**Supplementary material**

**Among-individual and within-individual variation in seasonal migration covaries with subsequent reproductive success in a partially-migratory bird**

Jane M. Reid, Moray Souter, Sarah R. Fenn, Paul Acker, Ana Payo-Payo, Sarah J. Burthe, Sarah Wanless & Francis Daunt

**Proceedings of the Royal Society B**

10.1098/rspb.xxxx.xxxx

**S1. Details of the field system**

European shags are seabirds that are restricted to coastal areas all year round, and have patchy breeding and wintering distributions due to their requirements for appropriate nest and night roost sites (typically on rocky islands and cliffs). They additionally use coastal rocks as day roosts between foraging bouts, where colour-ringed individuals can be readily observed all year round [1].

Figure S1 shows the locations of the Bullers of Buchan (‘BoB’) study area and adjacent day roosts and breeding colonies. During 2009-2018, samples of chicks and adults were ringed at BoB, using colour-rings with three-letter individual codes (licensed by British Trust for Ornithology). The ringed individuals comprised those that could be accessed without undue risk to researchers or other breeding seabirds in the cliff colony. Totals ringed per year varied through approximately 60-110 chicks and 2-12 adults, representing 5-10% of the total individuals present. Some ringing occurred within nine widely-distributed and accessible sub-colonies within the overall colony. Recruited and surviving colour-ringed adults could subsequently breed at any nest site, including many sites that were not accessible for ringing but could still be readily viewed from cliff tops to record reproductive success. Consequently, breeding attempts made by ringed adults are likely to be broadly representative of the whole colony. Reproductive success of these individuals was ascertained by systematically recording nest contents every 5-10 days throughout the breeding period following well-established and widely used seabird monitoring protocols [2]. Offspring were deemed to have reached fledging age once observed with fully grown wing and tail feathers, with little or no down remaining on the back and upper wings [2].

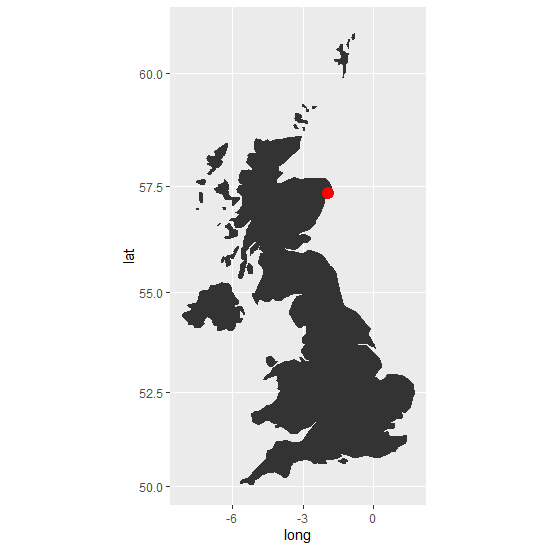
Most breeding shags that had been ringed at BoB and recruited locally bred within the main study area. However, adjacent breeding areas (~3km south and north) were also intensively surveyed to locate colour-ringed adults that had dispersed slightly further (Fig. S1). Beyond these areas, there are no further shag breeding areas for >40km (Fig. S1). Data from three immigrants that were originally ringed at colonies further south in Scotland (Isle of May and Craigleith, Firth of Forth), and then dispersed to breed at BoB, were included in analyses. There is no evidence of long-distance breeding dispersal of recruited adults.

European shags typically breed from age three years. Although they can attempt to breed aged two years, this was rarely observed at Bullers of Buchan in 2018 and 2019 (only two attempts by known ringed two year-olds). Individuals that would have been two years old in the summer following each focal winter were consequently excluded from current analyses of reproductive success. All surviving three year-olds were defined as adults and included (Table S1). The frequent breeding season surveys ensured that most colour-ringed individuals were resighted on multiple occasions during in each summer (Figure S2). Adults that were not seen in a particular breeding season were never seen subsequently, either in subsequent breeding seasons or winters, implying that breeding season resighting probability was close to 1 during 2017-2019.

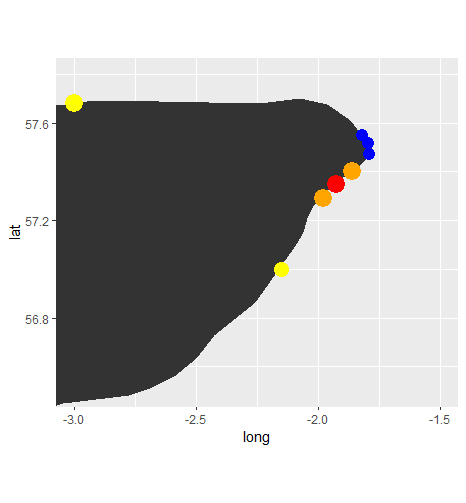
During 2017-2019, local (i.e. breeding area) winter resighting surveys of all local roost sites were undertaken by experienced observers, primarily JMR and MS, typically at times of day and tide that surveys undertaken during 2009-2017 had shown to maximise the probability of observing roosting shags (including pre-dusk gatherings). Roosting groups were repeatedly scanned with 60x magnification telescopes, and colour-ring codes recorded and checked. This core survey effort was supplemented with substantial additional effort from numerous other observers, who checked diverse local sites at diverse times of day and tide. This substantial combined effort reduced heterogeneity in local resighting probability per defined 10-day ‘occasion’, and increased the probability that colour-ringed shags that were locally present would be resighted on multiple occasions during the winter.

**Figure S1.** (A) Location of the main Bullers of Buchan study area, and (B) zoomed-in locations of the main breeding colony (red), two adjacent sub-colonies (orange, Whinnyfold to the south and Longhaven to the north), and three main winter day roost sites (blue) ranging up to 16km north of the main breeding area. The next nearest breeding colonies to the north and south are also shown (yellow). Wider surveys showed that few shags utilised winter day roost sites south of Bullers of Buchan, since the main foraging areas are slightly north.

