SAMBAH Code File 6
Density and Abundance Estimates

Len Thomas & Louise Burt, CREEM

January 2, 2022

1 Introduction

This document (the `.Rnw` version) contains the `R` code to produce density and abundance estimates for the SAMBAH survey, plus bootstrap estimates of uncertainty. It is based on SAMBAH internal reports; this version has been created to accompany the paper:

Amundin et al. In press. Estimating the abundance of the critically endangered Baltic Proper harbour porpoise (Phocoena phocoena) population using passive acoustic monitoring. Ecology and Evolution.

The document is a `Sweave` file – i.e., it is a mixture of `LaTeX` and `R` that is designed to be compiled into a report in pdf (or another format such as html). We have tested it using the `Knitr` package in `R version 4.1.1 (2021-08-10)`. Readers wishing to see the underlying code should view the version with the `.Rnw` suffix, and look for code chunks starting with <<.

Warning: running the `Sweave` file in `R` takes some time because it is executing the bootstrap.

2 Methods

Data collection and analysis methods are described in detail in Amundin et al. (In press); we give an overview here.

2.1 Estimating density and abundance

Density was estimated separately for each sampling location, month and diel phase (dawn, day, dusk and night) using the estimator:

$$\hat{D}_{i,m,d} = \frac{n_{i,m,d}(1-\hat{c})}{T_{i,m,d}\hat{\nu}_{i,m,d}} \tag{1}$$

where

- $\hat{D}_{i,m,d}$ is estimated density at sampling location i in month m and diel phase d;
- $n_{i,m,d}$ is the number of porpoise positive seconds (PPS, i.e., seconds where one or more porpoise click is detected) at sampling location i in month m and diel phase d;
- $\hat{c}$ is the proportion of porpoise positive seconds that are false positives – this is assumed to be zero in this study;

- $T_{i,m,d}$ is the number of seconds that a C-POD was recording at sampling location i in month m and diel phase d;
- $\hat{\nu}_{i,m,d}$ is the estimated effective survey area (ESA) at sampling location i in month m and diel phase d;

We outline in subsections below how each of these components was obtained.

Density per sampling location and month was estimated using a weighted average of the diel phase-speciic estimates:

$$\hat{D}_{i,m} = \sum_{d=1}^{4} w_{i,m,d} \hat{D}_{i,m,d} \tag{2}$$

where $w_{i,m,d}$ is the proportion of month m at location i that is diel period d. In practice, this proportion was approximated by taking the proportion of the 15th day of the month m at the location of sampling position i that was each diel period.

Density at higher levels of aggregation was estimated using simple averages of the location- and month-specific estimates. These estimates were made using only locations and months that were inside the main SAMBAH survey period. In particular, density was estimated at the following levels of aggregation:

- Estimates by country[1] and month were produced for graphical summary only;
- Estimates by country and season were produced, where the two seasons were "winter" (November-April) and "summer" (May-October);
- Estimates by region and season were produced, as follows: for winter, all countries were combined; for summer, the study was divided into northeast and southwest strata by a line running approximately from Nogersund in Sweden to Jaroslawic in Poland.

Abundance was estimated as density multiplied by the relevant survey area.

2.2 Estimating encounter rate (n and T)

A description of the methods used to obtain the number of porpoise-positive seconds n and the monitoring time T is given in SAMBAH Code File 1. That document also contains an analysis of encounter rate by diel phase, justifying the requirement to model detectability by diel phase. Note that the 1-second period is also called a "snapshot".

2.3 Estimating effective detection area, ν

The effective detection area (EDA) for each sampling location, month and diel phase was estimated using the equation:

$$\hat{\nu}_{i,m,d} = \frac{\hat{\nu}_d^* \hat{p}_c \hat{\xi}_{i,m}}{\hat{\xi}^*} \tag{3}$$

where

- $\hat{\nu}_d^*$ is the estimated EDA for porpoises in diel phase d estimated at the Great Belt tracking site near Kerteminde, using acoustic encounters with porpoises that were clicking (hence the need for p_c,below);
- $\hat{p}_c$ is the estimated probability that porpoises at the Great Belt tracking site produced one or more click during the time period of an acoustic encounter;
- $\hat{\xi}^*$ is the estimated EDA for an artificial click at the Great Belt tracking site;
- $\hat{\xi}_{i,m}$ is the predicted EDA for an artificial click at sampling location i and month m, estimated from the playback study in the SAMBAH area.

