

Dear Recommender and reviewer,

Thank you for your feedback on our manuscript. We have addressed all the comments, this is detailed below in a point by point response.

Also, we would like to inform you that we noticed an error for the data of environmental covariates for one nestbox. For nestbox 6815702 the values were not correct because in the dataset two nestboxes had incorrectly the same label and thus the areas of both nestboxes have been added together. This error has been corrected. Thus there is a new version of the dataset published on Zenodo data (<https://zenodo.org/record/8316158>). When we ran again the statistical analyses with the corrected version of the data the results did not change but we had a problem of collinearity between the environmental covariates that was not detected before. For this reason and also because of the comments of reviewer 2 we decided to change our statistical approach. We kept only one environmental covariate for each nestbox: the proportion of favourable habitat defined as the proportion of meadows and orchards in the buffer (see lines 154-157). Since with this new approach we have only one environmental covariate it was possible to perform only one statistical model that include the individual and environmental covariates. Thus, all the statistical models are now also taking into account the effects of hatching rank, sex, year of birth, nestling number, nestling age, the proportion of meadows and orchards, and the interaction between hatching rank and sex, and the interaction between hatching rank and the proportion of meadows and orchards (see lines 187-189). This modelling decision did not change the main results but clarified the and we hope you'll be convinced that it answers the related comments of the reviewers.

Dear François and co-authors

I have now received reviews from two referees. Based on these two reports and my own reading of the paper I am willing to recommend your paper. However, reviewers have raised important points that need to be addressed in a revised version. Taking these comments into account will definitely clarify the magnitude and the robustness of the relationships documented in your paper. This needs to be done before I can recommend the preprint.

I also have a few additional remarks. Some of these comments echo those of the referees

- Lines 90-92: Agree but this requires to test an interaction between laying order and environmental conditions which is not what has been done in your analyses.

Agree, we added the interaction  $\text{rank} \times \text{proportion of favourable habitat}$  (see above for the reason why we changed the analysis of environmental covariates).

- Lines 96- 96: Please rephrase as environmental conditions is unlikely to influence the sex of the individuals. More generally, I don't understand the rational of this analysis as the whole introduction is focused on the factors modulating telomere length. This comes rather out of the blue.

Done (lines 103-108).

- Lines 122-124: Please give the units

There is a linear relationship between the age and the length of the feather:  $\text{length of the feather} = 3,3 \times \text{age} - 36$ . The length of the feather is in mm, the age is in days, 3,3 is the slope (mm/day); -36 is the intercept (mm). This formula is valid between age 15 and age 35. The units have been added in the text of the article (Line 133)

- Line 127: I assume this contains broods with 1 chick, you should rephrase the sentence as: 'The original dataset contained 142 nestlings....'

No there is no broods with 1 chick. We have a larger dataset but for this study the molecular analyses were only performed on a subset and we only included broods with more than 1 chick in this subset. This was rephrased for more clarity lines 139-142.

- Line 167: Please give the formula of the ScaleMassIndex

Done, lines 183-186.

- Lines 171-172: What is the justification for performing a model averaging approach.

We had several models with comparable AICc. In this case using a model averaging approach allows to take into account the size and presence of effects among the set of top models. For simplicity we applied the same method for both analyses (analysis of SMI and analysis of telomere length). It is true that when the null model is the best model we can expect that no effects will be significant and it is not necessary to have a model averaging approach. We explain this in the new version of the manuscript (lines 191-192). With the new statistical approach, the best model is not the null model, even if the null model is still in the top model set. Thus, the results are not changed (there are no significant effects) but we rephrased the results to remind that the null model is in the top models set (lines 212-213).

- Lines 190-193: This is due to the fact that the constant model has the lowest AIC. This model should be retained. I don't see the justification for performing model averaging in such a situation.

As explained above we ran again all the models. The constant model does not have the lowest AIC but it is indeed in the top model set (delta AIC = 0.93) for the analysis of SMI.

- Lines 283-284: They likely face the trade-off but differences in individual quality might explain the pattern.

We added the "high quality explanation" (lines 332-334).