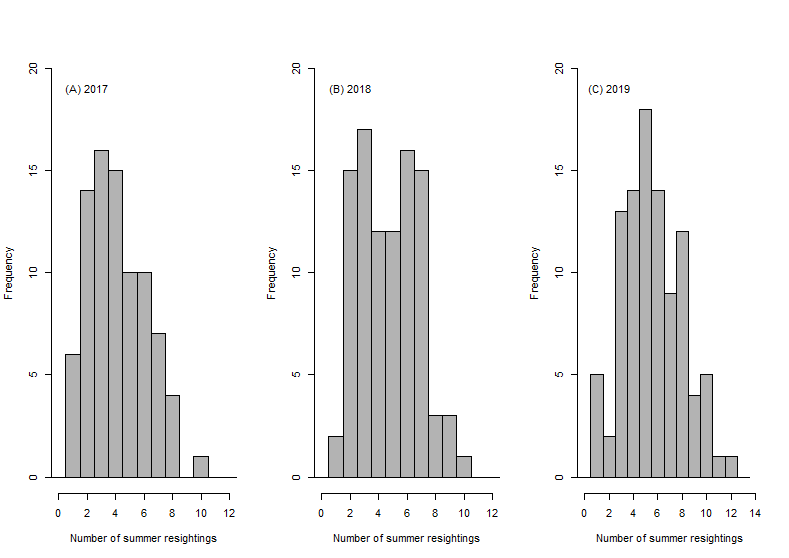
**(A)**



**(B)**

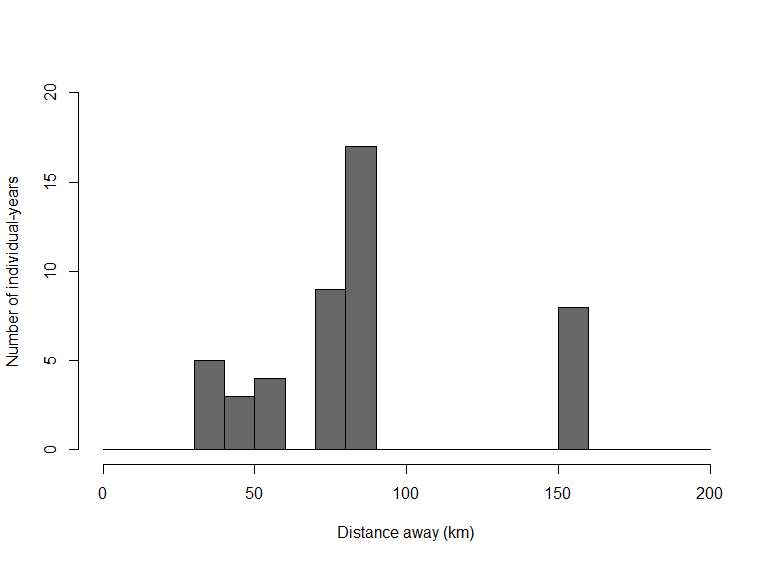


**Figure S2.** Frequency distributions of the number of times that an adult colour-ringed shag was resighted at Bullers of Buchan during the breeding season in (A) 2017, (B) 2018 and (C) 2019. Y-axes show the numbers of individuals. Mean numbers of resightings per individual were 4.1, 4.7 and 5.6 in 2017-2019 respectively, and variances were 4.1, 4.4 and 5.8 respectively. Consequently, there was therefore little or no overdispersion compared to expectation given a Poisson distribution of resightings.



Winter surveys for colour-ringed shags were also carried out at numerous other sites along the east coast of Scotland, primarily ranging from Northumberland to Caithness [1,2-4]. These surveys focussed on key night roost locations where shags could be observed pre-dusk, as well as day roosts adjacent to the night roosts. Totals of 17 and 29 colour-ringed individuals known to breed at Bullers of Buchan were resighted at night roosts elsewhere in winters 2017-2018 and 2018-2019 respectively (36 different individuals), and hence were directly observed to be migrants. These individuals were observed between ~35km and 160km away from Bullers of Buchan (Fig. S3). These individuals were seen at these locations on multiple occasions through the winter, with no intervening sightings in the Bullers of Buchan area. They consequently provide direct evidence that focal individuals had moved away for a protracted period, representing migration rather than solely daily foraging trips. The total of 46 individual-year encounter histories comprised 24 and 22 that the mixture models assigned as late migrants and early migrants respectively. Individuals that were never observed in any of the survey areas in either of the two winters are likely to have moved further away.

**Figure S3.** Distribution of distances (km) that migrant individual shags were observed away from Bullers of Buchan across winters 2017-2018 and 2018-2019. The discrete distribution reflects the distances to major surveyed night roost sites.



**Table S1.** Data summaries for 2017-2018 and 2018-2019: numbers of (A) focal individuals, (B) observed surviving individuals, (C) individuals with observed reproductive success and unique breeding events, (D) focal individuals originally ringed as chicks or adults, (E) individuals observed as young or older adults, (F) estimated initial class (i.e. state) probabilities (± 1 standard error) and 95% confidence interval, (G) surviving individuals assigned to their most likely class with probabilities PMax≥0.95 or ≥0.75, (H) individuals with observed reproductive success assigned to the classes interpreted as resident (R), late migrant (LM) or early migrant (EM). H shows means and 95% confidence intervals across 10000 realisations of individuals’ probabilistic class assignments. In B, C, G and H, proportions are shown in parentheses.

|  |  |  |
| --- | --- | --- |
|  | **2017-2018** | **2018-2019** |
| (A) Focal individuals | 121 | 127 |
| (B) Surviving individuals | 111 (0.92) | 122 (0.96) |
| (C) Observed reproductive success | Individuals: 105 (0.95)  Breeding events: 95 | Individuals: 118 (0.97)  Breeding events: 111 |
| (D) Chicks, adults | 72, 33 | 87, 31 |
| (E) Young, older | 41, 64 | 34, 84 |
| (F) Initial class probabilities | R: 0.42±0.05, 95%CI: 0.33,0.52  LM: 0.26±0.04, 95%CI: 0.19,0.35  EM: 0.32±0.05, 95%CI: 0.22,0.41 | R: 0.22±0.04, 95%CI: 0.16,0.30  LM: 0.39±0.05, 95%CI: 0.30,0.50  EM: 0.39±0.05, 95%CI: 0.29,0.51 |
| (G) Confident assignments | PMax≥0.95: 96 (0.86)  PMax≥0.75: 109 (0.98) | PMax≥0.95: 101 (0.83)  PMax≥0.75: 118 (0.97) |
| (H) Individuals assigned as R, LM or EM | R: 43.9 (0.42), 95%CI: 42,46  LM: 28.6 (0.27), 95%CI: 26,31  EM: 32.4 (0.31), 95%CI: 31,34 | R: 25.2 (0.21), 95%CI: 24,27  LM: 46.7 (0.40), 95%CI: 43,50  EM: 46.2 (0.39), 95%CI: 43,49 |

**S2. Details of capture-mark-recapture mixture models**

Diverse methods of extending basic capture-mark-recapture (CMR) models to estimate movement probabilities and account for temporary emigration have been devised, including utilising multiple secondary occasions (robust design), multi-site observations in multi-state models, and individual covariates that predict emigration [5-8]. Our analyses illustrate use of CMR mixture models to assign encounter histories, and hence individuals, to latent classes that represent different migratory strategies. We illustrate that individuals can be assigned to interpretable classes with high probability solely using frequent local (breeding area) resightings. This approach could be logistically feasible in diverse systems, and particularly useful for systems where individuals that have migrated away (i.e. temporarily emigrated) are not readily observable. Here, we provide further details of current analyses, and highlight some considerations for applying similar analyses to other datasets.

***Number of mixture classes:*** Support for modelling any particular number of classes within a mixture model can be formally assessed by using standard information criteria (including AIC) to compare candidate models with different numbers of classes [9-11]. This approach has been validated using simulated CMR datasets with two classes of detection probability, showing that AIC commonly identifies the correct model structure (i.e. two classes versus one or three classes, [9]; see also [10]). These simulations assumed no temporal (i.e. among-occasion) variation in detection probability, meaning that straightforward comparison between candidate models with constant detection probability within one, two or three mixture classes was appropriate. However, specifying an appropriate set of candidate models becomes more complex when detection probabilities vary among occasions and such variation can differ between mixture classes. A large number of models, with different class structures, could then be postulated.

Consequently, to validate the foundations of our inferences for shags, we first fitted mixture models that assumed two or three latent classes in resighting (i.e. detection) probability assuming constant probability across all occasions, and used AIC adjusted for small sample sizes (AICc) to assess support for these different structures. These models were fitted to separate datasets for 2017-2018 and 2018-2019. The three-class model was considerably better supported than the two-class model in both years (ΔAICc>20, Table S2). It has previously been suggested that it may rarely be useful to fit mixture models with more than three classes to typical CMR datasets, due to high resulting uncertainty [9]. Exploratory analyses of models with four mixture classes fitted to the current shag dataset supported this view. AICc values were similar to those for three-class models, but class assignments were highly uncertain and no longer easily interpretable. Further, simulations have shown that in these circumstances AIC may tend to indicate support for models with too many classes, especially for long-lived species with substantial heterogeneity ([9], see also [10]). Consequently, four-class models were not considered further in current analyses.

