

Supplementary material: Incorporating phenology to
estimate species' population trends from snapshot
citizen-science data

Table S1: Summary of BBC data by year: start and end dates (day of year) for the BBC sampling period, total number of 15-minute counts made and the number of UK 1km and 10km squares sampled. Counts made outside of the official sampling period were excluded. Counts where zero individuals were recorded were included.

Year	Start day	End day	Number of counts	1km	10km
2011	197	219	24957	14631	1904
2012	196	218	17586	11039	1872
2013	201	223	36350	17356	1985
2014	200	222	38114	17903	2057
2015	198	221	43593	19814	1973
2016	197	220	32366	15267	1935
2017	195	218	50971	23410	2116
2018	201	224	86303	31951	2296
2019	200	223	98383	35744	2479
2020	199	222	132531	36299	2262
2021	197	220	134423	34756	2300
2022	196	219	90395	27706	2190

Table S2: Summary of butterfly species featured in the BBC. Total represents the total number of individuals counted and 1km and 10km denote the number of squares the species has been recorded in by the BBC (up to 2022). Counts made outside of the sampling period were excluded. Three species with incomplete data for the full date range were excluded from analyses.

Species	Scientific name	Date range	Total	1km	10km
Small White	<i>Pieris rapae</i>	2011 - 2022	1573442	69452	2536
Large White	<i>Pieris brassicae</i>	2011 - 2022	1438953	65823	2465
Gatekeeper	<i>Pyronia tithonus</i>	2011 - 2022	1185687	53691	1939
Meadow Brown	<i>Maniola jurtina</i>	2011 - 2022	1057665	56933	2501
Peacock	<i>Aglais io</i>	2011 - 2022	869309	55481	2373
Red Admiral	<i>Vanessa atalanta</i>	2011 - 2022	516956	57494	2460
Painted Lady	<i>Vanessa cardui</i>	2011 - 2022	508877	37878	2478
Small Tortoiseshell	<i>Aglais urticae</i>	2011 - 2022	412511	50403	2499
Ringlet	<i>Aphantopus hyperantus</i>	2011 - 2022	326549	38828	2389
Green-veined White	<i>Pieris napi</i>	2011 - 2022	285133	40169	2471
Common Blue	<i>Polyommatus icarus</i>	2011 - 2022	231705	32606	2226
Speckled Wood	<i>Pararge aegeria</i>	2011 - 2022	227014	36929	2170
Comma	<i>Polygonia c-album</i>	2011 - 2022	223099	41740	1883
Holly Blue	<i>Celastrina argiolus</i>	2011 - 2022	115424	23977	1727
Small Copper	<i>Lycaena phlaeas</i>	2011 - 2022	105955	26880	2075
Marbled White	<i>Melanargia galathea</i>	2011 - 2022	104308	12810	1364
Brimstone	<i>Gonepteryx rhamni</i>	2011 - 2022	73552	20641	1541

Table S3: Summary of trends (mean and standard error, SE) and statistical power estimated from 1000 simulated BBC data sets (for each scenario), calculated (i) without and (ii) with phenology adjustment. The number of sites is the total number of sites, M , before sampling.

Phenology Scenario	Trend (%)	Number of sites	(i) without phenology adjustment			(ii) with phenology adjustment		
			Mean	SE	Power	Mean	SE	Power
Holly Blue	-25	1000	-16.1	0.7	0.4	-24.1	0.5	42.6
Holly Blue	-25	2500	-15.8	0.5	0.0	-24.1	0.4	64.9
Holly Blue	-25	5000	-17.7	0.4	0.0	-25.0	0.3	81.5
Holly Blue	-25	10000	-17.8	0.3	0.0	-24.8	0.2	88.7
Holly Blue	-25	50000	-18.2	0.1	0.0	-25.0	0.2	96.8
Gatekeeper	-25	1000	-9.2	0.7	2.8	-24.5	0.4	50.9
Gatekeeper	-25	2500	-10.6	0.5	1.6	-24.0	0.3	76.7
Gatekeeper	-25	5000	-11.3	0.4	0.6	-24.6	0.2	89.9
Gatekeeper	-25	10000	-11.4	0.3	0.3	-25.1	0.2	93.6
Gatekeeper	-25	50000	-11.4	0.2	0.1	-25.0	0.1	98.7
Marbled White	-25	1000	-71.0	0.3	2.5	-23.7	0.5	30.3
Marbled White	-25	2500	-71.4	0.2	0.6	-24.6	0.4	56.2
Marbled White	-25	5000	-71.6	0.1	0.6	-24.9	0.3	68.9
Marbled White	-25	10000	-71.4	0.1	0.2	-24.4	0.3	81.1
Marbled White	-25	50000	-71.8	0.1	0.0	-24.8	0.2	92.4
Holly Blue	0	1000	12.7	0.9	0.0	1.4	0.6	5.0
Holly Blue	0	2500	10.8	0.6	0.0	0.5	0.5	6.5
Holly Blue	0	5000	10.6	0.6	0.1	0.4	0.4	7.7
Holly Blue	0	10000	9.5	0.3	0.0	-0.1	0.3	7.3
Holly Blue	0	50000	9.2	0.2	0.0	-0.0	0.2	7.5
Gatekeeper	0	1000	19.3	1.0	2.8	0.4	0.5	6.2
Gatekeeper	0	2500	20.3	0.7	2.5	0.7	0.4	7.2
Gatekeeper	0	5000	18.1	0.5	0.7	0.1	0.3	5.0
Gatekeeper	0	10000	18.0	0.4	0.3	0.1	0.3	5.3
Gatekeeper	0	50000	18.0	0.2	0.3	0.0	0.1	5.7
Marbled White	0	1000	-61.4	0.3	0.4	1.4	0.6	5.0
Marbled White	0	2500	-62.0	0.2	0.2	0.5	0.5	5.3
Marbled White	0	5000	-62.2	0.2	0.0	-0.0	0.4	6.9
Marbled White	0	10000	-62.0	0.2	0.0	0.6	0.4	8.4
Marbled White	0	50000	-62.3	0.1	0.0	0.2	0.3	8.7

Table S4: Summary of the number of BBC sites (1km x 1km squares) modelled per year using the phenology adjustment approach (i) including imputed zeros and (ii) based on positive counts only.

