

## Response to reviewers

We would like to start by thanking Wolf Blanckenhorn, Frédéric Guillaume, and the two anonymous reviewers for their helpful and thoughtful comments. We address their questions and concerns below. There is no doubt their comments have significantly improved our manuscript.

### A population biological modelling approach to body size evolution

Body size evolution is a central theme in evolutionary Biology. Particularly the question of when and how smaller body sizes can evolve is of continuing interest within the field evolutionary ecology, because most life history models, and the empirical evidence, document that large body size is favoured by natural and sexual selection in most organisms and environments at most times.

The paper by Coulson et al. lifts this question to the level of the population, a novel approach, by using so-called integrated projection models (IPMs).

A minor correction here – we used integral projection models and not integrated projection models.

As well outlined by (anonymous) Reviewer 1, the authors assume the well-known carrying capacity ( $K$ ) of population biology as the fitness parameter to be maximized (rather than body size per se), and observe density-dependent (as well as density-independent), size-structured population growth trajectories in terms of age and size at maturity (including also other standard life history traits). Importantly and interestingly, life-history trade-offs are not assumed, as happens frequently in life history models, but emerge as a property from the modelling approach taken here. The authors find that often large body size indeed evolves, but under some (not overly rare) parameter combinations small size can also evolve, while yet other combinations lead to disruptive selection on body size. These results may ultimately explain the evolution of smaller body sizes from large body sizes at least under some environmental circumstances (despite common selection favouring larger individual body sizes).

All reviewers agree that the approach taken seems technically sound (as far as it can be evaluated), and that the results are interesting and worthy of publication after some revision. Nevertheless, at various places clarification and justification of e.g. some assumptions need to be provided as suggested by the reviewers.

Criticism centers on the often too technical descriptions of the model and its assumptions, especially if the targeted readership are general evolutionary ecologists. This should be changed in a revision of the manuscript, and especially reviewers 1 & 2 have made multiple concrete suggestions. One solution is to write the entire manuscript for a more general audience, and to relegate some of the more technical descriptions and justifications for the modelling specialists to an appendix (or the Methods).

We have simplified the more technical elements of the manuscript and in one case moved some technical elements to an Appendix. In particular, we have significantly restructured the methods and results to simplify the core message. We have also added in non-technical paragraphs to describe our approach before providing technical details.

In general, and related to the previous criticism of being too technical in writing, the precise focus of the paper needs clarification in the Introduction (again referring to reviewer 1s & 2s comments).

We have clarified that the main question is to identify the demographic circumstances when gigantism and dwarfism are expected to evolve. We have added an additional sentence to stress this is the objective in the introduction, methods, and discussion.

Reviewer 2 additionally points out the necessity of connecting the action of natural selection, in terms of mechanistic selection coefficients, to this overall phenomenological approach. This would help reconcile any differences in the results between this type of population biological model and the more traditional life history models.

An interesting comment, but not as straightforward to address as it might appear. First, the strength of selection is the derivative  $\frac{dK}{dz_m}$  that could be calculated from figures 1 and 2. Obviously the derivatives are sigmoidal, with a switch in sign at the point of lowest fitness. We could numerically calculate these easily enough but are unclear it adds much.

What the reviewer is asking is whether we can calculate selection differentials or selection gradients on the life history but to do that one would need a population of competing life history strategies. The selection differential on life history strategy could then be calculated as the covariance  $cov(z_m, \frac{K}{\bar{K}})$  where  $\bar{K}$  is the mean carrying capacity across competing life history strategies. However, one strength of our approach is we do not need to conduct an evolutionary game where strategies compete against one another. We gain all our insight by examining the dynamics of monomorphic populations consisting of a single strategy.

We can see, however, where the confusion might have arisen, and it is worth expanding on this. As noted by the reviewer, we use a population-level approach with each population consisting of individuals following a monomorphic life history strategy. What we mean by this is the survival, reproduction, development, and inheritance functions that determine the population model are probabilistic but with all individuals within the population being subject to the same set of equations. If we simulate individual life histories from these functions, the resulting life histories will not all be identical. We could simulate life histories by determining each individual's survival, growth, and reproduction at each age by randomly drawing numbers from distributions and comparing these numbers with the values predicted from the survival, reproduction, and growth functions. Depending on whether the randomly drawn number or the predicted number is larger, individuals either live or die, and reproduce or not. A similar process is used to determine how much each surviving individual will grow.