**Table S2.** Summaries of capture-mark-recapture (CMR) mixture models fitted to encounter histories observed in 2017-2018 and 2018-2019 with (A) two or three mixture classes in local resighting probability with constant class-specific probability, and (B) three mixture classes in local resighting probability with class-specific probabilities that varied among occasions. Here, probabilities were constrained either following Table S3 (model i), with additional independent variation in resighting probability for the postulated late migrant and resident classes (model ii), and with resighting probability for the final summer occasion fixed to one (model iii). Model deviance, number of parameters (Par.), AIC adjusted for small sample sizes (AICc) and the difference in AICc between each model and the best supported model in each section (ΔAICc) are shown. The model with three mixture classes and constraints on local resighting probability following Table S3 was best supported in both years (model i, bold). This model was consequently used to compute individual class assignment probabilities and infer individual migratory strategy, and hence to quantify relationships with subsequent reproductive success.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **2017-2018** | | | | **2018-2019** | | | |
|  |  | Deviance | Par. | AICc | ΔAICc | Deviance | Par. | AICc | ΔAICc |
| (A) | Two classes | 2511.5 | 4 | 2519.5 | +20.2 | 2609.3 | 4 | 2617.3 | +29.8 |
|  | Three classes | 2487.2 | 6 | 2499.3 | - | 2575.4 | 6 | 2587.5 | - |
|  |  |  |  |  |  |  |  |  |  |
| (B) | **Model i** | **1871.3** | **31** | **1935.4** | **-** | **1776.4** | **31** | **1840.9** | **-** |
|  | Model ii | 1865.3 | 34 | 1935.7 | +0.2 | 1771.3 | 34 | 1842.3 | +1.4 |
|  | Model iii | 1877.1 | 30 | 1939.0 | +3.6 | 1780.1 | 30 | 1842.3 | +1.4 |

***Parameter constraints:*** We took the well-supported three-class model and applied constraints on the form of temporal (i.e. among-occasion) variation in local resighting probability (PL) according to our biological understanding of the system and the postulated existence of three broad migratory strategies (i.e. resident, late migrant and early migrant, Table S3). These constraints were implemented primarily because exploratory analyses of models with full unconstrained occasion-dependence in PL for all three mixture classes showed that estimates of PL were bounded at zero for the putative migrant classes in some winter occasions. This makes biological sense; local resighting probability for migrant classes will be zero at times of year when all migrants are locally absent. However, estimated values of PL=0 prevented useful estimation of individual class assignment probabilities, or hence quantitative inference of individual migratory strategies. Consequently, values of PL for the migrant classes were constrained to be identical across restricted sets of winter occasions (Table S3). Since migrants were occasionally locally present and resighted in early and late winter (i.e. before departure or after return), the constrained values of PL were then estimated as slightly above zero (main text Figure 1C,H), thereby allowing individual class assignment probabilities to be computed. Further exploratory analyses with different parameter constraints showed that imposing some biologically-informed constraints also facilitated inference. This is because local resightings in different winter occasions are not equally informative regarding individual migratory strategy. For example, a local mid-winter resighting strongly implies that an individual is ‘resident’, but an early September resighting does not. Finally, imposing constraints that differed among mixture classes also meant that the order of identity of the three emerging mixture class was consistent across different analyses and datasets (rather than varying according to starting values and the input order of encounter histories, which occurs in the absence of any other constraints). This facilitated subsequent analyses and model comparisons. However, it is important to note that while parameter constraints are set *a priori*, parameter values are estimated from the encounter history data. Consequently, values will not necessarily concur with those that would be expected for the three postulated migratory strategies, and there are no *a priori* defined exact ‘cut-off’ dates that define strategies (e.g. early versus late migrants).

**Table S3.** Structure of occasion-specific local resighting probabilities (PL) for a three-class mixture model specified to capture postulated resident (R), late migrant (LM) and early migrant (EM) strategies. Occasions (Occ.) comprise 17 winter occasions (Wi) and the previous and subsequent summer breeding seasons (S1 and S2, hence 19 occasions in total). PL parameters for each winter are specified as Xy, where X is e, l or r denoting the EM, LM and R strategies respectively, and y is the first winter occasion for that parameter. Some Xy values for early and late migrants were constrained to be identical across multiple occasions, primarily to prevent estimation of occasion-specific PL=0. Since individuals only enter the dataset if observed in summer S1, initial PL=1 for all classes (i.e. conditioning on initial encounter). Since there was no expectation that PL in summer S2 would differ between classes, this parameter was set equal (s2). PL was also set equal for the LM and R strategies for occasions W1-W3, when all such individuals are locally present and are expected to be equally observable. Groups of adjacent parameters that were constrained to be equal are highlighted in shades of grey.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Occ.** | **S1** | **W1** | **W2** | **W3** | **W4** | **W5** | **W6** | **W7** | **W8** | **W9** | **W10** | **W11** | **W12** | **W13** | **W14** | **W15** | **W16** | **W17** | **S2** |
| **R** | 1 | r1 | r2 | r3 | r4 | r5 | r6 | r7 | r8 | r9 | r10 | r11 | r12 | r13 | r14 | r15 | r16 | r17 | s2 |
| **LM** | 1 | r1 | r2 | r3 | l4 | l5 | l6 | l7 | l8 | l9 | l9 | l9 | l9 | l9 | l9 | l15 | l15 | l15 | s2 |
| **EM** | 1 | e1 | e1 | e1 | e4 | e4 | e4 | e4 | e4 | e4 | e4 | e4 | e4 | e4 | e4 | e15 | e15 | e15 | s2 |

Observations from previous years suggested that migrants could start to return from late January, with most returning from late February. Consequently, the last winter occasion for which resighting data were collected (Feb 19th–28th) was excluded from encounter histories because sightings were no longer strongly informative of an individual’s migratory strategy. Local resighting probability was constrained to be constant across the three previous late winter occasions for both early and late migrants, but allowed to differ from mid-winter values (Table S3). This allows PL to increase in late winter, as migrants return. However, in practice, late winter PL was estimated to be only slightly higher than mid-winter PL for both migrant classes (main Figure 1). There was therefore no evidence that early and late migrants returned at substantially different times (up to mid-February) as well as departing at different times.

***Additional constrained models:*** We fitted two sets of additional models to verify whether conclusions were robust given slightly different parameter constraints from those defined in the primary model (Table S3).

**1.** *Different constraints on winter PL:* We fitted a set of 7 additional models with slightly different constraints on winter PL (codes in Box 1). These models were considerably less well supported than the primary model (ΔAICc>>2), except for a model where PL values for the postulated late migrant class for the first three winter occasions varied independently from those for the postulated resident class (rather than being constrained to be equal, as in Table S3). This model was only slightly less well supported than the primary model in both years (lower deviance, but more parameters, Table S2B model ii versus model i). Individual class assignment probabilities calculated from the two models were quantitatively very similar, meaning that the choice between these models did not alter subsequent inferences.

**2.** *Different constraints on final summer PL (i.e. parameter S2):* The primary model estimated the constant between-occasion survival probability *ɸI* to be 1.00 in both years. Consequently, local resighting probability in each subsequent summer (S2) was estimated as <1. In reality, some or all individuals that were unobserved in the subsequent summer will have died during winter, implying that *ɸI* was slightly overestimated and S2 slightly underestimated. These small biases probably arose because *ɸI* was constrained to be equal across all occasions for all three classes. This in turn was because estimation of fully time- and class-specific *ɸI* was not feasible given current sample sizes of individuals (~120 encounter histories per year) relative to the large number of survival parameters in a fully saturated model (3 classes x 18 intervals = 54 parameters), coupled with high overall annual survival probabilities (≥0.92) and no *a priori* expectation of when mortality would occur or that it should be concentrated into specific occasions. Consequently, to evaluate the implications of minor biases for individual class assignment probabilities and hence inference of individual migratory strategy, an additional model was fitted, parameterised following Table S3 but with s2 fixed to equal 1. This forced estimation of *ɸI* <1, in practice yielding estimates of 0.995 (95%CI: 0.991,0.997) and 0.997 (95%CI: 0.994,0.999) for 2017-2018 and 2018-2019 respectively). This model was slightly less well supported than the primary model in both years (Table S2B model iii versus model i). Class assignment probabilities estimated for surviving individuals were quantitatively very similar across both models. The primary model was consequently used for subsequent inference, but class assignment probabilities for individuals that were not observed during the second summer or subsequently, and hence most likely died during the focal winter, were not interpreted. These individuals anyway do not enter analyses of variation in reproductive success during the subsequent summer.

