s(dist_colony)3   3.280853   1.989413    4.689987   0.700     0.48425
s(dist_colony)4  -6.214575  14.924667    7.421612  -0.837     0.40243
s(dist_coast)    -0.119088   0.053660    0.083740  -1.422     0.15506
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 5.610195)

    Null deviance: 4102.2  on 4768  degrees of freedom
Residual deviance: 2752.1  on 4759  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 74
Number of Fisher Scoring iterations: 9
```

Figure A6-116 Code excerpt detailing the summary for the SALSA 1D model for common tern MRSea analysis.

SALSA 2D

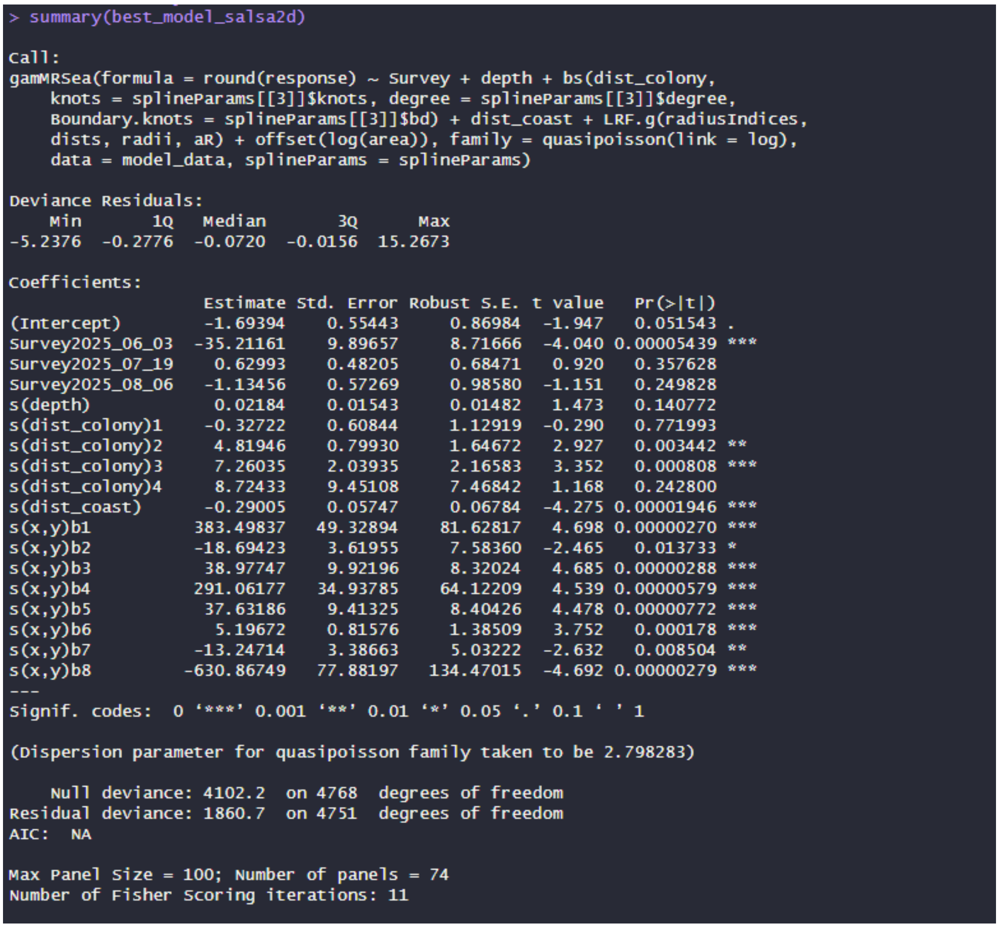
A.3.83 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.84 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-117 & Figure A6-118).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC'
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),   # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2doutput <-runSALSA2D(model =  salsa1doutput$bestModel, # ini
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-117 Code excerpt for SALSA 2D model setup for common tern MRSea analysis.

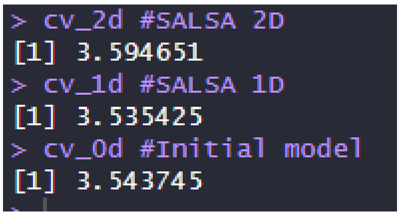


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + depth + bs(dist_colony, 
    knots = splineParams[[3]]$knots, degree = splineParams[[3]]$degree, 
    Boundary.knots = splineParams[[3]]$bd) + dist_coast + LRF.g(radiusIndices, 
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log), 
    data = model_data, splineParams = splineParams)

Deviance Residuals: 
    Min       1Q   Median       3Q      Max  
-5.2376  -0.2776  -0.0720  -0.0156  15.2673  

Coefficients:
                   Estimate Std. Error Robust S.E. t value   Pr(>|t|)    
(Intercept)        -1.69394    0.55443     0.86984  -1.947   0.051543 .  
Survey2025_06_03  -35.21161    9.89657     8.71666  -4.040 0.00005439 ***
Survey2025_07_19    0.62993    0.48205     0.68471   0.920   0.357628    
Survey2025_08_06   -1.13456    0.57269     0.98580  -1.151   0.249828    
s(depth)            0.02184    0.01543     0.01482   1.473   0.140772    
s(dist_colony)1    -0.32722    0.60844     1.12919  -0.290   0.771993    
s(dist_colony)2     4.81946    0.79930     1.64672   2.927   0.003442 ** 
s(dist_colony)3     7.26035    2.03935     2.16583   3.352   0.000808 ***
s(dist_colony)4     8.72433    9.45108     7.46842   1.168   0.242800    
s(dist_coast)      -0.29005    0.05747     0.06784  -4.275 0.00001946 ***
s(x,y)b1          383.49837   49.32894    81.62817   4.698 0.00000270 ***
s(x,y)b2          -18.69423    3.61955     7.58360  -2.465   0.013733 *  
s(x,y)b3           38.97747    9.92196     8.32024   4.685 0.00000288 ***
s(x,y)b4          291.06177   34.93785    64.12209   4.539 0.00000579 ***
s(x,y)b5           37.63186    9.41325     8.40426   4.478 0.00000772 ***
s(x,y)b6            5.19672    0.81576     1.38509   3.752   0.000178 ***
s(x,y)b7          -13.24714    3.38663     5.03222  -2.632   0.008504 ** 
s(x,y)b8         -630.86749   77.88197   134.47015  -4.692 0.00000279 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 2.798283)

    Null deviance: 4102.2  on 4768  degrees of freedom
Residual deviance: 1860.7  on 4751  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 74
Number of Fisher Scoring iterations: 11
```

Figure A6-118 Code excerpt detailing SALSA 2D summary for common tern MRSea analysis.

Cross-Validation

A.3.85 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a slightly higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-119).



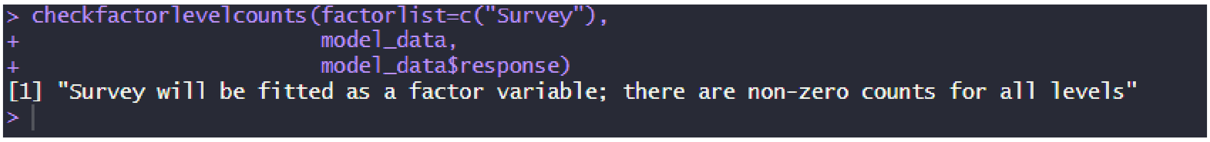
```
> cv_2d #SALSA 2D
[1] 3.594651
> cv_1d #SALSA 1D
[1] 3.535425
> cv_0d #Initial model
[1] 3.543745
```

Figure A6-119 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for common tern MRSea analysis.

Roseate Tern

Factor Level Check

A.3.86 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-120).

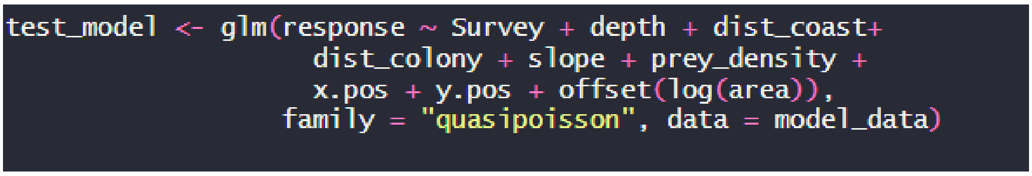


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                        model_data,
+                        model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-120 Code excerpt detailing factor level check for categorical variables considered within roseate tern MRSea modelling.

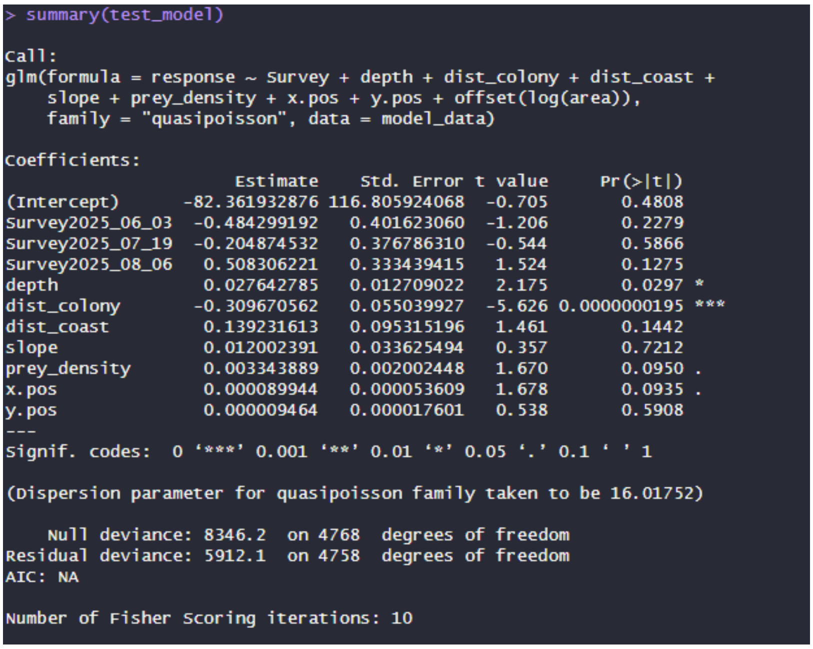
Initial GLM

A.3.87 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-121 and Figure A6-122).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-121 Code excerpt detailing the setup of the initial GLM considered within roseate tern MRSea analysis.



```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_colony + dist_coast + 
    slope + prey_density + x.pos + y.pos + offset(log(area)), 
    family = "quasipoisson", data = model_data)

Coefficients:
                    Estimate    Std. Error t value     Pr(>|t|)    
(Intercept)      -82.361932876 116.805924068  -0.705       0.4808    
Survey2025_06_03  -0.484299192   0.401623060  -1.206       0.2279    
Survey2025_07_19  -0.204874532   0.376786310  -0.544       0.5866    
Survey2025_08_06   0.508306221   0.333439415   1.524       0.1275    
depth              0.027642785   0.012709022   2.175       0.0297 *  
dist_colony       -0.309670562   0.055039927  -5.626 0.0000000195 ***
dist_coast         0.139231613   0.095315196   1.461       0.1442    
slope              0.012002391   0.033625494   0.357       0.7212    
prey_density       0.003343889   0.002002448   1.670       0.0950 .  
x.pos              0.000089944   0.000053609   1.678       0.0935 .  
y.pos              0.000009464   0.000017601   0.538       0.5908    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 16.01752)

    Null deviance: 8346.2  on 4768  degrees of freedom
Residual deviance: 5912.1  on 4758  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 10
```

Figure A6-122 Code excerpt detailing the summary of the initial GLM considered within roseate tern MRSea analysis.

Collinearity Check

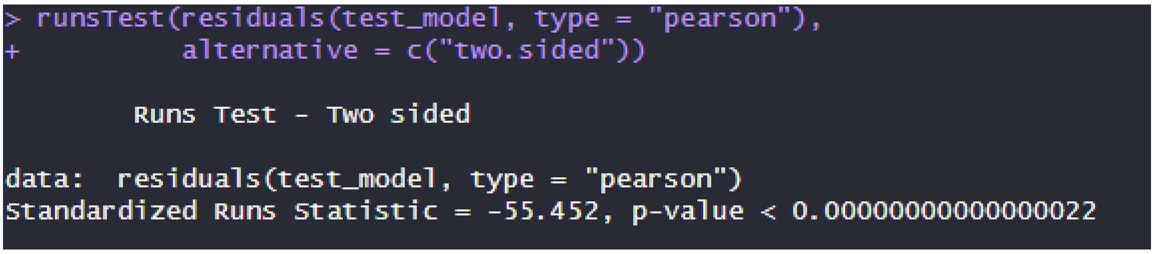
A.3.88 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold except for distance to coast which was dropped from further analysis (Figure A6-123).

```
> vif_out
                   GVIF Df GVIF^(1/(2*Df))
survey         1.052957  3        1.008638
depth          6.863480  1        2.619824
dist_colony   12.195737  1        3.492240
dist_coast    45.530551  1        6.747633
slope          1.076570  1        1.037579
prey_density   1.216540  1        1.102969
x.pos         15.979878  1        3.997484
y.pos          3.114135  1        1.764691
```

Figure A6-123 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within roseate tern MRSea analysis.

Runs Test

A.3.89 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-124).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -55.452, p-value < 0.00000000000000022
```

Figure A6-124 Code excerpt detailing runs test carried out on the initial GLM considered within roseate tern MRSea analysis.

A.3.90 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-125).

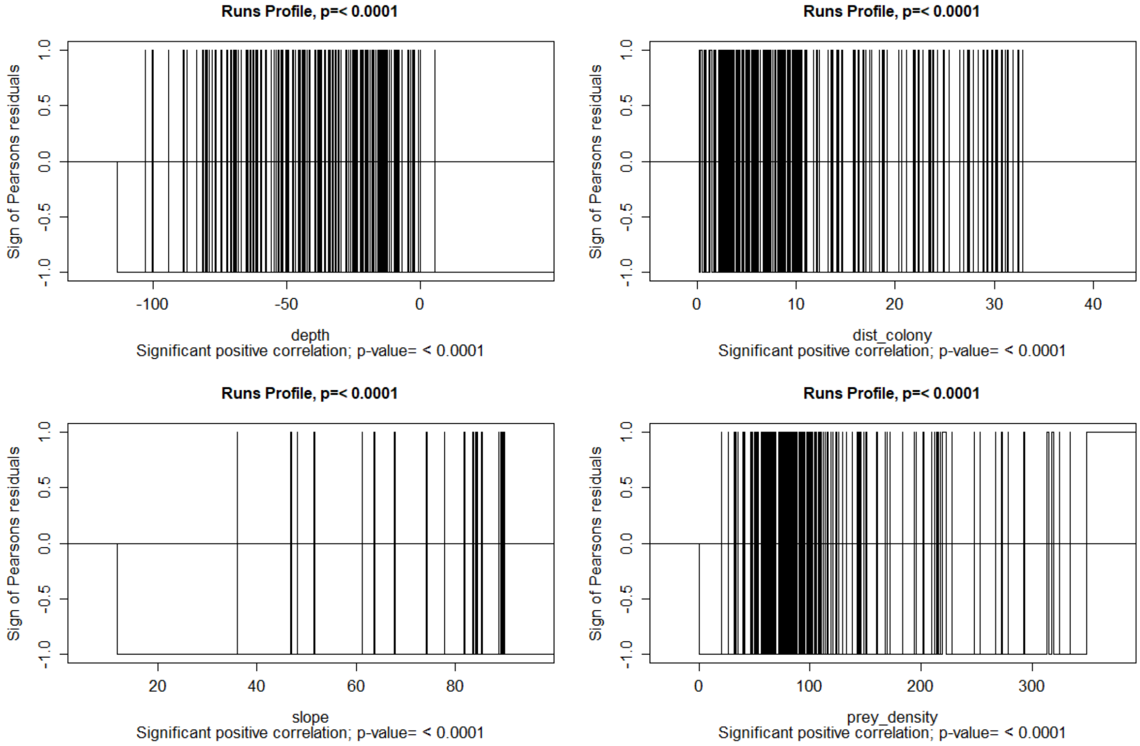


Figure A6-125 Runs profiles for explanatory variables considered within the initial GLM considered for roseate tern MRSea analysis.

Autocorrelation

A.3.91 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-126).

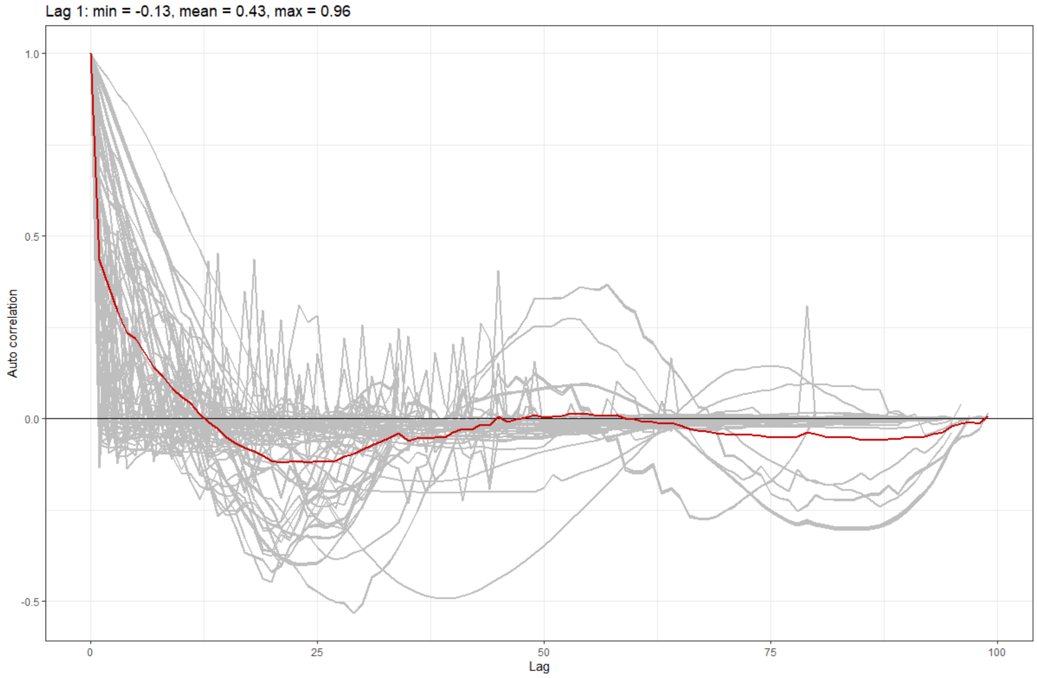


Figure A6-126 ACF plot for the initial GLM considered for roseate tern MRSea analysis.

Cumulative Residuals

A.3.92 Cumulative residuals for explanatory variables were plotted (Figure A6-127, Figure A6-128, Figure A6-129 & Figure A6-130). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

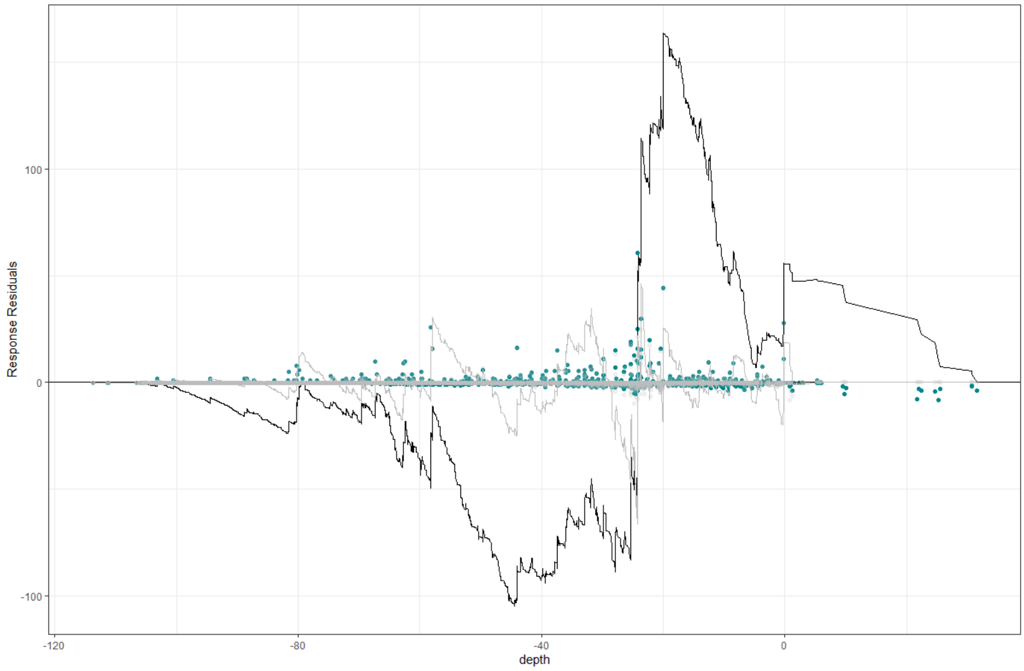


Figure A6-127 Cumulative residuals of depth plotted for the initial GLM considered for roseate tern MRSea analysis.

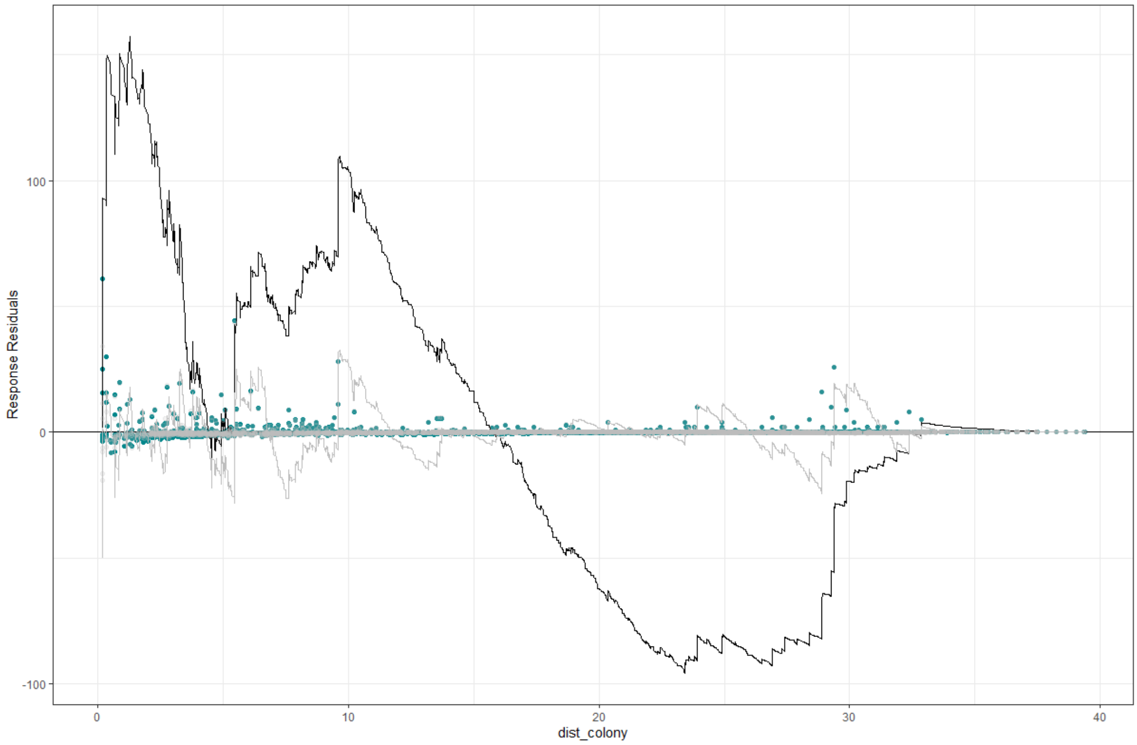


Figure A6-128 Cumulative residuals of distance to colony plotted for the initial GLM considered for roseate tern MRSea analysis.

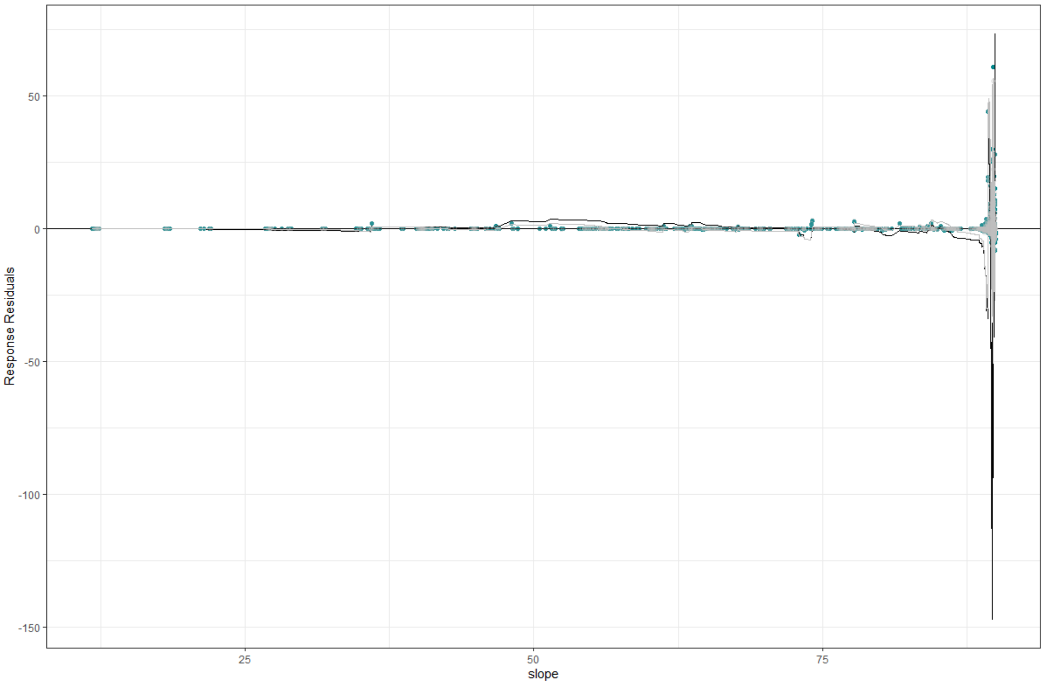


Figure A6-129 Cumulative residuals of seabed slope plotted for the initial GLM considered for roseate tern MRSea analysis.

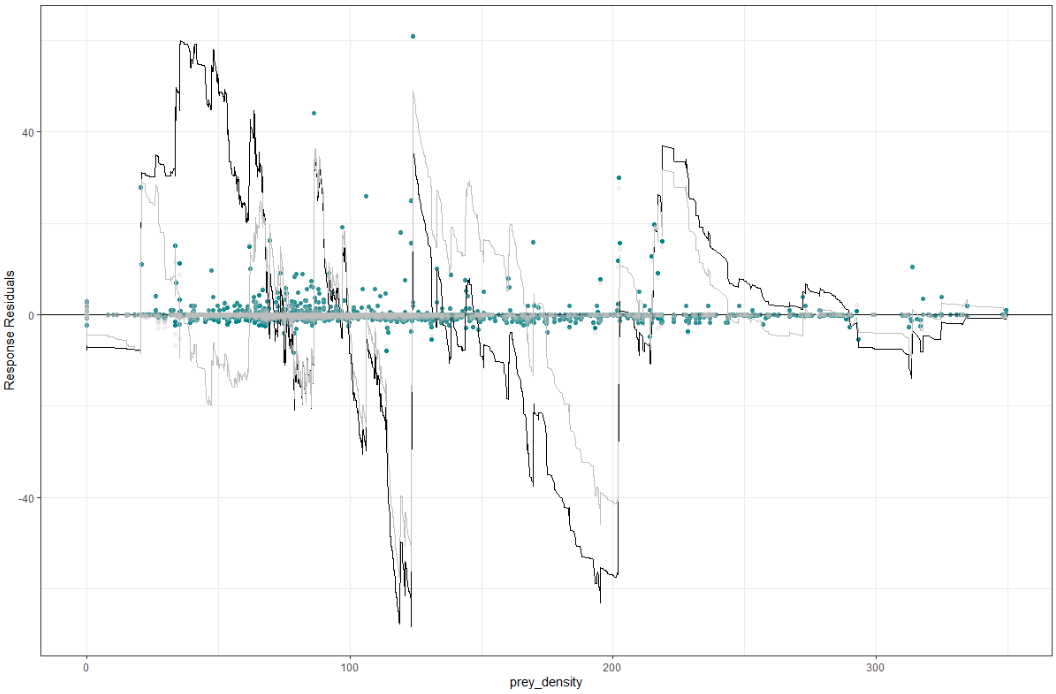
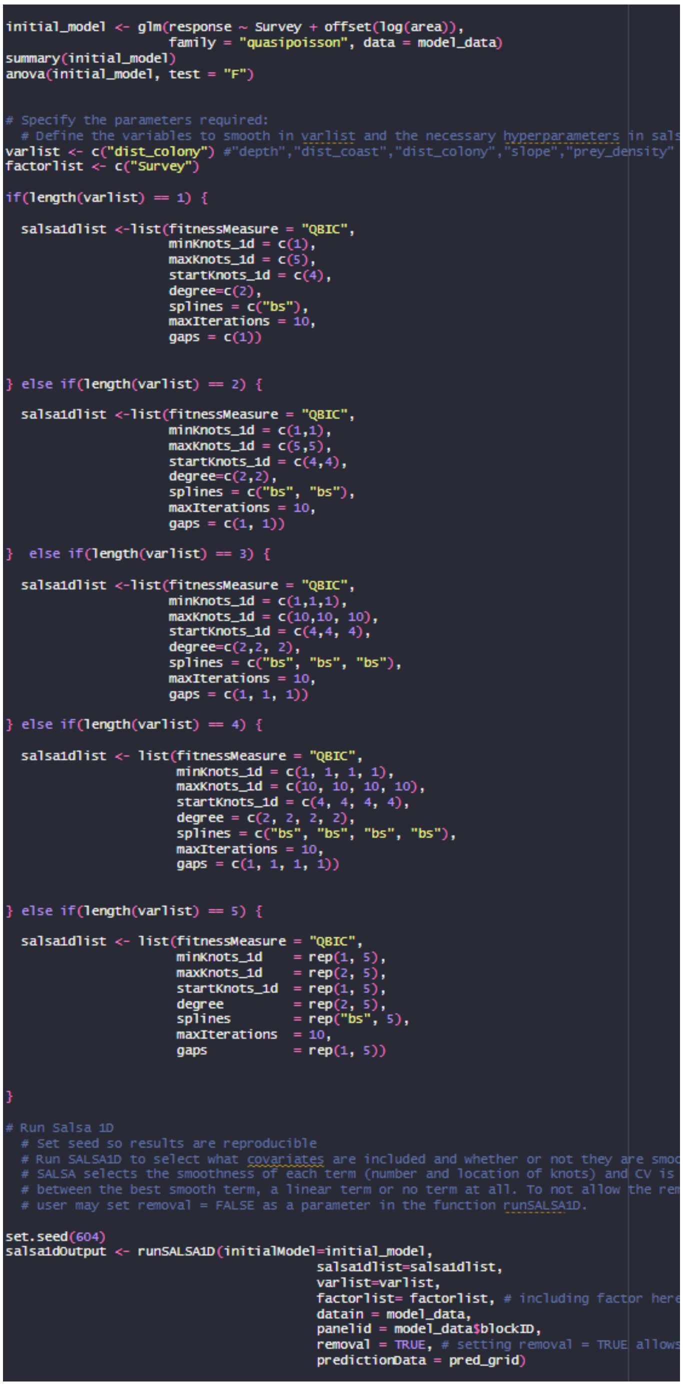


Figure A6-130 Cumulative residuals of prey density plotted for the initial GLM considered for roseate tern MRSea analysis.

SALSA 1D

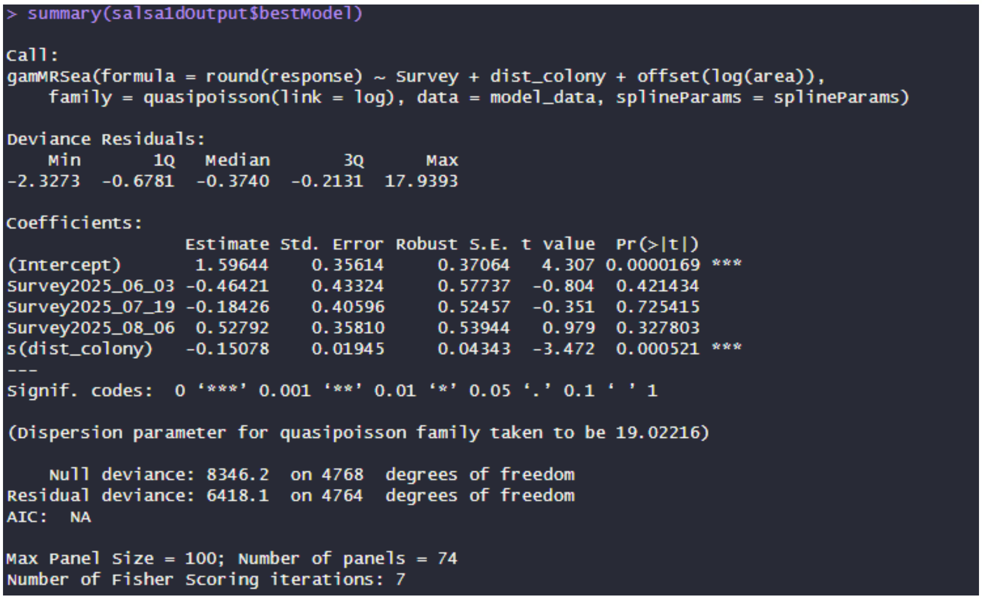
A.3.93 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.94 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). The following covariates: depth, distance to coast, seabed slope and prey density were dropped from the model at this point. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-131 & Figure A6-132).



