We thank the anonymous reviewers for their careful reading of our manuscript and their many insightful comments and suggestions. Below we respond to the comments of each reviewer in detail (in yellow). We are also providing a revised manuscript (citations in green) that reflects their suggestions and comments. We feel that this has resulted in a stronger manuscript.

Review by anonymous reviewer 1, 27 Aug 2024 06:23

The manuscript presents a very interesting approach to try to measure the effect of productivity and survival from count data conducted in the winter. To assess their new model the authors have tried to compare their model output to other external sources of data. The results of their model seems to align with these external sources of data so their "apparent sex ratio" method seems to hold promise.

The title of the manuscript clearly reflect the content of the article. The abstract present the main findings of the study but should be refined (see suggestions below).

Abstract has been modified, see below in the "comment section" for a detailed response.

While I find the approach very interesting, I think the authors should provide more information in their method section, so that the readers can better assess the validity of the proposed model. In particular I found difficult to figure out how many distinct models were used to complete this manuscript.

The section corresponding to the description of the data and the models is now more detailed and a new sub-section has been added for the sake of clarity ("data formatting"). Specific references to the outline scheme are also added in the text.

See also the next "comment section" for detailed responses to the comments on this topic.

81 2.1 Data formatting

82 The method requires the specification of three observable variables (see "Observations" in Figure 4). First,
83 the size of Ruddy duck populations was monitored in both Great Britain (GB) and France (FR) by conducting
84 extensive counts on the wintering grounds between 1 December and 31 January from 1960 to 2019. During
85 this period, the individuals in the wintering grounds are highly detectable as they gather in open water. At
86 least one count was carried out annually around mid-January. In some years, additional counts have also
87 been carried out in December in GB and FR. Although the counted numbers are generally similar (variation
88 coefficient of 10%), the peak counts are generally observed in January, suggesting that a small part of the
89 population arrives late at the wintering grounds in some years. This conclusion prevents the estimation of
90 detection probability using an N-mixture model (violation of the "closed population" assumption, see Costa
91 et al. (2021)) and led the authors to retain the maximum count value as an estimate of population size
92 (Figure 1).

93 Second, the sex ratio in adults (alternatively, the proportion of males among adults), which is presumed to
94 be stable over several years, is theoretically observable during spring when all males are sexually mature and
95 display their colorful plumage. However, the poor detectability of this species at this time of year prevented
96 from using such an approach. This variable was then determined directly on culled individuals, which can
97 be aged and sexed in the hand with certainty. The absence of the bursa of Fabricius was used as a reliable
98 criterion for determining an adult, and the presence of a penis indicates a male individual. In France, it was
99 not possible to obtain sufficiently reliable estimates of sex ratio in adults due to the small sample size. In
100 Great Britain, a preliminary analysis indicated no seasonal variation of the sex ratio in adults. Data from
101 culled adults collected throughout the whole biological cycle were then pooled to assess annual sex ratios.
102 There was also no statistical difference in adult sex ratio between years in which more than 500 adults were
103 culled. This is consistent with previous findings showing that sex ratios in adult ducks tend to be very stable
104 in the short term (Wood et al., 2021), but can fluctuate in the long term. Therefore, we pooled the data
105 from all adults culled in the control programme to estimate the sex ratio in adults.

106 Third, the apparent sex ratio when immature males look like females was assessed annually by pooling
107 all winter counts that distinguished between female-like and male-like individuals. It is assumed that this
108 apparent sex ratio directly reflects the proportion of adult males in the population. Years with winter counts
109 distinguishing apparent sex are a subset of the extensive count time series described previously (Figure 1).
110 These counts were carried out in Great Britain from 2006 to 2012, which corresponded to a period of sharp
111 population decline. In France, counts distinguishing apparent sex were carried out in 1999, 2001-2009, and
112 2012-2019. The French population grew rapidly in the first years of monitoring and then stabilized from
113 2006 onwards as a result of high culling pressure.

114 In both countries, culling was carried out under strict official control, so the exact number of individuals
115 culled was known. Although the age ratio of birds culled before 2009 could not be determined with certainty
116 in France, it was still possible to roughly categorize the culling pressure in both countries into three categories

117 (Figure 1): "no culling" before 1999, as the culling rate on adults was mostly zero and always below 10%
118 in both countries, "low culling" between 1999 and 2005, as the adult culling rate fluctuated around 20%,
119 and "high culling" from 2006 onwards as the adult culling rate fluctuated around 50% (see Figure 8 in the
120 Appendix section).