***Goodness of fit:*** General goodness of fit (GoF) tests for CMR mixture models are not yet available. We therefore used a group-structured approximation to evaluate GoF and overdispersion. We assigned each observed encounter history to its most likely class according to mixture model probabilities, then structured the encounter histories into three fixed groups defined by the assigned classes. Standard approaches to evaluating GoF of group-structured CMR models can then, in principle, be applied. However, in the current case, fully time-dependent group-structured CMR models could not be usefully fitted and evaluated because local resighting probability (PL) was estimated as zero for the migrant groups in some mid-winter occasions (as makes biological sense, see above). We therefore fitted models with constrained group-structured temporal variation in PL (following Table S3), and used simulation utilities in programme MARK to assess GoF and overdispersion (ĉ, full methodological details and explanations in [12]). In brief, we simulated 1000 independent datasets given estimated parameters of the group-structured model and with the same number of encounter histories as in the real datasets, refitted the model, and quantified the proportion of instances in which the deviance exceeded that estimated when the model was fitted to the real data. This gives a bootstrap probability that the fitted model does not fit the data; values >0.05 can be interpreted that the observed deviance can arise by chance. Overdispersion was estimated from these simulations as the ratios of observed to mean simulated deviance or ĉ. As a further measure, the value of ĉ for which the proportions of simulated values that were higher and lower were exactly 0.5 (i.e. ‘median ĉ’) was computed (see [12]).

In summary, these analyses suggested that a model with three fixed classes of resighting probability, fitted by assigning each observed encounter history to its most likely class as estimated by the mixture model, adequately fitted the data with little additional overdispersion (Table S4). This implies that the three-class mixture model also adequately fitted the data; indeed the mixture model likely fits even better than the group-structured approximation, since encounter histories can be assigned to multiple classes with some probability, better capturing remaining heterogeneity.

Note that this approach to assessing GoF for mixture models would not work so well for datasets and models where most individual encounter histories are not assigned to single classes with high probability (as they were in our current analyses). The simple group-structured approximation would then be less useful.

**Table S4.** Summary of (A) bootstrap goodness of fit (GoF) tests and (B) median ĉ simulations for group-structured CMR models, where each encounter history was assigned to its most likely class of resighting probability as estimated by mixture models. In (A) the GoF probability is the proportion of bootstrap replicates where the deviance of the model fitted to simulated data exceeded that of the model fitted to the real observed data. Overdispersion was estimated as the ratio of observed to simulated (i) deviance and (ii) ĉ.

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | 2017-2018 | 2018-2019 |
| (A) | GoF probability | 0.164 | 0.110 |
| (i) | Observed model deviance | 857.1 | 745.4 |
|  | Mean simulated deviance | 814.3 | 698.8 |
|  | Observed / mean simulated deviance | 1.05 | 1.07 |
| (ii) | Observed model ĉ | 14.3 | 13.3 |
|  | Mean simulated ĉ | 12.8 | 11.6 |
|  | Observed / mean simulated ĉ | 1.12 | 1.15 |
|  |  |  |  |
| (B) | Median ĉ | 1.07 | 1.09 |

**Box 1.** **E-SURGE specifications for a CMR mixture model with three mixture classes in detection (resighting) probability:**

Number of states: 4 (3 live states and dead)

***GEPAT code:***

Initial state probabilities:

|  |  |  |
| --- | --- | --- |
| π | π | \* |

State transition probabilities (assuming no transitions among classes):

|  |  |  |  |
| --- | --- | --- | --- |
| Ψ | \_ | \_ | \* |
| \_ | ψ | \_ | \* |
| \_ | \_ | ψ | \* |
| \_ | \_ | \_ | \* |

Event probabilities:

|  |  |
| --- | --- |
| \* | β |
| \* | β |
| \* | β |
| \* | \_ |

***GEMACO code:***

Initial state: to

Transition: i

Event (following Table S3):

firste+nexte.[from(1:2).t(2\_4)+from(1).t(5\_18)+from(2).t(5\_9,10:15,16:18)+from(3).t(2:4,5:15,16:18)+from(1:3).t(19)]

First detection probability was fixed to 1. Replicate models were fitted with multiple random starting values, yielding no evidence of local minima of the deviance. Individual class assignment probabilities were calculated (following [5,12]) and saved by selecting ‘View history state dependent probability’ in the ‘Run & See’ menu [13].

GEMACO event code for 7 additional candidate models, with slightly different constraints on occasion-specific local resighting probability, that were less well supported (all with firste+nexte.):

* [from+t]
* [from(1).t(2\_18) + from(2).t(2\_8,9:18) + from(3).t(2:18) + from(1:3).t(19)]
* [from(1).t(2\_18) + from(2).t(2\_8,9:18) + from(3).t(2:4,5:15,16:18) + from(1:3).t(19)]
* [from(1:2).t(2,3,4,5,6) + from(1).t(7\_18) + from(2).t(7,8,9) + from(3).t(1:9) + from(2:3).t(10:15,16:18) + from(1:3).t(19)]
* [from(1).t(2\_18) + from(2).t(2\_8,9:15,16:18)+ from(3).t(2:4,5:15,16:18) + from(1:3).t(19)]
* [from(1).t(2\_18) + from(2).t(2\_9,10:15,16:18)+ from(3).t(2:4,5:15,16:18) + from(1:3).t(19)]
* [from(1:2).t(2,3,4) + from(1).t(5\_18) + from(2).t(5\_8,9:15,16:18)+ from(3).t(2:4,5:15,16:18) + from(1:3).t(19)]

The use of CMR mixture models to assign encounter histories, and hence individuals, to different latent classes that represent different migratory strategies worked well with our dataset, but some points should be considered for application to other datasets or questions.

***Separate year models:*** We fitted separate models to separate encounter histories for 2017-2018 and 2018-2019 (hence, in Table S3, S1=2017 and S2=2018, or S1=2018 and S2=2019). This facilitated: (i) direct estimation of initial class probabilities for each year (i.e. the proportions of individuals entering each winter that were assigned to the classes interpreted as residents, late migrants and early migrants); (ii) direct estimation of individual class assignment probabilities for each year; (iii) comparison of class assignment probabilities for identical or similar encounter histories that were observed in both years, thereby validating whether class interpretations were consistent; and (iv) assessment of goodness of fit and overdispersion using tractable group structured approximations. Comparisons of 15 identical or very similar encounter histories showed that class assignment probabilities were indeed very similar across the two years (Table S5). ‘Very similar’ encounter histories were defined as those with the same number of positive occasions, with broadly similar distributions across each winter (Table S5).

**Table S5.** Comparisons of class assignment probabilities for (A) 10 identical and (B) 5 very similar encounter histories observed in both 2017-2018 and 2018-2019. The probability PA that each encounter history is assigned to the resident (R), late migrant (LM) or early migrant (EM) class in each year is shown, with the absolute magnitude of the between-year difference. The focal encounter histories comprise 5, 4 and 6 that were assigned to the R, LM and EM classes with high probabilities respectively (grey shading).



It would in principle be possible to fit a single CMR mixture model to individuals’ two-year encounter histories. By allowing class (i.e. state) transitions in the middle summer, the probability of changing class between years could then be estimated (at the population level). However, such two-year models would not necessarily define the same biologically interpretable mixture classes as the single-year models, and individual class assignment probabilities cannot be so easily estimated when individuals can change class. Such additional analyses therefore provide no major benefit given our current aims. Indeed, since no parameter values are shared between the two years and most individuals were assigned to one class in each year with very high probability, the frequencies of individual class transitions can be estimated *post hoc*, while propagating the small remaining uncertainty (see main text). However, CMR mixture models that directly consider multiple years may be appropriate for other datasets and questions.