```
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in sals
varlist <- c("dist_colony") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d    = rep(1, 5),
                      maxKnots_1d    = rep(2, 5),
                      startKnots_1d  = rep(1, 5),
                      degree         = rep(2, 5),
                      splines        = rep("bs", 5),
                      maxIterations  = 10,
                      gaps           = rep(1, 5))


}

# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smoo
  # SALSA selects the smoothness of each term (number and location of knots) and CV is
  # between the best smooth term, a linear term or no term at all. To not allow the rem
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                                        salsa1dlist=salsa1dlist,
                                        varlist=varlist,
                                        factorlist= factorlist, # including factor here
                                        datain = model_data,
                                        panelid = model_data$blockID,
                                        removal = TRUE, # setting removal = TRUE allows
                                        predictionData = pred_grid)
```

Figure A6-131 Code excerpt detailing the setup process for the SALSA 1D model considered for roseate tern MRSea analysis.



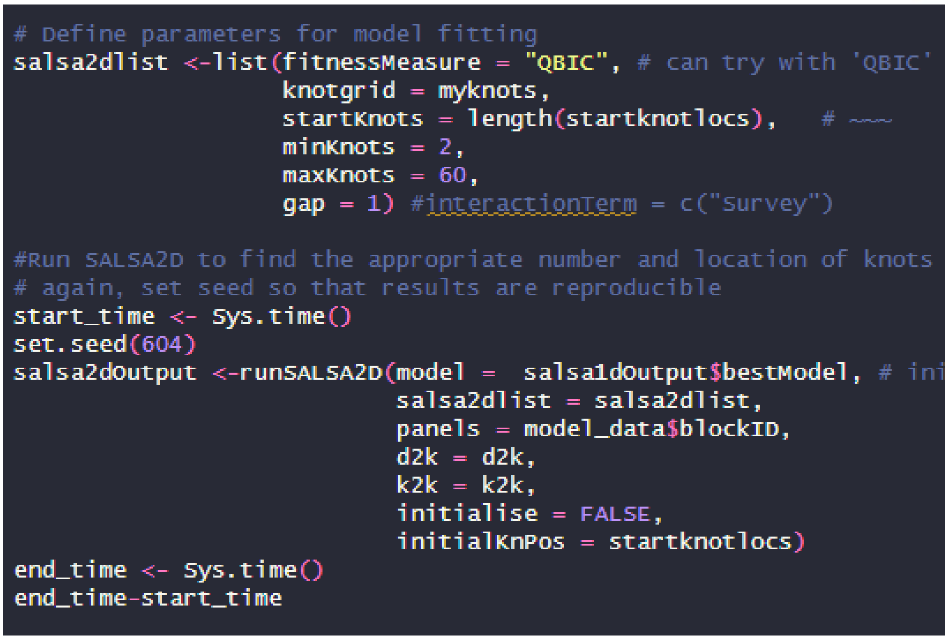
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + dist_colony + offset(log(area)),
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.3273  -0.6781  -0.3740  -0.2131  17.9393

Coefficients:
                 Estimate Std. Error Robust S.E. t value  Pr(>|t|)
(Intercept)       1.59644    0.35614     0.37064   4.307 0.0000169 ***
Survey2025_06_03 -0.46421    0.43324     0.57737  -0.804  0.421434
Survey2025_07_19 -0.18426    0.40596     0.52457  -0.351  0.725415
Survey2025_08_06  0.52792    0.35810     0.53944   0.979  0.327803
s(dist_colony)   -0.15078    0.01945     0.04343  -3.472  0.000521 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 19.02216)

    Null deviance: 8346.2  on 4768  degrees of freedom
Residual deviance: 6418.1  on 4764  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 74
Number of Fisher Scoring iterations: 7
```

Figure A6-132 Code excerpt detailing the summary for the SALSA 1D model for roseate tern MRSea analysis.

SALSA 2D

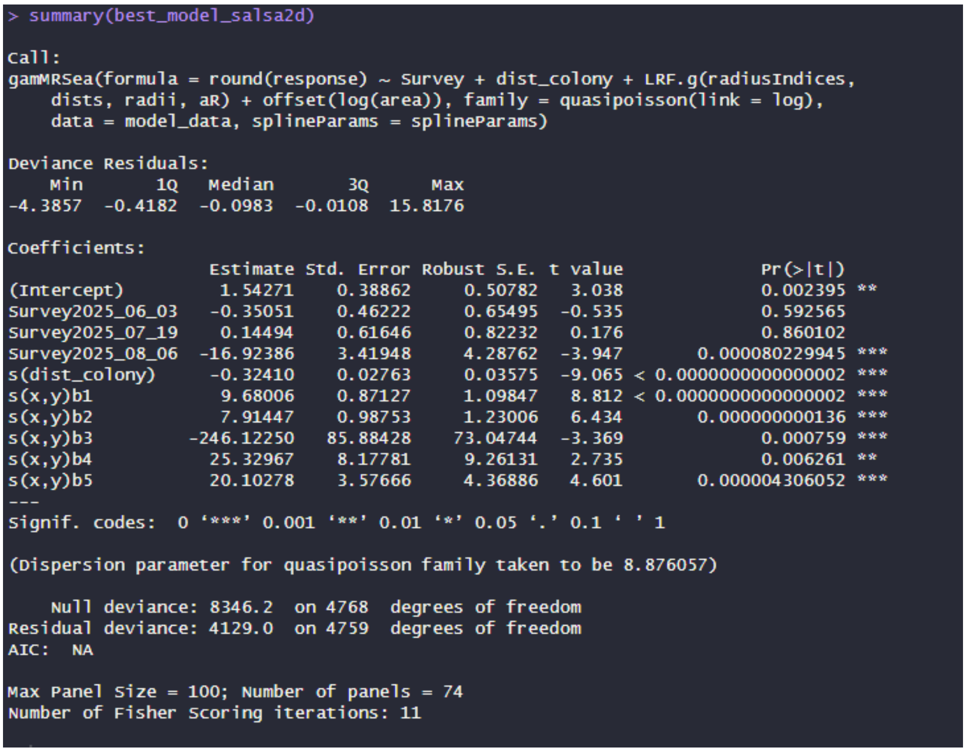
A.3.95 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.96 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-133 & Figure A6-134).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC'
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),   # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2dOutput <-runSALSA2D(model =  salsa1dOutput$bestModel, # ini
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-133 Code excerpt for SALSA 2D model setup for roseate tern MRSea analysis.

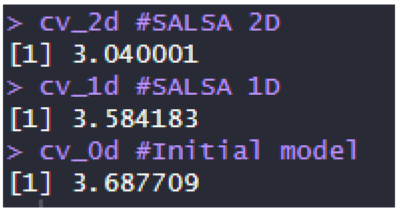


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + dist_colony + LRF.g(radiusIndices,
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log),
    data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.3857  -0.4182  -0.0983  -0.0108  15.8176

Coefficients:
                   Estimate Std. Error Robust S.E. t value             Pr(>|t|)
(Intercept)         1.54271    0.38862     0.50782   3.038             0.002395 **
Survey2025_06_03   -0.35051    0.46222     0.65495  -0.535             0.592565
Survey2025_07_19    0.14494    0.61646     0.82232   0.176             0.860102
Survey2025_08_06  -16.92386    3.41948     4.28762  -3.947         0.000080229945 ***
s(dist_colony)     -0.32410    0.02763     0.03575  -9.065 < 0.0000000000000002 ***
s(x,y)b1            9.68006    0.87127     1.09847   8.812 < 0.0000000000000002 ***
s(x,y)b2            7.91447    0.98753     1.23006   6.434         0.000000000136 ***
s(x,y)b3         -246.12250   85.88428    73.04744  -3.369             0.000759 ***
s(x,y)b4           25.32967    8.17781     9.26131   2.735             0.006261 **
s(x,y)b5           20.10278    3.57666     4.36886   4.601         0.000004306052 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 8.876057)

    Null deviance: 8346.2  on 4768  degrees of freedom
Residual deviance: 4129.0  on 4759  degrees of freedom
AIC:  NA

Max Panel Size = 100; Number of panels = 74
Number of Fisher Scoring iterations: 11
```

Figure A6-134 Code excerpt detailing SALSA 2D summary for roseate tern MRSea analysis.

Cross-Validation

A.3.97 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. Although the 2D model exhibited a slightly higher CV score than the 1D model, indicating a marginally poorer fit, it was retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-135).



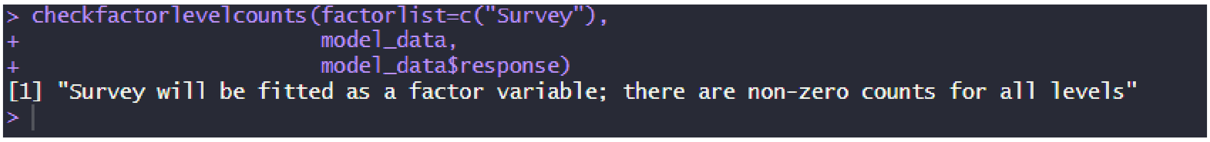
```
> cv_2d #SALSA 2D
[1] 3.040001
> cv_1d #SALSA 1D
[1] 3.584183
> cv_0d #Initial model
[1] 3.687709
```

Figure A6-135 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for roseate tern MRSea analysis.

Red-throated Diver

Factor Level Check

A.3.98 To fit the model, it was necessary for all levels of any categorical variables to have non-zero counts. The only categorical variable, “Survey” (a proxy for month), was checked and required no further action (Figure A6-136).

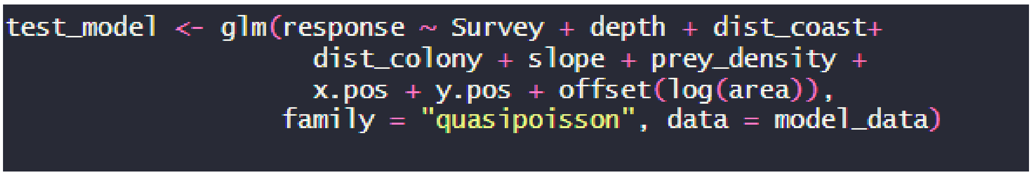


```
> checkfactorlevelcounts(factorlist=c("Survey"),
+                         model_data,
+                         model_data$response)
[1] "Survey will be fitted as a factor variable; there are non-zero counts for all levels"
> 
```

Figure A6-136 Code excerpt detailing factor level check for categorical variables considered within red-throated diver MRSea modelling.

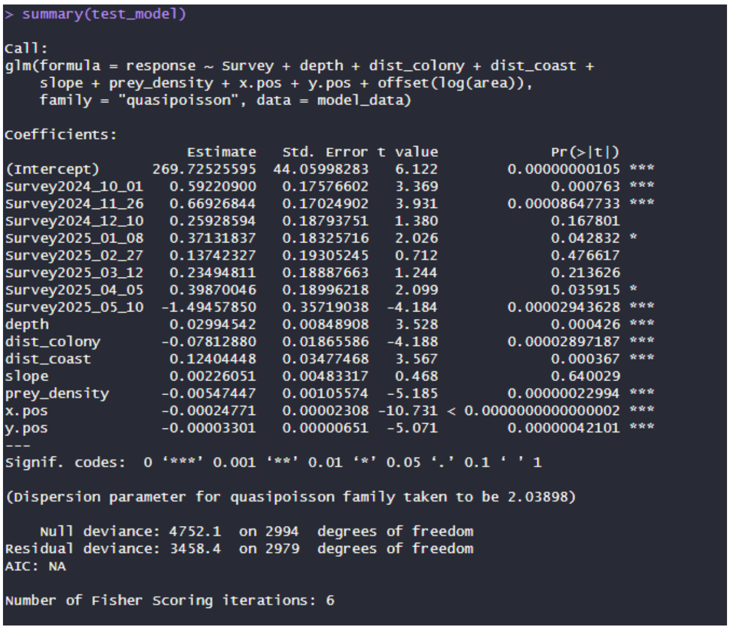
Initial GLM

A.3.99 Before creating more complex models, a simple Generalised Linear Model (GLM) was developed and run as an initial model. A code excerpts detailing the setup and summary of the initial GLM are provided below (Figure A6-137 and Figure A6-138).