121 2.2 Inferring adult survival and recruitment rates from changes in population
122 size and proportion of immatures

123 The method introduced in this study is called the "apparent sex ratio" method and is performed as a single
124 model structured in two levels (see "Parameters" in Figure 4).

125 First, the population growth rate, which is the ratio of the population size over two consecutive years,
126 must be determined. The method is also valid if only abundance indices are available. In parallel, the
127 originality of the method consists in inferring the immature/recruit proportion of a population by analyzing
128 the difference between the sex ratio among adults (which is stable over time), and the apparent sex ratio
129 observed in counts carried out during winter (see Equations (1) & (2)). These counts are pooled according
130 to the additive property of the binomial distribution.

131 Second, combining the population growth rate with the proportion of immatures yields estimates of adult
132 survival and recruitment rates, which are latent variables, see Equation (3) and Figure 4. If the absolute
133 value of the population size is known, the number of adults and the number of recruits can be estimated, see
134 Equation (4).

I was less convinced by the discussion about the effectiveness of the different culling strategies in relation to the population dynamic of the species. The culling strategies were used in two populations that were not on the same scale but also were probably very different in term of effectiveness (i.e., culling large groups vs culling breeding pairs and individuals).

The discussion has been reworked to take these aspects into account. See the next "comment section" for detailed responses to the different comments on this.

For instance:

379 […] However, it cannot be excluded
380 that population size also influenced the presumed effect of the culling strategy as the population size in GB
381 was much larger than in FR.

I was more puzzled than satisfied with the part of the discussion that tries to relate population dynamics under a cull order to regular harvest seasons. Figure 8 seems to imply than more than 95% of the population of Ruddy duck was culled in GB. It is hard to relate the population dynamics of such a system to a regular harvest framework were managers will aim to harvest population at the maximum sustainable yield (or any other strategy).

It is not 95% of the population but 95% of the adults. Nevertheless, the comment remains because such a harvest rate is not realistic for a usual harvested population. The comments on the harvested populations have been modified.

407 If investigations on game birds lead to the same observation, a target for effective management of
408 harvested waterfowl could be an earlier starting and ending of the hunting season.

&

410 A possible implication of our study concerns the timing of harvest to limit its impact on population growth
411 rates. Unexpectedly, our results indicate that the persistence of waterfowl populations may benefit from
412 earlier openings and closings of the hunting season (see above). However, this conclusion should be taken
413 with caution, as the harvesting process and the resulting pressure are very different between game birds and
414 introduced species. Moreover, an earlier opening would go against the key concept that hunting does not
415 open while young have not yet fledged.

I think some more general statement of the applicability of the proposed method for other species and other populations would be useful more.

Ok, this corresponds to the last paragraph of the Discussion. It has been improved to respond to this suggestion.

435 To apply the "apparent sex ratio" method, the time window of censuses must be chosen wisely so that
436 it fits into the period when immature males do not look like adult males yet, which can differ from a species
437 to another. Apart from this precaution, modifying standard monitoring protocols to distinguish between
438 male-like and female-like individuals is almost costless, but worth the effort as it would greatly increase the
439 efficiency of conservation/management actions (Nichols & Williams, 2006).

I have outlined some more targeted minor and majors concerns below.

#Minor and major comments

In general for the abstract, I would suggest that instead of focusing on the limitations of capture, mark, recapture programs the authors should give us more information on the data they used for their analysis, the assumption of their model, and how their model hold against these assumptions.

Ok with the reviewer, the abstract has been revised accordingly:

Monitoring the number of individuals is by far the most popular strategy for studying the environmental factors that determine population dynamics and for measuring the effectiveness of management actions aimed at population recovery, control or eradication. Unfortunately, population size monitoring is inefficient in identifying the mechanisms underlying demographic processes and, in particular, in assessing the extent to which population growth rate is influenced by changes in adult survival rather than variations in reproductive parameters. In many waterfowl species, sexual dichromatism is observed in adults, while immatures of both sexes display a plumage pattern similar to that of adult females. In these species, the apparent proportion of males increases as the female-like immature males gradually take on the plumage of adult males. The difference between the apparent sex ratio before and after the young reach sexual maturity then provides information about the age ratio of a population. Using winter counts that distinguished between female-like and male-like individuals of two non-native populations of Ruddy duck Oxyura jamaicensis, a species that exhibits such a plumage pattern, we present a noninvasive method based on the apparent sex ratio to split population growth rate into adult survival and recruitment rates (the latter also referred to as productivity). This method can correctly detect annual changes in vital rates, supporting the assumption that counts conducted in an appropriate time window reflect the age structure of a population. We exemplify how the respective contributions of survival and productivity to the population growth rate are essential for understanding the processes behind demographic dynamics. Finally, we point out some best practices to correctly apply the "apparent sex ratio" method described here.