Methods and and code for obtaining $\hat{\nu}_d^*$, the porpoise EDA at the Great Belt site, is given in SAMBAH Code File 2; for $\hat{\xi}^*$, the playback EDA the Great Belt site, is in SAMBAH Code File 3; for $\hat{\xi}_{i,m}$, the playback EDA in the SAMBAH area, is in SAMBAH Code File 4; and for $\hat{p}_c$ is given in SAMBAH Code 5.

[1]Denmark mainland and Denmark Borholm were treated separately in all cases where "country" is mentioned

2.4 Variance estimation

The distribution of density estimates was estimated by using a bootstrap procedure, where each component of the density estimate was generated from an independent bootstrap, as follows:

- Encounter rate (n and T). A non-parametric bootstrap was used, resampling sampling locations within countries within regions (NW or SE).
- EDA for porpoises at the Great Belt tracking site (ν^*). A non-parametric bootstrap was used, resampling porpoise encounters within diel phase. Details are given in SAMBAH Report 2a.
- EDA for playbacks at the Great Belt tracking site (ξ^*). A parametric bootstrap was used, resampling from the fitted detection function model. Details are given in SAMBAH Report 3a.
- EDA for playbacks in the SAMBAH area (ξ). A non-parametric bootstrap was used, resampling playback sessions (i.e., a set of playbacks performed at the same sampling location and date). Details are given in SAMBAH Report 3b.
- Proportion of time clicking (p_c). A parametric bootstrap was used, because there were not enough tags for a nonparametric bootstrap. The observed mean and variance from SAMBAH Report 4a was fitted to a beta distribution, by matching the moment. This distribution was then sampled from to produce a bootstrap realizations of p_c. (For the record, since it is not recorded elsewhere, the random seed used for this parametric bootstrap was 7049681.)

In all cases, 1000 bootstrap resamples were generated. For each bootstrap replicate, porpoise density at each site and month was then estimated, using equations 1, 2 and 3; these site and month estimates were combined as described in section 2.1 to produce 1000 bootstrap replicate estimates of density and abundance at the level of season and region. Estimates of variance in density and abundance were derived from the bootstrap replicates using the usual estimator, and confidence intervals were derived using the percentile method.

3 Results

3.1 Encounter rate

Mean encounter rates by country and month are shown in Figure 1.

3.2 Effective detection area

Mean effective detection area by country and month, calculated from equation 3 is shown in Figure 2.

For the paper, we wanted a plot of playback EDA by country and month, so here it is in Figure 3.

Here we also report results for p_c, the probability of clicking during an encounter. The mean proportion encounter intervals when animals were found not to have been silent, over the 6 tagged animals (weighted by record length), was 0.8215 with variance 0.003162. This was fitted to a beta distribution, which yielded beta parameters α =37.27, β =8.099; this distribution is in Figure 4.

3.3 Density and abundance

Density and abundance estimates by country and month are shown in Figures 6 and 7. The estimates were aggregated by country within region and season; these are shown in Table 1. Values aggregated by country within region are shown in Table 2.

3.4 Variance in density and abundance

Bootstrap results are presented as the bootstrap mean, 95% confidence limits (labelled LCL and UCL), and coefficient of variation (CV, the square root of the estimated variance divided by the original estimate, not the mean of the bootstrap estimate).

The estimates were aggregated by country within region and season; these are shown in Table 3. Values aggregated by country within region are shown in Table 4. Totals are shown in Tables 5 and 6.