```
test_model <- glm(response ~ Survey + depth + dist_coast+
                    dist_colony + slope + prey_density +
                    x.pos + y.pos + offset(log(area)),
                  family = "quasipoisson", data = model_data)
```

Figure A6-137 Code excerpt detailing the setup of the initial GLM considered within red-throated diver MRSea analysis.

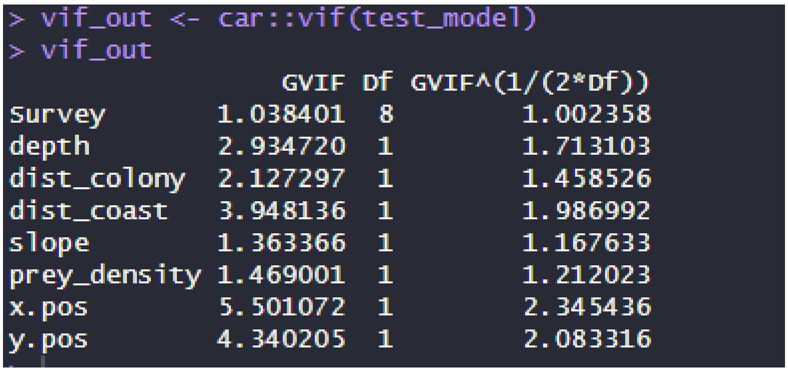


```
> summary(test_model)

Call:
glm(formula = response ~ Survey + depth + dist_colony + dist_coast + 
    slope + prey_density + x.pos + y.pos + offset(log(area)), 
    family = "quasipoisson", data = model_data)

Coefficients:
                    Estimate   Std. Error t value              Pr(>|t|)    
(Intercept)     269.72525595  44.05998283   6.122        0.00000000105 ***
Survey2024_10_01   0.59220900   0.17576602   3.369             0.000763 ***
Survey2024_11_26   0.66926844   0.17024902   3.931        0.00008647733 ***
Survey2024_12_10   0.25928594   0.18793751   1.380             0.167801    
Survey2025_01_08   0.37131837   0.18325716   2.026             0.042832 *  
Survey2025_02_27   0.13742327   0.19305245   0.712             0.476617    
Survey2025_03_12   0.23494811   0.18887663   1.244             0.213626    
Survey2025_04_05   0.39870046   0.18996218   2.099             0.035915 *  
Survey2025_05_10  -1.49457850   0.35719038  -4.184        0.00002943628 ***
depth              0.02994542   0.00848908   3.528             0.000426 ***
dist_colony       -0.07812880   0.01865586  -4.188        0.00002897187 ***
dist_coast         0.12404448   0.03477468   3.567             0.000367 ***
slope              0.00226051   0.00483317   0.468             0.640029    
prey_density      -0.00547447   0.00105574  -5.185        0.00000022994 ***
x.pos             -0.00024771   0.00002308 -10.731 < 0.0000000000000002 ***
y.pos             -0.00003301   0.00000651  -5.071        0.00000042101 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 2.03898)

    Null deviance: 4752.1  on 2994  degrees of freedom
Residual deviance: 3458.4  on 2979  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 6
```

Figure A6-138 Code excerpt detailing the summary of the initial GLM considered within red-throated diver MRSea analysis.

Collinearity Check

A.3.100 The collinearity of explanatory variables was initially assessed by observing Generalised Variance Inflation Factors (GVIFs). Covariates were removed if strong collinearity was detected (GVIF value over 5). All adjusted GVIF values were below this threshold (Figure A6-139).

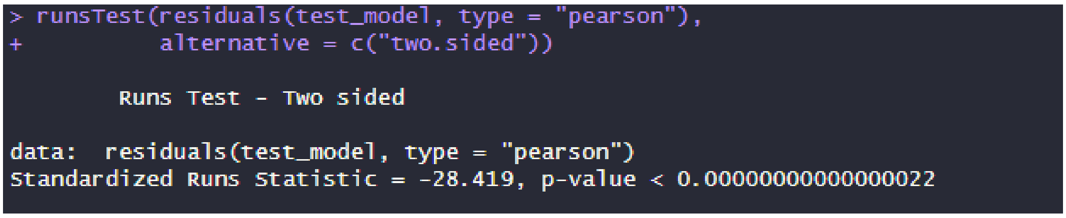


```
> vif_out <- car::vif(test_model)
> vif_out
                 GVIF Df GVIF^(1/(2*Df))
Survey       1.038401  8        1.002358
depth        2.934720  1        1.713103
dist_colony  2.127297  1        1.458526
dist_coast   3.948136  1        1.986992
slope        1.363366  1        1.167633
prey_density 1.469001  1        1.212023
x.pos        5.501072  1        2.345436
y.pos        4.340205  1        2.083316
```

Figure A6-139 Code expert detailing collinearity check of explanatory variables considered for the initial GLM considered within red-throated diver MRSea analysis.

Runs Test

A.3.101 A runs test was conducted on the initial GLM, indicating significant residual correlation due to the highly significant p-value. The code excerpt detailing the runs test is provided below (Figure A6-140).



```
> runsTest(residuals(test_model, type = "pearson"),
+          alternative = c("two.sided"))

        Runs Test - Two sided

data:  residuals(test_model, type = "pearson")
Standardized Runs Statistic = -28.419, p-value < 0.00000000000000022
```

Figure A6-140 Code excerpt detailing runs test carried out on the initial GLM considered within red-throated diver MRSea analysis.

A.3.102 This is supported by a visual examination of the runs profile plots which highlight the presence of non-randomness in the residuals within the initial GLM (Figure A6-141).

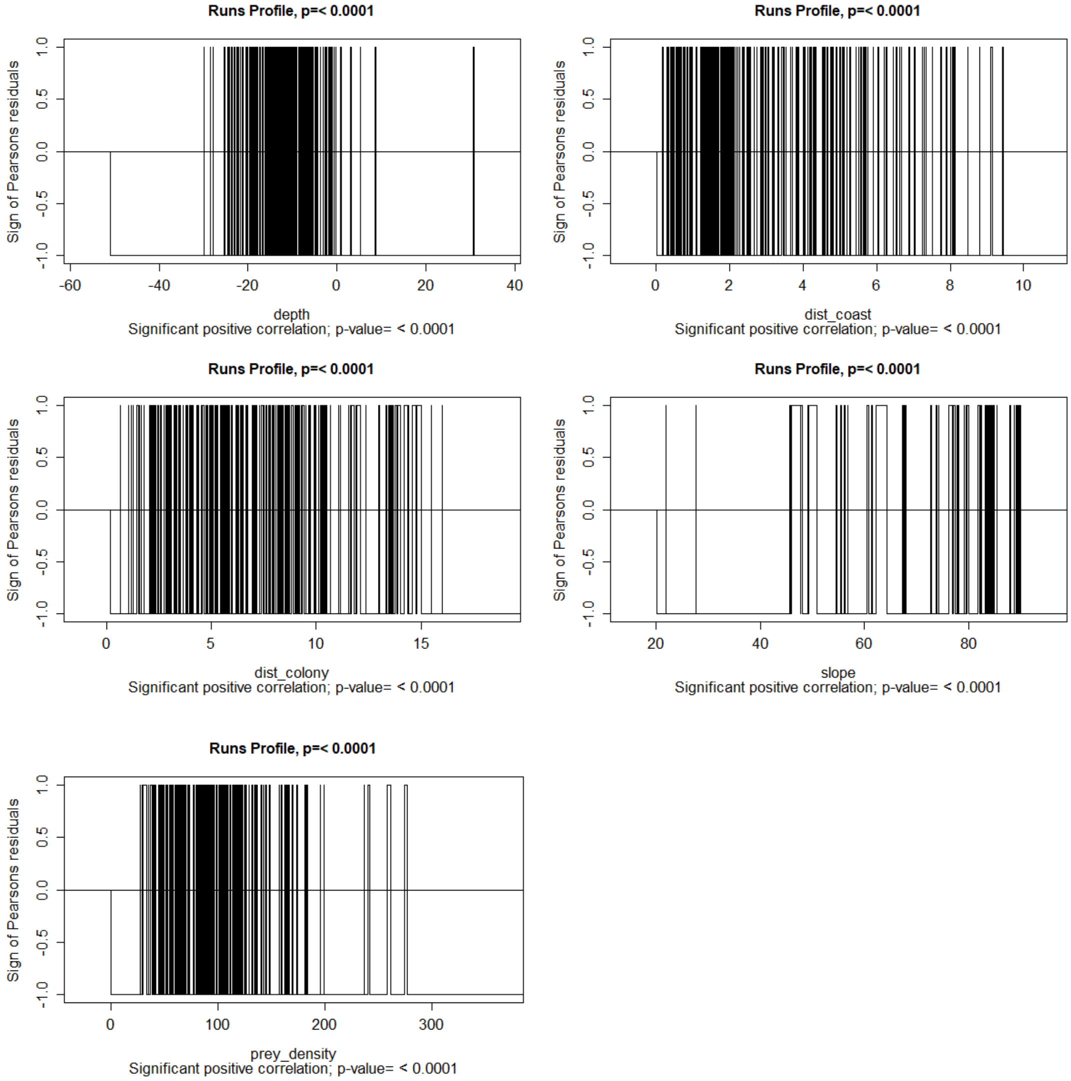


Figure A6-141 Runs profiles for explanatory variables considered within the initial GLM considered for red-throated diver MRSea analysis.

Autocorrelation

A.3.103 Given the observed correlation, a blocking structure was introduced to account for within-survey dependencies. This structure combined Survey ID and Transect ID, enabling the model to treat observations from each transect within a survey as correlated while assuming independence across different transects and surveys. Assessment using an Auto-Correlation Function (ACF) confirmed its effectiveness, as both the mean residual correlation and within-block correlations rapidly approached zero, indicating the blocking structure was appropriate (Figure A6-142).

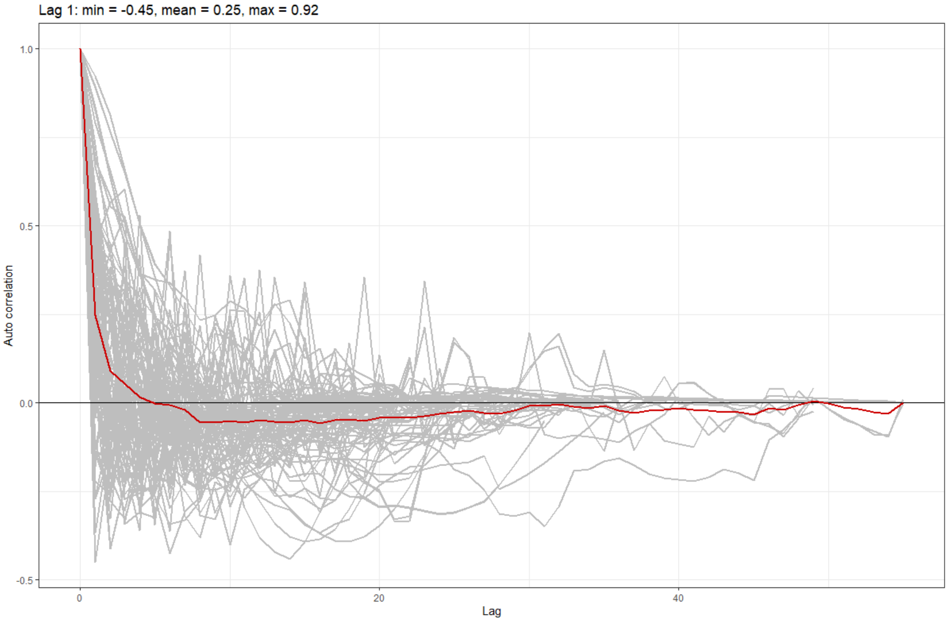


Figure A6-142 ACF plot for the initial GLM considered for red-throated diver MRSea analysis.

Cumulative Residuals

A.3.104 Cumulative residuals for explanatory variables were plotted (Figure A6-143, Figure A6-144, Figure A6-145, Figure A6-146 & Figure A6-147). The black line indicates the modelled cumulative residuals, while the grey line represents the expected fit. Systematic patterns of over- and under-prediction were observed across multiple variables which highlighted the need for a more complex, non-linear, modelling approach.

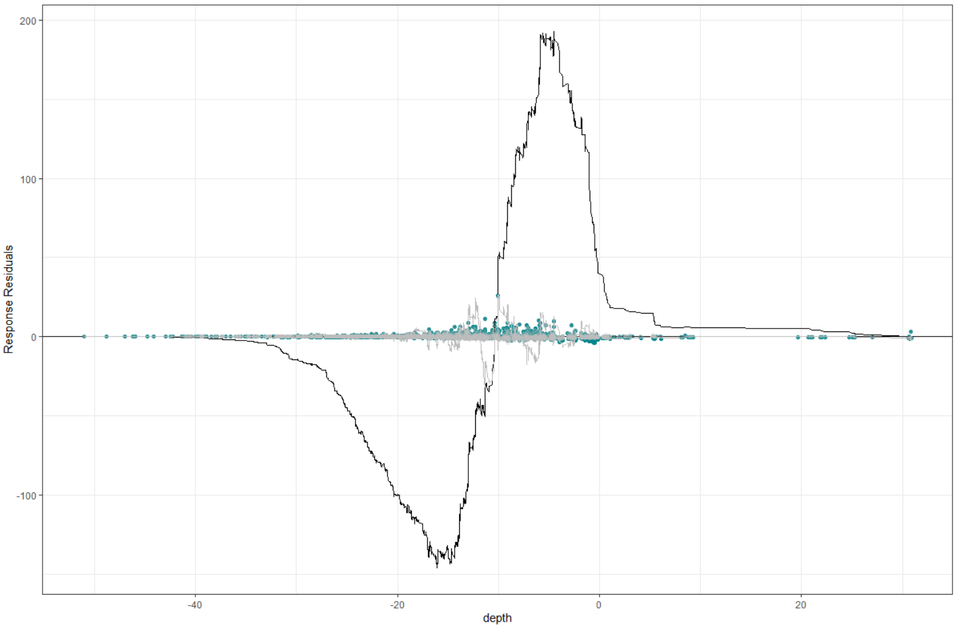


Figure A6-143 Cumulative residuals of depth plotted for the initial GLM considered for red-throated diver MRSea analysis.

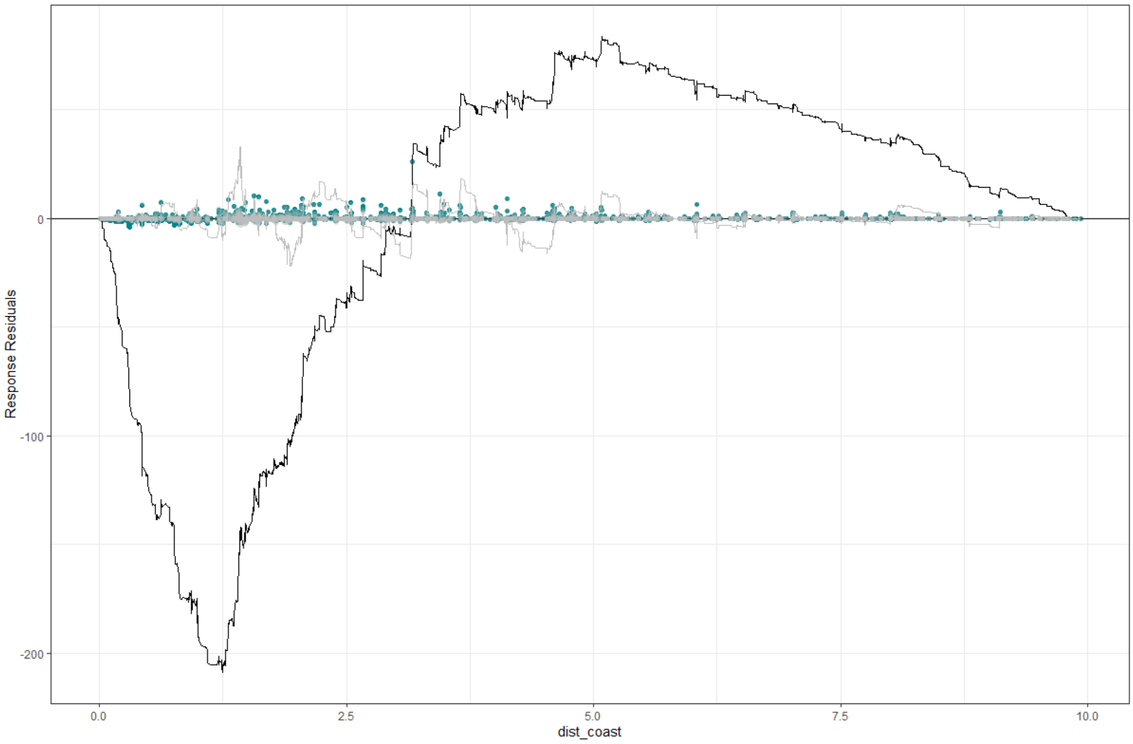


Figure A6-144 Cumulative residuals of distance to coast plotted for the initial GLM considered for red-throated diver MRSea analysis.

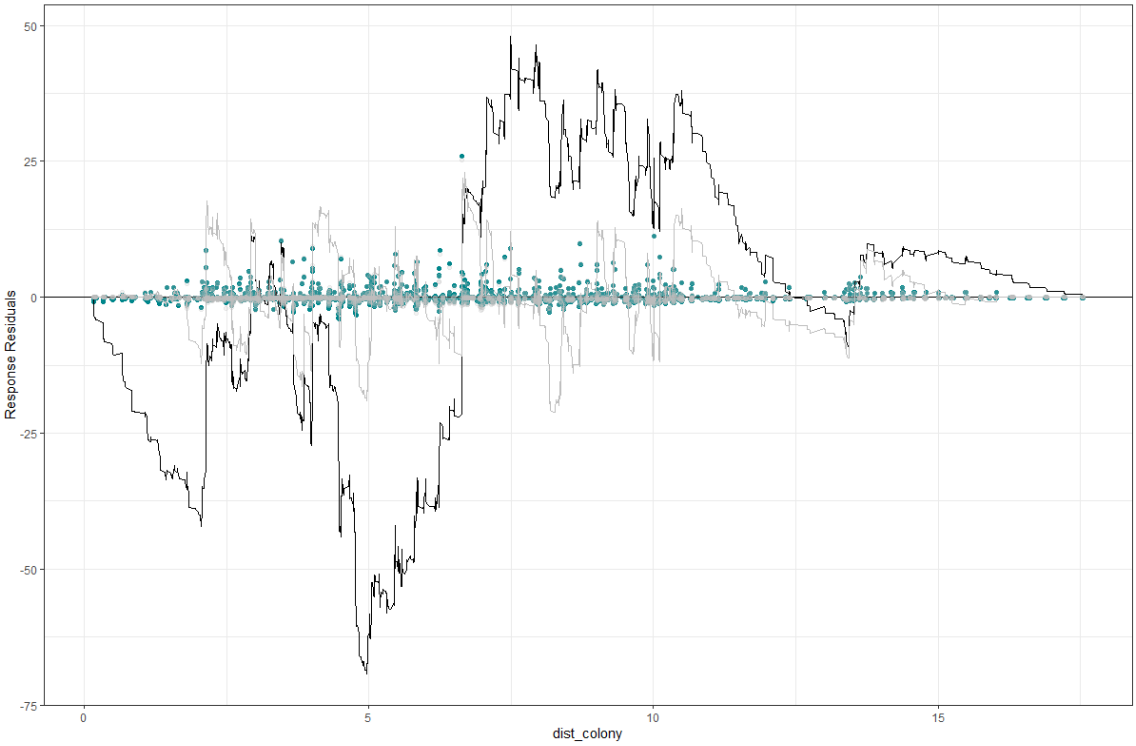


Figure A6-145 Cumulative residuals of distance to colony plotted for the initial GLM considered for red-throated diver MRSea analysis.

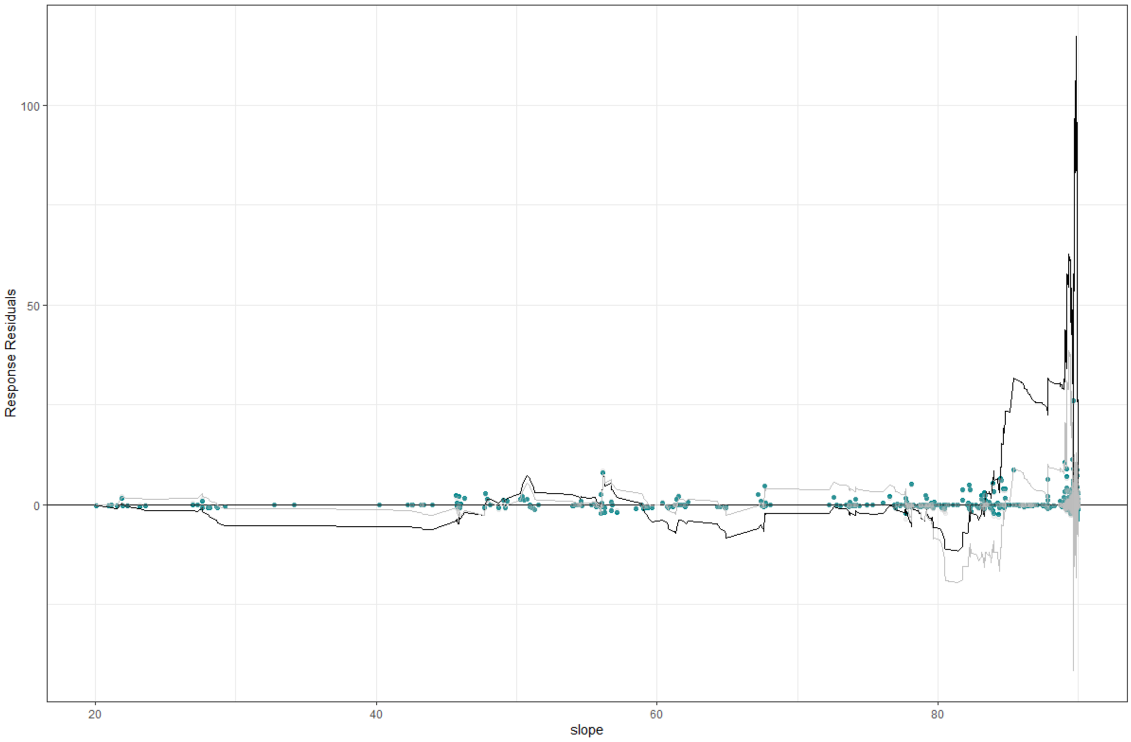


Figure A6-146 Cumulative residuals of seabed slope plotted for the initial GLM considered for red-throated diver MRSea analysis.

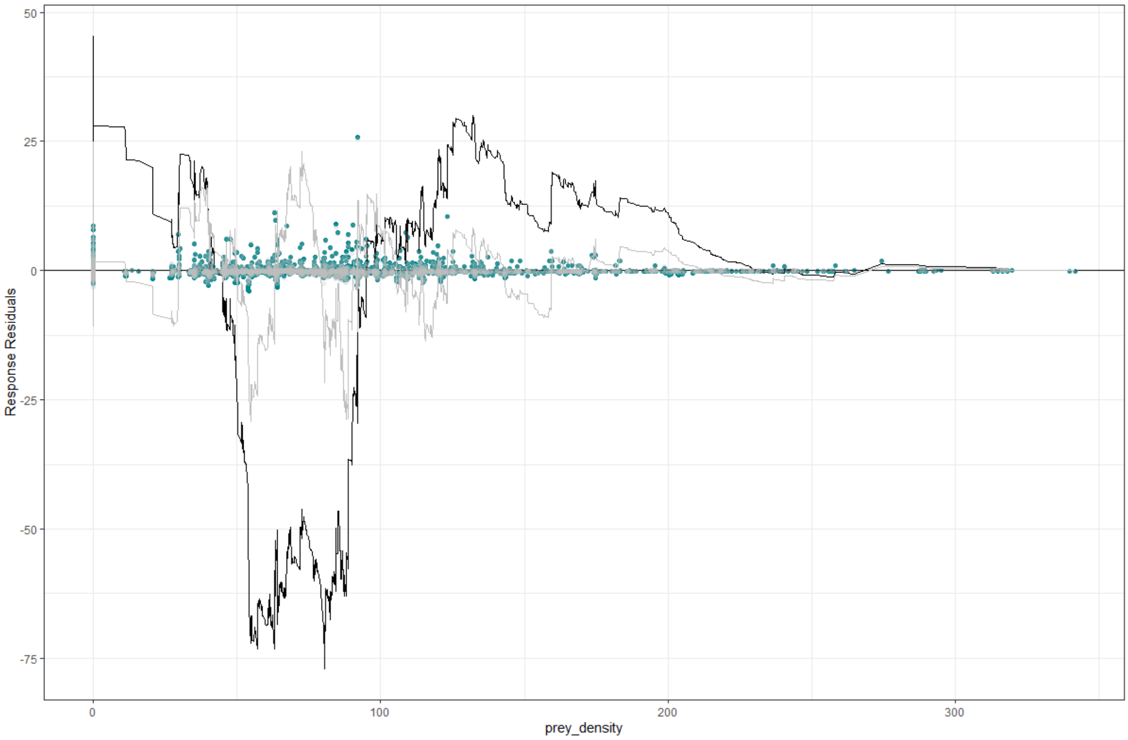
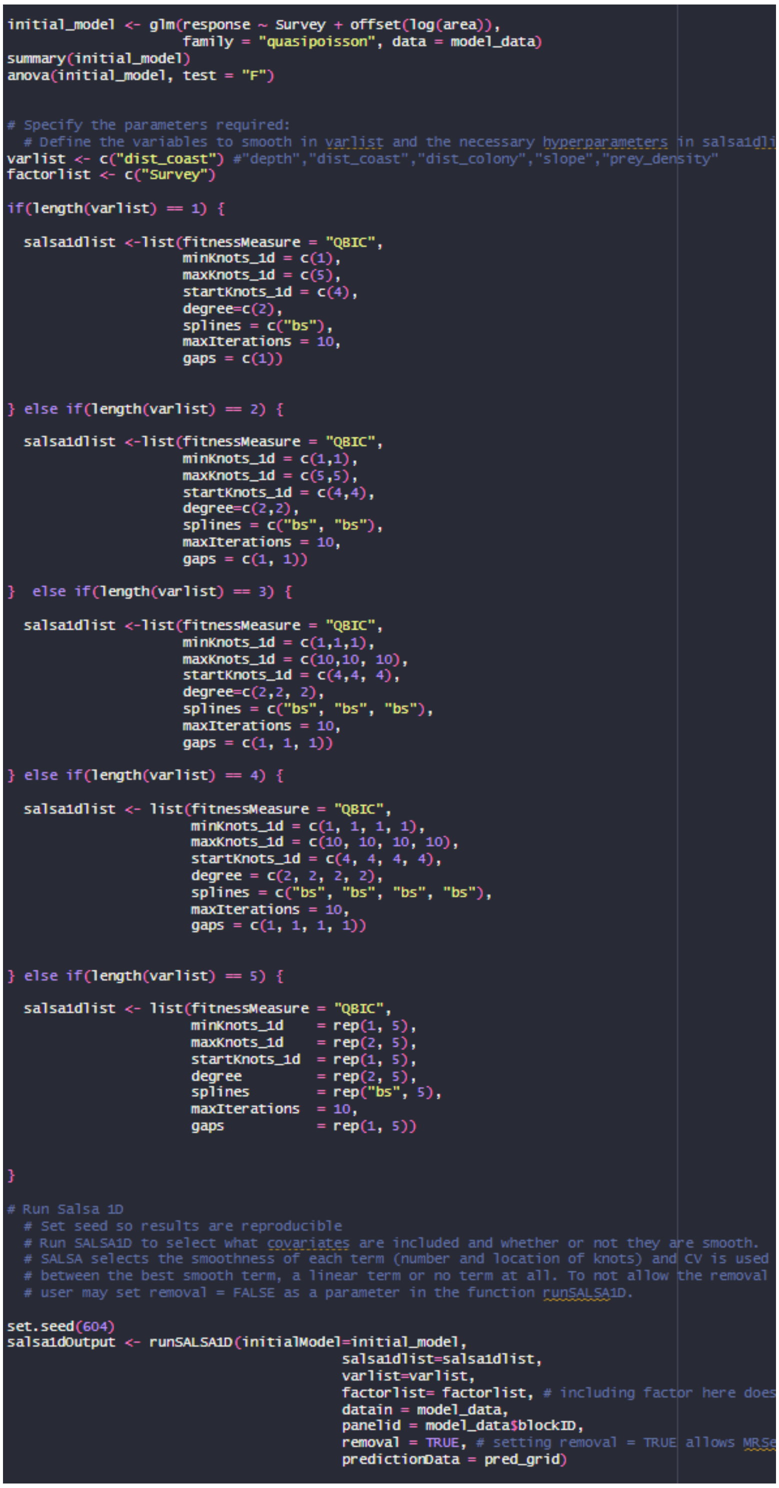


Figure A6-147 Cumulative residuals of prey density plotted for the initial GLM considered for red-throated diver MRSea analysis.

SALSA 1D

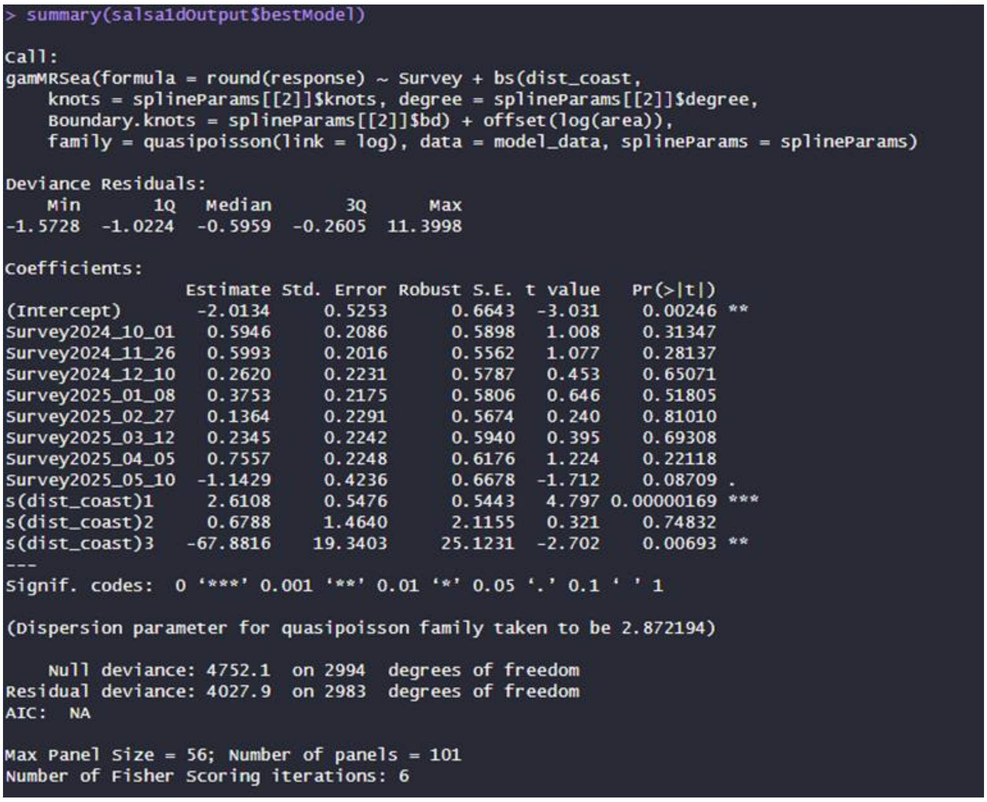
A.3.105 After establishing an appropriate blocking structure and confirming the need for a non-linear modelling approach, the 1D SALSA algorithm was applied using the parameters specified below. Spline parameters were generated with the makesplineParams function (default degree = 2).

A.3.106 The SALSA 1D function fits multiple candidate models and evaluates them using the Quasi-Bayesian Information Criterion (QBIC). The following covariates: depth, distance to colony, seabed slope and prey density were dropped from the model at this point. The model with the lowest QBIC, representing the optimal trade-off between fit and complexity, is retained as the "best model." Code excerpts for model setup and summary are provided below (Figure A6-148 & Figure A6-149).