L3 – Capitalize « Unfortunately »

Ok, done

Figure 1: Are those winter or spring survey? A mix of both?

We added a sentence in the caption accordingly:

Figure 1: Temporal changes in the number of Ruddy ducks (logarithmic scale) in Great Britain (GB) and France (FR) from 1960 to 2019, with references to periods when apparent sex ratios were reported; data were collected exclusively during winter

L70: Some of the assumptions of the model are described below but it would probably be beneficial to regroup them in a section or a table. If you use a table, you could point out how you tested those assumptions with other datasets and if the assumptions are always respected.

Good point, we regrouped and detailed the assumptions more explicitly:

135 2.3 Validating the "apparent sex ratio" method

136 The method is based on three assumptions. The first assumption is that the population size is known exactly.
137 As explained in the Data formatting Section, the detection probability of birds occupying wintering grounds
138 is reasonably close to 100%, so detection issues should be negligible in the present case study. A negative
139 bias in population size may nevertheless occur if not all birds arrived at the wintering grounds at the time of
140 the exhaustive count survey. If the bias is constant over years, this is not an issue as the population growth
141 rate is a relative quantity and can be estimated using abundance indices. However, if some birds arrive late
142 only in some years, the population growth rate may be positively or negatively affected, depending on the
143 year in which such a bias occurs. This assumption is discussed further, but remains specific to this case study.
144 This bias can be avoided in other studies by conducting repeated counts around the peak of occupancy of
145 wintering grounds. In the case of imperfect detection, true population size can be assessed using repeated
146 counts within an N-mixture framework (see Costa et al., 2021 for details).

147 The second assumption is the stability of the adult sex ratio over time. Violation of this assumption would
148 lead to an increasing bias in the vital rates with the drift of the adult sex ratio over time. An examination
149 of the stability of the sex ratio over the time series available in GB confirms that this assumption of is valid
150 (see Data formatting Section). The stability of the adult sex ratio was observed at least over a decade, which
151 is in line with the literature (e.g. Wood et al., 2021).

152 The third and most important assumption is that the male-like individuals observed during winter correspond
153 exclusively to adult males, as immature males are mistaken for females at this time of year. A violation of
154 this assumption occurs when a significant proportion of immature males have acquired male plumage prior
155 to the counts considering apparent sex. This violation would be detected through estimates of adult survival
156 rate that are higher than expected for this species and, conversely, a recruitment rate biased low. A strong
157 corruption of this assumption would even lead to adult survival estimates higher than one and negative

158 recruitment rates for closed populations. A second violation occurs when adult males are less detectable
159 than other individuals, but nothing indicates that such an issue would occur. This would be detected through
160 an overestimation of recruitment rates and low or even negative adult survival rates.

Figure 3: Are those data from the Great Britain, France, both country?

We modified the caption to describe this explicitly:

Figure 3: Proportions of males estimated by internal examination of individuals culled in a given year under the eradication programmes (Male % in cullings), and proportions of individuals showing male-like plumage counted during winter (Male % in counts); data from Great Britain and France are pooled together; discrepancies between the two estimates are due to delayed dichromatism (immature males looking like females before molting); we used such discrepancy to estimate survival and productivity/recruitment

Figure 4: I like this figure, but I am not sure which data set inform with parameter. Can you subset the first table to clearly link the observation to each one of your dataset that you used?

A new sub-section referring to this figure has been added: see “Data formatting” section in material and method

Figure 4: What is the difference between C[i,t] and N[i,t]?

This has been clarified in the table and in the text:

129 These counts are pooled according
130 to the additive property of the binomial distribution.

L79 -80: Would it be possible to report more information on the magnitude of the cull in both country in the text rather than in an Annex?

Ok, done.

114 In both countries, culling was carried out under strict official control, so the exact number of individuals
115 culled was known. Although the age ratio of birds culled before 2009 could not be determined with certainty
116 in France, it was still possible to roughly categorize the culling pressure in both countries into three categories
117 (Figure 1): “no culling” before 1999, as the culling rate on adults was mostly zero and always below 10%
118 in both countries, “low culling” between 1999 and 2005, as the adult culling rate fluctuated around 20%,
119 and “high culling” from 2006 onwards as the adult culling rate fluctuated around 50% (see Figure 8 in the
120 Appendix section).