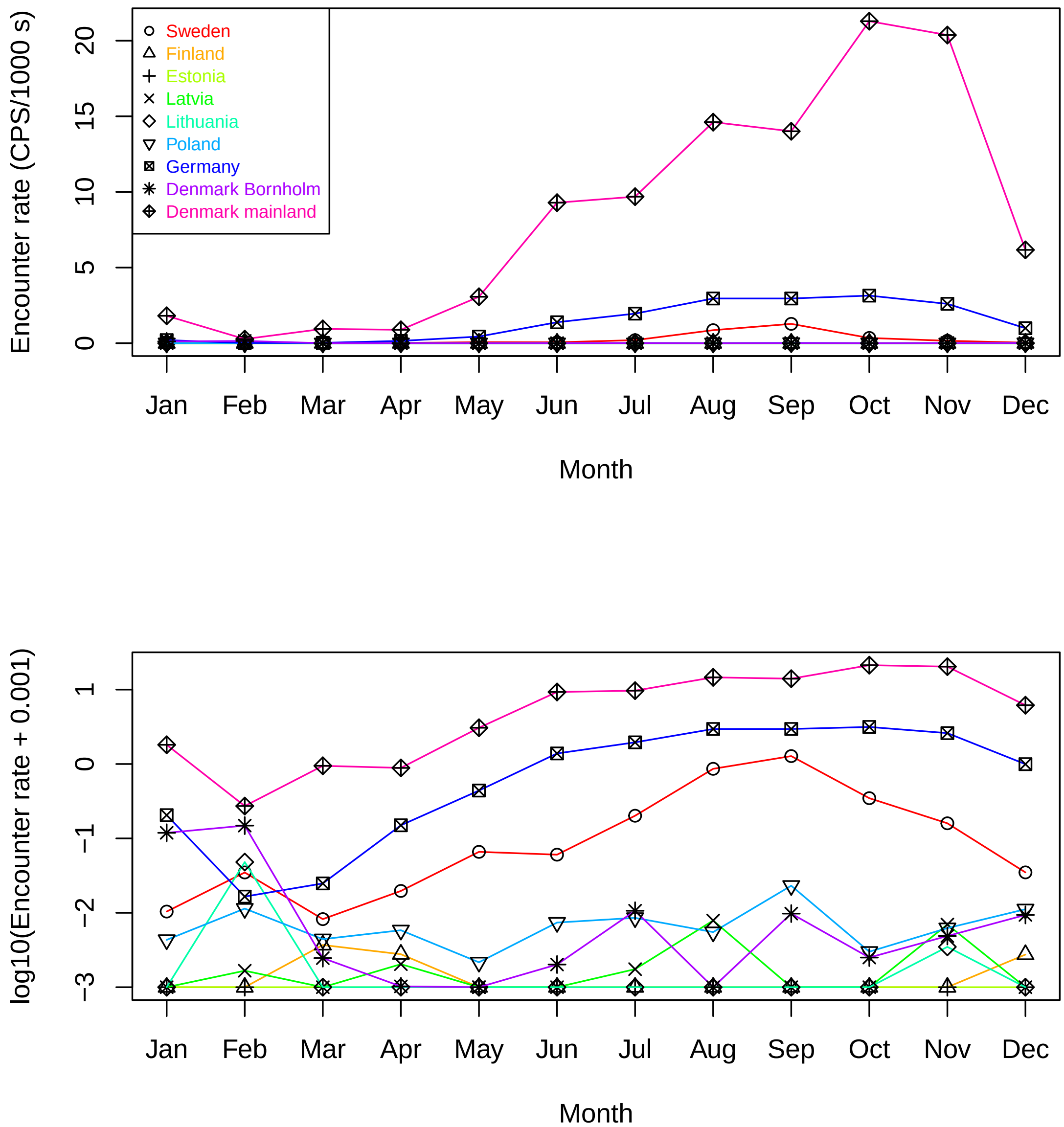

Figure 1: Average encounter rate by month, by country. Bottom plot is on the log10 scale, to show patterns in low density countries.