The simulation process is like repeatedly tossing a coin 10 times and recording the sequence of heads and tails. Let's assume the first sequence we record is HTHHTHTTTH. We now toss the same coin in the same way another ten times and record the sequence. You will not be surprised if the sequence differs. Rather you would be amazed if you got the same sequence twice in succession. The chance of doing so would be  $0.5^{10}=0.00098$ . Similarly, randomly simulating whether an individual lives or dies, and reproduces or not, will be highly unlikely to result in two identical life histories if only two are simulated.

We can consequently generate many different life histories from the same life history generating process – the monomorphic IPM. This may not be surprising, but what perhaps is more surprising is that if you were to calculate selection differentials and gradients on body size from the model you would discover body size is under selection. The reason for this is those individuals that survive to older ages by chance will have grown larger and will have higher survival and reproductive rates given the shape of the survival and reproduction functions. However, body size cannot evolve in response to this selection because there is no selection on life history strategy within the population because it is monomorphic. The individual life histories are generated by random realisations around a single generating process.

Most IPMs parameterised from field data describe the average life history of the population. These IPMs can be used to calculate selection differentials and gradients and to explore how they arise. Of course, there may be multiple competing life history strategies within the population used to parameterise the IPM, but a standard IPM does not capture this. They either need to be analyzed by asking what would happen if the life history were to evolve using, for example, sensitivity analyses, or the IPM needs to be extended to explicitly incorporate heritable variation. Both approaches are valid.

There is a significant literature on these topics. First, various authors have pointed out that purely random sampling of a single life history generating function can result in distributions of lifetime reproductive success that suggest significant opportunities for selection on phenotypic traits such as body size, when in fact no evolution is possible because all life histories were generated from the same process (Steiner and Tuljapurkar, 2010; Snyder and Ellner, 2018). It is also possible to generate exact distributions of LRS from a single IPM (Tuljapurkar *et al.* 2020, 2021).

Second, to estimate a selection differential on a life history strategy, one needs multiple different life history strategies competing against one another within a population at a point in time. We would need more than one life history generating function within a population: i.e. two or more IPMs. Once again, this has been the subject of several papers (e.g. Kentie *et al.* 2020; Childs *et al.* 2016). In these cases, each genotype (or strategy) can result in a slightly different distribution of lifetime reproductive successes and phenotypic trait distributions and evolution can occur. It is possible to estimate a selection differential on life history strategies when there are multiple strategies competing within a population, but that is not the approach we use.

Put another way, in the approach we use, we do not need to directly compete different life history generating functions (i.e. IPMs) against one another as was done by Kentie *et al.* and Childs *et al.*

We have tried to stress these key points in the revamped manuscript, but given reviewers have requested shortening the paper, we believe that a new paper discussing these nuances may be a better place to delve into these details.

Finally, all reviewers made some more specific, minor suggestions on how to improve the paper even further that should be addressed in a revision.

We have addressed these issues.

I am looking forward to seeing a revised version of this manuscript in light of the reviewer comments.

Wolf Blanckenhorn, University of Zürich

May 2022

## Reviews

*Reviewed by anonymous reviewer, 07 Apr 2022 06:36*

In this manuscript, the authors study body size evolution, with a particular focus on identifying what promotes the evolution of extremes in body size.

A very succinct and elegant summary of the topic of the paper. Thank you.

They build a size-structured integral projection model (IPM), use carrying capacity as fitness, assume no a priori life history trade-offs, and model some life history traits as density-dependent and others as density-independent. They show that 1) some parameter combinations and scenarios can result in disruptive selection and lead to the evolution of extreme body sizes,

A minor correction – we show that across ranges of parameter values we observe disruptive selection. We see disruptive selection by comparing across populations of monomorphic life history strategies. We have reworked bits of the manuscript text that may have been confusing to clarify this point.

2) under disruptive selection, if the cost of delaying maturity is compensated by a benefit to adults (increased reproduction or lifespan), large bodies will be favoured, otherwise small bodies will evolve,

Another minor clarification – disruptive selection occurs because these costs and benefits vary as we modify developmental trajectories across models. We do not impose disruptive selection on our model, it is an emergent property.

and 3) life history trade-offs can emerge without any a priori trade-off assumptions, simply due to the fact that (negative) density-dependence leads to minimization of density-dependent traits, while density-independent traits are maximized.

An excellent summary.

General comments:

Overall, I think the topic of body size evolution is very interesting. Body size is a fundamental trait linked to various life history traits, evolutionary dynamics, and ecological interactions, and I think it is always fascinating to read a new study that approaches body size evolution from a different angle. Although I am a theoretician and study life history evolution, I do not have expertise in IPMs, having never used them myself, but it seems to be an adequate framework to tackle the questions the authors study. These said, I do have some suggestions for improvement and some questions.