L73-74: Those are the same count that were conducted in Great Britain from 2006 to 2012 and in France in 1999, 2001-2009, and 2012-2019? I am wondering if there are surveys without the apparent sex-ratio and some years were the sex-ratio was recorded.

Yes, from 1960 to 2019, winter counts were made systematically without the apparent sex ratio being recorded. Years with counts accounting for the apparent sex ratio are a subset of the previous time series. (see figure 1)

We added the details on this in the data formatting section:

82 […] First,
83 the size of Ruddy duck populations was monitored in both Great Britain (GB) and France (FR) by conducting
84 extensive counts on the wintering grounds between 1 December and 31 January from 1960 to 2019.

&

108 Years with winter counts
109 distinguishing apparent sex are a subset of the extensive count time series described previously (Figure 1)

L78-80: Once again I am not sure if you mean that winter count were not made in France in 2000 and 2010-2011 or that there were winter surveys but that the apparent sex was not recorded during those surveys.

There were total count surveys but the apparent sex was not recorded in those counts. The previous sentences and the new reference to figure 1 aims to clarify this.

L79: I would suggest a paragraph break before “In both countries, […]”. To help the reader understand that you are now addressing a new dataset.

Yes, good point. It is done.

L93-94: “A preliminary analysis indicated” that the proportion of males also did not differ statistically between years?

See next comment

L94-95: Is the sex ration among juvenile and adults similar? Or do you mean to say that the sex ratio among adults did not differ significantly among years?

The corresponding section did not look clear enough. We modified it for more clarity:

99[…] In
100 Great Britain, a preliminary analysis indicated no seasonal variation of the sex ratio in adults. Data from
101 culled adults collected throughout the whole biological cycle were then pooled to assess annual sex ratios.
102 There was also no statistical difference in adult sex ratio between years in which more than 500 adults were
103 culled. This is consistent with previous findings showing that sex ratios in adult ducks tend to be very stable
104 in the short term (Wood et al., 2021), but can fluctuate in the long term. Therefore, we pooled the data
105 from all adults culled in the control programme to estimate the sex ratio in adults.

PS: Sex ratio among juvenile is uncertain but seems to be balanced between males and females. We never mention the sex ratio in juveniles because it is not used in this analysis.

L110: (Nichols et al. 1997)

Ok, modified

L141-142: So to be clear you used 10 years of data in GB and 5 years in FR to evaluate the population maximum growth rate and you fit the model for each country independently?

Yes, the only thing in common is a hierarchical prior as justified in the statistical framework (same species in comparable ecosystems so the maximum growth rates are expected to be similar).

L150-151: It is not clear to me if you run both sub model in the same model or did you compare the posterior distribution of both model after running them independently.

It is a single model. We added some words for this:

123 The method introduced in this study is called the “apparent sex ratio” method and is performed as a single
124 model structured in two levels (see “Parameters” in Figure 4).

L153 – 156: You refer to the proportion of adults in both country which makes me wonder if the adults were deliberately targeted? Is there a reason why juvenile were not included in the cull or in the proportion you presented?

Juveniles are also targeted, but we do not know how many are produced, whereas the number of adults is known (winter counts).

A few words were added in the legend of Figure 8: *“(the culling rate on juveniles is not available)”*

L160-162: How many year of culling data from GB did you compare to FR?

Reference to figures are added to visualize that point in details:

203 During the period when culling pressure was high (i.e. from 2006 onwards, see Figures 1 and 8), eradication
204 strategies in Great Britain and France differed.

L164: How many submodel did you run exactly? It could be good idea to name them so that we can track all of them.

Good point, it was unclear in our manuscript. It is now explicitly stated:

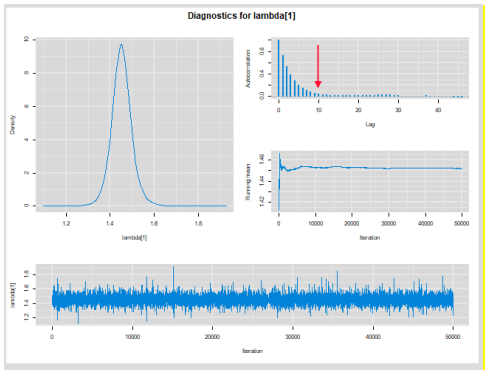
214 We used the Bayesian framework to implement the core model, which estimates vital rates using the “apparent
215 sex ratio” method, and both validation models estimating maximum growth rates and vital rates from culled
216 individuals.