***Encounter history structure:*** Given our datasets for 2017-2018 and 2018-2019, the mixture models assigned most individuals to one of three classes with very high probability (Fig. 1). However, in general, such assignments, and hence the success of the mixture model approach, will be sensitive to the structures of the defined encounter history. For example, exploratory analyses showed that defining encounter histories with fewer longer winter resighting occasions yielded less clear assignments for some individuals. This is because there was less information to distinguish late migrants from unobserved residents, or to distinguish early migrants from unobserved late migrants. However, defining more shorter occasions also caused problems since local detection probability PL for all classes was then very low for some occasions with little field effort, impeding discrimination. The specification of 17 10-day winter occasions represents a heuristic balance; similar exploratory analyses and decisions should be undertaken to successfully apply similar analyses to other datasets. In principle, encounter occasions do not necessarily have to be of uniform duration. For example, when two successful surveys were undertaken within the same 10-day period these could have been split into two occasions, potentially further increasing power for class assignments. However, since the fine-scale distribution of survey dates differed somewhat between our two focal years, further targeted sub-division of encounter occasions would have impeded between-year comparisons and consistency.

***Individual assignments and survival:*** Individual class assignments can be uncertain for individuals that died at some (unknown) early time in winter, for example because there is little information to distinguish a dead (and hence unobserved) potential resident from a migrant that survived until late winter but cannot be locally observed (see also [10]). The very high breeding season resighting probability in our system likely contributed substantially to the high individual assignment probabilities, since it allowed clear distinction between winter mortality and low winter-long local detection probability. [6] note that individual class assignments will also be highly uncertain for individuals that enter a dataset at later occasions. This does not affect our current analyses, which focus on multiple occasions within one year, since all individuals enter the dataset in the initial summer. Note that individual class assignments should not be used to estimate class-specific survival probabilities *post hoc*. Rather, such probabilities should be directly estimated within the CMR mixture model. This is technically feasible, but was not attempted in our current analyses because overall adult survival probabilities across both focal years were very high (see above). This implies that there will be low power to detect variation in survival among the relatively large number of between-occasion intervals, but also implies that survival is unlikely have varied substantially among strategies across the current dataset.

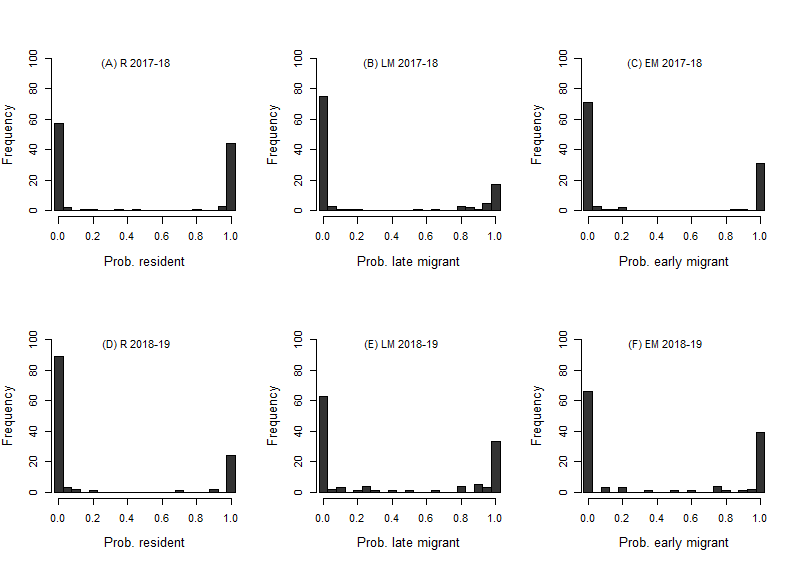
***Between-class transitions:*** A further possibility would be to consider two mixture classes of detection probability relating to resident and migrant states, and then allow transitions between them (thereby capturing ‘late migrants’ that switch between resident and migrant states during mid-winter). However, exploratory analyses of the current datasets suggested that this approach did not work as well as specifying three mixture classes. It required further constraints on the transition process to allow transitions away from the resident state at the end of the initial summer and transitions back to that state before the subsequent summer, and allow transitions away from the resident state in late autumn (to capture ‘late migrants’), but prevent estimation of repeated transitions between resident and migrants states within winters (which system-wide observations suggest very rarely occur). This required tight *a priori* specification of time windows when ‘early migrants’ and ‘late migrants’ could depart and hence what classes of individuals existed. In contrast, the three-class mixture model allowed these classes to emerge without requiring strong *a priori* definitions.

Note that for other kinds of datasets and structures, for example with sparse local winter sightings but larger numbers of sightings across one or multiple other sites, relationships between winter location and subsequent reproductive success might be best estimated using other approaches, for example using multi-state models that consider multiple observable (and/or unobservable) winter locations and reproductive outcomes as explicit states. Our clear interpretation of different mixture classes as early and late migrants and residents also assumes that there is no further major heterogeneity in resighting (i.e. detection) probability within classes; for example that there were no residents that were consistently not observed. This assumption is likely to be valid in our system, as evidenced by the goodness of fit tests that showed no major lack of fit of the group-structured approximation. This is facilitated because individual shags using day roost sites are highly visible, and these sites are highly spatially restricted (meaning that all individuals can in principle be observed). However, caution may be required in other systems whose ecologies could foster greater individual heterogeneity in local detection.

**S3. Details of class assignments**

Figure S4 summarises the probabilities with which individual shags were assigned to the three mixture model classes interpreted as residents, late migrants and early migrants in 2017-2018 and 2018-2019, further illustrating that individuals could be assigned to any of the three strategies with high probabilities. Table S6 provides examples of individual encounter histories that were assigned to each class with high probability, and with lower probability. All encounter histories are provided in the supporting datafiles.

**Figure S4.** Frequencies of the probabilities with which individuals were assigned to the mixture model classes interpreted as (A and D) resident (R), (B and E) late migrant (LM) or (C and F) early migrant (EM) in (A-C) 2017-2018 and (D-F) 2018-2019.



**Table S6.** Illustrative examples of individual encounter histories (*i-xi*) from 2017-2018 assigned with high probability (PA≈1) to the mixture classes interpreted as (A) resident (R), (B) late migrant (LM) and (C) early migrant (EM), or (D) assigned with lower probability. For each of the 19 occasions (S1, W1-W17 and S2, representing the 17 winter occasions and the preceding and subsequent summers), 1 (grey shading) and 0 respectively denote that an individual was or was not locally resighted. Assignment probabilities (PA) are shown, where PA(R)+PA(LM)+PA(EM)=1.



Testing whether patterns of change in individual migratory strategy between years differ from null expectation requires definition of an appropriate expectation. Table S7 shows the expected frequencies of two-year migratory strategies given different null models. The conclusion that more surviving individuals than expected retained the same strategy across both years, and that plasticity was directional, was consistent across all considered null models. It was also consistent given a further null expectation that all strategies occurred with an equal probability of 0.33 in 2018-2019 (data not shown).

**Table S7.** Frequencies of two-year migration strategies for 103 individual European shags included in analyses for both focal years, given (A) observed (probabilistic) strategies for both years, (B) a strategy randomly sampled (with replacement) from the observed strategies in both years, (C) observed (probabilistic) strategy for 2017-2018 and a strategy for 2018-2019 randomly sampled with replacement from the strategies observed in 2018-2019, and (D) observed (probabilistic) strategy for 2017-2018 and a strategy for 2018-2019 randomly sampled with replacement from the strategies observed in 2017-2018. Block C matches the results presented in main Table 1. In addition, (E) shows the expected null frequencies of two-year strategies conditional on change, assuming that new strategies occur with equal probability of 0.5. Rows and columns index strategies assigned in 2017-2018 and 2018-2019 respectively, with the mean frequency and 95% confidence interval across 10000 realisations. In A, dark grey shading (leading diagonal) identifies strategies where individuals were assigned to the same class in both years. Light grey and white identifies strategies where individuals became more and less migratory in 2018-2019 respectively. In B-D, italics show the mean and 95% confidence interval of the difference between the observed and randomised frequencies. Yellow and blue shading indicate two-year strategies that occurred more and less frequently than expected respectively, and green shading indicates strategies whose frequencies did not differ from overall expectation (but were more frequent than expected conditional on change). Comparison between A and E shows that individuals that changed strategy between years became more migratory than expected by chance (mean χ2: 20.1, 95%CI: 15.5-24.8, all p<0.005 across 10000 realisations).