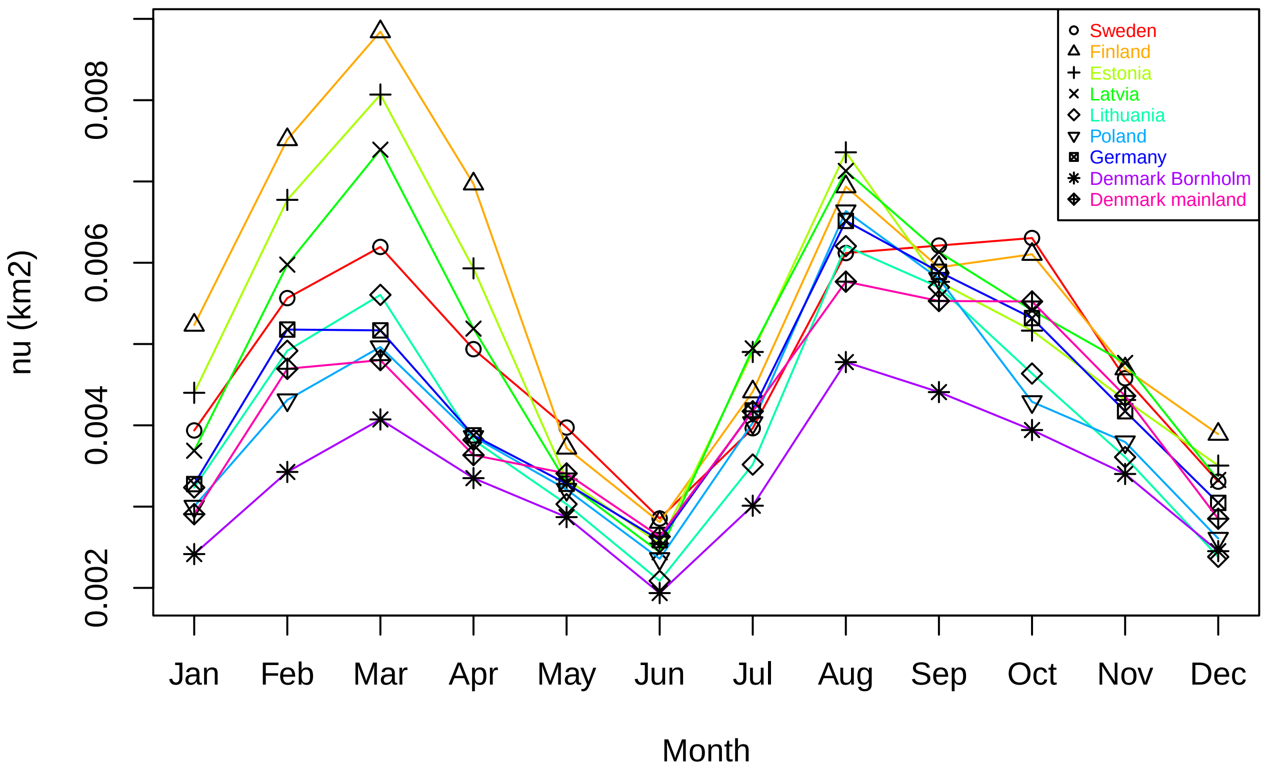

Figure 2: Average effective detection area by month, by country.

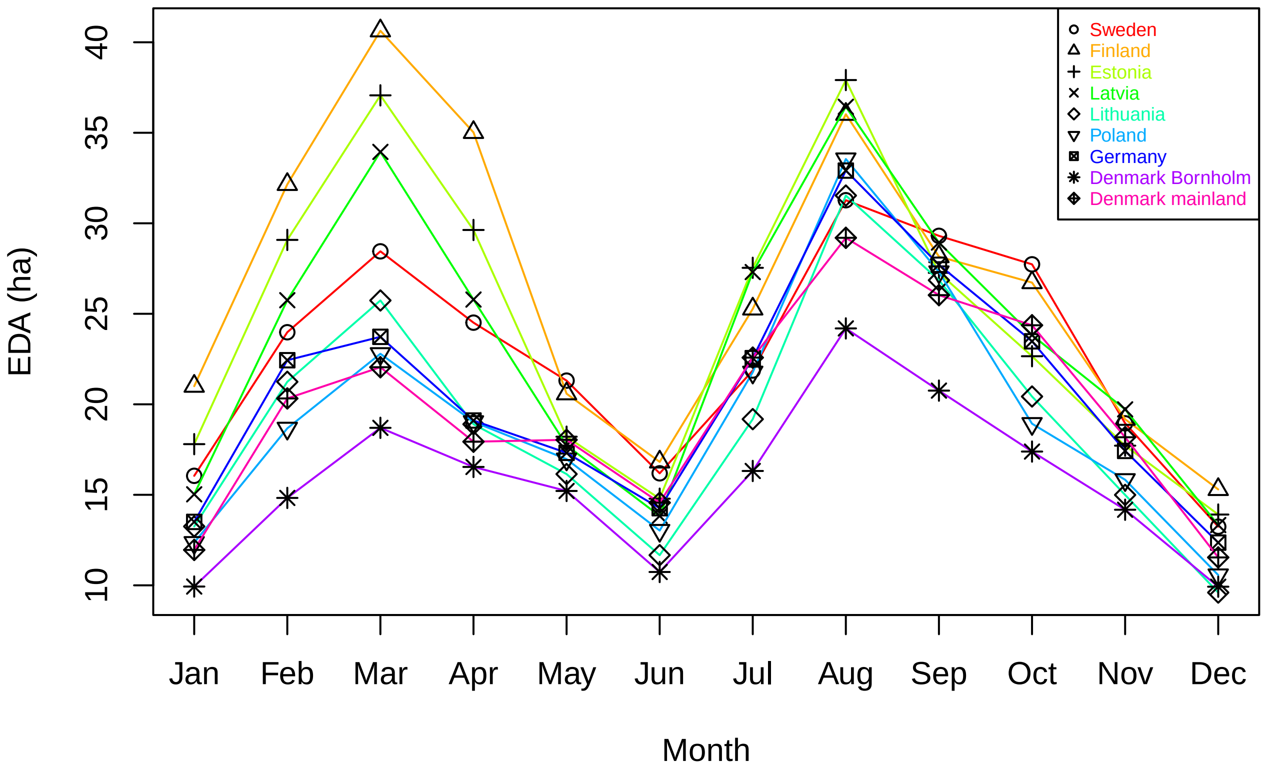

Figure 3: Average effective detection area for playbacks, from the playback experiment by month, by country.