Species	(i) Sites			(ii) Sites with positive counts		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
Small White	21898	10000	33313	16595	5398	27382
Large White	21427	9822	32795	15669	4801	27066
Red Admiral	19955	9189	29798	11469	2777	20528
Peacock	19418	7788	29608	10793	958	21871
Meadow Brown	19186	9518	29169	10684	5073	19587
Gatekeeper	18734	9121	28338	12248	5139	19781
Small Tortoiseshell	18192	8623	26783	8728	2157	14711
Comma	16745	8115	24367	7743	2349	12379
Green-veined White	15868	7948	22296	6091	1890	10708
Painted Lady	15244	6988	27186	4948	471	22424
Speckled Wood	14995	7355	21449	6071	1692	10355
Common Blue	14032	6993	20117	5202	1177	10729
Ringlet	13813	7792	21150	5432	2545	12273
Holly Blue	11831	5945	16938	3896	980	8189
Small Copper	11747	5895	17196	3401	716	7862
Brimstone	9639	4695	13728	2988	454	5416
Marbled White	5231	3311	8151	1626	671	4356

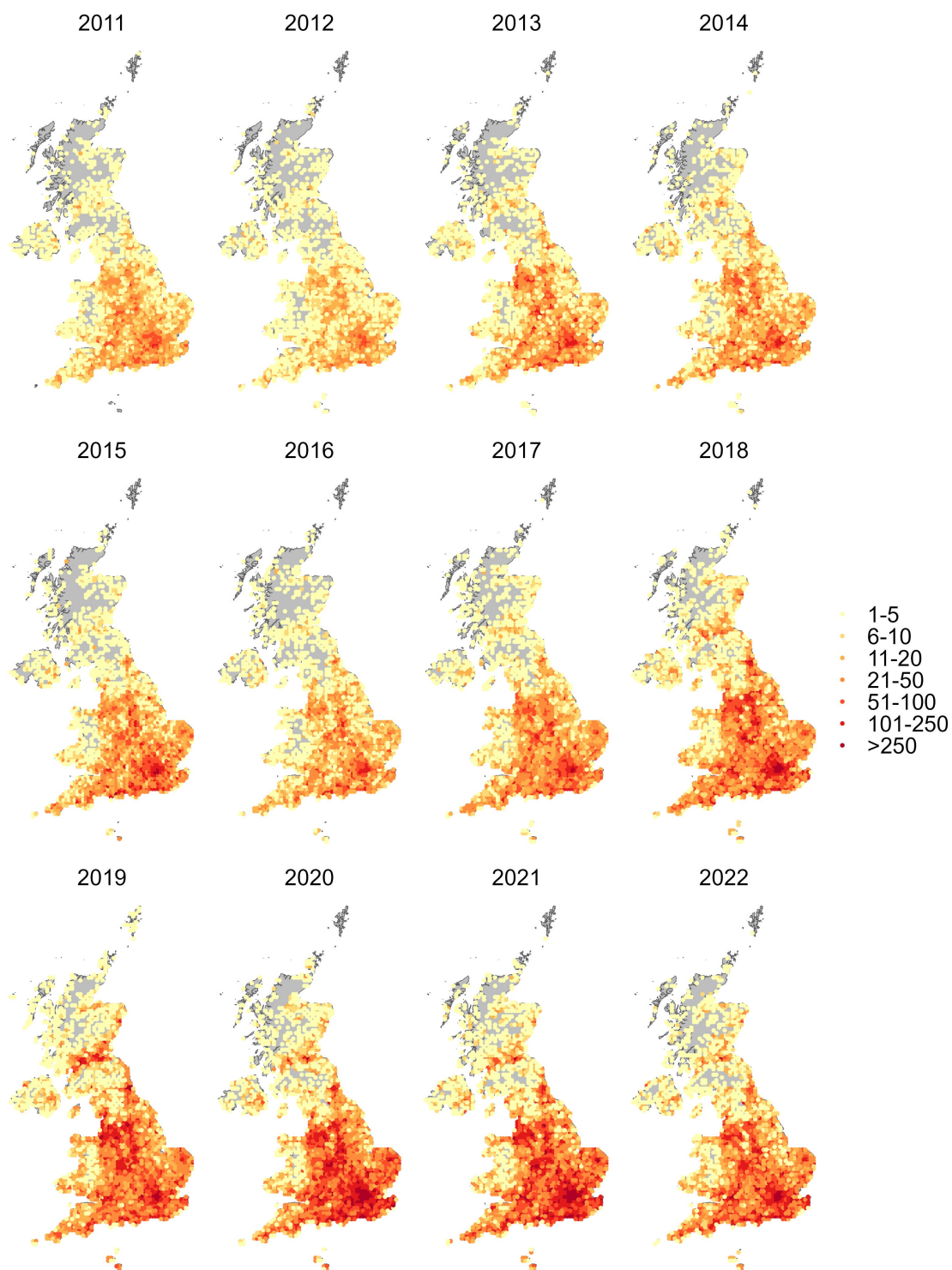


Figure S1: Maps showing the number of 15-minute BBC counts per 10km square per year.

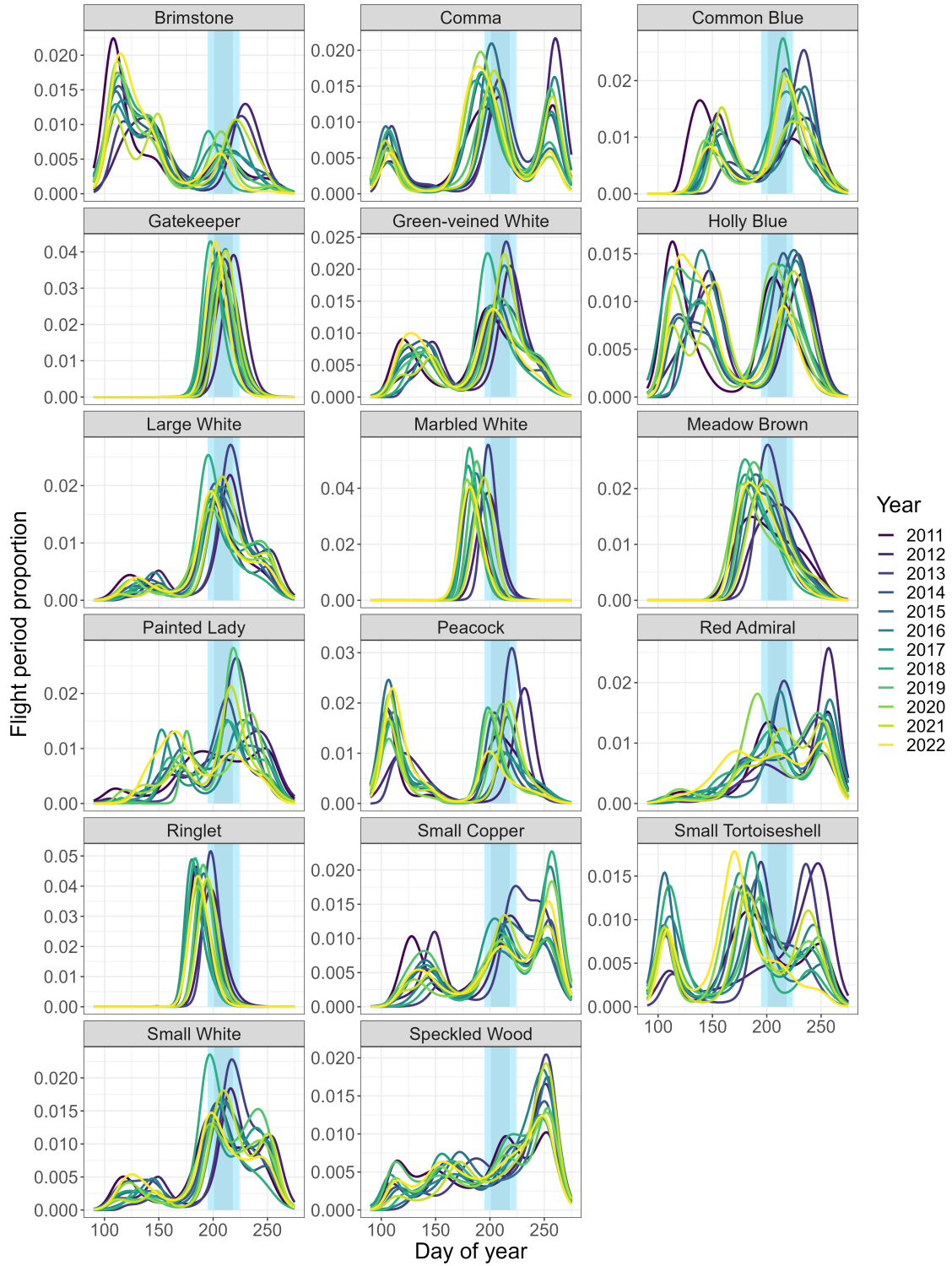


Figure S2: Annual flight periods for each of 17 BBC species for 2011-2022, estimated by fitting the GAI approach to UKBMS data. Each annual curve sums to one. The shaded blue areas represent the BBC sampling period, with shading to indicate the variation in dates across years.