```r
initial_model <- glm(response ~ Survey + offset(log(area)),
                     family = "quasipoisson", data = model_data)
summary(initial_model)
anova(initial_model, test = "F")


# Specify the parameters required:
  # Define the variables to smooth in varlist and the necessary hyperparameters in salsa1dli
varlist <- c("dist_coast") #"depth","dist_coast","dist_colony","slope","prey_density"
factorlist <- c("Survey")

if(length(varlist) == 1) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1),
                     maxKnots_1d = c(5),
                     startKnots_1d = c(4),
                     degree=c(2),
                     splines = c("bs"),
                     maxIterations = 10,
                     gaps = c(1))


} else if(length(varlist) == 2) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1),
                     maxKnots_1d = c(5,5),
                     startKnots_1d = c(4,4),
                     degree=c(2,2),
                     splines = c("bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1))

}  else if(length(varlist) == 3) {

  salsa1dlist <-list(fitnessMeasure = "QBIC",
                     minKnots_1d = c(1,1,1),
                     maxKnots_1d = c(10,10, 10),
                     startKnots_1d = c(4,4, 4),
                     degree=c(2,2, 2),
                     splines = c("bs", "bs", "bs"),
                     maxIterations = 10,
                     gaps = c(1, 1, 1))

} else if(length(varlist) == 4) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d = c(1, 1, 1, 1),
                      maxKnots_1d = c(10, 10, 10, 10),
                      startKnots_1d = c(4, 4, 4, 4),
                      degree = c(2, 2, 2, 2),
                      splines = c("bs", "bs", "bs", "bs"),
                      maxIterations = 10,
                      gaps = c(1, 1, 1, 1))


} else if(length(varlist) == 5) {

  salsa1dlist <- list(fitnessMeasure = "QBIC",
                      minKnots_1d     = rep(1, 5),
                      maxKnots_1d     = rep(2, 5),
                      startKnots_1d   = rep(1, 5),
                      degree          = rep(2, 5),
                      splines         = rep("bs", 5),
                      maxIterations   = 10,
                      gaps            = rep(1, 5))


}

# Run Salsa 1D
  # Set seed so results are reproducible
  # Run SALSA1D to select what covariates are included and whether or not they are smooth.
  # SALSA selects the smoothness of each term (number and location of knots) and CV is used
  # between the best smooth term, a linear term or no term at all. To not allow the removal
  # user may set removal = FALSE as a parameter in the function runSALSA1D.

set.seed(604)
salsa1dOutput <- runSALSA1D(initialModel=initial_model,
                                        salsa1dlist=salsa1dlist,
                                        varlist=varlist,
                                        factorlist= factorlist, # including factor here does
                                        datain = model_data,
                                        panelid = model_data$blockID,
                                        removal = TRUE, # setting removal = TRUE allows MRSe
                                        predictionData = pred_grid)
```

Figure A6-148 Code excerpt detailing the setup process for the SALSA 1D model considered for red-throated diver MRSea analysis.



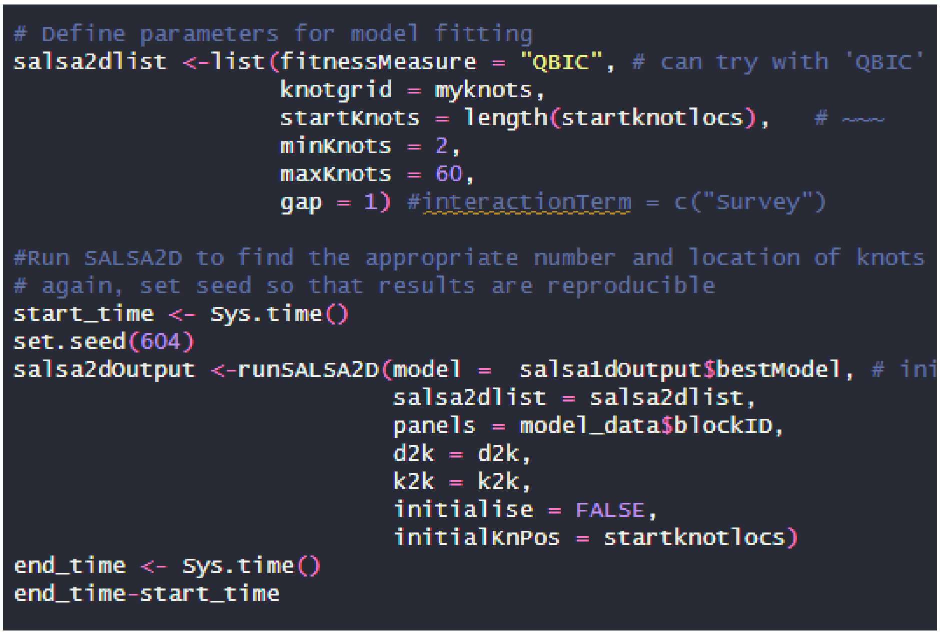
```
> summary(salsa1dOutput$bestModel)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(dist_coast,
    knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree,
    Boundary.knots = splineParams[[2]]$bd) + offset(log(area)),
    family = quasipoisson(link = log), data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5728  -1.0224  -0.5959  -0.2605  11.3998

Coefficients:
                 Estimate Std. Error Robust S.E. t value   Pr(>|t|)
(Intercept)       -2.0134     0.5253      0.6643  -3.031    0.00246 **
Survey2024_10_01   0.5946     0.2086      0.5898   1.008    0.31347
Survey2024_11_26   0.5993     0.2016      0.5562   1.077    0.28137
Survey2024_12_10   0.2620     0.2231      0.5787   0.453    0.65071
Survey2025_01_08   0.3753     0.2175      0.5806   0.646    0.51805
Survey2025_02_27   0.1364     0.2291      0.5674   0.240    0.81010
Survey2025_03_12   0.2345     0.2242      0.5940   0.395    0.69308
Survey2025_04_05   0.7557     0.2248      0.6176   1.224    0.22118
Survey2025_05_10  -1.1429     0.4236      0.6678  -1.712    0.08709 .
s(dist_coast)1     2.6108     0.5476      0.5443   4.797 0.00000169 ***
s(dist_coast)2     0.6788     1.4640      2.1155   0.321    0.74832
s(dist_coast)3   -67.8816    19.3403     25.1231  -2.702    0.00693 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 2.872194)

    Null deviance: 4752.1  on 2994  degrees of freedom
Residual deviance: 4027.9  on 2983  degrees of freedom
AIC:  NA

Max Panel Size = 56; Number of panels = 101
Number of Fisher Scoring iterations: 6
```

Figure A6-149 Code excerpt detailing the summary for the SALSA 1D model for red-throated diver MRSea analysis.

SALSA 2D

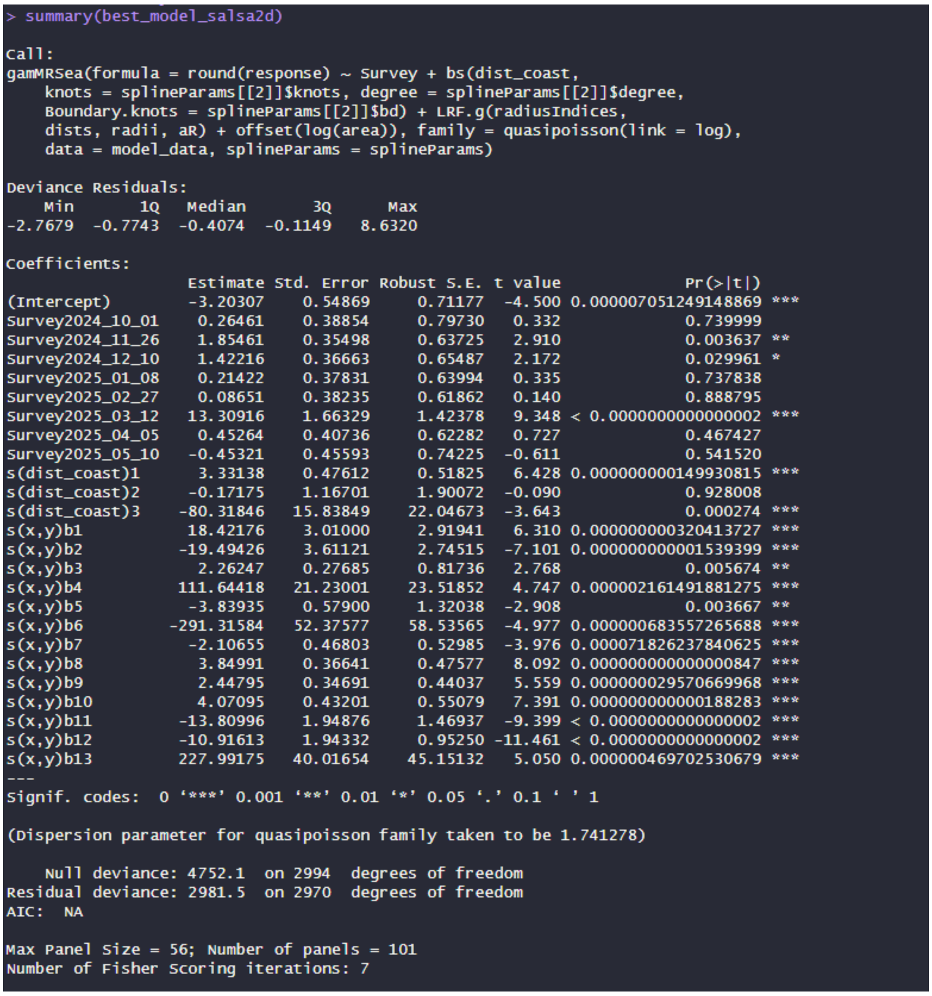
A.3.107 The SALSA 2D framework was applied to extend the best-fitting 1D model by fitting a Complex Region Spatial Smoother (CReSS). Bird count was specified as the response variable, with x.pos and y.pos as spatial coordinates and log(area) included as an offset. The model assumed a quasi-Poisson error distribution with a log link and incorporated Transect ID as a panel identifier.

A.3.108 SALSA 2D evaluated multiple candidate models and selected the one with the lowest QBIC score as the “best-fitting 2D model.” Code excerpts for model setup and summary are provided below (Figure A6-150 & Figure A6-151).



```
# Define parameters for model fitting
salsa2dlist <-list(fitnessMeasure = "QBIC", # can try with 'QBIC'
                   knotgrid = myknots,
                   startKnots = length(startknotlocs),    # ~~~
                   minKnots = 2,
                   maxKnots = 60,
                   gap = 1) #interactionTerm = c("Survey")

#Run SALSA2D to find the appropriate number and location of knots
# again, set seed so that results are reproducible
start_time <- Sys.time()
set.seed(604)
salsa2doutput <-runSALSA2D(model =  salsa1doutput$bestModel, # ini
                           salsa2dlist = salsa2dlist,
                           panels = model_data$blockID,
                           d2k = d2k,
                           k2k = k2k,
                           initialise = FALSE,
                           initialKnPos = startknotlocs)
end_time <- Sys.time()
end_time-start_time
```

Figure A6-150 Code excerpt for SALSA 2D model setup for red-throated diver MRSea analysis.

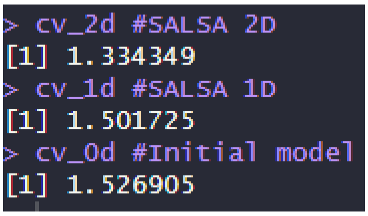


```
> summary(best_model_salsa2d)

Call:
gamMRSea(formula = round(response) ~ Survey + bs(dist_coast,
    knots = splineParams[[2]]$knots, degree = splineParams[[2]]$degree,
    Boundary.knots = splineParams[[2]]$bd) + LRF.g(radiusIndices,
    dists, radii, aR) + offset(log(area)), family = quasipoisson(link = log),
    data = model_data, splineParams = splineParams)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7679  -0.7743  -0.4074  -0.1149   8.6320

Coefficients:
                   Estimate Std. Error Robust S.E. t value             Pr(>|t|)
(Intercept)        -3.20307    0.54869     0.71177  -4.500 0.000007051249148869 ***
Survey2024_10_01    0.26461    0.38854     0.79730   0.332             0.739999
Survey2024_11_26    1.85461    0.35498     0.63725   2.910             0.003637 **
Survey2024_12_10    1.42216    0.36663     0.65487   2.172             0.029961 *
Survey2025_01_08    0.21422    0.37831     0.63994   0.335             0.737838
Survey2025_02_27    0.08651    0.38235     0.61862   0.140             0.888795
Survey2025_03_12   13.30916    1.66329     1.42378   9.348 < 0.0000000000000002 ***
Survey2025_04_05    0.45264    0.40736     0.62282   0.727             0.467427
Survey2025_05_10   -0.45321    0.45593     0.74225  -0.611             0.541520
s(dist_coast)1      3.33138    0.47612     0.51825   6.428 0.000000000149930815 ***
s(dist_coast)2     -0.17175    1.16701     1.90072  -0.090             0.928008
s(dist_coast)3    -80.31846   15.83849    22.04673  -3.643             0.000274 ***
s(x,y)b1           18.42176    3.01000     2.91941   6.310 0.000000000320413727 ***
s(x,y)b2          -19.49426    3.61121     2.74515  -7.101 0.000000000001539399 ***
s(x,y)b3            2.26247    0.27685     0.81736   2.768             0.005674 **
s(x,y)b4          111.64418   21.23001    23.51852   4.747 0.000002161491881275 ***
s(x,y)b5           -3.83935    0.57900     1.32038  -2.908             0.003667 **
s(x,y)b6         -291.31584   52.37577    58.53565  -4.977 0.000000683557265688 ***
s(x,y)b7           -2.10655    0.46803     0.52985  -3.976 0.000071826237840625 ***
s(x,y)b8            3.84991    0.36641     0.47577   8.092 0.000000000000000847 ***
s(x,y)b9            2.44795    0.34691     0.44037   5.559 0.000000029570669968 ***
s(x,y)b10           4.07095    0.43201     0.55079   7.391 0.000000000000188283 ***
s(x,y)b11         -13.80996    1.94876     1.46937  -9.399 < 0.0000000000000002 ***
s(x,y)b12         -10.91613    1.94332     0.95250 -11.461 < 0.0000000000000002 ***
s(x,y)b13         227.99175   40.01654    45.15132   5.050 0.000000469702530679 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for quasipoisson family taken to be 1.741278)

    Null deviance: 4752.1  on 2994  degrees of freedom
Residual deviance: 2981.5  on 2970  degrees of freedom
AIC:  NA

Max Panel Size = 56; Number of panels = 101
Number of Fisher Scoring iterations: 7
```

Figure A6-151 Code excerpt detailing SALSA 2D summary for red-throated diver MRSea analysis.

Cross-Validation

A.3.109 A ten-fold cross-validation was performed to compare the final 2D model with the 1D and GLM models. The 2D model exhibited a lower CV score than the 1D model, indicating a better fit. It was therefore retained due to its superior capacity to capture environmental complexity and incorporate spatial smoothing. The code excerpt detailing the cross-validation examination is provided below (Figure A6-152).