Thank you – we appreciate your kind comments and your thoughtful advice. We explain now why IPMs are the ideal framework in which to explore body size evolution.

One of my main problems with this manuscript was that I found it very difficult to identify what the main question was. Is it studying body size evolution under density-dependence? Is it looking at what promotes the evolution of body size extremes? Is it how life history trade-offs can emerge when body sizes evolve under density-dependence? Is it how these trade-offs can affect selection for body size extremes? Is it the demographic patterns that result from these body sizes and life histories?

We have now more clearly stated that our work is motivated by understanding when we expect selection for extremes of body size and life history – gigantism and dwarfism. To do this, we assume a density-dependent framework, and within this framework we pioneer a novel approach. Some interesting results are how trade-offs and disruptive selection emerge as a consequence of the functions used to construct the models. All the questions listed are addressed, but our motivation is to understand when extremes evolve, and to understand that we need to know when small and large body size is selected.

There are some suggestions throughout the manuscript (e.g. lines 59-61, 187-189), but it was very difficult to pinpoint where the main emphasis was. At times, it seemed like the authors were more interested by the methodology and wanted to study what happens when one constructs a size-structured

IPM and includes density-dependence, without any specific question in mind. In sum, while I found the results quite interesting, I think the main question(s) should be clearer and the manuscript can be more focused. As it is, it feels unorganized, and at times, it was difficult to read.

On re-reading we agree that the results and discussion, in particular, were hard going. We have completely rewritten these. We have also modified the introduction to stress that the motivating question is the evolution of extremes. We stress that in doing this we generate new theory, a new approach, and new biological insight. The reviewer is wrong in thinking that we developed a density-dependent IPM to see what might happen. We wish we had time to do that! Instead we were motivated by a suspicion that we could use IPMs to gain insight into the evolution of life history and body size extremes. Our hunch paid off.

I found the emergence of life history trade-offs quite interesting as a result, and it was great to see how these trade-offs can result in different life histories and population structures. However, I wonder whether the fact that density-dependent rates are minimized whereas density-independent ones are maximized is rather trivial.

Again, a minor correction. Specific terms in demographic rate functions are minimised and maximized. In some models, this can result in the rate being minimised or maximized, but if there are multiple terms in the demographic rate functions, the rates themselves may not necessarily be minimised (see the density-dependent juvenile survival scenario). Is this trivial? Perhaps, and once it is pointed out that this happens, it does indeed seem obvious. Yet we are unaware that this point has ever been made before. It is particularly pleasing when a pattern emerges that then appears obvious but has previously been missed. We consider this to be an important finding of our paper and we are delighted that it seems to make so much sense to the reviewer.

Showing that these trade-offs emerge, and that different population dynamics and life histories can result from them is important and very interesting, but I think it is also important to acknowledge that these trade-offs are rather intuitive based on the modelling approach and assumption, particularly because the authors use carrying capacity as fitness. But perhaps I am missing something.

A minor clarification – we did not “use” carrying capacity as fitness. Carrying capacity IS fitness in these models, and that is why we can pioneer our novel approach. This has been proven analytically and via simulation for this class of model and we state this more directly in the revised manuscript with appropriate references in support.

On a similar vein, I found the emphasis on disruptive selection a little too strong. If I understood correctly, at least in the scenario 1, disruptive selection occurs only at a specific part of the parameter space, which coincides with the parameters that the authors use (see Figure 6).

This is not quite right. Disruptive selection occurs across the full range of parameter values we use. It is not specific to the parameter values. Disruptive selection emerges as a consequence of changing development functions and this alters the various age-specific trade-offs between density-dependent and -independent rates. It is the disruptive selection that emerges that enables the evolution of extremes.

It is not a limited part of the parameter space, and the results are nevertheless very interesting, but I think this could be acknowledged more clearly in the manuscript. Now, it looks like disruptive selection is the main result from their model.

In many ways it is the main result as it answers our motivating question. Because we are interested in when extremes emerge, we would expect disruptive selection to be a consequence of altering parameter values. This is a key component of the paper.

They could say that depending on the parameterization of the survival function, one can observe directional selection for large sizes or small sizes, as well as disruptive selection.

Yes, we see selection for small or large sizes depending upon the developmental function, but this is, by definition, disruptive selection. We never observe disruptive selection within a population with a monomorphic life history strategy (see above). Instead, within a monomorphic population we see selection for directional selection for larger body sizes.