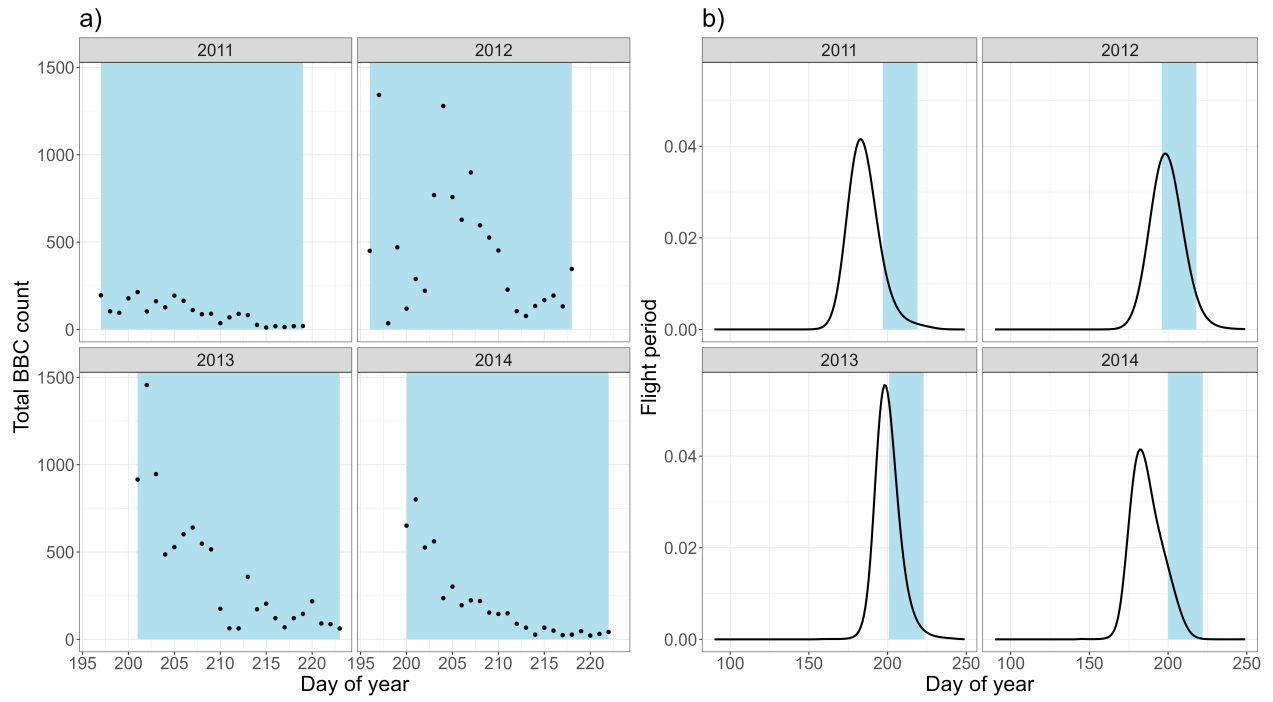


Figure S3: Marbled White example for four years: a) the total number of individuals counted by BBC per day; b) flight periods estimated from UKBMS data. The blue shaded area represents the BBC sampling period in each year.

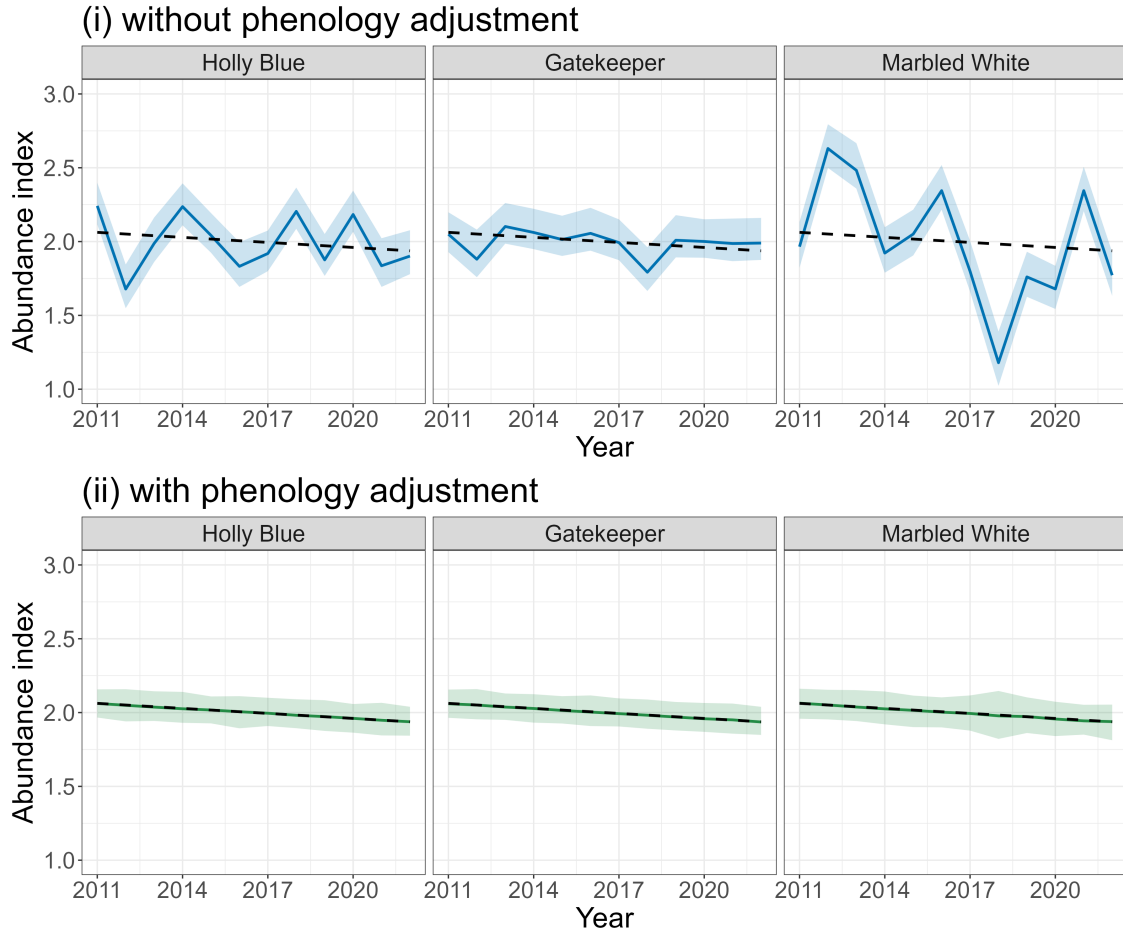


Figure S4: Estimated abundance indices (on the log10 scale) from 1000 simulations for three phenology scenarios (i) without and (ii) with phenology adjustment, where the solid lines indicate the median value and 95% quantiles are shown as shaded intervals. Dashed lines indicate the true underlying index (averaged across simulations). Based on scenarios for $M = 2500$ sites.