I hope that all these comments will be helpful,

Kind regards,

Jean-François Lemaître

We appreciated your feedback and hope to have answered with adequacy to the points addressed.

### **Review 1:**

Comment to the Authors:

This paper presents an analysis of the effect of hatching rank, body condition, sex and nest environment on telomere length using 4 years of data on a raptor species, *Athene noctua*. The results suggest that telomeres are longer in females and in individual in better body condition. In addition, authors found, following their prediction, a negative effect of hatching rank on telomere length. This effect was due to shorter telomere in the last-hatched individuals in large clutches. Finally, while no

effect of the environment was found on telomere length, telomere length seems to decrease over the 4 years of sampling. The manuscript is nice to read and presents interesting results, however, I feel like the result section lacks basic information (as described in my detailed comments to the authors). Relationships seem to be weak and authors should be more cautious about their findings. The result section should be more detailed and more complete (e.g. adding a table with estimates, SE and confidence interval for explanatory variable). My detailed comments are presented below, and I hope they will be helpful to the authors in revising their manuscript.

Florentin Remot (please note that I sign all my reviews)

Thank you for your comments. As the reviewer is right in saying that the present draft report a first descriptive study on telomere dynamics in our population, we tried to modulate our conclusions over the discussion and in the abstract. Please see lines 30, 302, 310-313, 348-350, 353.

My main concerns are:

- All your results are based on model averaging, however, your supplementary tables are the list of the best models (with a  $\Delta AICc < 2$ ) and you never present the results from the model averaging. Presenting estimates and confidence interval from model averaging is mandatory here. Since effect are small, they are not clear on some figures (see comments below). Results would be better highlighted if they were presented in a table in the main manuscript.

We agree with you that it is important to show the estimates of the average model. For the model on telomere length, the estimates and confidence interval are shown in Figure 1. We did not add a Table since it will be redundant information with the figure that already presents the value of the estimates in the main manuscript. For the model on SMI, since there are no significant effects the figure with the value of the estimates was presented in appendix (Figure S1).

- As qPCR is a relative measure of telomere length, it is always difficult to interpret the biological relevance of an effect. For instance, looking at Table S3, slope for the cohort effect is around -0.053. While the cohort effect may be statistically significant, one could wonder whether this effect is biologically significant. Same comment for sex differences, what is the magnitude of the sex difference in telomere length? How much longer are telomeres in females than in males? This question needs to be addressed in the manuscript.

Following a next comment from the same reviewer, we modified our statistical analysis by using a qualitative method to assess the cohort effect. Consequently, there is no slope values anymore for the cohort effect in the corrected draft. Nevertheless, the reviewer is right in asking whether the sex or cohort difference has any biological meaning in our case. As we did not test for relative telomere length vs. survival rate of the chicks, it is impossible to conclude yet. Still, there are several past studies looking at sex differences in telomeres length (using different methods) that did find estimates around 0.1 and 0.2 in species where it is known that there is a sex difference in lifespan (e.g. Barrett et al. 2011). We already added in our discussion lines 307-308 that the biological significance of our sex difference in telomere length or the putative link with survival need to be further tested.

- Regarding the cohort effect, the year of birth was tested as a quantitative variable, meaning that a linear relationship was tested. However, with only 4 years of data, do we expect a linear relationship to occur? Looking at Figure 3, telomere length seems to be shorter only in 2017. In my opinion, it would be interesting to test the cohort as a qualitative variable (like hatching order).

We agree with your comment and decided to test the effect of cohort with a qualitative variable. And indeed it is only the last year (2017) where telomeres are shorter. The results have been modified (see for example Figure 1) and this has been taken into account in the discussion (Lines 302, 310-312).

- Regarding hatching order, looking at the raw data, it seems that for many nestboxes the exact hatching order within brood was not known (individuals with the same rank value within brood). This information and how you dealt with it need to be mentioned in your materials & methods section.

It can happen that some hatching had the same size. In this case they were estimated to have the same age and thus they had tied rank. This information has been added (Lines 136-137).