```
> cv_2d #SALSA 2D
[1] 1.334349
> cv_1d #SALSA 1D
[1] 1.501725
> cv_0d #Initial model
[1] 1.526905
```

Figure A6-152 Code excerpt detailing CV values for SALSA 2D, SALSA 1D and the initial model considered for red-throated diver MRSea analysis.

A.4 Final Model Validation

A.4.1 Outputs are provided to illustrate the validation process of the final model created for each species.

Guillemot

A.4.2 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in guillemot MRSea analysis (Figure A6-153, Figure A6-154, Figure A6-155, Figure A6-156, Figure A6-157, Figure A6-158 & Figure A6-159).

Cumulative Residuals

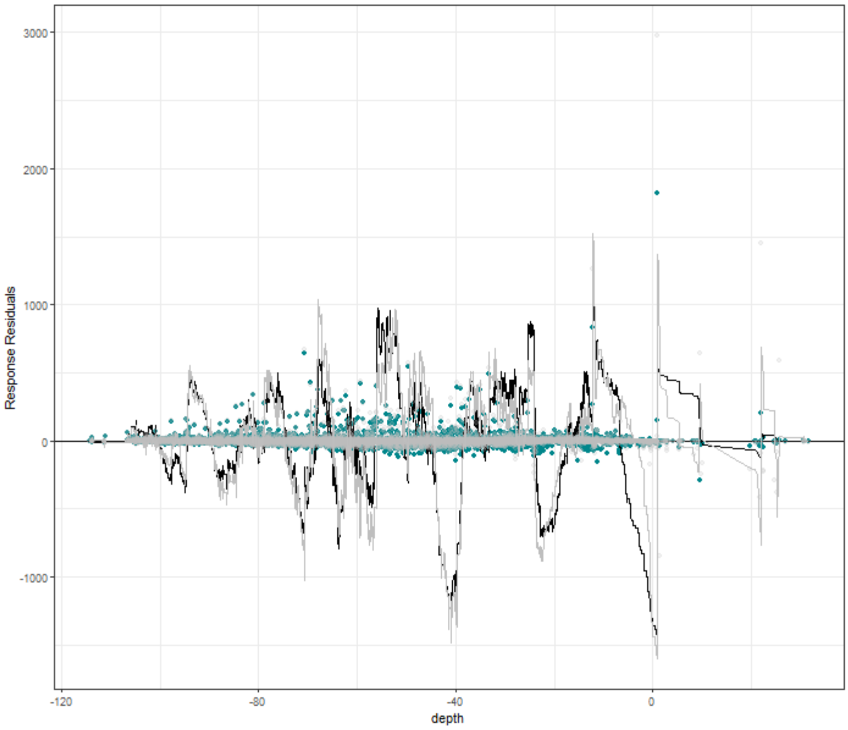


Figure A6-153 Cumulative residual plot of depth within the final SALSA 2D model selected for guillemot MRSea analysis.

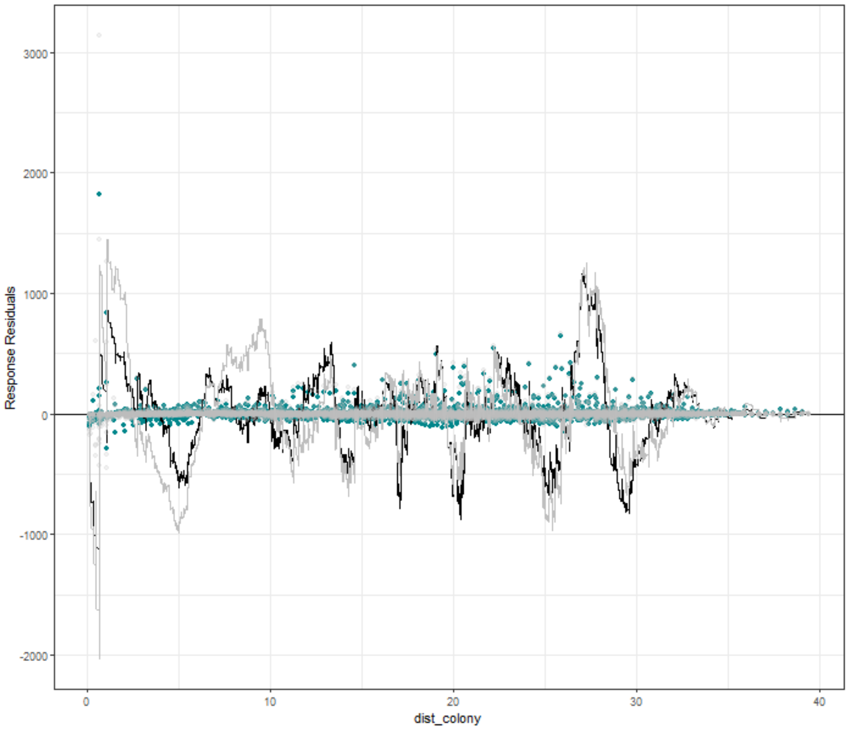


Figure A6-154 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for guillemot MRSea analysis.

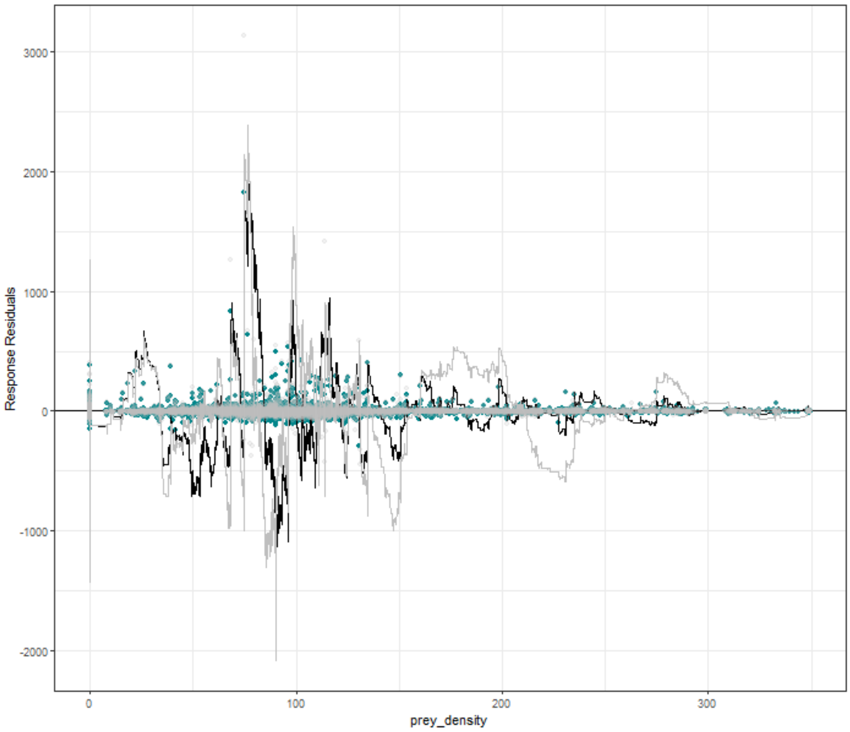


Figure A6-155 Cumulative residual plot of prey density within the final SALSA 2D model selected for guillemot MRSea analysis.

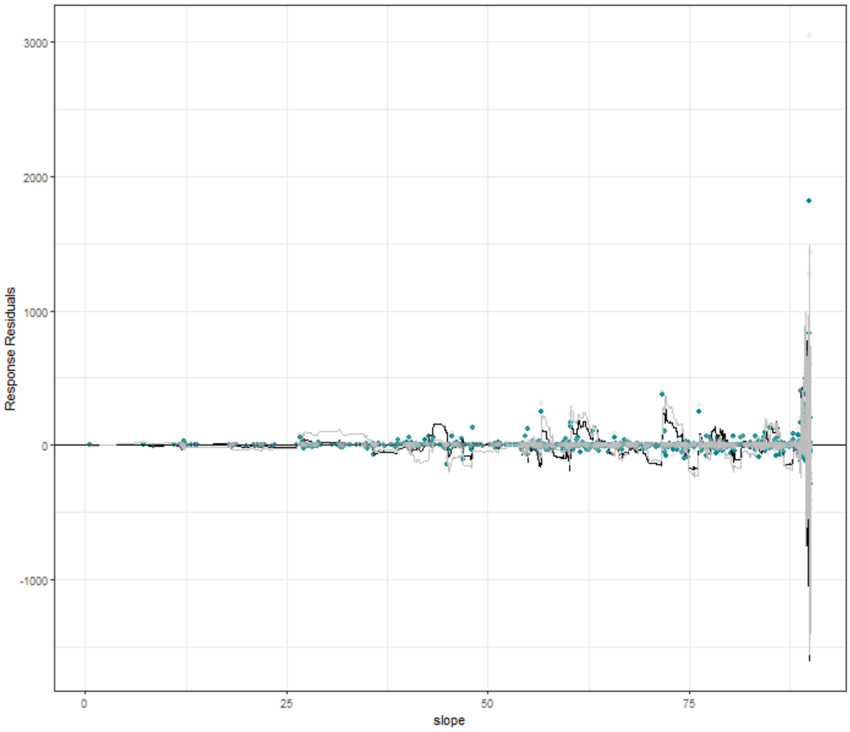


Figure A6-156 Cumulative residual plot of seabed slope within the final SALSA 2D model selected for guillemot MRSea analysis.

Observed vs Predicted

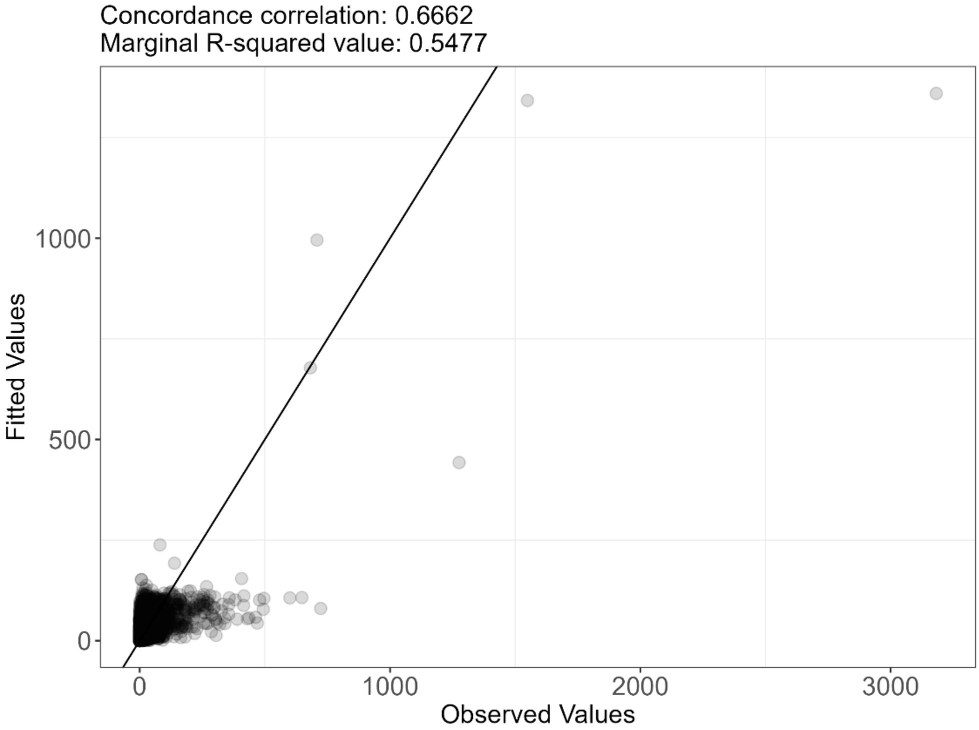


Figure A6-157 Observed vs fitted values plot for the final SALSA 2D model selected for guillemot MRSea analysis.

Pearson Residuals v Fitted

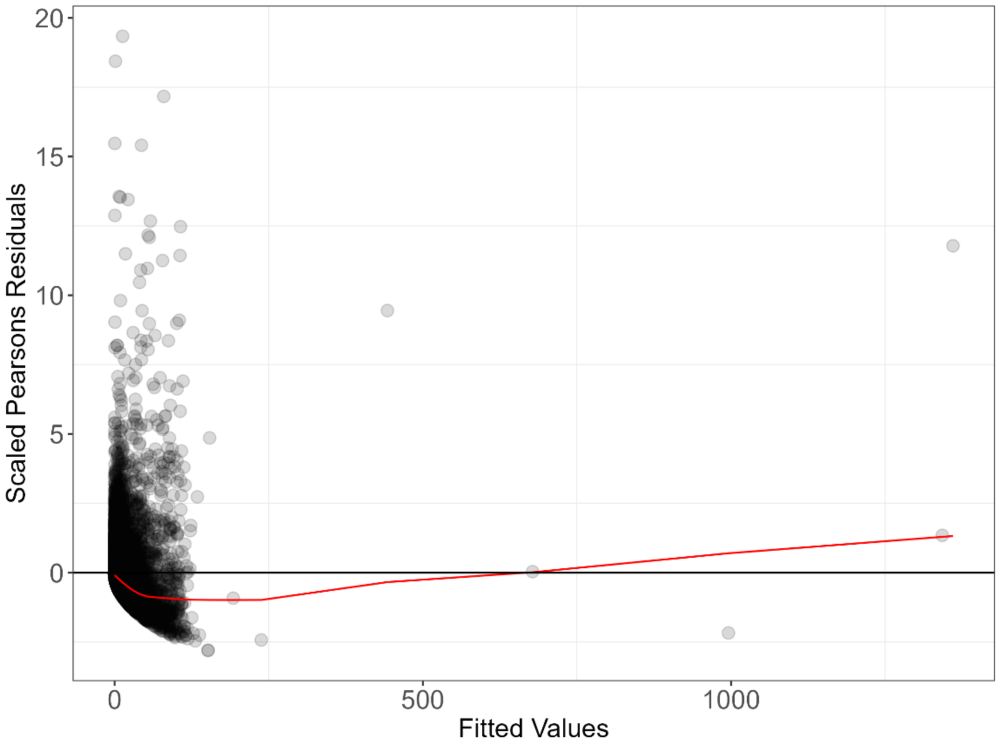


Figure A6-158 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for guillemot MRSea analysis.

Mean-variance Relationship

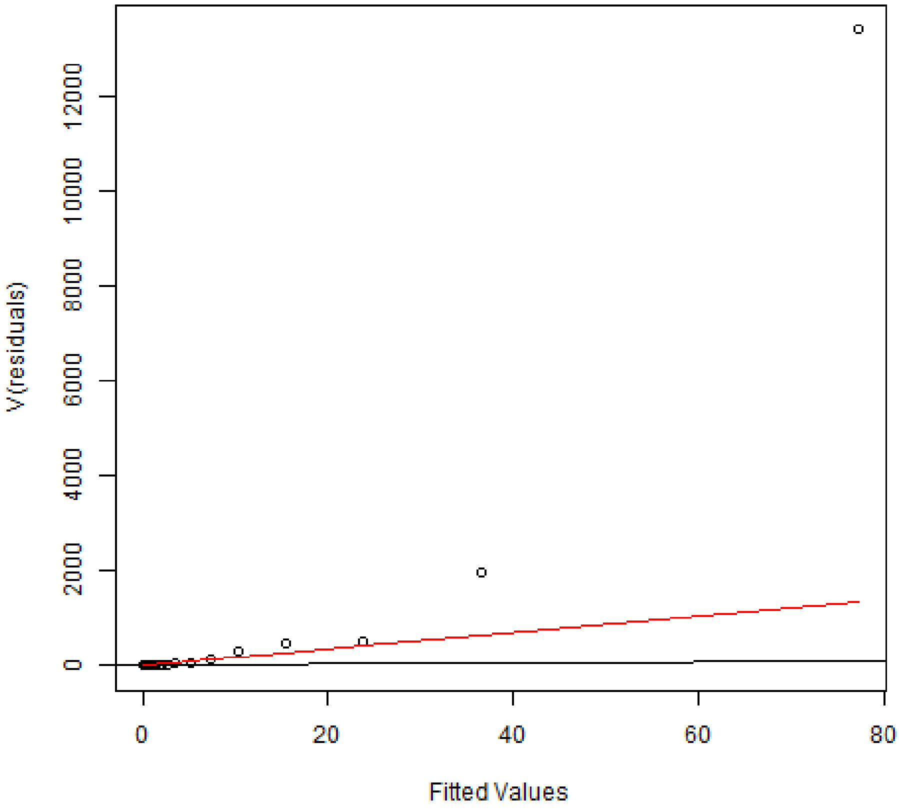


Figure A6-159 Mean-variance plot for the final SALSA 2D model selected for guillemot MRSea analysis.

Razorbill

A.4.3 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in razorbill MRSea analysis (Figure A6-160, Figure A6-161, Figure A6-162, Figure A6-163, Figure A6-164 & Figure A6-165).

Cumulative Residuals

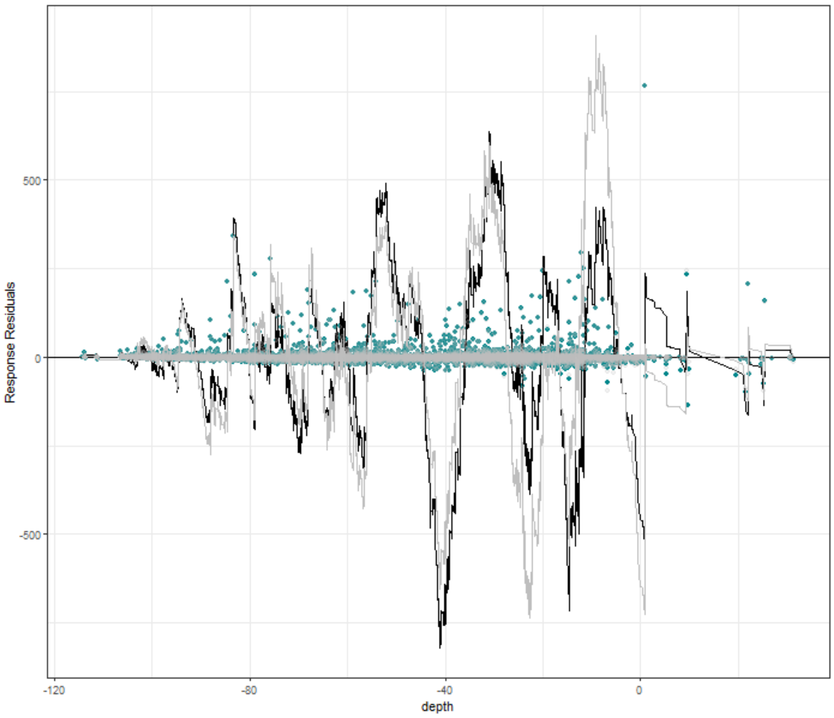


Figure A6-160 Cumulative residual plot of depth within the final SALSA 2D model selected for razorbill MRSea analysis.

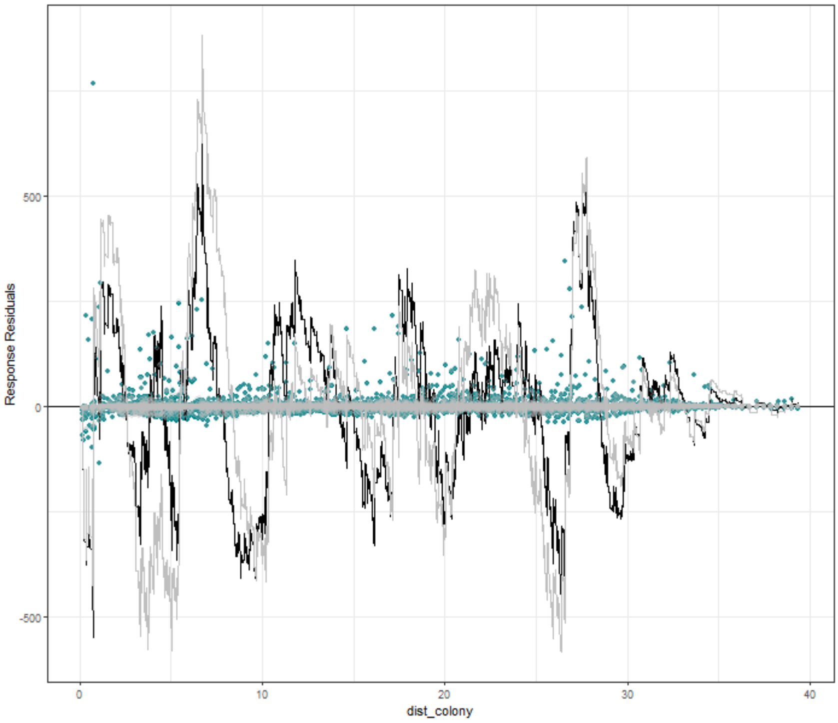


Figure A6-161 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for razorbill MRSea analysis.

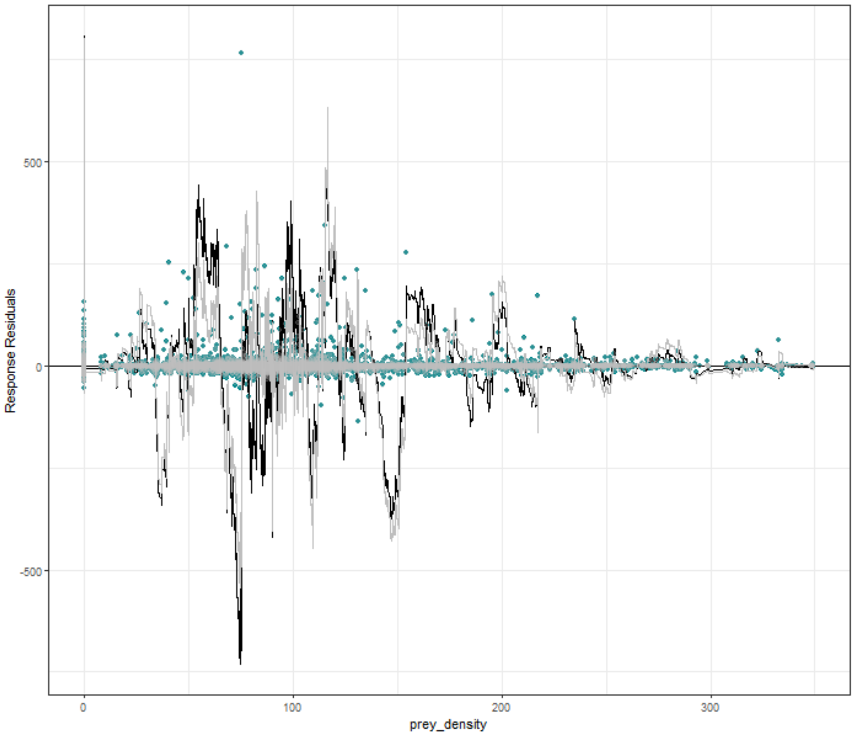


Figure A6-162 Cumulative residual plot of prey density within the final SALSA 2D model selected for razorbill MRSea analysis.

Observed vs Predicted

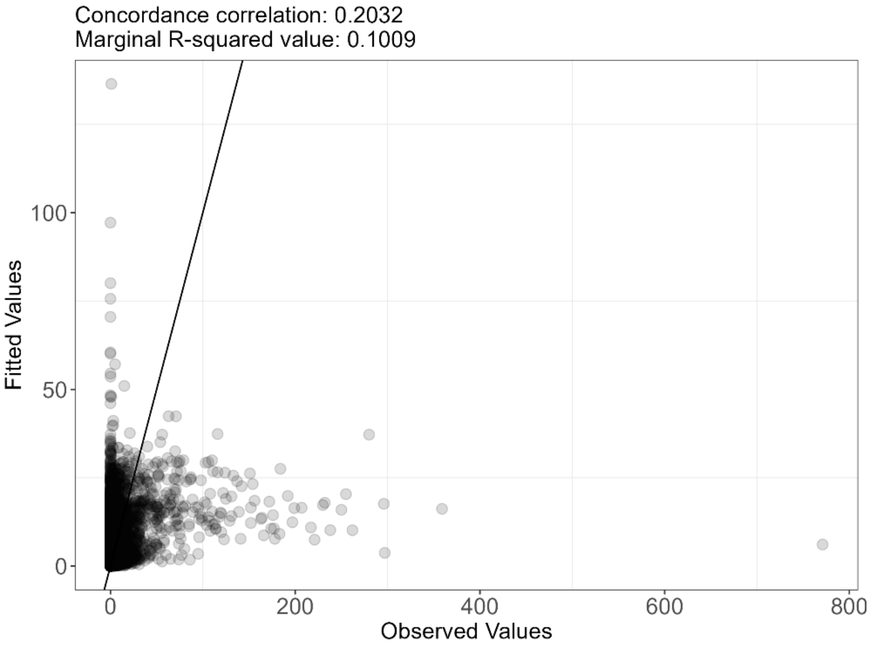