The former cases might be too “obvious” to discuss at length and I understand opting for leaving them out and focusing on the parameter space where both extremes can occur. However, I think it is important to acknowledge this choice of focusing on one particular set of parameters, and do it earlier in the results section than at the very end. Also, on a related note, if the idea is to look at when extreme body sizes evolve, showing when there is directional selection for large or small sizes is also an answer the authors’ question, or am I wrong?

We have reworded the manuscript to stress that under some circumstances we see directional selection for large sizes, under other circumstances we see selection for small sizes, and together this generates disruptive selection across the parameter space we explore.

A mix between being very accessible and very technical persists throughout the paper. For instance, I found the Introduction very clear and accessible, until it suddenly became a bit more technical (~line 71). This made me wonder what the target audience of the paper is. Sometimes it read like a paper written for a general (life history) evolution audience, whereas some bits seemed more oriented towards those who are specifically interested in demography and modelling. I wouldn’t necessarily say it is bad to do both in one paper per se, but at times the technical explanations came before those that are less jargon-y, which made the paper difficult to read, at least for myself. Several times I found myself looking at the description of what happens in the model and trying to get a biological intuition, only to realize that it followed soon after (e.g. lines 306-328). I think the manuscript would benefit from rethinking a little bit how to present the model and the results, as to make sure the readers are not stuck trying to think what do increasing rates and derivatives mean biologically.

We found this comment interesting. We are trying to write for a broad audience, but also felt the readership needed to understand the technical aspects of the paper to understand how the results arose. If we were to exclude the technical aspects, the risk is that those who read the paper would not be able to understand why the results arose. So, we believe the paper needs to include both. Nonetheless, the suggestion from the reviewer that we give the biological insights before the technical ones is a good idea, and we have modified the text to do this. We have also moved the discussion of carrying capacity that the reviewer found too technical (see next comment) to the appendix.

A little more specific comment, but since it is related to my confusion about what the paper is about and who the target audience is: the section about carrying capacity at the Discussion also made me wonder what the aim and audience of the paper is. I think it was interesting to read from a methodological point of view, and to see how this approach can be used or adapted to study interactions between conspecifics, interspecific interactions, responses to environmental change in communities, coexistence etc. But I think if the main goal of the paper is studying body size evolution and life history evolution (?), this section derails it from that goal. For instance, I would have liked the next section that is about the empirical

considerations and body size evolution to have a more prominent place in the discussion, possibly with more discussion about how the four functions that the authors used might vary interspecifically and be linked to different sizes we observe in different lineages. There could also be more discussion about how to use existing data or collect data to see how these functions in the nature are. And, perhaps more speculatively, there could also be a more concrete mention of eco-evolutionary dynamics, and how these functions might change when a species responds to environmental change. I particularly liked the discussion of different rules related to body size evolution and the example of sauropods, but I just wished there was more of that; a more prominent and extensive discussion of body size evolution, which to me seems to be the main question of this paper, rather than carrying capacity as fitness, which seems more methodological. And if the paper is indeed a methodological paper inspired by a biological question (instead of vice versa: a novel approach used to study a biological question), then this is not clear and it should be.

Taken with the previous comments, it seems that the reviewer did not clearly follow our approach. This is our fault, not the reviewer's, and we have tried to explain more clearly exactly what was done. We have consequently re-arranged the methods, and completely rewritten the results and discussion. We hope this enables the reviewer to understand our approach.

More specific comments:

I wonder whether the manuscript would benefit from a figure that explains different scenarios and shows the model structure. I found it difficult to constantly remember what was density-dependent and what was body-size dependent, and what was not, in two different scenarios. Not sure if this is a good idea, but maybe even making a big figure showing the model structure, and inserting the density-independent plots in Fig 1 in there, to show how these functions behave.

Figure 1 does show model functions. We now refer to the density-dependent reproduction and density-dependent juvenile survival scenarios throughout.

The authors refer to slow and fast life histories, but I think they don't really define what they are or explain what rates and what values of them are associated with "slowness" or "fastness" of life histories, which I think would be useful in general, but particularly in the context of their model.

These are widely used terms in the life history literature since Steve Stearns' pioneer paper (Oikos 1983) and we do not believe a discussion of their definition is necessary. Obviously, a fast life history will be described using a short life cycle graph parametrized with fast growth, early sexual maturity, high annual fecundity and short lifespan, whereas a slow life history will be described using a long life cycle graph parametrized with slow growth, late sexual maturity, low annual fecundity and long lifespan. This has been repeatedly reported since the eighties (e.g. Gaillard et al. Oikos 1989, Promislow & Harvey 1990, Bielby et al. 2007, Swanson & Dantzer 2014, Bakewell et al. 2020, Del Giudice 2020)

Although it is clear from the context, I think it would be good if the authors clarified earlier on that they talk about "negative density dependence".