Minor comments:

Lines 127-129: Sample size are quite small for each level (especially for broods with 2 and 5 chicks). In my opinion, it is hard to tell whether the absence of effect of nestling number is biologically meaningful or is due to low power. In addition, the effect of nestling number could be tested separately by also using clutches containing 1 chick.

While we totally agree that it would have been interesting to have also clutches with only 1 chick to investigate the effect of nestling number on chick telomere length, this was not the main focus of the present study. This is why we did not measure telomere length of chicks in clutches containing only 1 chick. Thus, we do not have the data to add this analysis in the manuscript. In this study, we aimed to test the effect of hatching rank and this is why we included only clutches with more than 1 chicks (as explained above). Indeed, hatching rank has been previously shown to be related to chicks' telomere length, and this is this question that we wanted to address here. Still, as the reviewer pointed out, we added nestling number as a covariate to control for potential effects but only with the aim to control for in analysing the effect of hatching rank on our response variables.

Line 172: Could you be more precise regarding the model averaging procedure? Which package/function did you use? Your estimates were from conditional or full averaging?

We used package MuMin and conditional average. We added the information line 192.

Line 210: Is the positive effect of SMI on RTL statistically significant? It is not clear on the figure. As I previously said in the general comment above, you present as supplementary tables the list of the best models selected based on AICc, however you then use a model averaging procedure for your results. Thus, it'll be more informative to add in the main manuscript the table with the estimate and the 95% CI calculated from the model averaging.

This information is on the Figure 1. In order to have a clearer figure we added significant asterisks.

Figure 2: The sex difference is not clear on the figure; my suggestion would be to remove the figure or to represent the sex difference differently (e.g., with a pairwise contrast plot).

Done, Figure 2 was removed as requested.

Figure 3: The cohort effect is not clear on the plot. I would suggest plotting the cohort effect and the hatching order effect separately.

Done, Figures 2 and 3 were modified as requested.

Line 233: The use of the word 'longitudinal' here is a bit confusing as it is not a longitudinal follow-up. I would suggest dropping this term from the sentence since it could be misleading for the reader.

Done (Line 258).

I would suggest authors to be consistent with variable names, for instance 'hatching order' in Figure 1 but 'Rank' in table S3, as it might be confusing for the reader.

We carefully changed hatching order to use always the same term: hatching rank or rank.

## **Review 2:**

Thank you for the opportunity to review "Telomere length vary with sex, hatching order and year of birth in little owls". In this study, variation in TL due to sex, hatching rank, body condition and environmental factors was investigated in a wild population of little owls. The study found females to have longer TLs compared to males. Nestlings with higher body condition also had longer TLs, independent of the sex of the individual. Finally, in large clutches, it was found that last-hatched nestlings had shorter TLs. These results present an interesting case study of how variation in biological factors such as sex and body condition influence TL in the wild. This paper is a nice addition to the burgeoning literature on how early-life conditions experienced by wild populations and how key biological variables such as sex and body condition affect TL dynamics.

My major comments are 1) the lack of information provided in the methods section relating to the decisions made during statistical modelling as well as 2) the models investigating how environmental factors influence TL where sex, body condition, year etc were not fitted as fixed effects in these models when they should have been. Re-running these models to include these biological variables as predictors would be required to infer the effects of environmental factors on TL. Some clarity in the ideas expressed in the introduction and discussion sections of the manuscript would also improve the readability of the paper.

We thank the reviewer for his/her constructive comments. In particular, note that all models have been checked and rerun according to both reviewers advises (see above and below for more details).

Major Comments:

L.157 – Were models including year/cohort as a random effect tested? Not accounting for year-to-year variation could potentially have significant effects on the final results and inference. Please explain why year was not included in the models and how much variation is explained by year across the different models.

Year (i.e. cohort) is in fact included in the models as a fixed factor, as we aim to establish the dynamics of chicks' telomere length and SMI over consecutive years. See lines 186-189 and 201. In addition, as a global response to both reviewers' comments on our original statistical analysis, all models have been rerun, in particular according to the use of year as a qualitative explanatory fixed factor or to the use of environmental variables.