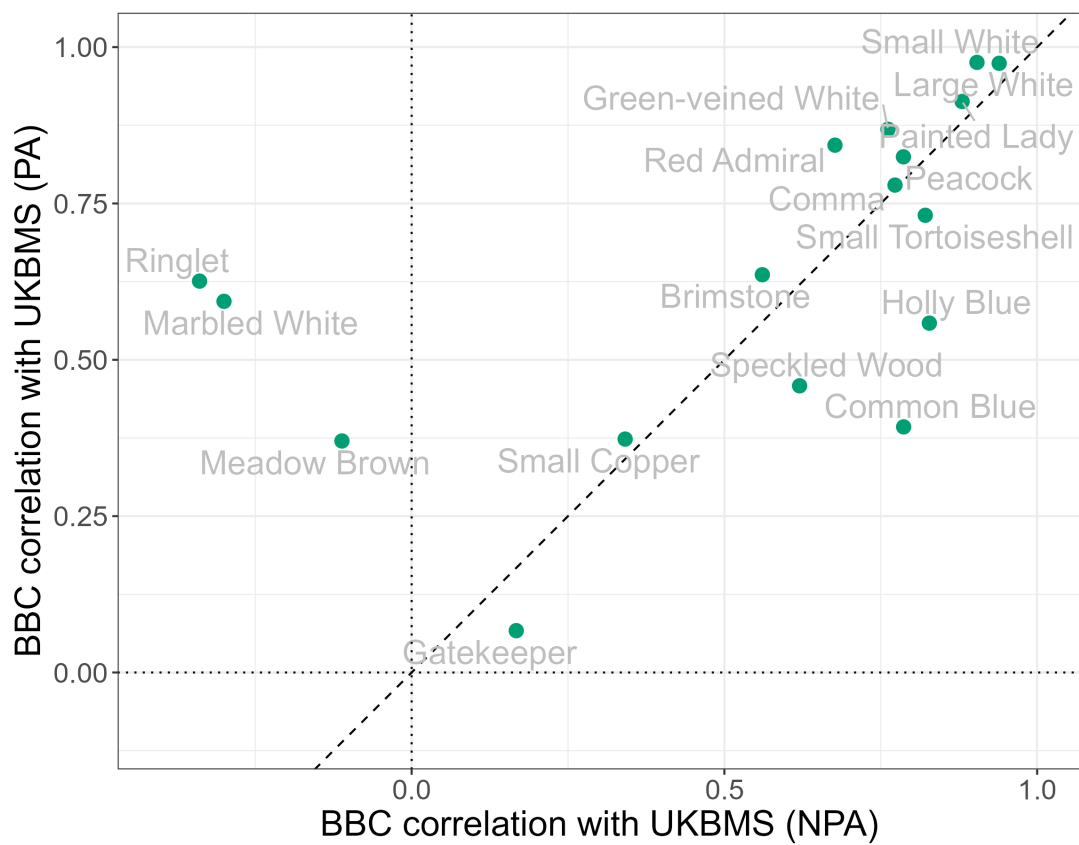


Figure S5: Comparison of Pearson correlations between BBC and UKBMS abundance indices without (NPA) and with (PA) phenology adjustment. Dashed line indicates 1-1 line and dotted lines indicate zero.

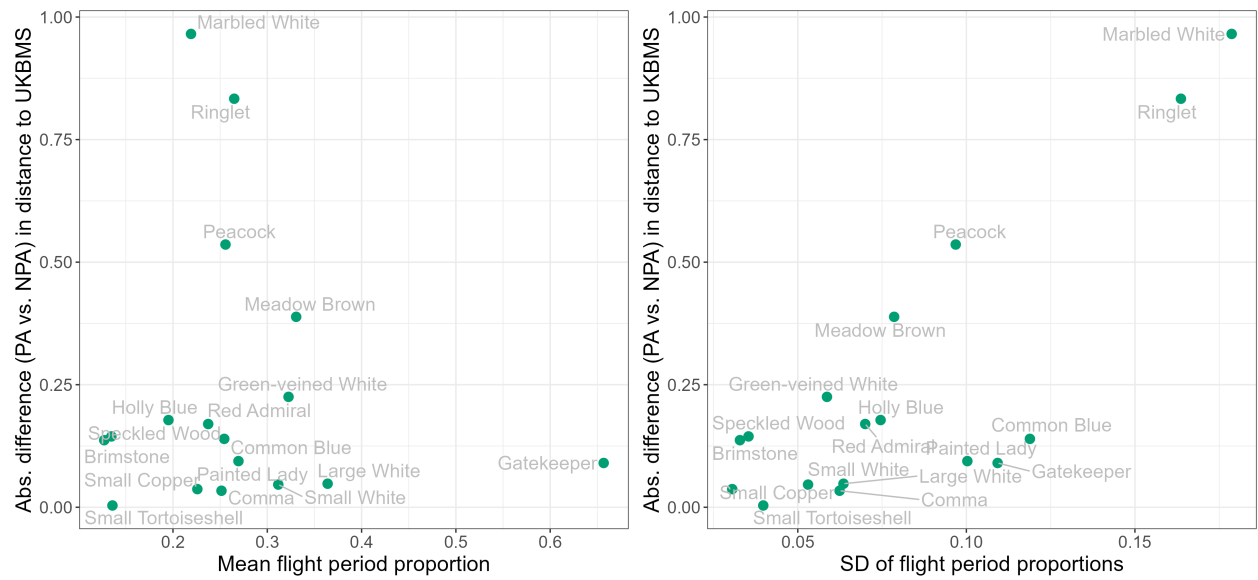


Figure S6: Comparison of the mean (left) and standard deviation (SD, right) of the proportion of each species' flight period covered by the BBC sampling period each year (see Figure 1 of the paper) against the absolute difference in the distances of BBC abundance indices with (PA) and without (NPA) phenology adjustment to UKBMS indices.