L172: Did you also monitor your effective sample size? You don’t need to report it, but you should make sure that the effective sample size is high enough for all the parameters of interest.

We did not, because nimble does not directly estimate the effective sample size. We used another method by setting a thinning to avoid autocorrelation in samples. This was done by

visually checking the autocorrelation at different lags using the function *mcmcplots::mcmcplot()* (see code available here: https://zenodo.org/records/11471723).

Only lambda max (from the validation model) displayed autocorrelation to lag = 10:



So, beyond a thinning set to 10, the effective sample size and the sample size matched and we assume no issue in estimation of the median and the credible interval. (see section "15.4.4 Thinning Samples" in the stan manual: *https://mc-stan.org/docs/2_18/reference-manual/effective-sample-size-section.html)*

173: Which CI did you use? You mentioned that posteriors are skewed so did you use High Density Intervals?

The distributions of the parameters of interest are not skewed in this study. Please replace the line 690 by *mcmcplot(mcmcout = JuvOut, parms = JuvMon)* for a full check. We used 2.5 and 97.5% quantiles, but thanks for the "high density intervals" tip, we did not know this nice way to define credible intervals for skewed distributions. We removed the part mentioning potential skewed distributions to avoid confusion, it was only mentioned by precaution.

L174: It would be nice to present a recap of the all the dataset you used. The average population size in GB and in France. The numbers of ducks that were culled and used in each of the analysis, etc. Doing so gives us a more general idea of the sample size involved in your analysis.

Good point, we added a section on this:

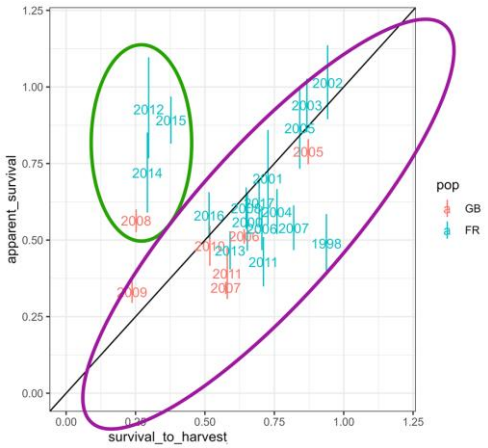
227 Over the period 1960-2019, population size fluctuated between 0 and 5,473 in Great Britain, and between 0
228 and 259 in France. During this period, 12,316 Ruddy ducks were culled in GB, and 2,246 in FR. Of these,
229 8,440 adults were used to determine the male proportion among adults: 0.60 [0.59; 0.61]. In GB and FR, 12
230 and 55 independent count events considering apparent sex ratio were used to estimate vital rates over 7 and
231 18 years, respectively.

L178? How does one come to the conclusion that the Adult survival, recruitment rates, and population growth rate are correctly estimated? Probably because they have produced realistic values that are aligned with external sources of data?

The idea here was just to say that these parameters were estimated (convergence) with a satisfying uncertainty (comparisons possible between years), we modified the sentence:

232 The "apparent sex ratio" method successfully provided estimates and associated uncertainties for the latent
233 variables: proportion of immatures, and then, adult survival and recruitment rates (Figure 5).

Figure 5: I have trouble reconciling the very high cull rate presented in Figure 8 and the survival rates presented in Figure 5 in France. Maybe it is simply something that I misunderstood. Did you try to assess the correlation between survival and the cull rate?



Good point, we did not check this carefully before, here you have details about it:

On the previous graph, you have a comparison between two variables: the apparent adult survival, which is derived from the apparent sex ratio method (adult survival + migration), and 1 – adult culling rate, which can be labelled as adult survival to harvest; the latter does not consider natural mortality and migration. By definition, these two "survival rates" are different. Without immigration, the apparent survival is expected to be equal to or lower than the survival to harvest, which is the case most of the time (purple circle). A higher apparent survival (green circle) is expected if the immigration rate is higher than the natural mortality + emigration, or if the population growth rate is overestimated. Here we cannot distinguish between the two hypotheses, but there is a strong suspicion that the GB population is the source of the different continental populations in Europe through migration events. Therefore, the immigration hypothesis cannot be rejected in the present case study.

As this comparison is not straightforward, we avoid developing it in the manuscript, but we now emphasise the consequences of a biased population growth rate in the discussion and consequences of immigration.