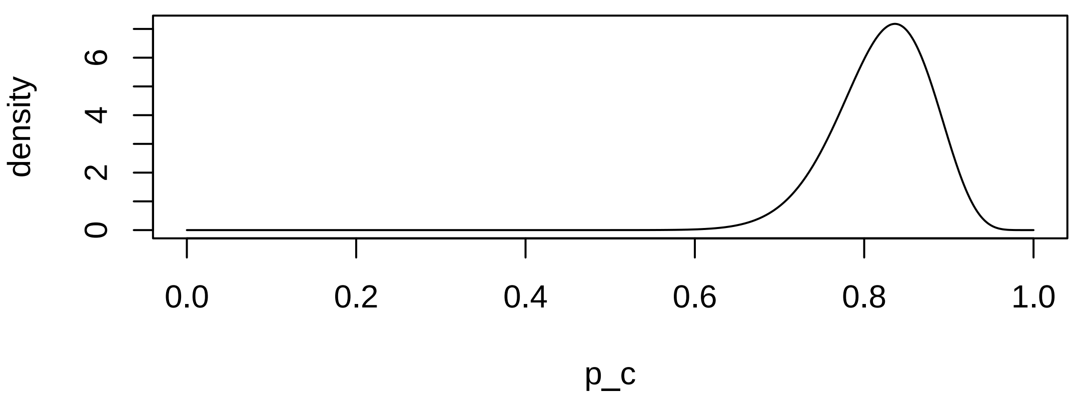

Figure 4: Beta density used to generate samples for p_c from tag data.

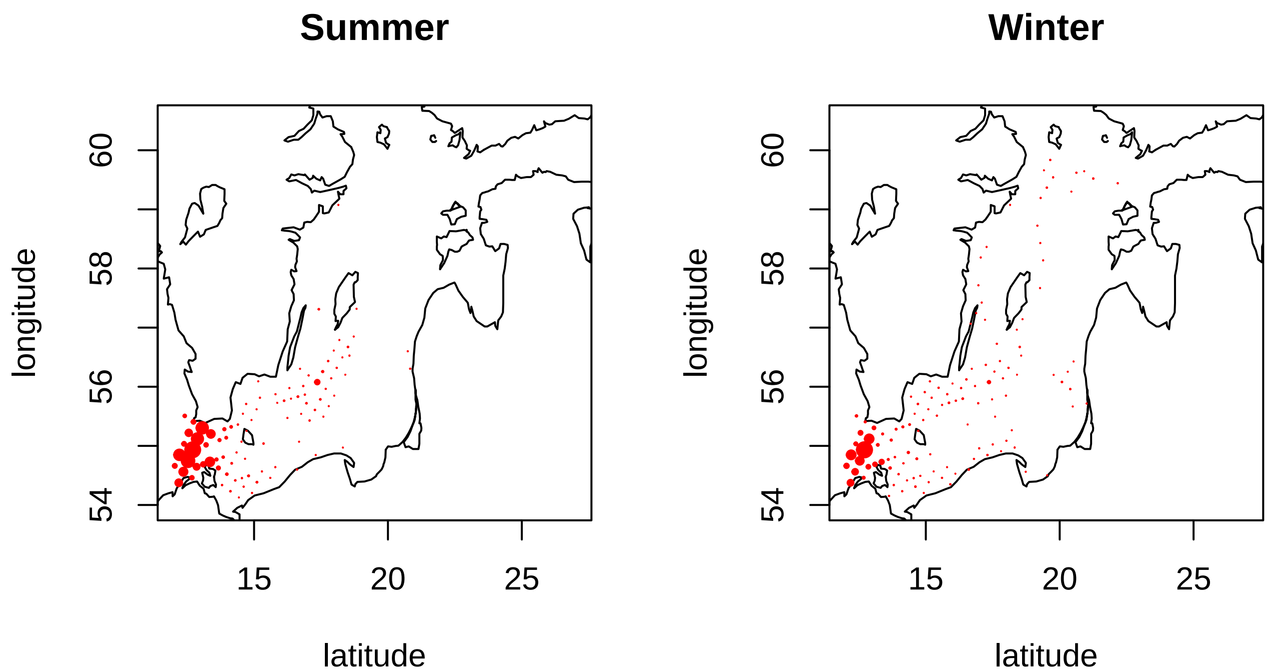

Figure 5: Density map in Summer and Winter. Area of circle is porportional to density.

A 95% confidence interval on the difference in abundance between winter and summer by region was requested. This was calculated from the bootstraps. The interval for the NE region is (-812, 317) and for the SW region is (-27160, 3874).