L.165 – Why does this model not include cohort as a fixed/random effect but the other models (RTL, environment) account for variation in birthyear? Please justify the exclusion of birthyear in some models but not others.

This model was initially used to check whether there were not more males (with on average shorter telomeres) in rank 5, as an explanation of shorter telomeres in rank 5. However, as sex is included in

the model testing for ranking effects on telomeres, we just deleted this part. This means that now all the models presented in the paper include cohort as a fixed effect.

L.173 – Please specify why this model was run in the first place? Presumably, to test for environment-dependent effects on body condition. Also, in these models, were other fixed effects (such as sex, body condition, cohort) also included? Please explain why these predictors were excluded from this model. Is it because a previous model with SMI as a response variable found none of these variables to explain variation in the response. If so, then clearly state that. Re-running this model investigating how environmental variation explains variation in body mass, over and above effects due to sex, body condition, cohort etc would be great to show.

As explained above we changed all the statistical analyses. We integrated all the covariates in one analysis. Thus, the effect of the nestbox environment is now tested taking into account the effect of sex, SMI and cohort (see lines 186-189).

L.190 – How can it be explained that none of the biological variables (such as sex, hatching order, rank, cohort etc) explain variation in body condition/SMI but explain variation in TL? Isn't it more obvious, biologically speaking, to expect these factors to influence body condition/SMI more than TL? Please discuss the lack of consistency in results here in the discussion. I was also wondering if body mass is used as a response variable instead of SMI, whether the results are consistent?

Telomere length may record short-term stress that is not visible on phenotypic variables like body mass. For instance, short reduction in growth rate followed by a growth compensation when feeding conditions are re-established may not be visible in body mass or condition at day 20, while telomeres may be significantly affected (Metcalf and Monaghan 2001, 2003). We added this point shortly in the discussion lines 284-288.

We also performed the analysis with body mass as the response variable. The results are the same. In this case there are 5 models selected and the covariates are: age, cohort, number of nestlings, and the proportion of favourable habitat. However, there is only a significant effect of age (positive effect), which is expected since the chicks are growing. There is also an indication that chicks are lighter for the cohort 2016 only. Thus the results are consistent whether we use SMI or body mass. As explained below in response to another comment we prefer to use SMI.

L.214 – I think it would only make sense to interpret any environment-related effects on TL after accounting for variation due to biological factors such as sex, hatching rank and cohort which were shown in the previous model to influence TL.

This is now the case with the new statistical models (see lines 199-203).

L.312 – Ideally, these models would need to be re-run after accounting for variation due to sex, hatching order and cohort and checking to see whether any environmental variables still crop up as significantly explaining variation in TL, over and above these obvious biological factors which were previously shown to affect TL in this population.

Done (see above).

Minor Comments:

L.45 – Was the primary interest in evolutionary biologists to study TLs solely driven by stress-related effects on TL? I surmised it was also the fact that TL declines with age and therefore being involved with lifespan/the ageing process at the cellular level (Monaghan., 2010). Rephrasing the sentence to convey this would be great.

Done lines 46-51.

L.54 – I find this sentence confusing – is the manuscript referring to pleiotropy in the context of the antagonistic pleiotropy theory of ageing and therefore are being explicit about genetics (Williams., 1957)? Or is the paper referring to tradeoffs between life- history traits (i.e. investing resources in growth vs other traits; Stearns., 1992)? Please provide more clarity.

We replaced the word pleiotropy with a more explicit term (life history trade-off) (line 56).

L.57 – Which physiological traits are being referred to as both here? I take that one of them is growth but what's the other – cell division? It is not clear.

We simplified the sentence, but indeed the two physiological traits were the high metabolic rate and rate of cell division. (Line 59).

L.58 – I find this sentence confusing and too long, would it be more straightforward to say, “Studies have shown juveniles exposed to challenging conditions in early life to have shorter TL. This could be due to reduced investment in somatic maintenance as a consequence of low resource availability when conditions are harsh”.

Done (lines 63-65).

L.66 – Instead of saving telomere length, would preserving TL work as a better alternative?