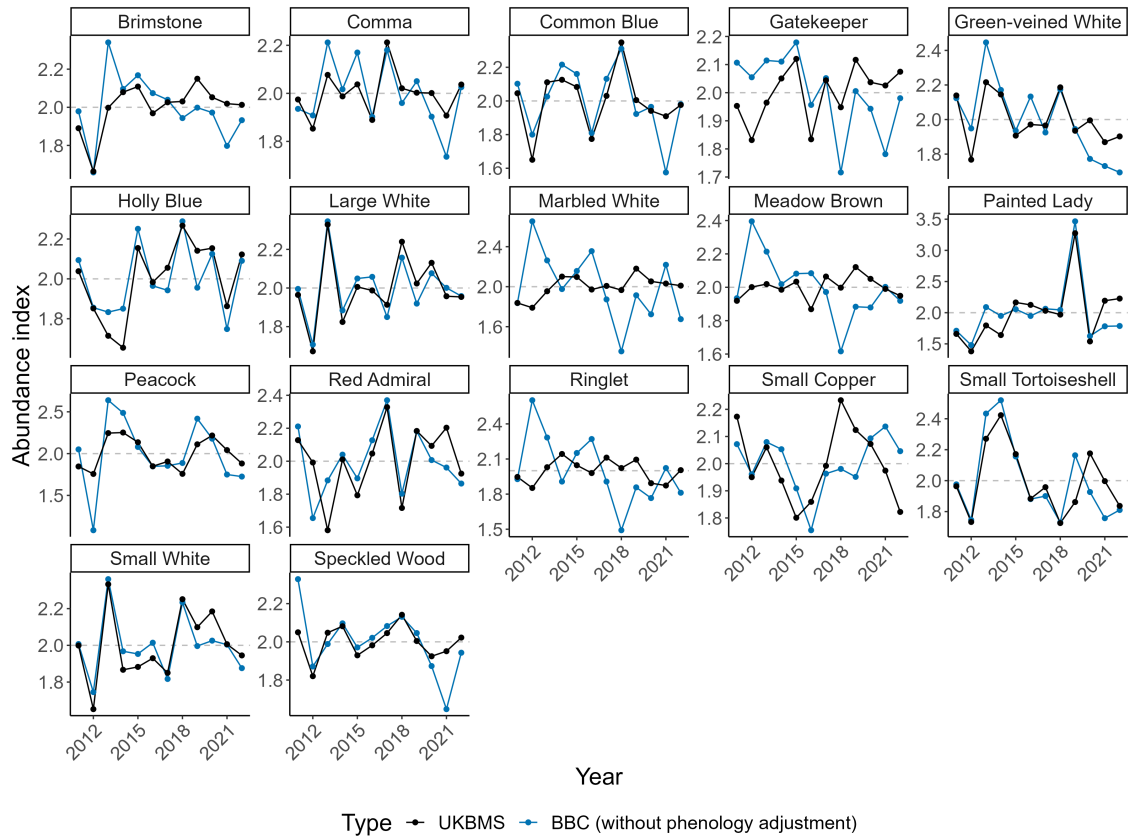


Figure S7: Relative abundance indices produced from BBC data without phenology adjustment (blue) and from UKBMS data (black). Indices are on the log10 scale with a mean value of 2 (indicated by the horizontal dashed lines).

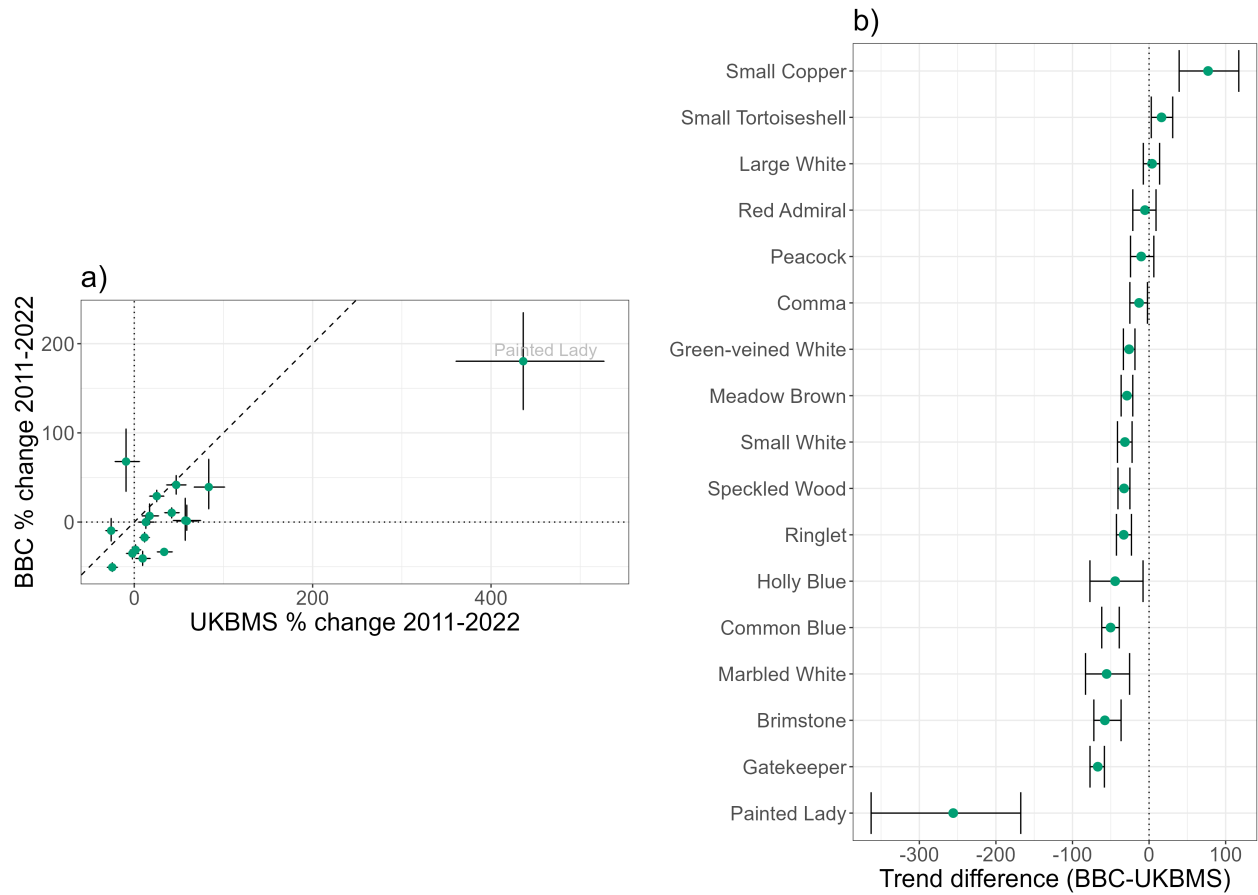


Figure S8: Comparison of (a) and difference between (b) abundance trend estimates for 2011-2022 from UKBMS (using the GAI approach) and BBC (with phenology adjustment approach), including Painted Lady. Error bars represent 95% confidence intervals produced using non-parametric bootstrapping. The dashed line represents the 1-1 line and dotted lines represent zero. Figure 6 of the main paper shows the same figure but with Painted Lady excluded.

S1 Exploring site size specification for BBC data

In the paper, a BBC site is defined by the 1km square that counts were recorded in. Here we apply the phenology adjustment approach to estimate abundance indices and trends from BBC data, but instead defining sites by 100m or 10km. The choice of site size will influence the number of sites analysed per year, and the level of replication for each site. BBC abundance indices for varying site definitions are broadly similar (Figure S9), but indices are generally closest to UKBMS when based upon the 1km site definition (Figure S10), hence this site size was adopted for all other analyses.