Figure A6-163 Observed vs fitted values plot for the final SALSA 2D model selected for razorbill MRSea analysis.

Pearson Residuals v Fitted

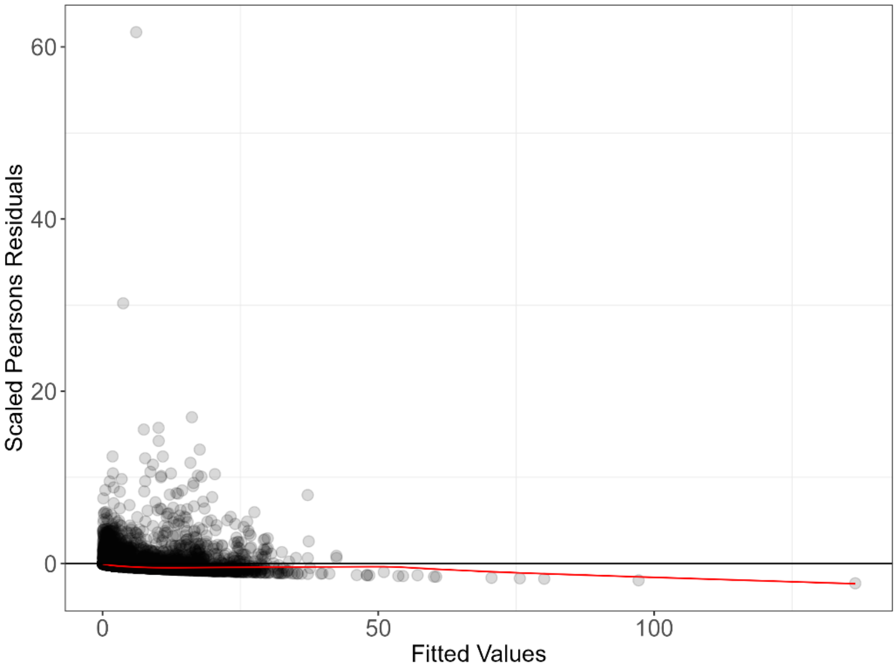


Figure A6-164 Scaled Pearsons Residuals vs fitted values plot for the final SALSA 2D model selected for razorbill MRSea analysis.

Mean-variance Relationship

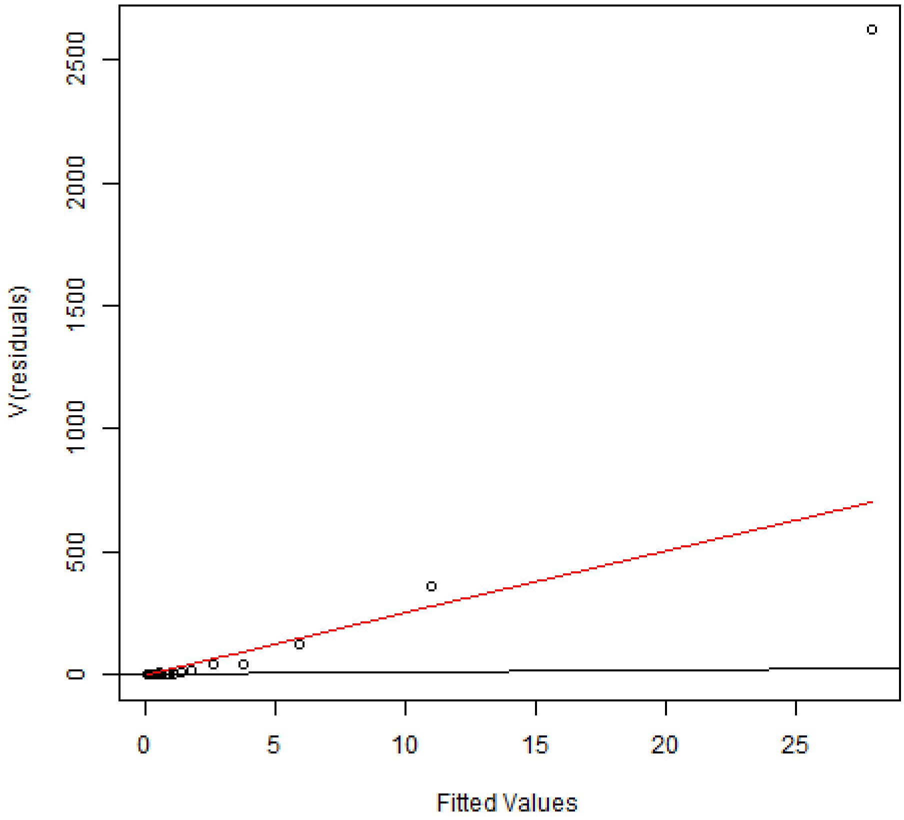


Figure A6-165 Mean-variance plot for the final SALSA 2D model selected for razorbill MRSea analysis.

Kittiwake

A.4.4 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in kittiwake MRSea analysis (Figure A6-166, Figure A6-167, Figure A6-168 & Figure A6-169).

Cumulative Residuals

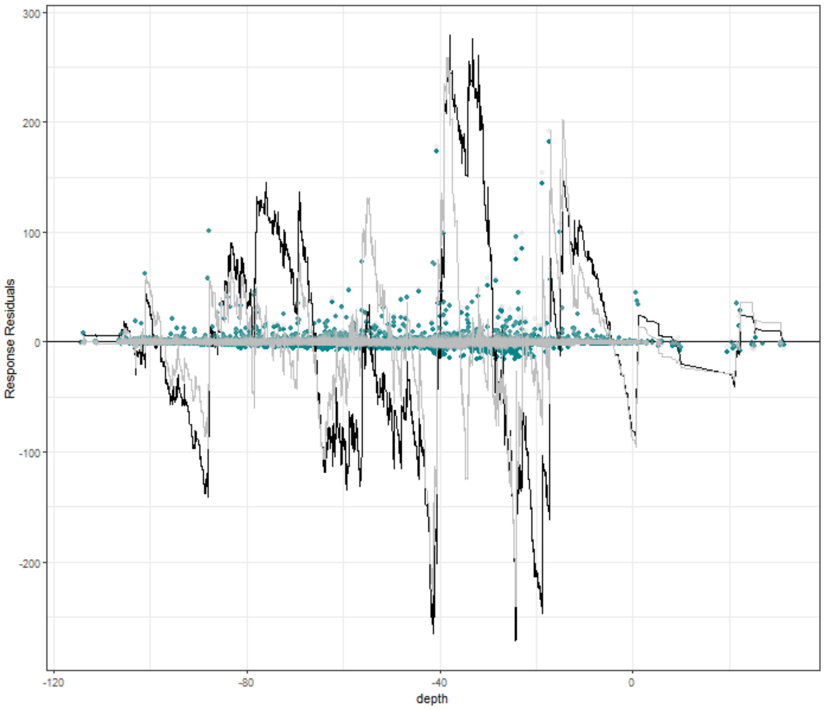


Figure A6-166 Cumulative residual plot of depth within the final SALSA 2D model selected for kittiwake MRSea analysis.

Observed vs Predicted

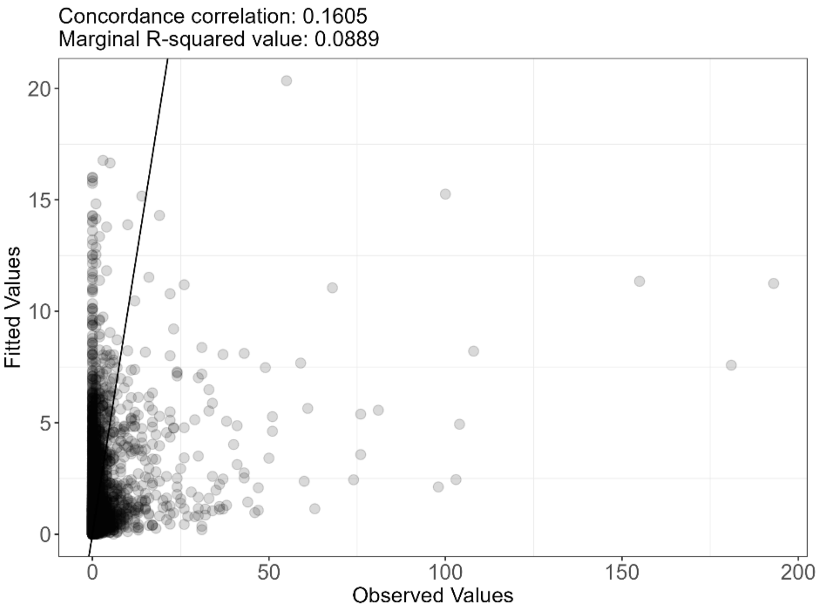


Figure A6-167 Observed vs fitted values plot for the final SALSA 2D model selected for kittiwake MRSea analysis.

Pearson Residuals v Fitted

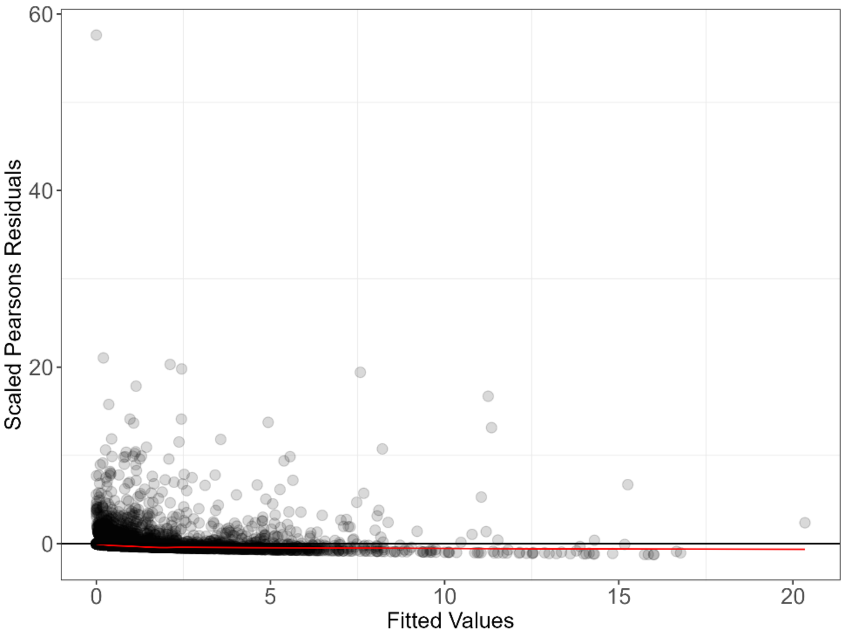


Figure A6-168 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for kittiwake MRSea analysis.

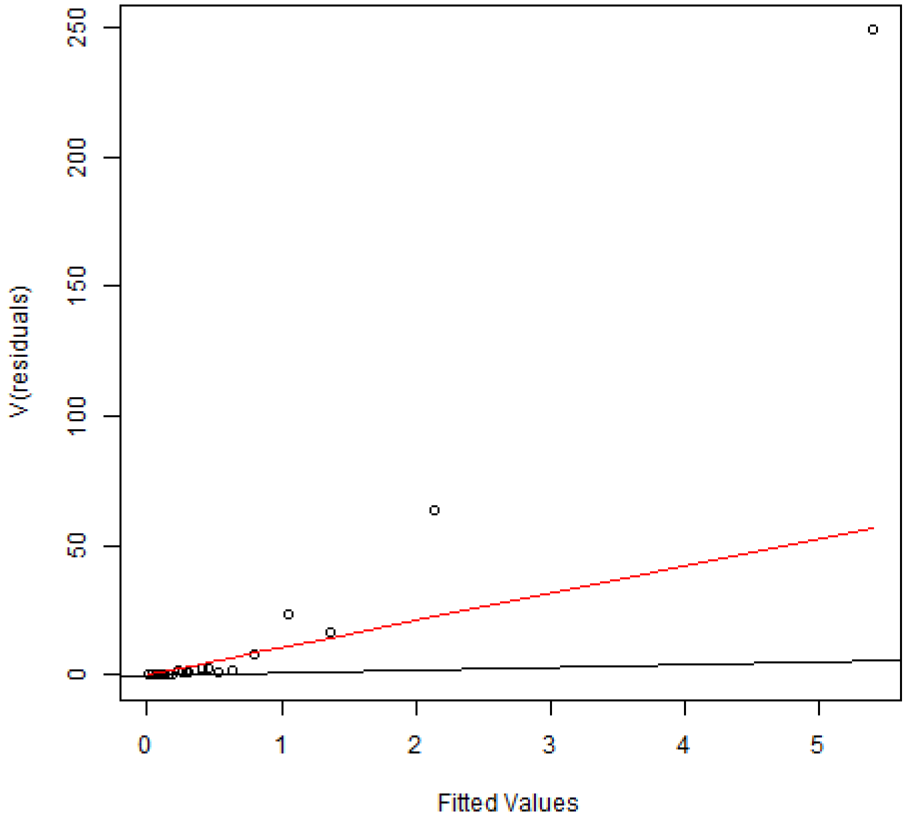


Figure A6-169 Mean-variance plot for the final SALSA 2D model selected for kittiwake MRSea analysis.

Herring Gull

A.4.5 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in herring gull MRSea analysis (Figure A6-170, Figure A6-171, Figure A6-172 & Figure A6-173).

Cumulative Residuals

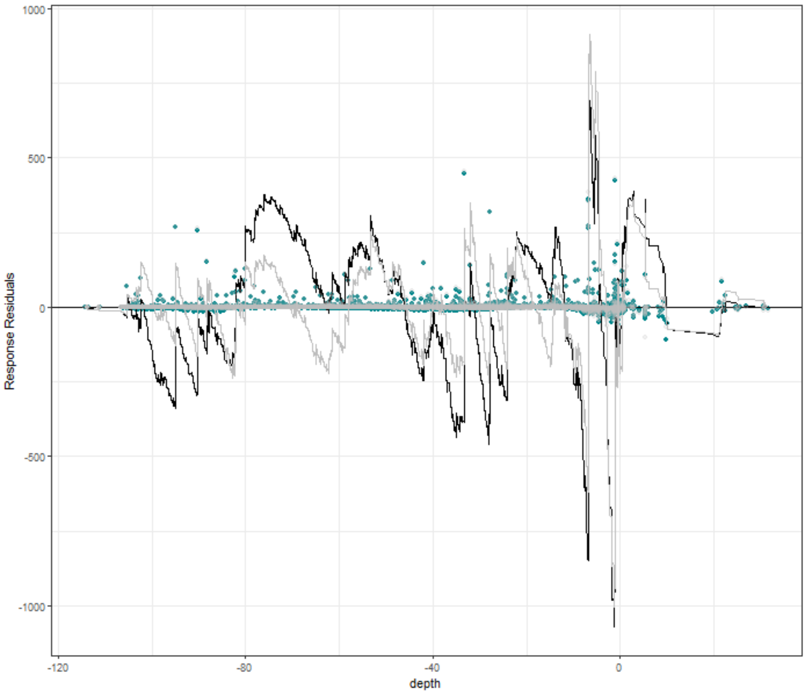


Figure A6-170 Cumulative residual plot of depth within the final SALSA 2D model selected for herring gull MRSea analysis.

Observed vs Predicted

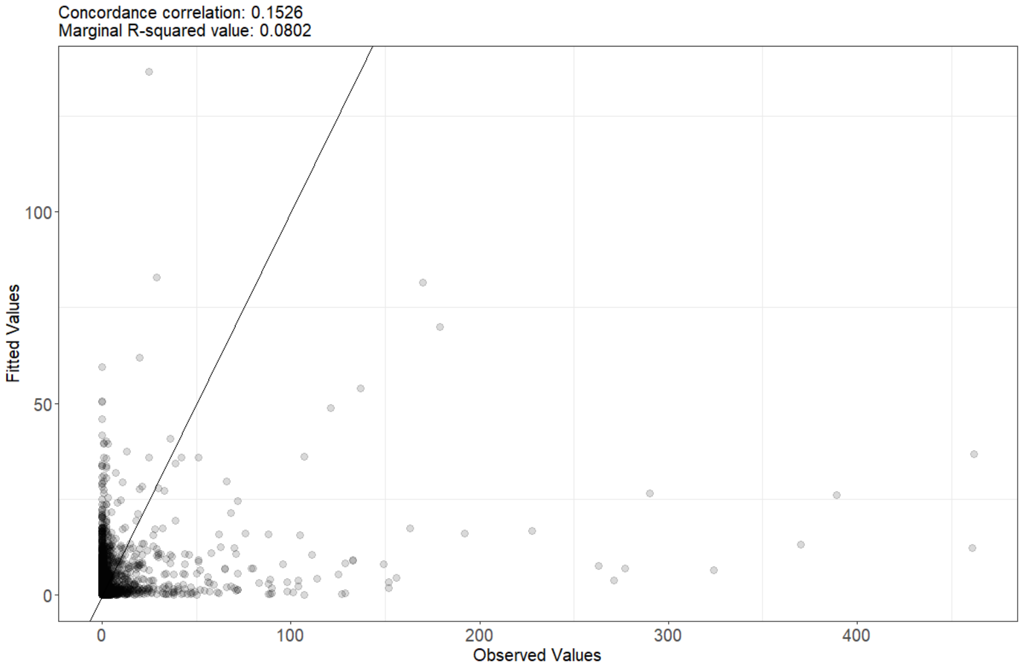


Figure A6-171 Observed vs fitted values plot for the final SALSA 2D model selected for herring gull MRSea analysis.

Pearson Residuals v Fitted

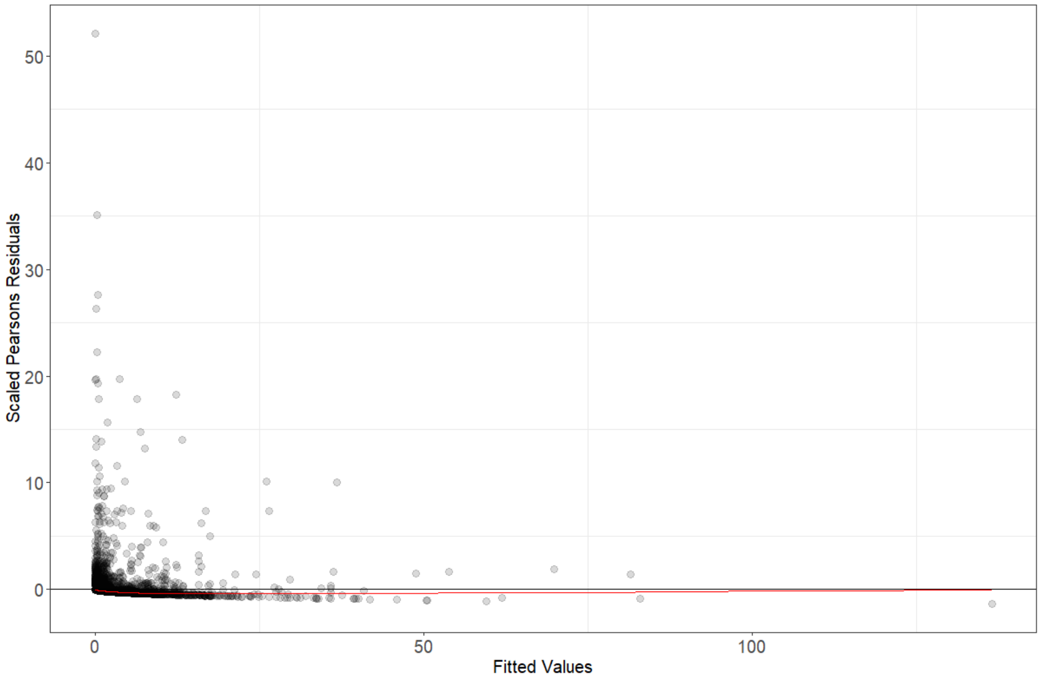


Figure A6-172 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for herring gull MRSea analysis.

Mean-variance Relationship

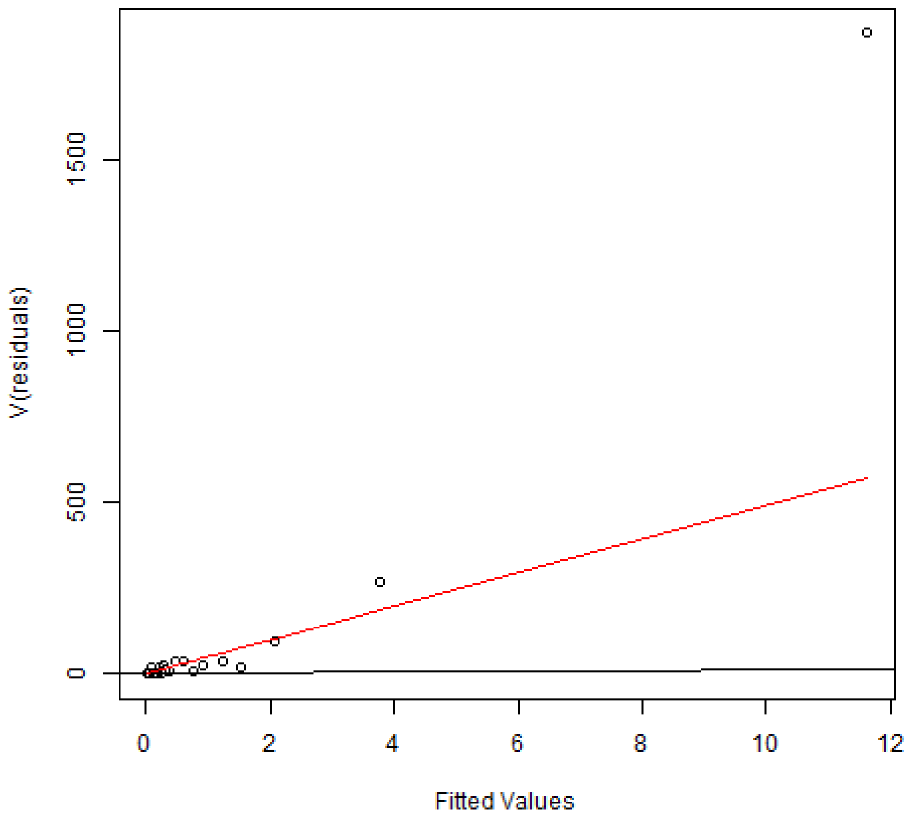


Figure A6-173 Mean-variance plot for the final SALSA 2D model selected for herring gull MRSea analysis.

Great Black-backed Gull

A.4.6 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in great black-backed gull MRSea analysis (Figure A6-174, Figure A6-175, Figure A6-176, Figure A6-177 & Figure A6-178).

Cumulative Residuals

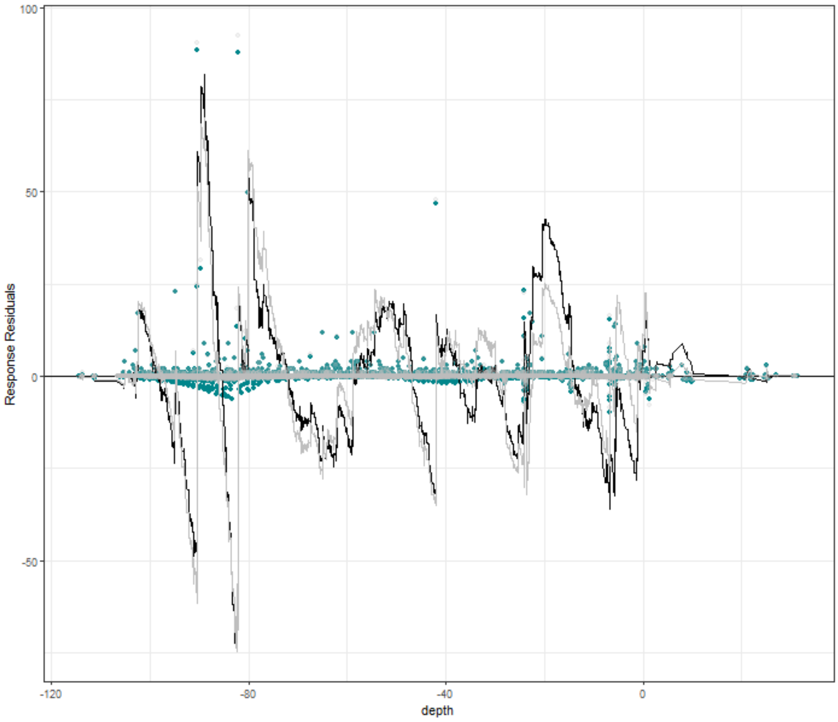