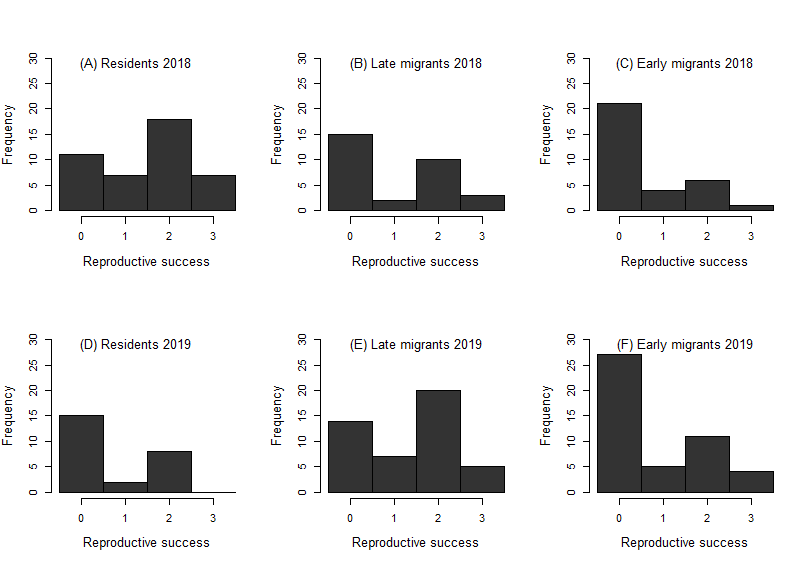
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  | **2018-2019** | | |
|  |  |  |  | Resident | Late migrant | Early migrant |
| (A) | **2017-2018** | Resident | Obs. | 22.4 (21,24) | 19.1 (17,21) | 3.1 (2,4) |
|  | Late migrant | Obs. | 2.0 (1,3) | 17.3 (14,20) | 8.7 (6,11) |
|  | Early migrant | Obs. | 0.0 (0,0) | 4.1 (2,6) | 26.3 (24,29) |
|  |  |  |  |  |  |  |
| (B) | **2017-2018** | Resident | Samp.  *Diff.* | 5.8 (2,11)  *16.6 (12,21)* | 9.6 (4,16)  *9.5 (3,15)* | 9.1 (4,15)  *-6.0 (-12,-1)* |
|  | Late migrant | Samp.  *Diff.* | 9.6 (4,16)  *-7.5 (-14,-2)* | 15.9 (9,24)  *1.4 (-7,9)* | 14.9 (8,22)  *-6.2 (-14,1)* |
|  | Early migrant | Samp.  *Diff.* | 9.1 (4,15)  *-9.1 (-15,-4)* | 14.9 (8-22)  *-10.8 (-18,-4)* | 14.1 (7,22)  *12.2 (5,19)* |
| (C) | **2017-2018** | Resident | Samp.  *Diff.* | 10.6 (5,17)  *11.8 (6,16)* | 17.5 (11,24)  *1.6 (-3,8)* | 16.5 (10,23)  *-13.4 (-20,-7)* |
|  | Late migrant | Samp.  *Diff.* | 6.6 (2,11)  *-4.6 (-9,0)* | 11.0 (6,16)  *6.3 (1,12)* | 10.4 (5,16)  *-1.7 (-7,4)* |
|  | Early migrant | Samp.  *Diff.* | 7.2 (3,12)  *-7.2 (-12,-3)* | 12.0 (7,17)  *-7.9 (-13,-2)* | 11.2 (6,17)  *15.1 (10,20)* |
| (D) | **2017-2018** | Resident | Samp.  *Diff.* | 10.6 (5,16)  *11.8 (6,17)* | 17.4 (11,24)  *1.6 (-5,8)* | 16.5 (10,23)  *-13.4 (-20,-7)* |
|  | Late migrant | Samp.  *Diff.* | 6.6 (3,11)  *-4.6 (-9,0)* | 11.0 (6,16)  *6.3 (1,12)* | 10.4 (6,16)  *-1.7 (-7,4)* |
|  | Early migrant | Samp.  *Diff.* | 7.2 (3,12)  *-7.2 (-12,-3)* | 11.9 (7,17)  *-7.8 (-13,-2)* | 11.3 (6,17)  *15.1 (9,20)* |
|  | Early migrant | Samp. | 2.1 (1,3.5) | 2.1 (1,3.5) | - |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| (E) | **2017-2018** | Resident | Samp. | - | 11.1 (10,12) | 11.1 (10,12) |
|  | Late migrant | Samp. | 5.4 (4,6.5) | - | 5.4 (4,6.5) |
|  | Early migrant | Samp. | 2.1 (1,3.5) | 2.1 (1,3.5) | - |

**S4. Details of reproductive success analyses**

Raw distributions of annual reproductive success of individual European shags classified as most likely to be resident, late migrant and early migrant are shown in Figure S5. All focal individuals that were resighted in each summer were included in analyses. The only exceptions were 6 and 4 individuals in 2018 and 2019 respectively whose nests were unobservable (in sea-caves); these outcomes were assumed ‘missing completely at random’ with respect to migratory strategy. Full results of models fitted to estimate variation in reproductive success as a function of inferred migratory strategy are shown in Table S8. The mean expected reproductive success for each possible two-year migratory strategy is shown in Table S9.

**Figure S5.** Frequency distributions of reproductive success of shags classified as (A and D) resident, (B and E) late migrant and (C and F) early migrant in (A-C) 2018 and (D-F) 2019.



**Table S8.** Summaries of models estimating effects of assigned class and hence migratory strategy (resident, late migrant or early migrant) and age category (young or older adults) on reproductive success in (A) 2018, (B) 2019 and (C) both years combined. Estimated effect sizes (on the latent log scale) and associated standard errors (SE) and probabilities (p) are presented. χ2 is the overall test statistic comparing models (A,B) with and without effects of migratory strategy, or (C) with and without effects of the strategy by year interaction. χ2Diff is the difference in χ2 between models with observed and randomised reproductive success. All statistics are presented as means and 95% confidence intervals over 10000 realisations of each individual’s probabilistic migratory strategy. P(χ2Diff<0) is the proportion of realisations where χ2Diff was negative. Late migrant, early migrant and older age effects are estimated as contrasts from young residents. There was no evidence of a migratory strategy by age class interaction in either 2018 or 2019.