Done (Line 70).

L.72 – What is the multivariate egg concept? Please elaborate and also provide clarity in this sentence. Is the paper intending to convey that egg-associated traits could be positively or negatively correlated with each other and may influence future offspring phenotype? If so, please rephrase sentence to reflect that.

This is exactly the message, we have rephrased the sentence (Lines 76-78). We removed the reference to the multivariate egg concept that is not necessary in this paper (see also Lines 87-88).

L.85 – Please include reference to the study.

This is the study Noguera et al., 2016 (Line 87).

L.85 – I do not understand this sentence – is the paper trying to convey that many studies have shown negative consequences of telomere erosion during growth on future individual fitness. Please state that clearly if so.

L.89 – This sentence is very long and difficult to follow. And it is not clear to me why variability in TL within clutch is or is not an epiphenomena? Please use simple and plain language to convey the main ideas.

Following both of your comments we have reformulated this sentence. We inverted the order of the sentence to be more straightforward. The idea is to explain that many studies have shown negative consequences of telomere erosion on future individual fitness and thus if there is variability in telomere length at hatching it may be important to take into account this difference in the studies of telomere length dynamics (Lines 91-96).

L.123 – How accurately does this formula capture the actual age of the bird? Please provide more information.

As also requested by the recommender, we added more details on this formula (Lines 131-136).

L.125 – What is ventral covert? Is this blood sampled from covert feathers?

We collected ventral feathers. The sentence has been corrected. (Line 137).

L.128 – Did all 39 broods have more than 1 chick? Please specify the number of broods that were excluded/part of the final analysis.

Yes, all the broods had more than 1 chick: this is how we selected the brood to include in the molecular analysis.

L.138 – Was the final nest environment category assigned based on which of the groupings (buildings, meadows, crops... etc) covered the most area (in m<sup>2</sup>) within the buffer zone? Was it always clear that one category covered more area than another category? If not, what was done in such a scenario?

Concerning the environmental covariates we calculated the surface covered by each category (buildings, meadows...). Then we calculated the proportion of favourable habitat for the little owl (the reason for this change in this analysis has been explained at the beginning of this letter). Thus, in this new analysis we have only one variable to characterize the nest environment (this is a continuous variable) (see lines 154-157).

L.167 – Why was SMI used as a measure of individual body condition instead of the body mass of nestlings? Please justify the use of this measure instead of body mass. Were the results consistent when body mass was used instead?

We followed the conclusions of Peig and Green 2009, where they showed that using SMI gives a better measure of the relative size of body energy reserves than other methods and obviously of body mass. We agree with the reviewer that body condition in growing chicks, particularly when they are approaching fledging is a complicated question, for instance because of variation in water body content and the final feather growth (see Durant et al. 2008). Still, with our single-measure of chicks, we think that using SMI is better than body mass to assess nutritional status of our birds. As explained above, we checked and the results are consistent with body mass.

L.171 – In total, how many models were compared using AICc? Please include this information.

For the model on SMI we compared 104 models and for telomere length we compared 208 models.

L.172 – How many models were part of the set of top models within 2 AICc of each other? Additionally, how was model averaging performed? Were specific R packages (MuMin etc) used to obtain model-averaged estimates? Please provide these important details.

The list of the models part of the set of top models is given in appendix (Table S1 and Table S2). We added the number of models in the main text (see line 210 and line 228). We added more details on the R package, see line 192.

L.174 – Was the environment a single predictor consisting of six different categories? Or were they all fit as separate predictors (each environment type a different predictor)? It is not clear to me how the environment experienced by nestlings was modelled.

See our previous answers, this analysis is different in this new version of the paper.

L.174 – How correlated were the different environment variables with each other? Was the VIF < 3? Please include this information.



We agree with the reviewer that this information was not originally clearly provided. Also, as we explained to the recommender, we noticed an error for the data of environmental covariates for one nestbox that was corrected and lead us to restructure the way environmental data were analysed, notably to avoid collinearity. This is now clearly presented lines 154-157 and 175-176.