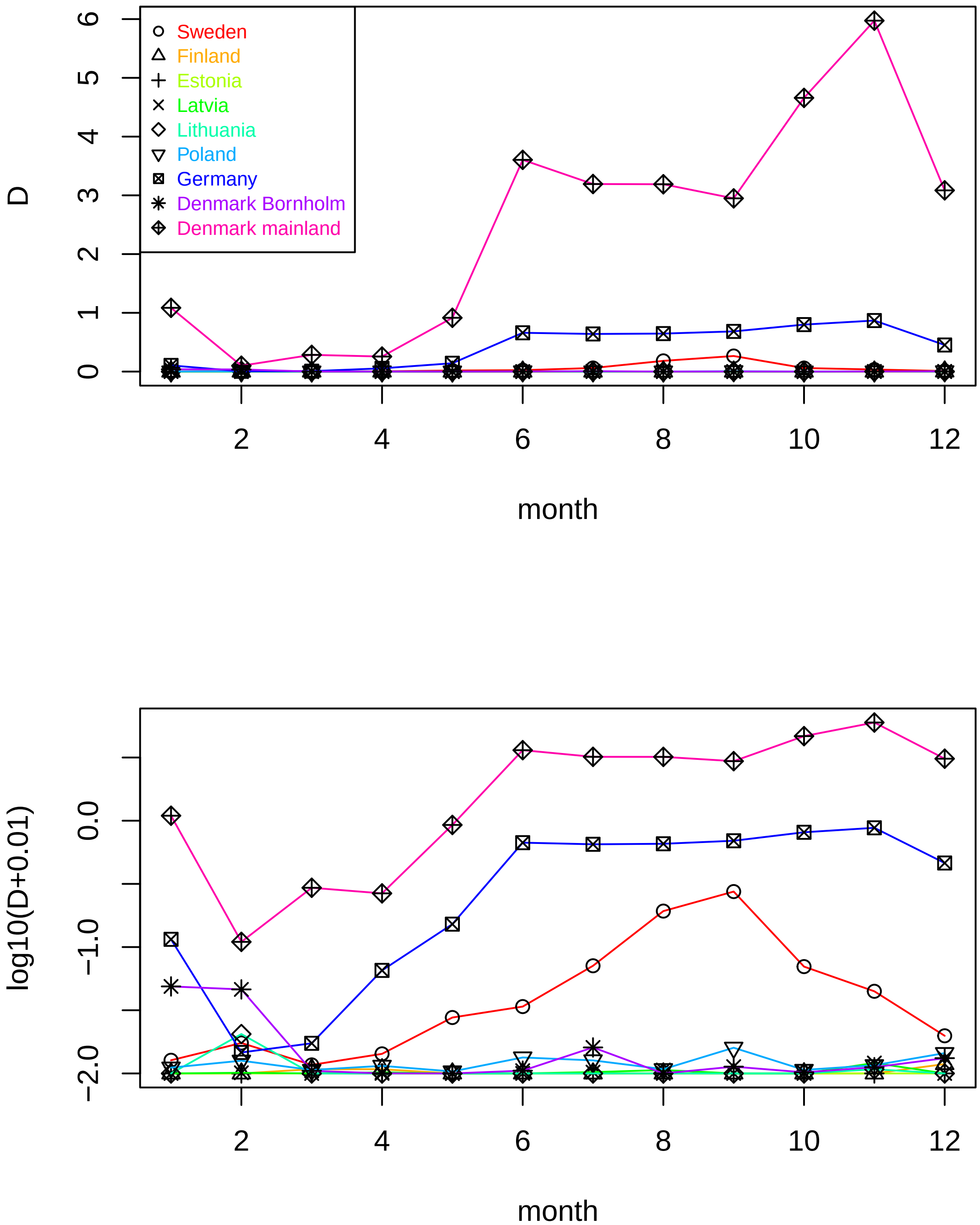

Figure 6: Average density by month, by country. Bottom plot is on the log10 scale, to show patterns in low density countries.