|  |  |  |  |
| --- | --- | --- | --- |
| **(A) 2018** | **Estimate** | **SE** | ***p*** |
| Resident | 0.17 (0.10,0.25) | 0.22 (0.21,0.23) | 0.45 (0.26,0.66) |
| Late Migrant | -0.30 (-0.48,-0.15) | 0.25 (0.24,0.27) | 0.25 (0.06,0.55) |
| Early Migrant | -0.88 (-0.98,-0.75) | 0.27 (0.25,0.27) | 0.0014 (0.0005,0.0046) |
| Older age | 0.31 (0.22,0.35) | 0.23 (0.22,0.23) | 0.18 (0.12,0.27) |
| χ2 | 13.6 (10.5,16.3) |  | 0.0025 (0.0005,0.0087) |
| χ2Diff | 11.5 (5.3,15.4) |  | P(χ2Diff < 0): 0.002 |
| **(B) 2019** |  |  |  |
| Resident | -0.84 (-0.92,-0.74) | 0.32 (0.31,0.34) | 0.011 (0.007,0.021) |
| Late Migrant | 0.67 (0.55,0.78) | 0.28 (0.26,0.29) | 0.018 (0.007,0.044) |
| Early Migrant | 0.20 (0.07,0.33) | 0.30 (0.28,0.31) | 0.50 (0.28,0.80) |
| Older age | 0.61 (0.57,0.64) | 0.24 (0.24,0.25) | 0.015 (0.010,0.021) |
| χ2 | 9.3 (5.9,13.0) |  | 0.019 (0.0021,0.065) |
| χ2Diff | 7.0 (0.45,11.9) |  | P(χ2Diff < 0): 0.020 |
| **(C) Both years** |  |  |  |
| Resident | 0.06 (-0.002,0.11) | 0.19 (0.18,0.19) | 0.76 (0.53,0.98) |
| Late Migrant | -0.25 (-0.42,-0.10) | 0.24 (0.23,0.26) | 0.32 (0.09,0.67) |
| Early Migrant | -0.87 (-0.96,-0.75) | 0.27 (0.25,0.27) | 0.002 (0.0005,0.004) |
| Year | -0.76 (-0.85,-0.66) | 0.28 (0.26,0.29) | 0.007 (0.003,0.014) |
| Late Migrant:Year | 0.92 (0.72,1.12) | 0.37 (0.35,0.38) | 0.016 (0.003,0.045) |
| Early Migrant:Year | 1.06 (0.88,1.23) | 0.40 (0.38,0.41) | 0.010 (0.003,0.025) |
| Older age | 0.45 (0.42,0.48) | 0.16 (0.16,0.17) | 0.007 (0.004,0.012) |
| χ2 | 10.5 (8.1,13.2) |  | 0.009 (0.002,0.024) |
| χ2Diff | 8.3 (2.2,12.2) |  | P( χ2Diff < 0): 0.009 |

**Table S9.** Expected mean reproductive success (±1SE) for two-year migratory strategies comprising combinations of resident (R), late migrant (LM) and early migrant (EM) for (A) young (aged 3-4 years) adults and (B) older (aged 5+ years) adults. Rows and columns identify strategies in 2017-2018 and 2018-2019 respectively. Expectations are simply calculated as the means of predicted strategy-specific reproductive success in 2017-2018 and 2018-2019 and are consequently cross-sectional rather than individual-based. Standard errors are calculated using the Delta method. Darker colours highlight strategies with higher expected reproductive success. Thick borders highlight the most frequently observed two-year strategies (Table 1).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **(A) Young adults** | | |  | **(B) Older adults** | | |
|  |  | **2018-2019** | | |  | **2018-2019** | | |
|  |  | **R** | **LM** | **EM** |  | **R** | **LM** | **EM** |
| **2017-2018** | **R** | 0.78±0.12 | 1.01±0.13 | 0.83±0.12 |  | 1.22±0.14 | 1.58±0.15 | 1.30±0.14 |
| **LM** | 0.66±0.11 | 0.89±0.12 | 0.71±0.11 |  | 1.03±0.17 | 1.40±0.17 | 1.11±0.16 |
| **EM** | 0.47±0.09 | 0.70±0.11 | 0.52±0.09 |  | 0.74±0.13 | 1.10±0.13 | 0.82±0.12 |

**Age and sex effects**

The dataset included individuals that had been ringed as chicks and hence were of exact known age, and individuals that had been ringed as full breeding adults and hence were ≥3 years old at first capture (Table S1). Since local ringing commenced in 2009 and shags regularly live up to ~15+ years, the available data do not yet cover the full likely age range, or hence allow detailed analyses of age-specific variation in migratory strategy or reproductive success. However, to capture coarse age structure, and thereby examine whether observed relationships between reproductive success and migratory strategy could potentially be confounded by age, we classified all focal individuals in each year as young adults (aged 3-4 years) or older adults (aged 5+ years). This classification was designed to capture expected lower reproductive success in young adults, as was apparent in exploratory analyses of a larger dataset on shags breeding on Isle of May, Scotland.

Classification was straightforward and accurate for known-age individuals ringed as chicks. All individuals ringed as adults were classified as older adults. This classification is correct for individuals observed ≥2 years after ringing (i.e. ringed before 2017 and 2018 for analyses of reproductive success in 2018 and 2019 respectively), but could be incorrect for some individuals observed the year after ringing (which could be aged 4 years). In practice, such mis-categorisation is likely to be infrequent, since individuals ringed as adults are most likely to have been >3 years old at capture. Conclusions were unchanged when analyses of age effects were repeated after excluding the few individuals that could potentially have been misclassified. Note that no individuals were included in analyses of reproductive success in their ringing year, since their migratory strategy in the preceding winter is unknown.

Individuals were assigned as male or female based on call during ringing of adults or subsequent field encounters (following [15]). There were no conflicts across 20 instances where individuals were independently assigned twice. When a known-sex individual was paired with another ringed individual that individual was assigned as the opposite sex. This process yielded totals of 56 assigned-sex focal individuals in both 2018 and 2019, comprising 21 females and 35 males in both years. This slight bias towards assigned males likely arises because male calls are more audible.

These additional analyses showed that the frequencies with which individuals were assigned as resident, late migrant or early migrant differed between young and older age classes in 2017-2018. Here, late migrants were more likely to be young, and less likely to be older, than expected by chance (Table S10). However, there was no evidence of a similar pattern in 2018-2019 (Table S10), meaning that there was little evident age-specific variation in migratory strategy over the whole dataset. Similar patterns were evident when analyses were restricted to individuals whose age category was known with certainty. Sex-specific frequencies of the different migratory strategies also did not differ from random expectation in either year (Table S10).

There was no evidence that mated pairs systematically stay together over winter. For 10 and 7 pairs observed in 2018 and 2019 where both adults were colour-ringed, both adults were assigned to the same most-likely class in only 1 and 2 instances respectively (always both assigned as early migrants).

**Table S10.** Frequencies of assigned migration strategies for (A and B) young versus older individuals and (C and D) females versus males in (A and C) 2017-2018 and (B and D) 2018-2019, and χ2 test statistics and associated p values. Mean statistics and 95% confidence intervals calculated across 10000 realisations of each individual’s probabilistic migratory strategy are shown.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Residents** | **Late migrants** | **Early migrants** | **χ2 test statistic**  **p value** |
| **(A) 2017-2018** | **Young** | 11.2 (10,12) | 18.5 (16,21) | 11.3 (10,13) | χ2: 11.8 (8.1,16.1) |
|  | **Older** | 32.8 (32,34) | 10.1 (8,12) | 21.1 (20,22) | p: 0.005 (0.001,0.02) |
| **(B) 2018-2019** | **Young** | 5.0 (5,5) | 14.1 (12,15) | 14.9 (14,17) | χ2: 1.4 (0.9,2.2) |
|  | **Older** | 20.2 (19,22) | 32.6 (30,36) | 31.2 (29,34) | p: 0.52 (0.34,0.70) |
|  |  |  |  |  |  |
| **(C) 2017-2018** | **Female** | 11.4 (11,12) | 3.4 (2,4) | 6.1 (6,7) | χ2: 1.2 (0.5,2.2) |
|  | **Male** | 14.2 (14,15) | 7.8 (6,9) | 13.0 (12,14) | p: 0.61 (0.38,0.83) |
| **(D) 2018-2019** | **Female** | 4.7 (4,5) | 7.7 (6,9) | 8.6 (7,10) | χ2: 0.5 (0.1,1.5) |
|  | **Male** | 6.2 (6,7) | 15.3 (13,17) | 13.5 (12,15) | p: 0.83 (0.53,1.00) |