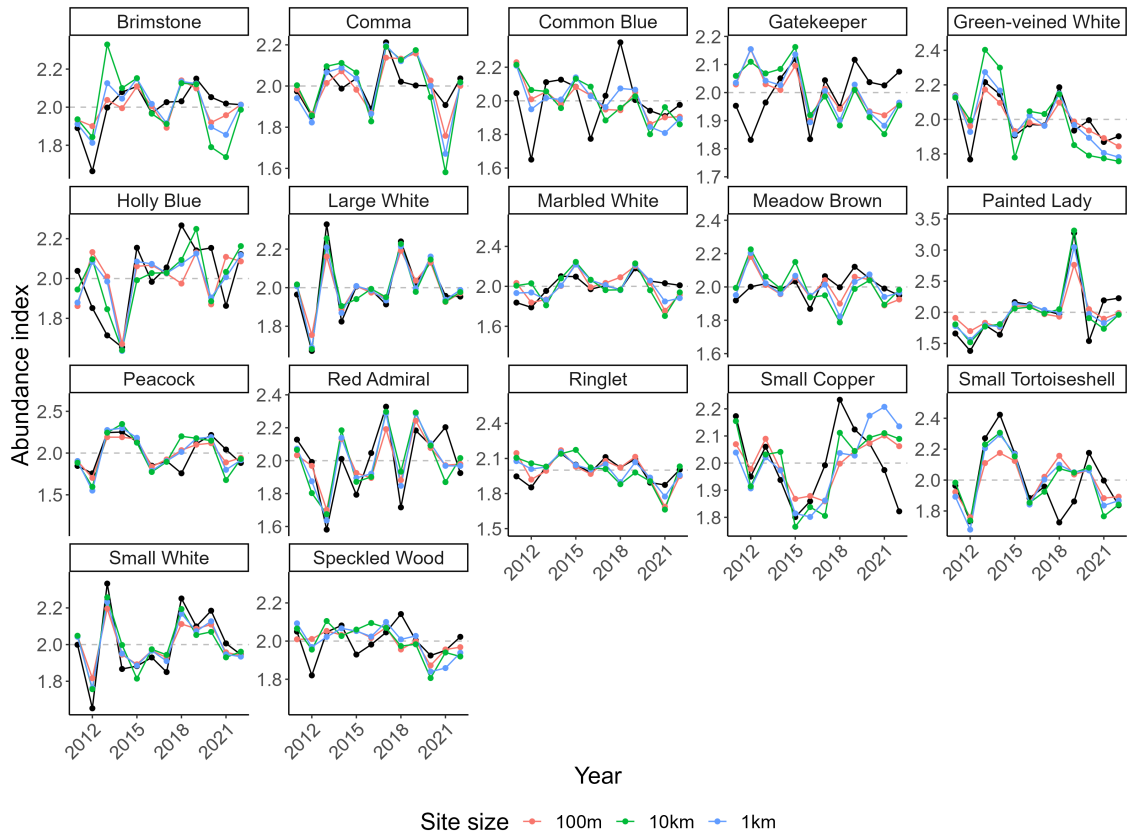


Figure S9: Relative abundance indices produced from UKBMS data (black), and from BBC data with phenology adjustment for three definitions of site size. Indices are on the log10 scale with a mean value of 2 (indicated by the horizontal dashed lines).

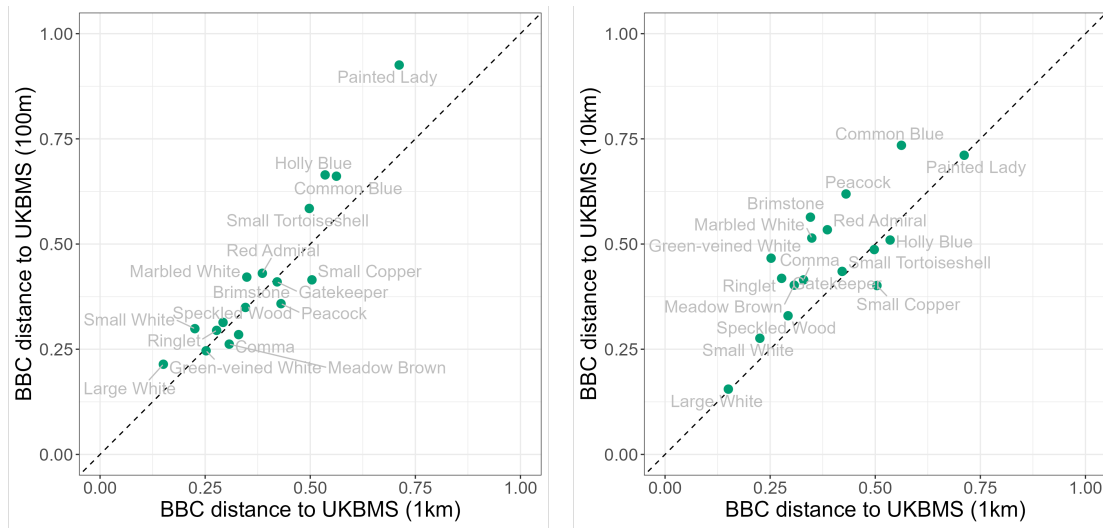


Figure S10: Comparison of Euclidean distances between BBC and UKBMS abundance indices, where BBC indices were produced using phenology adjustment for three definitions of site size. Dashed line indicates 1-1 line.

S2 Simulation study

Further details of the simulation study undertaken are provided here. As described in Section 5 of the paper, for a given scenario site abundance parameters $\{N_{m,k}\}$ were simulated for each site m of a total of M sites for each year $k = 1, \dots, 12$.

The data were then degraded to account for the fact that in practice most sites are not revisited in multiple years (based upon defining a site by a 1km square for BBC data). For each site a number of years were randomly selected for retention, based upon sampling probabilities from the actual BBC data (Table S5), such that 38% of sites were only sampled in one of the 12 years, 17% were sampled in two (random) years, and so on.

Table S5: The number of years that each BBC site (defined by 1km square) was sampled for 2011-2022.

Number of years	Proportion of sites
1	0.38
2	0.17
3	0.11
4	0.08
5	0.06
6	0.05
7	0.04
8	0.03
9	0.03
10	0.02
11	0.02
12	0.01

For each site and year combination, a number of visits (days with a count made) was then simulated - again randomly assigning a number of days to each site/year based upon sampling for BBC data (Table S6). In most cases (56%) a site only has counts for a single day in a given year. The days at which the counts were made were then selected randomly

26 from within official BBC count period dates.

Table S6: Summary of the number of days that BBC sites (defined by 1km square) were sampled in a year (within the official BBC sampling period) for 2011-2022. Cases (0.2%) where more than 20 days were sampled were excluded.

Number of days	Proportion of site/years
1	0.56
2	0.18
3	0.09
4	0.05
5	0.03
6	0.02
7	0.02
8	0.01
9	0.01
10	0.01
11-20	0.02

27 Finally a number of entries (15-minute counts) was assigned to each sampled site/date,
28 based upon sampled frequencies from BBC data (Table S7). In most cases only a single
29 15-minute count is made for a given 1km square and date.

30 Summaries of the number of sites sampled per year across simulations are given in Table
31 S8. Comparisons can be made to sampling levels for the actual BBC data (Table S4).