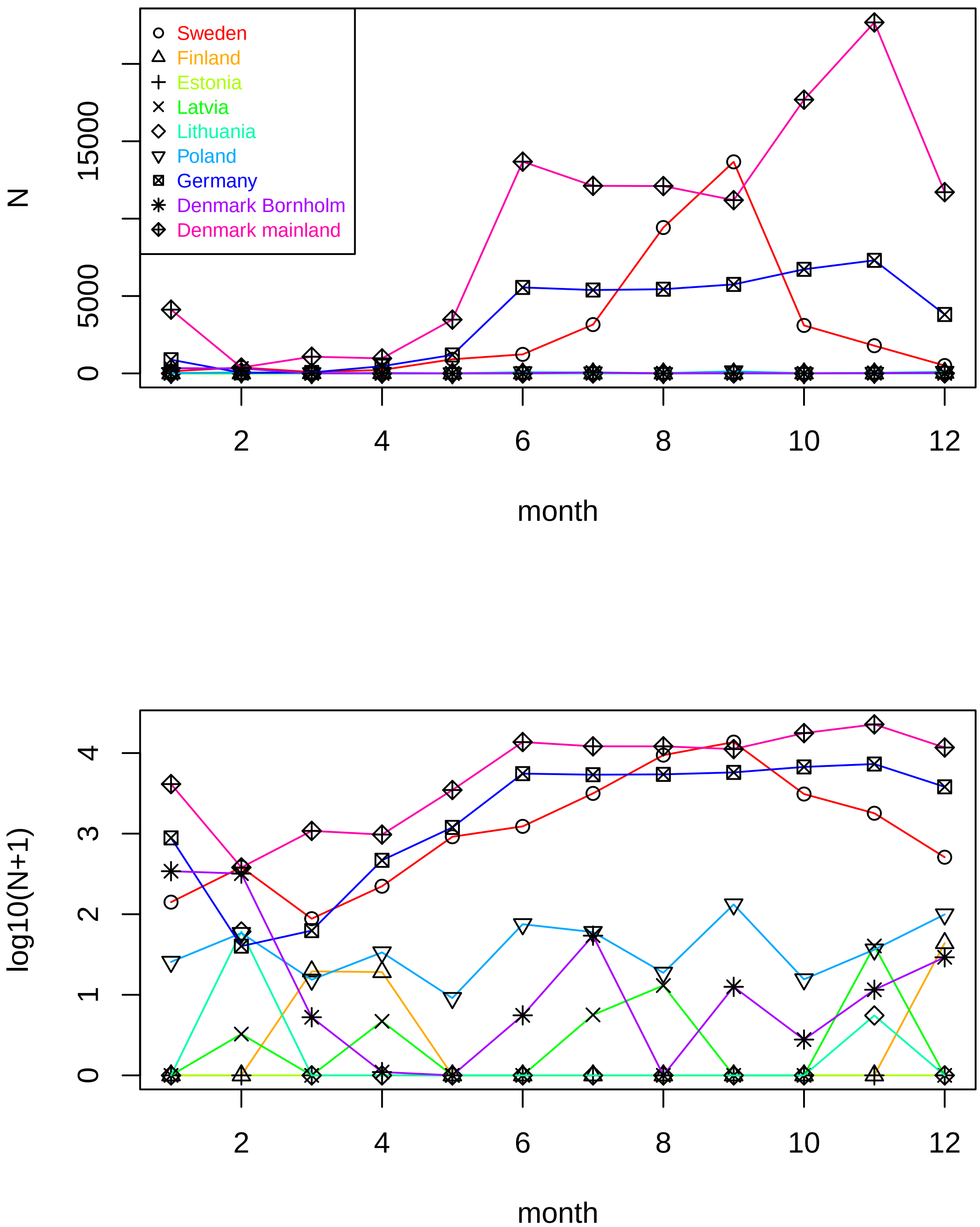

Figure 7: Average abundance by month, by country. Bottom plot is on log10 scale.

	season	region	name	n.stations	A	D	N
1	Summer	NE	Sweden	87	44860	10.74	482
2	Summer	NE	Finland	46	23005	0.00	0
3	Summer	NE	Estonia	34	24588	0.00	0
4	Summer	NE	Latvia	32	18735	0.10	2
5	Summer	NE	Lithuania	8	5719	0.00	0
6	Summer	NE	Poland	25	14849	0.48	7
7	Summer	NE	Denmark Bornholm	1	848	0.00	0
8	Summer	SW	Sweden	12	6750	674.87	4555
9	Summer	SW	Poland	14	7092	5.18	37
10	Summer	SW	Germany	16	8412	593.57	4993
11	Summer	SW	Denmark Bornholm	11	7932	1.63	13
12	Summer	SW	Denmark mainland	8	3797	3039.14	11538
13	Winter	NE	Sweden	85	44860	4.19	188
14	Winter	NE	Finland	45	23005	0.55	13
15	Winter	NE	Estonia	30	24588	0.00	0
16	Winter	NE	Latvia	27	18735	0.49	9
17	Winter	NE	Lithuania	8	5719	1.46	8
18	Winter	NE	Poland	23	14849	1.53	23
19	Winter	NE	Denmark Bornholm	1	848	2.40	2
20	Winter	SW	Sweden	12	6750	54.56	368
21	Winter	SW	Poland	14	7092	2.31	16
22	Winter	SW	Germany	16	8412	261.85	2203
23	Winter	SW	Denmark Bornholm	11	7932	10.08	80
24	Winter	SW	Denmark mainland	8	3797	2126.32	8073