L.180 – Why was RTL log-transformed? Please explain the rationale behind this modelling decision.

Log transformation was done to get a normal distribution of model residuals.

L.235 – What does the indication of an erosion of nestlings' RTL over years need to be replaced with? It is not clear to me. Is the paper suggesting that the inference needs to be placed in the context of recent studies investigating physiological markers of quality in conservation biology? If so, what have these studies found – please elaborate.

The whole paragraph of this part of the discussion (starting line 260) is dealing with the point raised by the reviewer: what may significate the loss of telomere length in our species in terms of population dynamics, or differently said, is it a good proxy for that purpose. However, as our new analysis of the year effect (qualitative and not continuous as previously, see reviewer 1) now indicates that telomeres are shorter only the last year of the study (and not over the full length of the study), we changed accordingly this part of the discussion (see lines 260-261).

L.242 – It is not clear to me how increased cell division in young, growing individuals means that they are impacted most by environmental stressors? Both of these could be simultaneously occurring as separate causes influencing TL (i.e. increased cell divisions and oxidative stress due to growth leading to TL shortening (Boonekamp et al., 2017; Reichert & Stier, 2017). Additionally, exposure to environmental stressors in early-life may affect senescence rates and by extension, TL dynamics (Lemaitre et al., 2015; Watson et al., 2015)). But there is no direct line of causality linking exposure to environmental stress and increased cell division in growing individuals.

To be clearer beside the classical explanation for growth, we extended this idea lines 270-282.

L.247 – Is it being suggested that nestlings born prematurely may be physiologically older due to the harsh early-life conditions they may experience? If so, please state that clearly. Does premature birth constitute harsh early-life conditions? Are nestlings born earlier in the hatching order considered premature?

Our draft was likely misleading. We were not talking about premature birth but just of accelerated aging. We corrected the line 284.

L.272 – I think sex and gender refer to two different things (<https://orwh.od.nih.gov/sex-gender>) and being consistent with what is being referred to in this context (i.e. sex) is important, even if it is repetitive.

Done (Line 316)

L.275 – It is not clear to me what a consensual general pattern is?

Corrected, lines 318-319.

L.279 – More information regarding general patterns of sexual dimorphism in growth rates in bird populations would be great to include.

We added a reference underlying this point (lines 327-328).

L.295 – The suggestion that parents adaptively manipulate sex ratios or invest more in female offspring compared to male comes with some issues (see Hasselquist and Kempenaers, 2002). Please acknowledge this when providing such an adaptive explanation for the results especially when investigations linking fitness and TL have not been explored here.

Done (lines 348-350).

L.299 – I think it is important to highlight here that the sample size of the largest clutches (n=6) is small.

Done (line 353)

L.280 – I think the paper is intending to say “even” and not “event”?

Corrected (Line 324)

L.296 – “Benefit” instead of beneficiate may be more straightforward to understand?

Corrected (Line 346)

## References

Barrett, E.L.B. & Richardson, D.S. 2011. Sex differences in telomeres and lifespan. *Aging Cell* **10**: 913–921.

Durant, J. M., Landys, M. M., & Handrich, Y. (2008). Composition of the body mass overshoot in European barn owl nestlings (*Tyto alba*): insurance against scarcity of energy or water?. *Journal of Comparative Physiology B*, *178*, 563-571.

Metcalfe, N.B. & Monaghan, P. 2001. Compensation for a bad start: grow now, pay later? *Trends Ecol. Evol.* **16**: 254–260.

Metcalfe, N.B. & Monaghan, P. 2003. Growth versus lifespan: perspectives from evolutionary ecology. *Exp. Gerontol.* **38**: 935–940.

Noguera, J.C., Metcalfe, N.B., Reichert, S. & Monaghan, P. 2016. Embryonic and postnatal telomere length decrease with ovulation order within clutches. *Sci. Rep.* **6**: 25915. Nature Publishing Group.

Peig, J. & Green, A.J. 2009. New perspectives for estimating body condition from mass/length data: the scaled mass index as an alternative method. *Oikos* **118**: 1883–1891.