Table S7: Summary of the number of entries (15-minute counts) made to BBC sites on a given date. Based on data from the official sampling period for 2011-2022. Cases (0.005%) where more than 20 entries were made were excluded.

Number of entries	Proportion of site/dates
1	0.87
2	0.10
3	0.02
4-20	0.01

Table S8: Summary of the number of simulated sites modelled per year using the phenology adjustment approach (i) including zeros and (ii) based on positive counts only. Based on simulated data for the 25% decline for each total site scenario (M), which are averaged across phenology scenarios and years.

M	(i) Sites			(ii) Sites with positive counts		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
50000	12777	6880	13684	4515	1154	6175
10000	2556	1321	2846	903	209	1296
5000	1278	596	1464	452	96	666
2500	639	294	750	226	42	351
1000	256	104	321	90	12	161

S3 Applying phenology adjustment with UKBMS data from an incomplete season

As mentioned in Section 6 of the paper, in practice annual reporting of BBC data precedes the completion of the UKBMS sampling season (at the end of September). Here we therefore explore application of the phenology adjustment approach to BBC data, but based upon flight period estimates from UKBMS data up to mid-August only for the most recent year (2022), since BBC reporting typically occurs in late August.

Annual flight periods for the 17 BBC species were estimated by applying the GAI approach to UKBMS data for 2011-2022 (see Section 3.1 of the paper), but for 2022 we restrict the UKBMS data to up to mid August only (day 226 of the year). The flight period estimates are generally similar to those produced from the full season data set (Figure S11), but flight periods for species which fly later in the UKBMS season, for example Small Copper and Speckled Wood, are understandably not so well estimated from data up to mid August only. There is the potential to therefore over-estimate the proportion of the flight period sampled by the BBC in these instances. Scaling the estimates by the proportion of the flight period sampled up to mid August (day 226 of the year), averaged from the preceding five years, results in more reasonable flight period estimates (Figure S11).

The effect of this additional scaling (for 2022) on the BBC abundance indices (produced with phenology adjustment) can be seen in Figure S12, and the associated absolute differences for the 2022 index value (relative to phenology adjustment using full season UKBMS data) suggest the scaling is particularly beneficial for species that are on the wing in late August and September (Figure S13), for example Speckled Wood and Small Copper.

These results suggest that the phenology adjustment approach can still be applied when full season UKBMS data are not available, but scaling of flight period estimates is recommended for later-flying species. In practice, data from all UKBMS transects are unlikely to be available by mid August, since recorders may not enter data “live”, for example they may wait until the end of the UKBMS sampling season to submit data, however exploratory analyses (results not shown) have suggested flight periods similar to those shown in Figure S11 could be achieved from data based on 50% UKBMS sites.

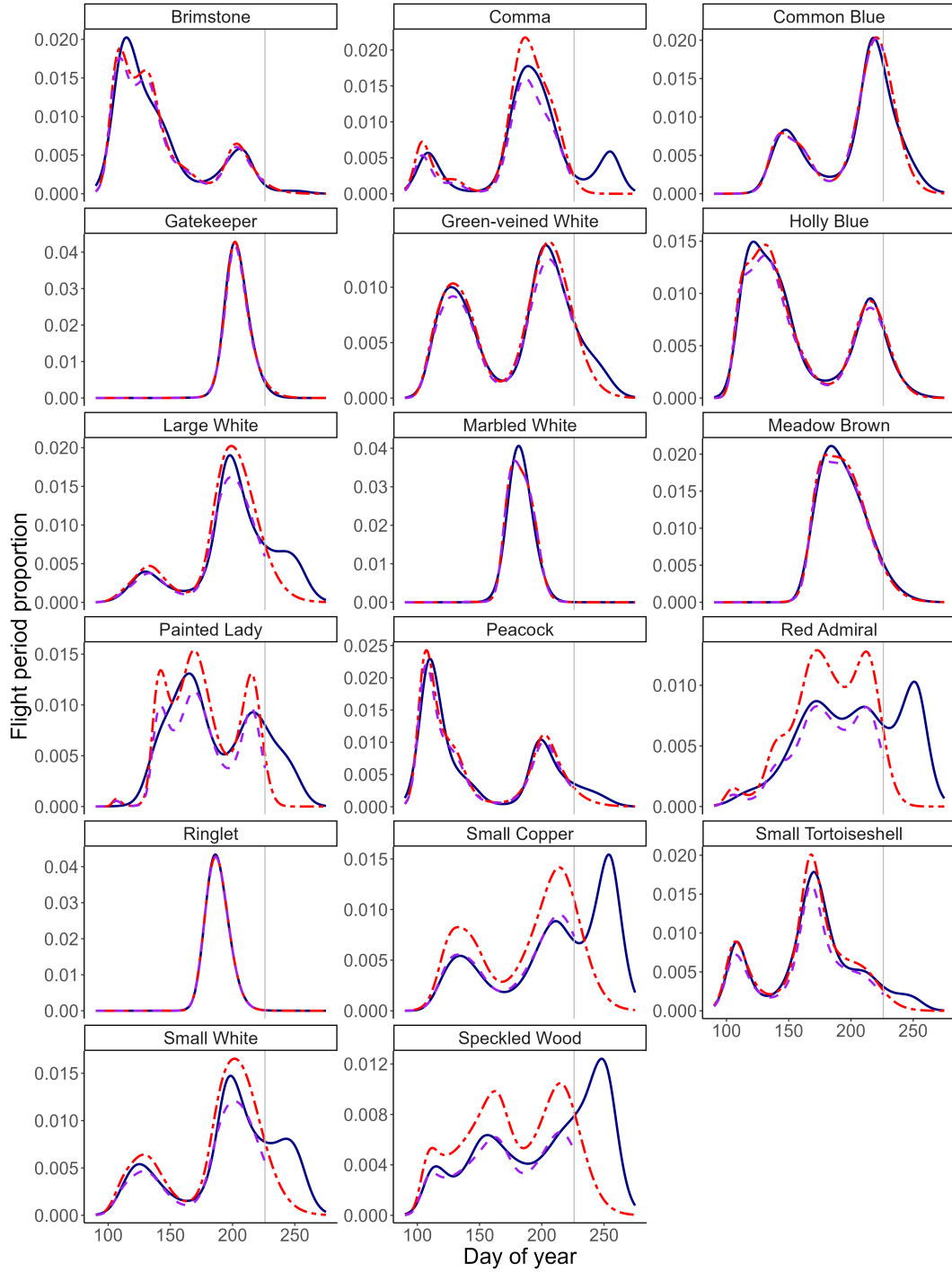


Figure S11: Flight periods for 2022 for each of 17 BBC species, estimated by fitting the GAI approach to UKBMS data, based upon full season data (dark blue), or data sampled up to mid August only (unscaled, dashed red) or scaled with respect to data from preceding years (dashed purple). The vertical grey lines indicate mid August (day 226 of the year).