315 Errors in the estimation of the population growth rate would also lead to absurd values of vital rates. This
316 could typically happen if the extensive winter census used to estimate population abundance takes place
317 before the arrival of the entire population in some years. In the present case study, the systematic count
318 conducted in mid-January seemed to avoid this pitfall. When applying the method to other species, the
319 approximate date on which all individuals reach their wintering grounds should be determined (e.g., from
320 ringing or satellite-tagging) to ensure an unbiased estimate of population size.

&

321 In some years, the estimated values for adult survival were outside the range of expected values for a closed
322 population, although they were never significantly above 1. One possible explanation for such discrepancies
323 is that individuals from other populations moved to France in those years. Although ringing data suggest
324 that Ruddy ducks in GB do not generally undertake long-distance seasonal migrations (Henderson, pers.
325 comm.), we know that migration events have occurred in at least some years, as continental populations of
326 Ruddy ducks in Europe are genetically closely related to the introduced GB population (Muñoz-Fuentes et
327 al., 2006). Consistent with this, the outliers in survival and recruitment rates observed in FR in 2002 and
328 2012 suggest that immigration events may have occurred.

L180-182: I suggest that your emphasis in your results which parameter is directly estimated from the data and which parameters is latent in your analysis. For example: The proportion of males among adults, which is a prerequisite for inferring the proportion of immatures, was estimated at 0.60 "in the GB population between XXXX and XXXX".

Another example: The proportion of immatures was "estimated" to range between 0.16 [0.07; 0.24] and 0.54 0.44; 0.62], depending on the population and year "by the model".

Ok, the section has been revised accordingly:

227 Over the period 1960-2019, population size fluctuated between 0 and 5,473 in Great Britain, and between 0
228 and 259 in France. During this period, 12,316 Ruddy ducks were culled in GB, and 2,246 in FR. Of these,
229 8,440 adults were used to determine the male proportion among adults: 0.60 [0.59; 0.61]. In GB and FR, 12
230 and 55 independent count events considering apparent sex ratio were used to estimate vital rates over 7 and

231 18 years, respectively.

232 The "apparent sex ratio" method successfully provided estimates and associated uncertainties for the latent
233 variables: proportion of immatures, and then, adult survival and recruitment rates (Figure 5).

190: Would it not be simpler to say "consistent" rather than "not inconsistent"?

Thanks, it has been revised

L92-193: Could this be because you estimated only seven years of data for GB but 18 years of data from FR?

Yes possibly. We added a sentence to notice this.

246 No trend was observed for the GB population over the seven years available, but the recruitment
247 rate decreased for the FR population, although it was noisier than the proportion of immatures

L201-202: True. But at the same time the apparent decline in the GB population is tied to the high survival rate observed in 2005-2006 so it is difficult to extrapolate to the entire time series.

Even if we remove the survival in 2005-2006, the survival variability in GB is more than twice as high as the productivity variability, and the decline in GB is consistent over the 7 years (see Figure 1), so this statement seems quite robust to the authors.

L211-213: Minor caveat but this sounds more like a discussion item (which you address below) rather than a result.

Ok, as it is redundant with the discussion, we removed this section. See discussion:

357 The "apparent sex ratio" method consistently yielded recruitment rates that were twice as low
358 as the method based on culling data, a result consistent with higher vulnerability of immatures to shooting
359 (Fox et al., 2014). In North America, for example, immature waterfowl were found to be 1.3 to 2.6 times
360 more vulnerable to hunting than adults (Bellrose, 1980).

L226: You may want to put "see 3rd Materials & Methods section" in parentheses.
Ok, modified as suggested.

L234-236: I am wondering if the differences observed in two the culling programs are really related to the timing of the cull, as implied in the results and discussion, or if the issue is related to the size of the cull. It seems to me you will be more efficient if you try to cull a population of a few thousand individuals than trying to cull a population of a few hundred individuals.

Good point, we moderate this conclusion now:

In Results, we describe facts:

285 Overall,
286 the culling effort in GB had a strong impact on both vital rates, leading to a greater population depletion
287 than in FR.

In Discussion, we added details on this hypothesis:

379, it cannot be excluded
380 that population size also influenced the presumed effect of the culling strategy as the population size in GB
381 was much larger than in FR.

L257-263: In North America most counts take place during the breeding season as many winter count have been phased out. It would be nice if you could suggest other species/situation to which your model could be applied to in the future.