Table 1: Estimates of density (per 1000km2) and abundance by country and region

	season	region	A	D	N
1	Summer	NE	132603	3.70	491
2	Summer	SW	33982	621.98	21136
3	Winter	NE	132603	1.83	243
4	Winter	SW	33982	316.05	10740

Table 2: Estimates of density (per 1000km2) and abundance by country and region

	season	region	name	mean.D	lcl.D	ucl.D	cv.D	mean.N	lcl.N	ucl.N
1	Summer	NE	Sweden	11.50	1.41	24.36	0.691	516	63	1093
2	Summer	NE	Finland	0.00	0.00	0.00		0	0	0
3	Summer	NE	Estonia	0.00	0.00	0.00		0	0	0
4	Summer	NE	Latvia	0.10	0.00	0.23	0.691	2	0	4
5	Summer	NE	Lithuania	0.00	0.00	0.00		0	0	0
6	Summer	NE	Poland	0.49	0.07	0.96	0.539	7	1	14
7	Summer	NE	Denmark Bornholm	0.00	0.00	0.00		0	0	0
8	Summer	SW	Sweden	744.43	118.96	1720.40	0.657	5025	803	11612
9	Summer	SW	Poland	5.59	0.66	11.74	0.558	40	5	83
10	Summer	SW	Germany	638.83	250.92	1214.57	0.408	5374	2111	10216
11	Summer	SW	Denmark Bornholm	1.81	0.15	4.26	0.703	14	1	34
12	Summer	SW	Denmark mainland	3247.95	1386.81	5795.17	0.367	12331	5265	22002
13	Winter	NE	Sweden	4.62	1.15	10.01	0.628	207	52	449
14	Winter	NE	Finland	0.72	0.08	2.38	1.127	17	2	55
15	Winter	NE	Estonia	0.00	0.00	0.00		0	0	0
16	Winter	NE	Latvia	0.57	0.01	1.96	1.147	11	0	37
17	Winter	NE	Lithuania	1.58	0.04	3.98	0.795	9	0	23
18	Winter	NE	Poland	1.66	0.57	3.14	0.435	25	9	47
19	Winter	NE	Denmark Bornholm	0.00	0.00	0.00	0.000	0	0	0
20	Winter	SW	Sweden	61.92	21.11	127.88	0.497	418	143	863
21	Winter	SW	Poland	2.64	0.44	6.69	0.737	19	3	47
22	Winter	SW	Germany	304.56	89.66	617.08	0.512	2562	754	5191
23	Winter	SW	Denmark Bornholm	11.34	1.00	28.96	0.761	90	8	230
24	Winter	SW	Denmark mainland	2420.08	823.98	4998.17	0.510	9188	3128	18976

Table 3: Estimates of density (per 1000km2) and abundance by season and country-within-region, with measures of uncertainty. mean.D and mean.N are means from the bootstrap. CV uses the standard error from the bootstrap, but the estimate from the original analysis.

	season	region	D	mean.D	lcl.D	ucl.D	cv.D	N	mean.N	N.20	lcl.N	ucl.N
1	Summer	NE	3.70	3.96	0.54	8.33	0.680	491	525	138	71	1105
2	Summer	SW	621.98	670.45	363.43	1143.21	0.334	21136	22783	17043	12350	38849
3	Winter	NE	1.83	2.02	0.71	4.22	0.541	243	268	149	94	560
4	Winter	SW	316.05	361.27	155.24	702.10	0.453	10740	12277	8130	5275	23859

Table 4: Estimates of density (per 1000km2) and abundance by season and region with measures of uncertainty

	season	A	N	mean.N	lcl.N	ucl.N	cv.N
1	Summer	166585	21627	23309	12937	39816	0.330
2	Winter	166585	10983	12545	5525	24546	0.448

Table 5: Estimates of abundance by season with measures of uncertainty. mean.N is the mean from the bootstraps. CV uses the standard deviation from the bootstraps but the N from the original analysis.

	season	D	mean.D	lcl.D	ucl.D
1	Summer	129.826	139.919	77.661	239.015
2	Winter	65.930	75.307	33.165	147.346

Table 6: Estimates of density (per 1000km2) by season with measures of uncertainty. mean.D is the mean from the bootstraps