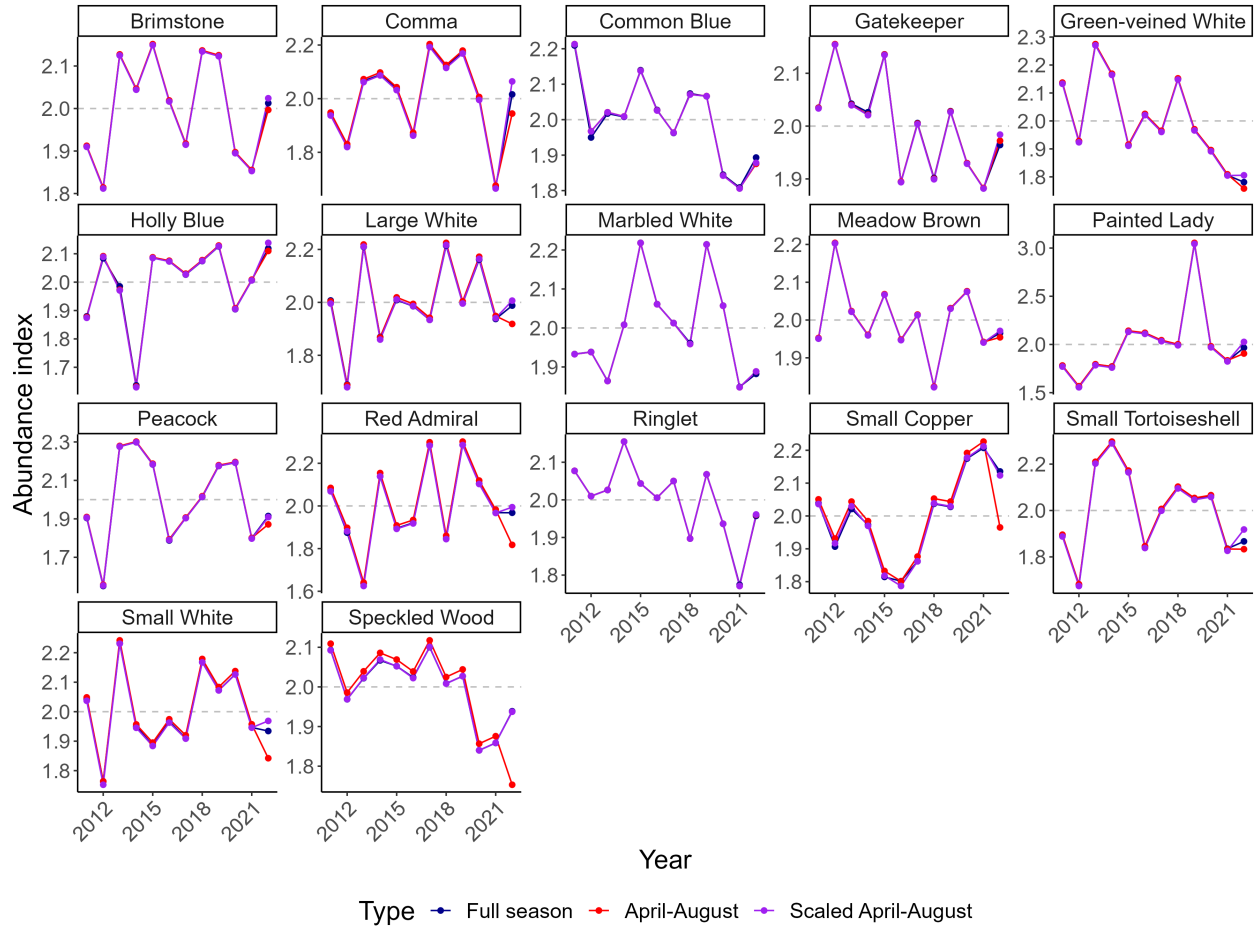


Figure S12: Relative abundance indices produced from the BBC data using phenology adjustment, where the 2022 flight period was estimated from full season UKBMS data (dark blue), or data sampled up to mid August only, either unscaled (red) or scaled (purple). Indices are on the log10 scale with a mean value of 2 (indicated by the horizontal dashed lines).

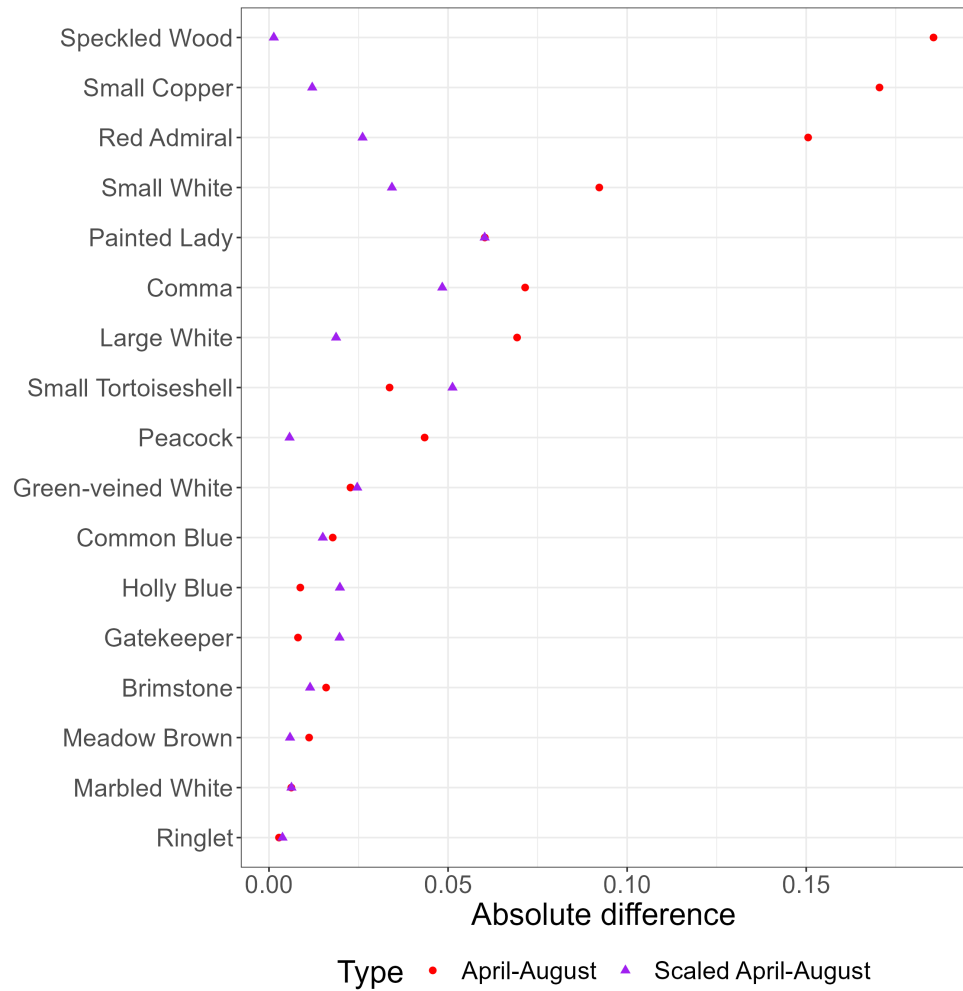


Figure S13: Absolute differences in relative abundance index estimates for 2022 produced from the BBC data using phenology adjustment, where the 2022 flight period was estimated from UKBMS data sampled up to mid August only compared to using full season UKBMS data. Differences are shown based on unscaled (red) or scaled (purple) flight period estimates up to mid August.