We conclude this paper with a comment on this:

428 Finally, we would like to emphasize that the main aim of our approach was to decompose population growth
429 into its two main components, and not to provide unbiased estimates of adult survival or recruitment rates.
430 However, if necessary, count surveys could be designed to do this, at least in theory, for species that exhibit
431 observable delayed dichromatism. This trait is very common in dabbling and diving ducks, but the delay
432 in dichromatism is especially large in some species, making the method particularly suited for them, e.g. in
433 stiff-tailed duck, Oxyura sp. (Johnsgard & Carbonell, 1996), Tufted duck, Aythya fuligula, Black scotter,
434 Melanitta nigra americana, Common Goldeneye, Bucephala clangula americana (Bellrose, 1980; Johnsgard,
435 1978). To apply the "apparent sex ratio" method, the time window of censuses must be chosen wisely so that

436 it fits into the period when immature males do not look like adult males yet, which can differ from a species
437 to another. Apart from this precaution, modifying standard monitoring protocols to distinguish between
438 male-like and female-like individuals is almost costless, but worth the effort as it would greatly increase the
439 efficiency of conservation/management actions (Nichols & Williams, 2006).

L269: Are there many populations that are monitored by winter survey that can be safely assumed to be "closed" in Europe? Or is this new method doomed to be used only on a few cases study?

The method can be used for open populations. See M&M:

175 In an
176 open population, adult survival and recruitment rates are confounded with adult and recruit migration rates,
177 respectively, but this does not change the equation.

This is actually an advantage for detecting significant migration events between two populations (e.g. an apparent survival that reaches 200% means that the number of adults doubled and then, at least an equal number of adults have migrated into the studied population).

We changed the sentence:

"One possible explanation for these discrepancies is a violation of the "closed population" assumption."

to:

321 In some years, the estimated values for adult survival were outside the range of expected values for a closed
322 population, although they were never significantly above 1. One possible explanation for such discrepancies
323 is that individuals from other populations moved to France in those years. Although ringing data suggest
324 that Ruddy ducks in GB do not generally undertake long-distance seasonal migrations (Henderson, pers.
325 comm.), we know that migration events have occurred in at least some years, as continental populations of
326 Ruddy ducks in Europe are genetically closely related to the introduced GB population (Muñoz-Fuentes et
327 al., 2006). Consistent with this, the outliers in survival and recruitment rates observed in FR in 2002 and
328 2012 suggest that immigration events may have occurred.

L285- 287: Was that an hypothesis that you wanted to test? I guess I am not too sure why you expected those results.

The recruitment rate is expected to be higher because culling adults in the post-breeding season let them the opportunity to breed and then produce recruits. We have removed this sentence as it is explained in detail in the next section of the discussion:

382 As expected, the pre-breeding culling strategy had the same effect on adult survival and recruitment rates,
383 as a bird killed before the breeding season reduces the breeding population (lower adult survival rate) and
384 prevents the bird from reproducing (lower recruitment rate).

L296-297: You mean the results of the preliminary analysis conducted on the GB data? Or the results of your "apparent sex ratio" model?

I have the impression that if sex-ratio was variable in your model (instead of being fixed) the confidence intervals in your models would have been slightly higher.

Yes, you are right because the limited amount of data on a yearly basis leads to quite uncertain adult sex ratio, and pooling data from all years resulted in a more robust estimate. This approach was justified as no difference in the adult sex ratio was observed between years (see the new "Data formatting" section). The strong correlation with vital rates estimated using the "culling-based method" confirms that this approach is not flawed.

L307: Given that you were unable to assess this parameter maybe use "could" instead of "would"?

Right -> modified as suggested

L365-366: I can live with "biased" estimates of adult survival or recruitment rates from the proposed model but what are the consequences of using the parameters in a management framework?

If there is a bias, the prediction of the effects of management measures may be flawed. However, when bias is suspected, the estimated fluctuations in vital rates remain a detectable response to management measures, and provide valuable information for managers.

L371-373: I would need to have more information on the methodology of the winter surveys conducted in Europe (or elsewhere) to assess the validity of this recommendation. Are there a lot of surveys right now where the sex of the observed ducks are not recorded?

Unfortunately, in the most extensive European mid-winter census, separating male and female plumage classes is not mandatory and rarely done ("This is usually done as a part of detailed demographic studies which are currently beyond the scope of IWC" https://www.wetlands.org/publication/iwc-guidance-field-protocol-for-waterbird-counting/).