Figure A6-174 Cumulative residual plot of depth within the final SALSA 2D model selected for great black-backed gull MRSea analysis.

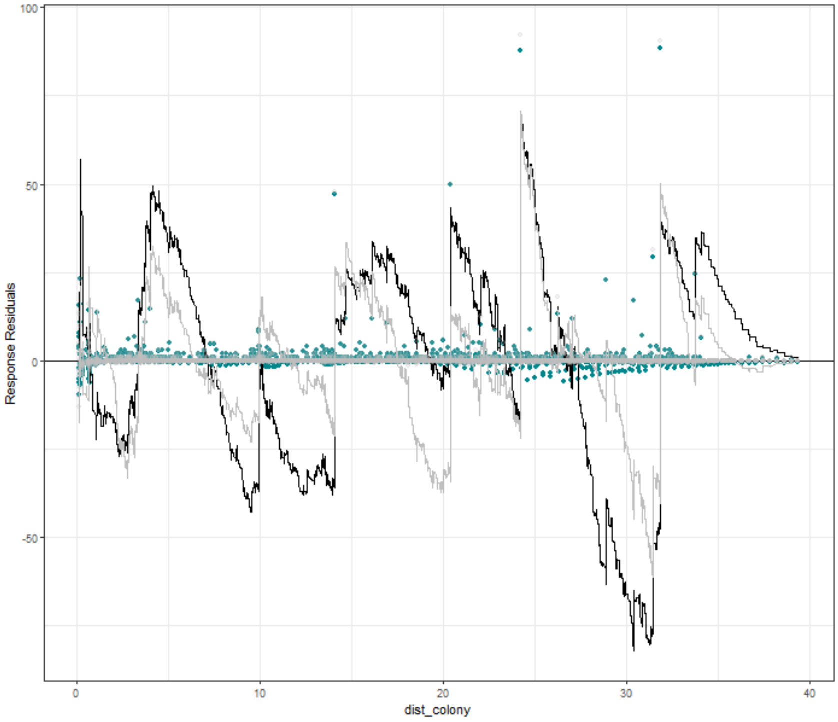


Figure A6-175 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for great black-backed gull MRSea analysis.

Observed vs Predicted

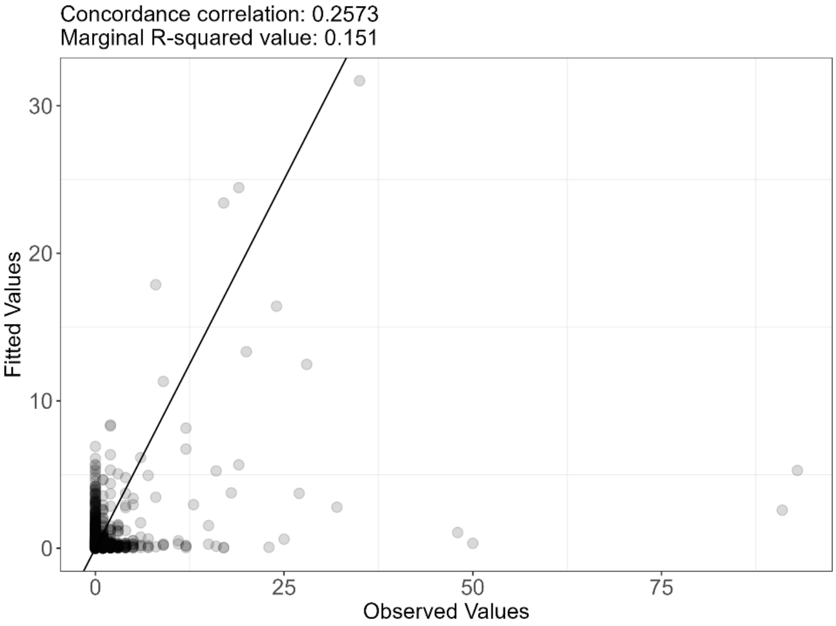


Figure A6-176 Observed vs fitted values plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis.

Pearson Residuals v Fitted

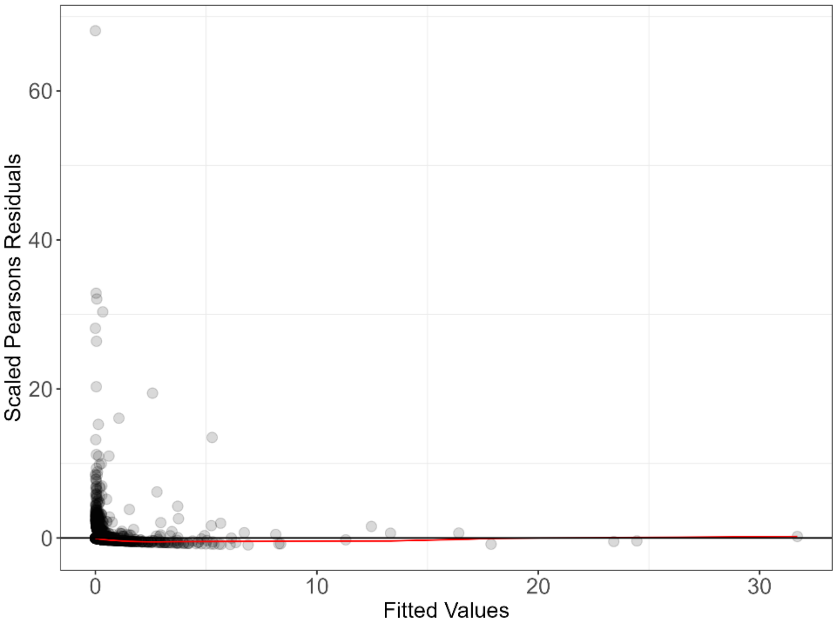


Figure A6-177 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis.

Mean-variance Relationship

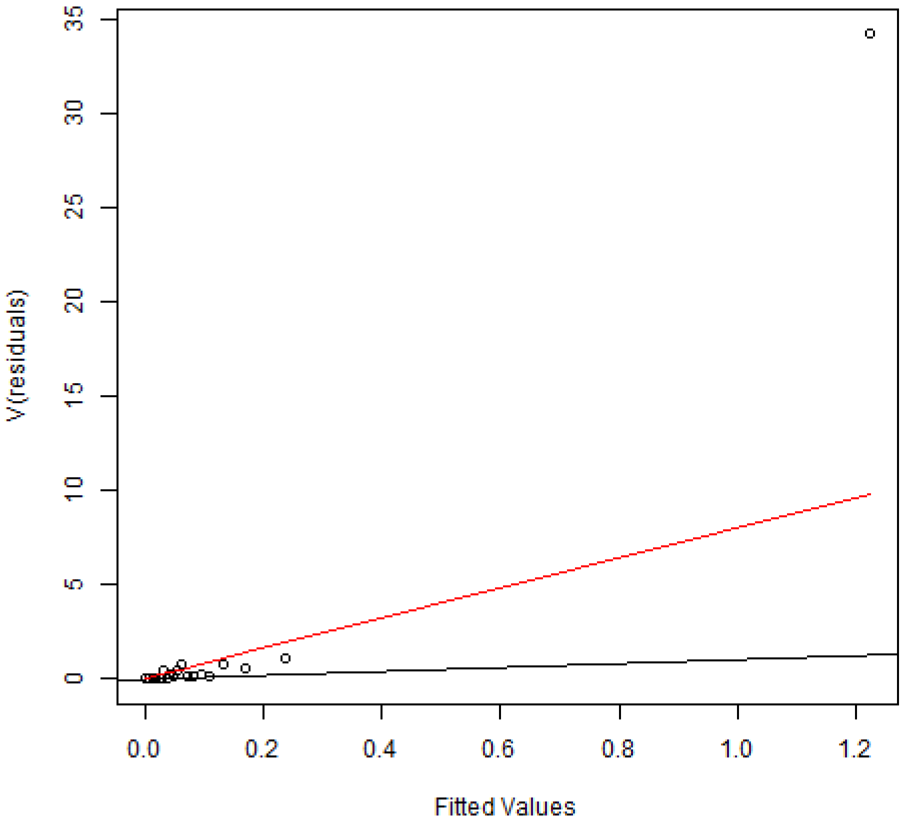


Figure A6-178 Mean-variance plot for the final SALSA 2D model selected for great black-backed gull MRSea analysis.

Gannet

A.4.7 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in gannet MRSea analysis (Figure A6-179, Figure A6-180, Figure A6-181, Figure A6-182 & Figure A6-183).

Cumulative Residuals

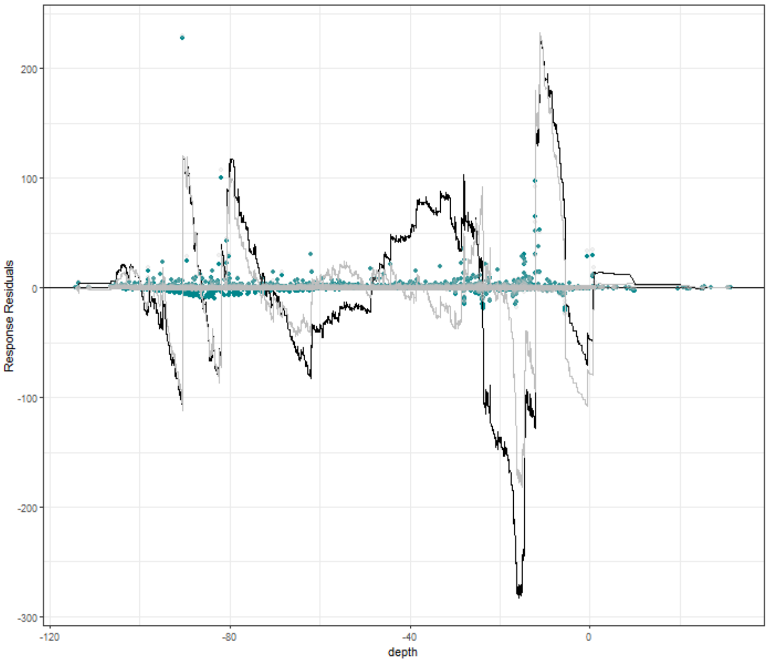


Figure A6-179 Cumulative residual plot of depth within the final SALSA 2D model selected for gannet MRSea analysis.

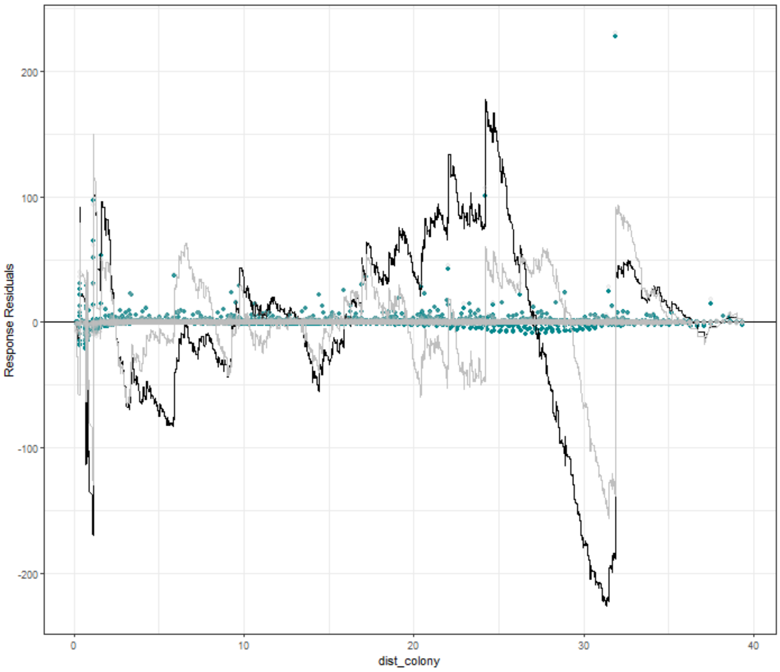


Figure A6-180 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for gannet MRSea analysis.

Observed vs Predicted

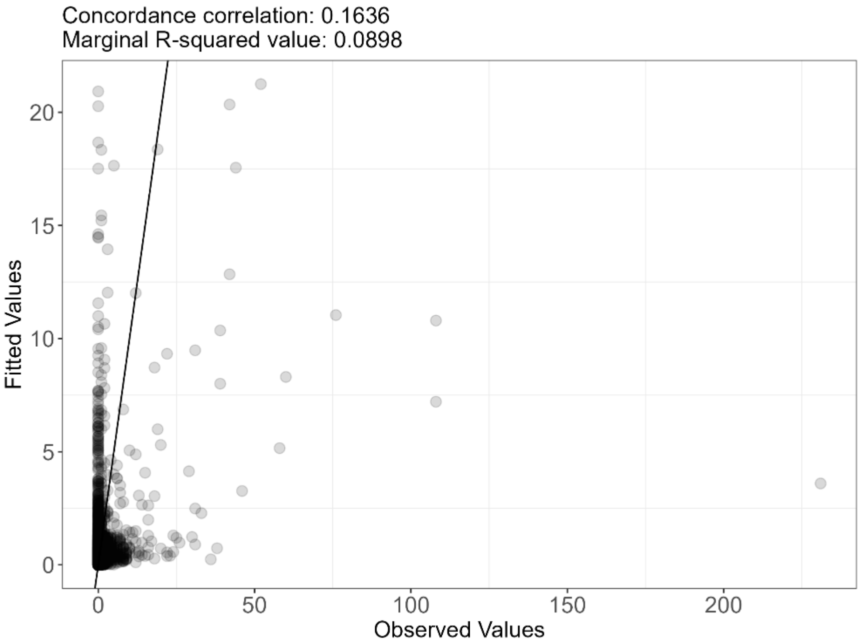


Figure A6-181 Observed vs fitted values plot for the final SALSA 2D model selected for gannet MRSea analysis.

Pearson Residuals v Fitted

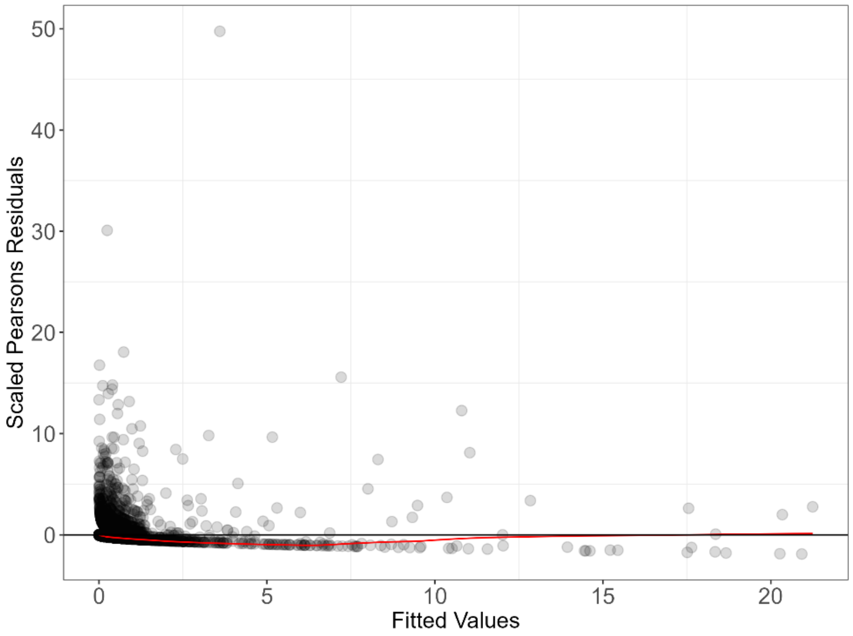


Figure A6-182 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for gannet MRSea analysis.

Mean-variance Relationship

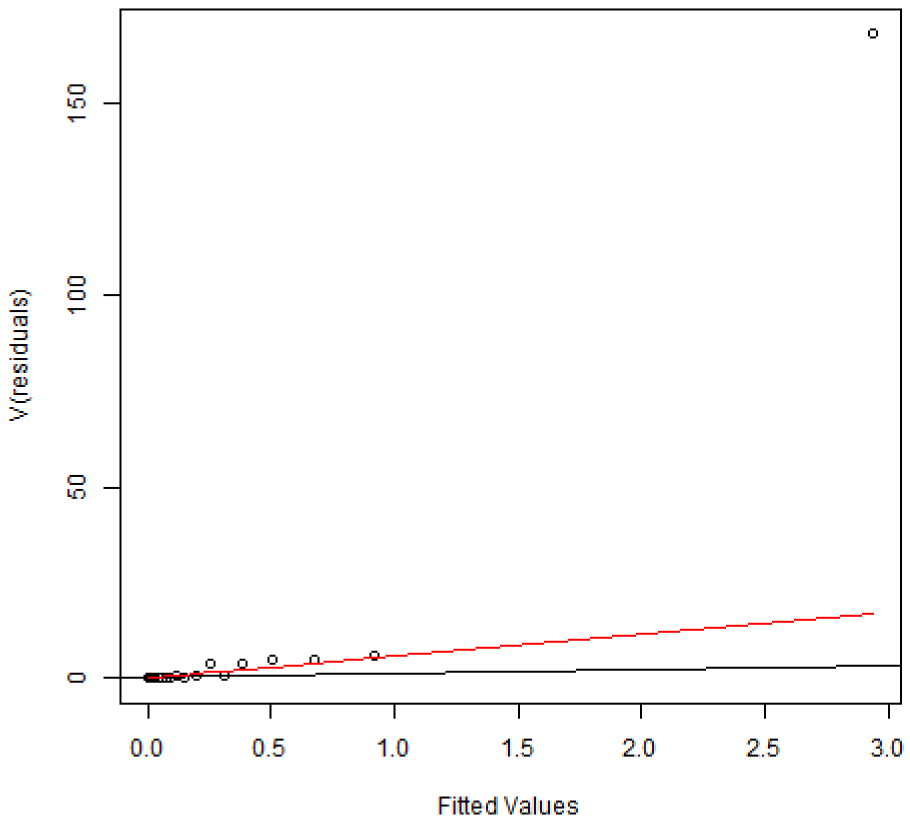


Figure A6-183 Mean-variance plot for the final SALSA 2D model selected for gannet MRSea analysis.

Common Tern

A.4.8 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in common tern MRSea analysis (Figure A6-184, Figure A6-185, Figure A6-186, Figure A6-187, Figure A6-188 & Figure A6-189).

Cumulative Residuals

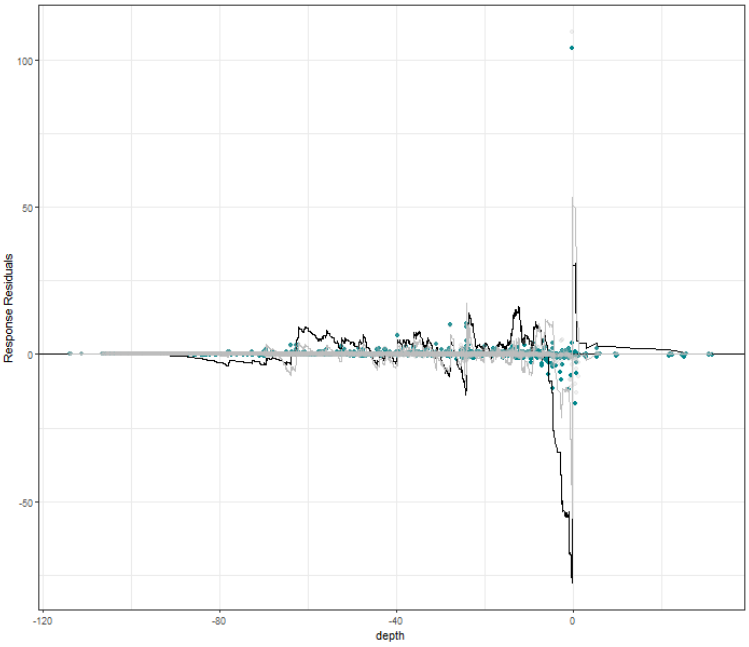


Figure A6-184 Cumulative residual plot of depth within the final SALSA 2D model selected for common tern MRSea analysis.

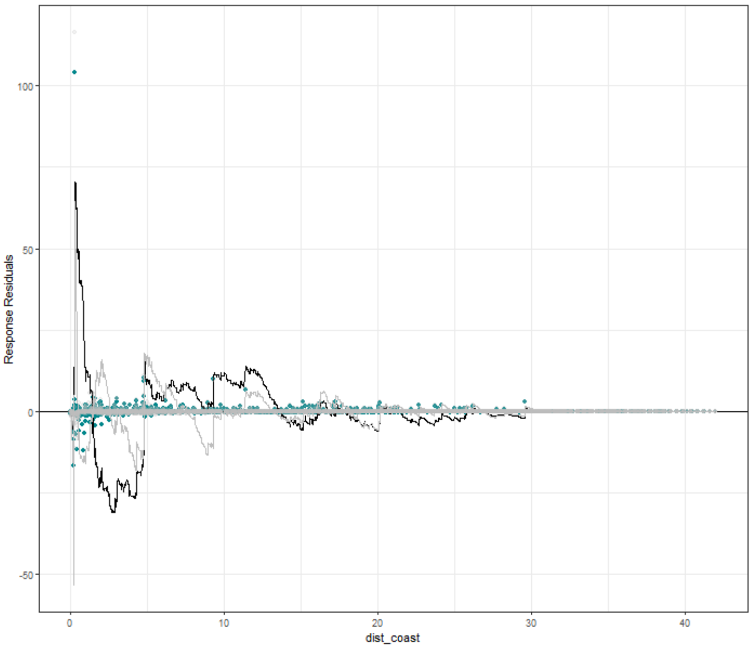


Figure A6-185 Cumulative residual plot of distance to coast within the final SALSA 2D model selected for common tern MRSea analysis.

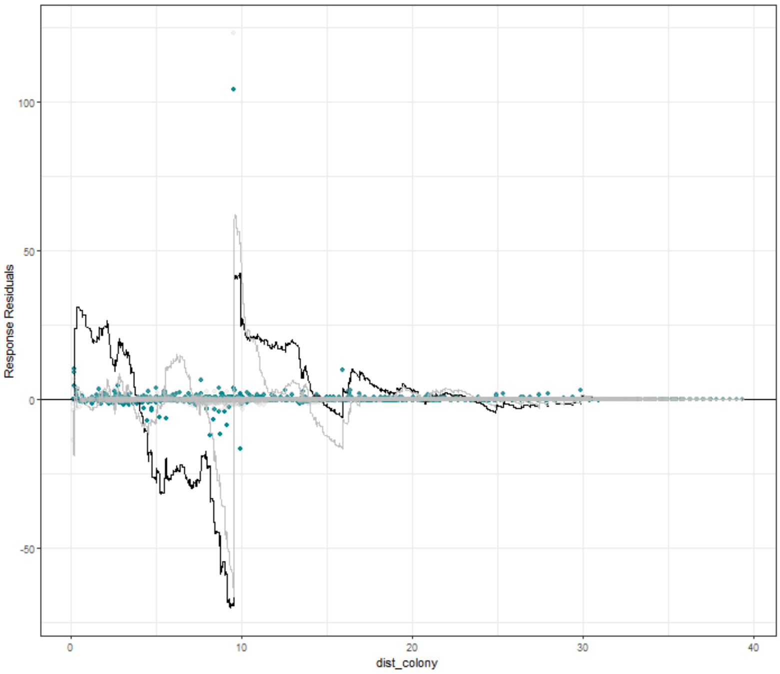


Figure A6-186 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for common tern MRSea analysis.

Observed vs Predicted

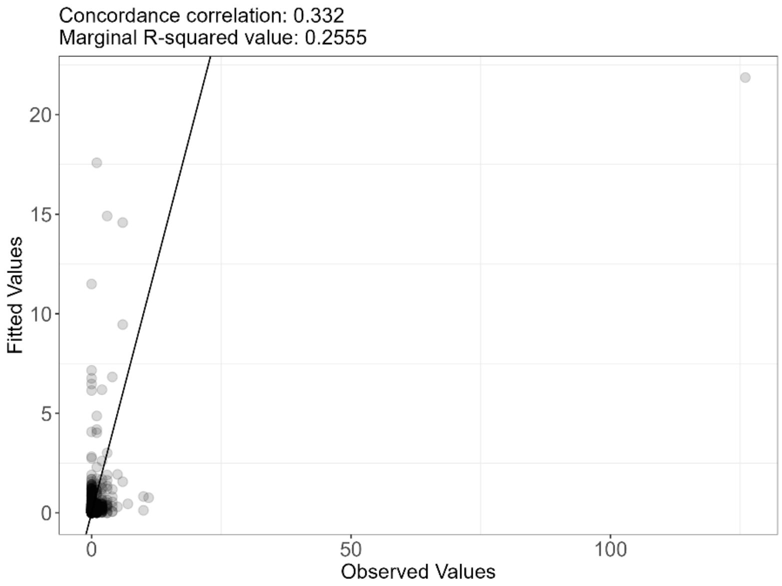