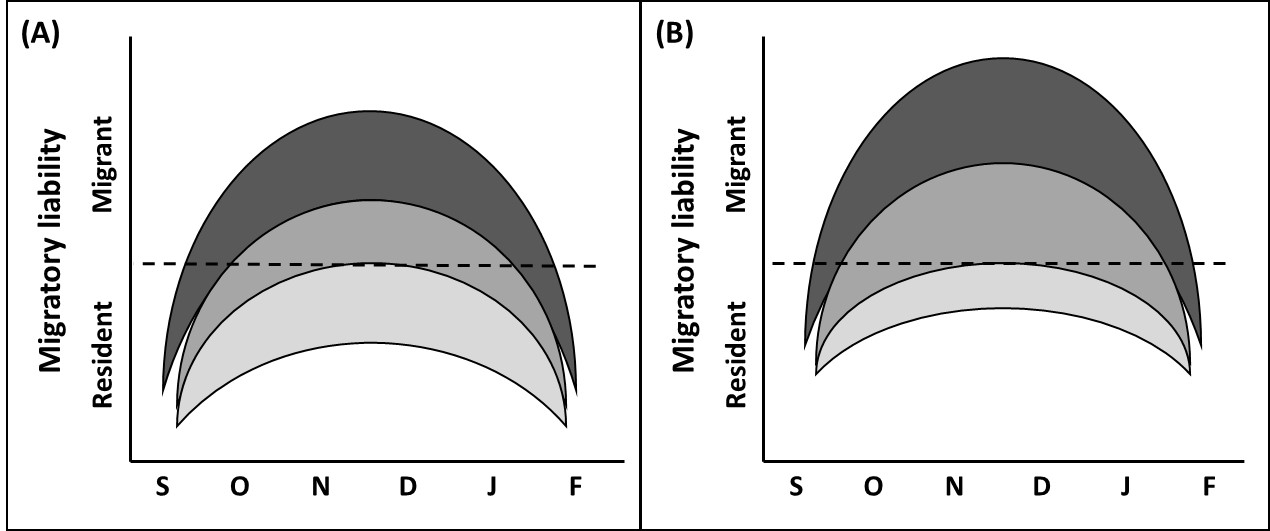
**S5. ‘Threshold model’ for partial migration**

The ‘threshold model’ of quantitative genetics proposes that discrete phenotypes, such as residence versus migration, are expressed as a function of a continuous underlying ‘liability’, which is best viewed as a compound latent trait that is influenced by some combination of genetic and permanent and/or temporary environmental effects [16]. Different phenotypes are expressed when liability values are above versus below some threshold(s). The existence of variation in liability within a population (whether genetic and/or environmental) then generates phenotypic variation. In the current context, variation in liability for seasonal migration can generate partial migration [17].

The threshold model can rationalise the occurrence of different annual ‘migratory strategies’, such as early migration, late migration and residence, and hence within-year phenotypic plasticity. Conceptually, early migrants comprise individuals whose liabilities exceed the migration threshold in early autumn, late migrants comprise individuals whose liabilities exceed the threshold later in autumn, and residents comprise individuals whose liabilities never exceed the threshold (Figure S5A). Such among-individual variation could arise, for example, because different individuals have different physiological responses to changing daylength, or to seasonally changing environmental conditions.

Further, the threshold model can rationalise how the relative frequencies of different migratory strategies could change between years (i.e. the form of between-year phenotypic plasticity). For example, poorer local environmental conditions in one year compared to the previous year could increase liability values across all or most population members. This would generate increased proportions of early and late migrants and a decreased proportion of residents (Figure S5). Moreover, if there are permanent individual effects on liability (either genetic or environmental, the latter representing developmental plasticity and/or subsequent phenotypic canalisation), individual changes would be directional, such that residents become late migrants and late migrants become early migrants, but early migrants do not become residents. These changes qualitatively match those observed in the European shag system.

**Figure S5.** Illustration of hypothetical forms of variation in individual liability for seasonal migration versus residence (‘migratory liability’) given (A) relatively good and (B) relatively poor local (i.e. breeding area) environmental conditions. Dark grey, mid grey and light grey bands illustrate potential patterns of within-winter temporal variation in migratory liability spanning September-February (sequential months abbreviated S, O, N, D, J and F) in relation to a threshold (dashed line). Individuals whose liability values exceed the threshold at different times become early (dark grey) or late (mid grey) migrants, while individuals whose liability values remain less than the threshold stay resident (light grey). Given poorer local environmental conditions all liabilities are increased, yielding increased proportions of early and late migrants and a decreased proportion of residents in B compared to A (illustrated by wider dark and mid grey bands, and a narrower light grey band). Note that the direction of liability that leads to migration is arbitrary, and among-year variation could be similarly conceptualised as a change in threshold rather than a change in liability. The depicted liabilities are hypothetical and designed to illustrate a general concept; in reality they could be different shapes, and are not necessarily symmetrical.

****

**Supplementary references**

1. Grist H, Daunt F, Wanless S, Nelson EJ, Harris MP, Newell M, Burthe S, Reid JM. 2014 Site fidelity and individual variation in winter location in partially migratory European shags. *PLoS One* 9,e98562.

2. Walsh PM, Halley DJ, Harris MP, del Nevo A, Sim IMW, Tasker ML. 1995 Seabird monitoring handbook for Britain and Ireland. JNCC, Peterborough.

3. Grist H, Daunt F, Wanless S, Burthe SJ, Newell MA, Harris MP, Reid JM. 2017 Reproductive performance of resident and migrant males, females and pairs in a partially migratory bird. *J. Anim. Ecol.* **86**,1010-1021.

4. Acker P, Daunt F, Wanless S, Burthe SJ, Newell MA, Harris MP, Grist H, Sturgeon J, Swann RL, Gunn C, Payo-Payo A, Reid JM. Strong selection on seasonal migration versus residence induced by extreme climatic events. In review.

5. Kendall WL, Nichols JD, Hines JE. 2017 Estimating temporary emigration using capture-recapture data with Pollock’s robust design. *Ecology* **78**,563-578.

6. Pledger S, Phillpot P. 2008 Using mixtures to model heterogeneity in ecological capture-recapture studies. *Biometrical Journal* **50**,1022-1034.

7. Lebreton JD, Nichols JD, Barker RJ, Pradel R, Spendelow JA. 2009 Modeling individual animal histories with multistate capture-recapture models. *Adv. Ecol. Res.* **41**,87-173.

8. Péron G, Crochet PAC, Choquet R, Pradel R, Lebreton JD, Gimenez O. 2011 Capture-recapture models with heterogeneity to study survival senescence in the wild. *Oikos* **119**,524-532.

9. Cubaynes S, Lavergne C, Marboutin E, Gimenez O. 2012 Assessing individual heterogeneity using model selection criteria: how many mixture components in capture-recapture models? *Methods Ecol. Evol.* **3**,564-573.

10. Hamel S, Yoccoz NG, Gaillard JM. 2017 Assessing variation in life-history tactics within a population using mixture regression models: a practical guide for evolutionary ecologists. *Biol. Rev.* **92**,754-775.

11. Gimenez O, Cam E, Gaillard JM. 2018 Individual heterogeneity and capture-recapture models: what, why and how? *Oikos* **127**,664-686.

12. Cooch E, White G. 1998 Program MARK – a ‘gentle introduction’. <http://www.phidot.org/software/mark/docs/book/>

13. Pledger S, Pollock KH, Norris JL. 2003. Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics* **59**,786-794.

14. Choquet R, Nogué E. 2011 E-SURGE 1.8 User's Manual. CEFE, UMR 5175, Montpellier, France. <http://ftp.cefe.cnrs.fr/biom/soft-cr/>.

15. Snow BK. 1963 The behaviour of the shag. *British Birds* 56,77–103.

16. Lynch M, Walsh B. 1998 Genetics and Analysis of Quantitative Traits. Sinauer, Sunderland.

17. Pulido F. 2011 Evolutionary genetics of partial migration – the threshold model of migration revis(it)ed. *Oikos* **120**,1776-1783.