Review by anonymous reviewer 2, 14 Sep 2024 02:07

I reviewed a pre-print entitled “Delayed dichromatism as a convenient tool to disentangle the effects of survival and productivity on the population dynamics in waterfowl”. The work is intended to investigate how, in dimorphic species, delayed sexual maturity of males can be used to estimate adult survival and recruitment rates by distinguishing male-like and female-like individuals in winter counts. The authors used the “apparent sex ratio” method to estimate adult survival and recruitment rates and evaluate the effects of two different eradication strategies used in Great Britain and France, respectively. The study is interesting and could be a valuable contribution. There is a dire need to develop tracking methods to effectively measure changes in population size and population growth rate to assess the relevance of management actions.

Title and abstract

Does the title clearly reflect the content of the article? No

Here is a shorter title if it suits better:

Delayed dichromatism in waterfowl as a convenient tool for assessing vital rates

The abstract section needs significant improvement, mainly highlighting the limitations of previous methods, the authors need to present their key findings here.

Ok, modified accordingly

Monitoring the number of individuals is by far the most popular strategy for studying the environmental factors that determine population dynamics and for measuring the effectiveness of management actions aimed at population recovery, control or eradication. Unfortunately, population size monitoring is inefficient in identifying the mechanisms underlying demographic processes and, in particular, in assessing the extent to which population growth rate is influenced by changes in adult survival rather than variations in reproductive parameters. In many waterfowl species, sexual dichromatism is observed in adults, while immatures of both sexes display a plumage pattern similar to that of adult females. In these species, the apparent proportion of males increases as the female-like immature males gradually take on the plumage of adult males. The difference between the apparent sex ratio before and after the young reach sexual maturity then provides information about the age ratio of a population. Using winter counts that distinguished between female-like and male-like individuals of two non-native populations of Ruddy duck Oxyura jamaicensis, a species that exhibits such a plumage pattern, we present a noninvasive method based on the apparent sex ratio to split population growth rate into adult survival and recruitment rates (the latter also referred to as productivity). This method can correctly detect annual changes in vital rates, supporting the assumption that counts conducted in an appropriate time window reflect the age structure of a population. We exemplify how the respective contributions of survival and productivity to the population growth rate are essential for understanding the processes behind demographic dynamics. Finally, we point out some best practices to correctly apply the “apparent sex ratio” method described here.

Introduction

Are the research questions/hypotheses/predictions clearly presented? [ ] Yes, [P] No (please explain), [ ] I don’t know
The introduction section lacks hypothesis and proper research questions. This sections sufficient evidence, and briefly explains the limitations of what has done earlier. At the end the authors have also provided scientific rationale of the current study. Authors are suggested to state a clear and concise hypothesis and specifically mention the research questions which are solved in the current research.

Ok, we now explicitly detail the research question and the tested hypotheses:

28 However, only few species, at least in
29 waterfowl, exhibit morphological differences between age groups that can be recognized from a distance, and
30 then recorded during counts. Alternatively, assessment of age structure in hunting bags has been used to infer
31 the role of decreasing reproductive success in population declines in a number of game species, including ducks
32 and geese, but this suffers from intractable biases (Fox & Cristensen, 2018). Clearly, the latter approach is
33 also not appropriate for protected/endangered species. This emphasizes the need to develop non-invasive
34 alternatives that can correctly track vital rates.

35 Here we test the hypothesis that delayed sexual maturity in dimorphic waterfowl species is a reliable source of
36 indirect information on the age structure of a population, allowing the tracking of changes in vital rates over
37 time. To this end, we first implemented a model based on winter counts that distinguishes between male-like
38 and female-like individuals to infer these vital rates. We then compared the outputs with other sources of
39 information on vital rates to check that the estimates coming from the tested method are not biased and
40 accurately reflect changes in vital rates over time. We used two non-native […]

Does the introduction build on relevant research in the field? [P] Yes, [] No (please explain), [ ] I don’t know

Materials and methods

Are the methods and analyses sufficiently detailed to allow replication by other researchers? [P] Yes, [ ] No (please explain), [ ] I don’t know

Are the methods and statistical analyses appropriate and well described? [P] Yes, [ ] No (please explain), [ ] I don’t know

Results

In the case of negative results, is there a statistical power analysis (or an adequate Bayesian analysis or equivalence testing)? [P] Yes, [ ] No (please explain), [ ] I don’t know

Are the results described and interpreted correctly? [P] Yes, [ ] No (please explain), [ ] I don’t know

Discussion

Have the authors appropriately emphasized the strengths and limitations of their study/theory/methods/argument? [P] Yes, [ ] No (please explain), [ ] I don’t know

Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [P] Yes, [ ] No (please explain), [ ] I don’t know