Figure A6-187 Observed vs fitted values plot for the final SALSA 2D model selected for common tern MRSea analysis.

Pearson Residuals v Fitted

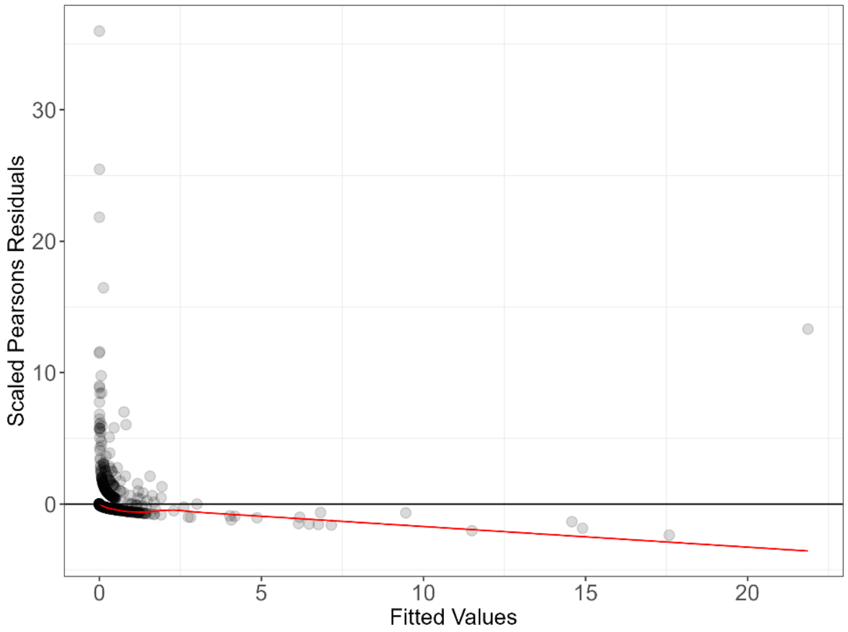


Figure A6-188 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for common tern MRSea analysis.

Mean-variance Relationship

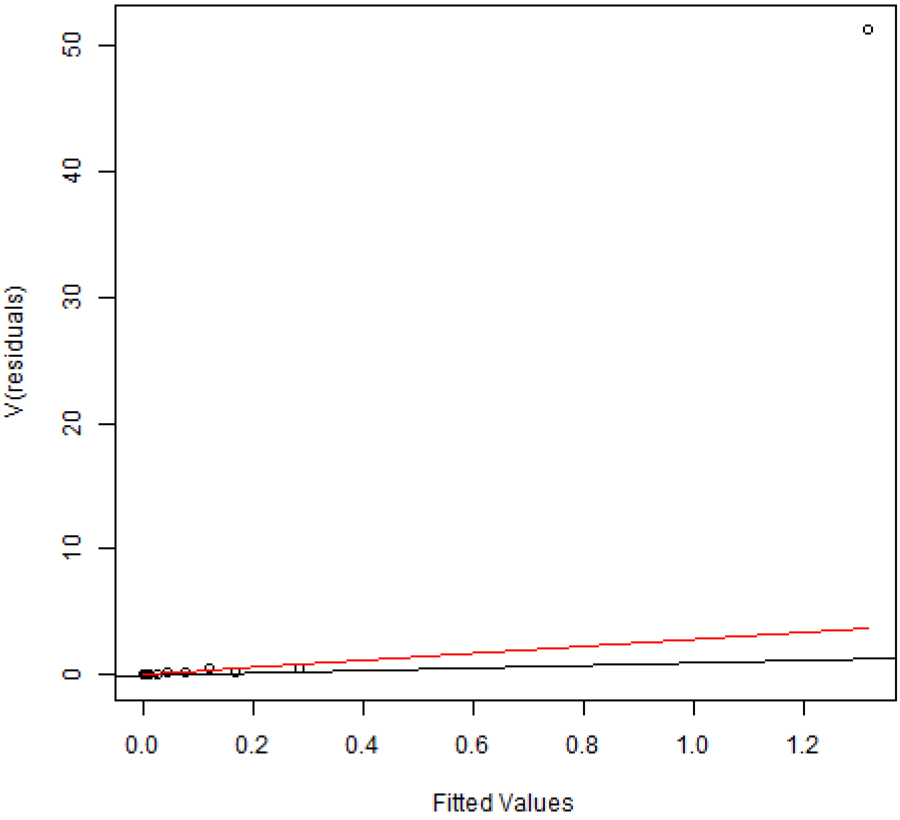


Figure A6-189 Mean-variance plot for the final SALSA 2D model selected for common tern MRSea analysis.

Roseate Tern

A.4.9 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in roseate tern MRSea analysis (Figure A6-190, Figure A6-191, Figure A6-192 & Figure A6-193).

Cumulative Residuals

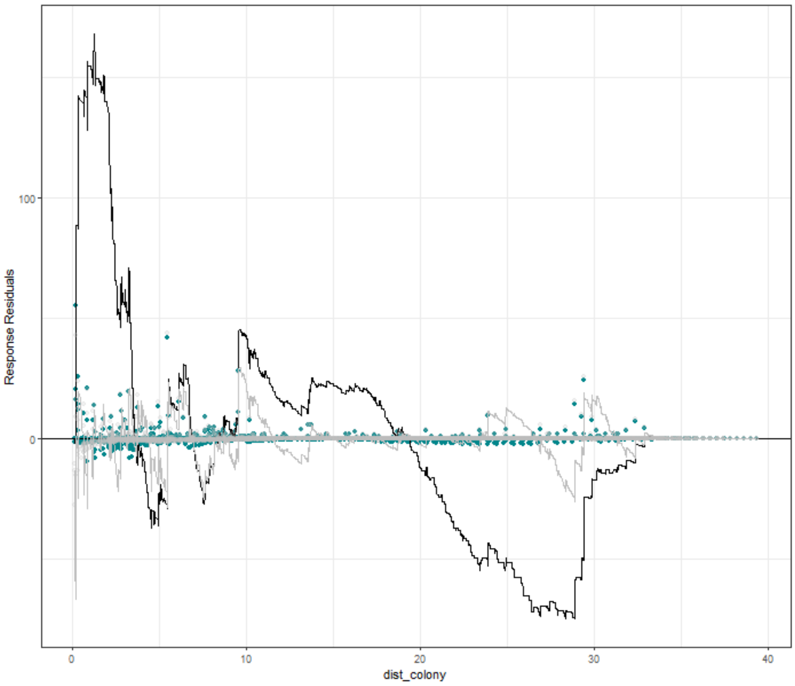


Figure A6-190 Cumulative residual plot of distance to colony within the final SALSA 2D model selected for roseate tern MRSea analysis.

Observed vs Predicted

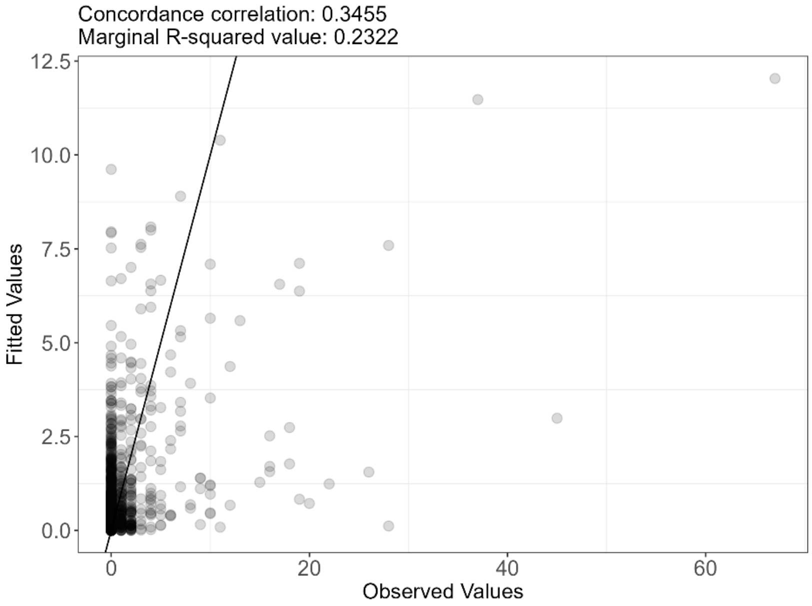


Figure A6-191 Observed vs fitted values plot for the final SALSA 2D model selected for roseate tern MRSea analysis.

Pearson Residuals v Fitted

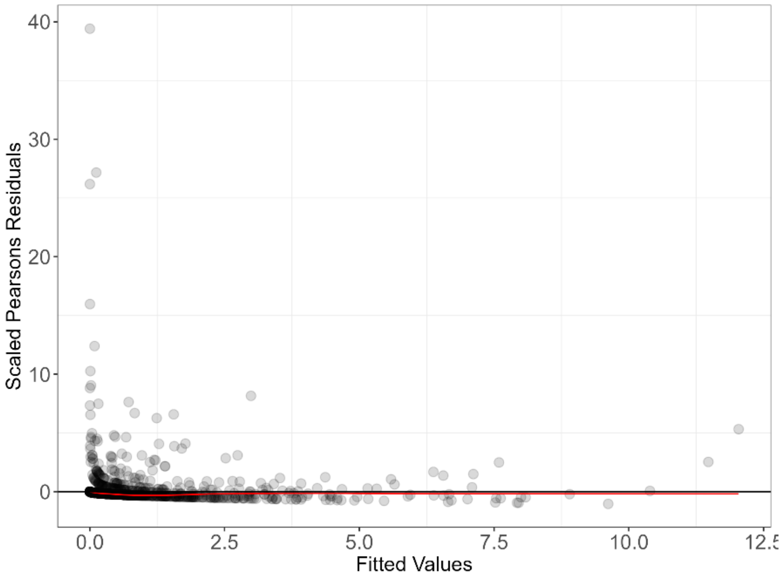


Figure A6-192 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for roseate tern MRSea analysis.

Mean-variance Relationship

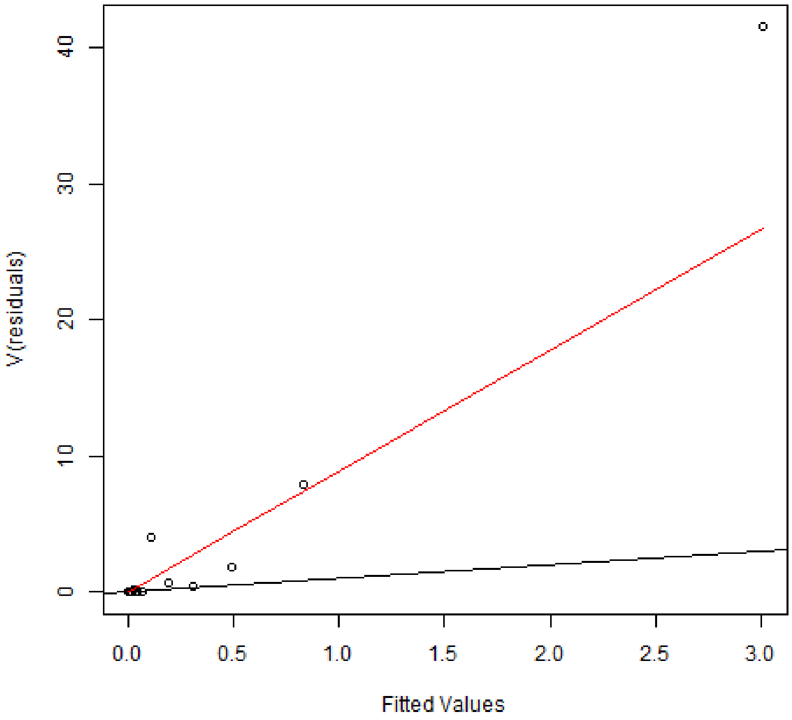


Figure A6-193 Mean-variance plot for the final SALSA 2D model selected for roseate tern MRSea analysis.

Red-throated Diver

A.4.10 Plots are provided below that include outputs of cumulative residuals, observed vs predicted, Pearson residuals v fitted and mean-variance relationships for the final 2D SALSA model used in red-throated diver MRSea analysis (Figure A6-194, Figure A6-195, Figure A6-196 & Figure A6-197).

Cumulative Residuals

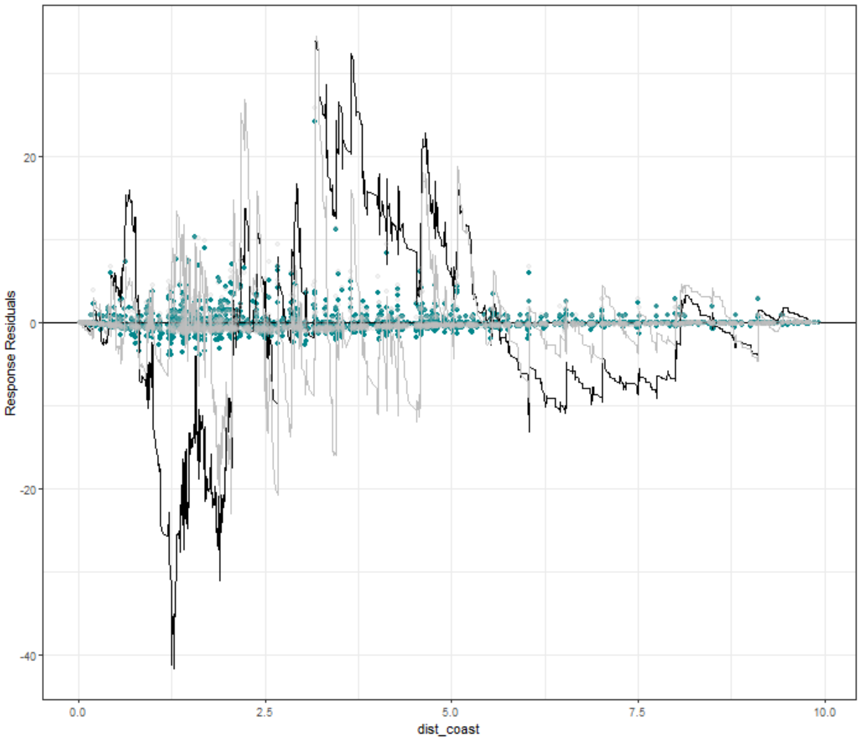


Figure A6-194 Cumulative residual plot of distance to coast within the final SALSA 2D model selected for red-throated diver MRSea analysis.

Observed vs Predicted

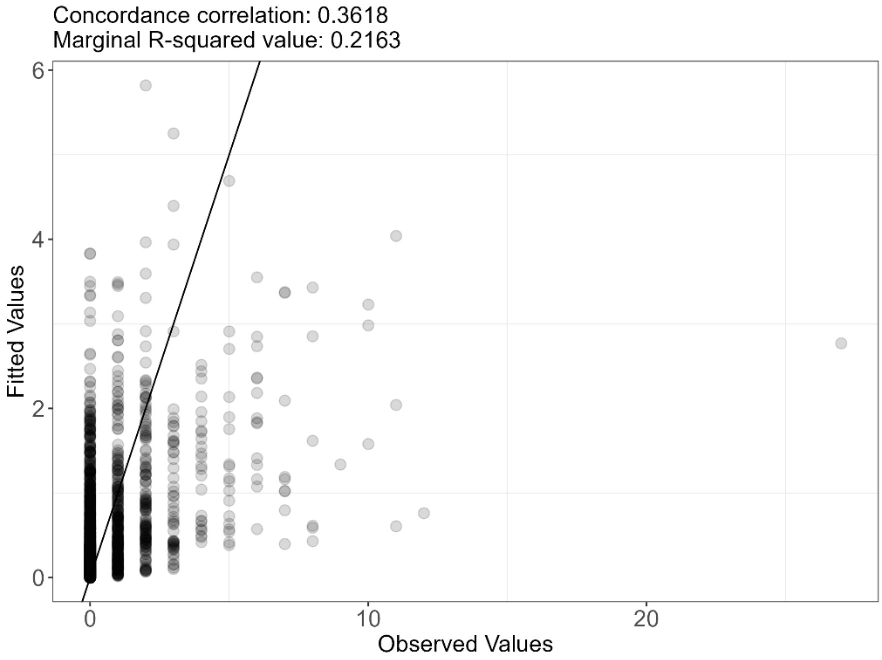


Figure A6-195 Observed vs fitted values plot for the final SALSA 2D model selected for red-throated diver MRSea analysis.

Pearson Residuals v Fitted

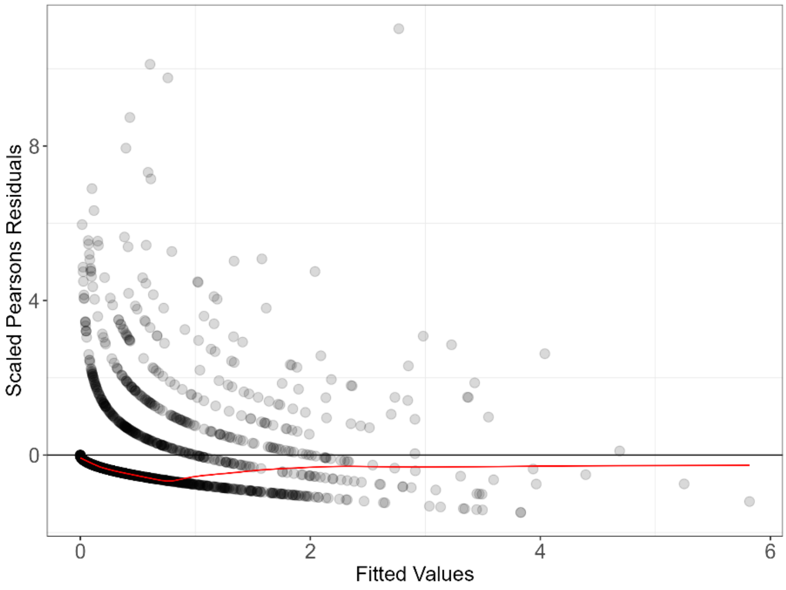


Figure A6-196 Scaled Pearsons Residuals v fitted values plot for the final SALSA 2D model selected for red-throated diver MRSea analysis.

Mean-variance Relationship

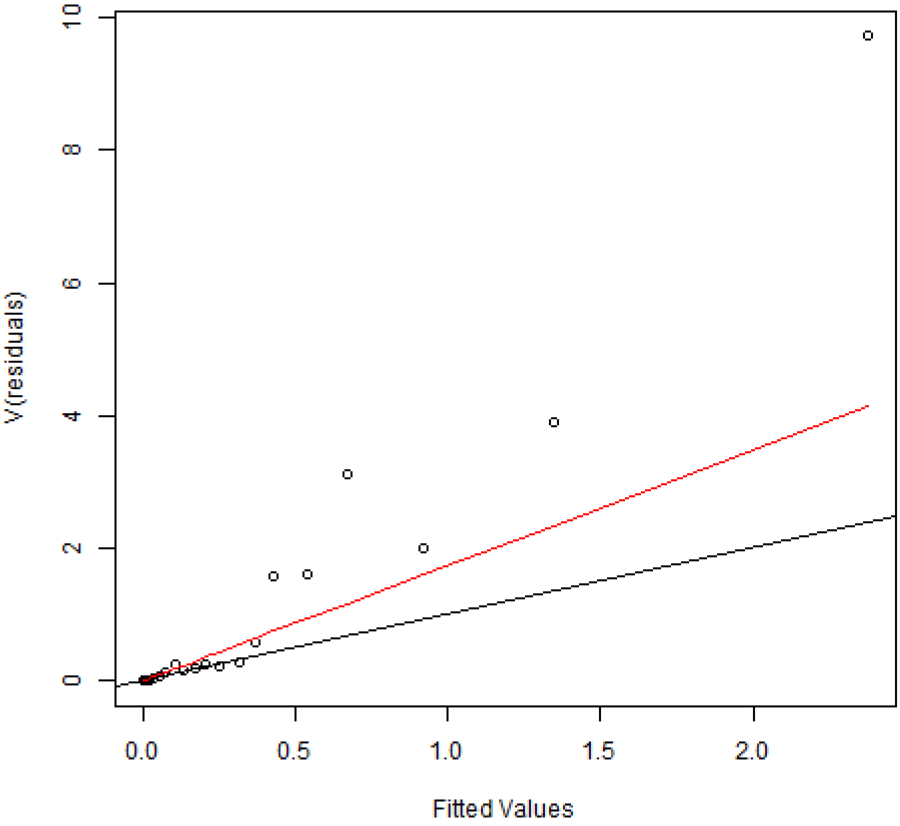


Figure A6-197 Mean-variance plot for the final SALSA 2D model selected for red-throated diver MRSea analysis.

GoBe Consultants Ltd
Suites B2 & C2, Higher Mill
Higher Mill Lane
Buckfastleigh
Devon
TQ11 0EN