An excellent idea – we have made this clearer.

Line 149: This mathematical notation is incorrect. Survival is not equal to beta, but beta is different in juveniles and adults.

Thanks, and well spotted! We have corrected the notation.

I found that the results and methods were written in a way that it is not clear what is a method and what is a result obtained from the model (e.g. lines 206-215).

We have significantly restructured the methods, and rewritten the results. We have been careful to ensure the methods and results are now appropriately separated.

Line 271: typo, “adult reproduction”?

Thank you! Corrected.

Line 276: Figure ref for adult reproductive rate?

The results have been rewritten and the offending sentence removed.

In scenario 2, larger body sizes and slower life histories performed much better compared to smaller body sizes and fast life histories, compared to scenario 1. Could this be discussed, and overall, would it make sense to compare these scenarios a little more?

We have now reworked the paper, following the comments from the referee, to focus on when selection on body size extremes occurs. Discussing nuances such as this would likely detract from our main message.

Line 306: It should be made clearer in the text that this is the point at which fitness is minimum, just like in the figure caption.

We have rewritten the results and make this point more clearly.

Lines 357-370: A very lengthy discussion and presentation of Fig 6 might not be essential for the manuscript text, which is already complicated and long, especially since these parameters were not explored. I would suggest acknowledging the variety of results that can be obtained based on different parameterization of the survival function more clearly in the text (as I suggested above), but moving the non-essential bits to the supplementary material.

In rewriting the results, we have made it clearer why we conducted this perturbation analyses. The reworked paper is now 25% shorter than before, so we have kept this text in.

Line 361: intercept instead of (or in addition to) elevation?

Elevation is the appropriate word because it captures both the intercept and the effect of other terms in the function.

Line 388: Again here, for instance, it seems to me that the authors completely ignore the fact that their models can also be used to show directional selection for either extreme in size.

Please see our response to the same point made above.

Line 391-393: Where is this shown?

This is shown in Figure 6 (now figure 5).

Simulation code:

I had a brief look at the code used in the model. The comments in the scripts were quite helpful, but it would be really nice if there was also more explanation on how to use the code, if possible (e.g. a “readme” text explaining which files lead to which figures).

A very good idea! Thanks.

Figures and tables:

Figure 1: Figure 1 was not very easy to understand during a first read. It is more accessible after reading through the manuscript once and after having seen the other figures. For instance, it is not necessarily clear what fitness is (it becomes clearer later) and the reader could be reminded in the caption. In A), what are the bars and are they necessary? In C), would it be possible to place the dots elsewhere, e.g. at the end of the curves, instead of on the part of the curves where life histories start to differ from each other in terms of growth?

Instead of altering the figure, we have extended the figure caption. It now reads “Figure 1. Density-independent functions in the density-dependent reproduction scenario. (A) association between parental size at time  $t$  and offspring size at time  $t+1$ , where the black lines are contours representing the probability distribution of offspring sizes, (B) development functions, (C) monomolecular growth functions showing size as a function of age, (D) mean of the body size-survival function. Unlike in A) we do not show the contours of the probability distributions. Each line represents one of the 20 life histories, and each dot represents size (or age in the case of C) at sexual maturity. The redder the colour of a point, the fitter (greater carrying capacity) the life history strategy.”

Figure 2: I found Figure 2C and D really complicated to unpack; there are two axes, showing three and two different things for scenarios 1 and 2, respectively, two colours for size-distribution of each life history, which are then overlaid creating even more colours. In C, dark colours are reproduction rate, whereas in D they are survival. In the size distribution, adults and juveniles are separated by a vertical line and different colours, whereas for the rates shown with dots, one needs to infer that they are mean rates for juveniles and adults by the positioning of the dots along the x-axis, which is not very consistent. Also, regarding the points showing density-dependent rates (figure caption); aren't all points solid? Do the authors mean lighter vs darker colours? Also the panels are so close that right y-axis label of C and left y-axis label of D are merged to become essentially one label. And what is the dashed line on panel C? The survival function from 1D? It should be explained. What do the lines that connect darker dots represent in panels C and D? Overall, I was very confused when I saw this figure for the first time, and I think this figure and its caption needs some work to make it easier for the readers to understand it.

We have removed the figures. They clearly were rather hard to follow.

Figure 3: Should have a figure main title and say this is scenario 1 in the caption.

The caption now reads: “Figure 3. Model dynamics for the density-dependent reproduction scenario. (A) survivorship functions for each life history, (B) survivorship to sexual maturity as a function of size at sexual maturity, (C) life expectancy as a function of body size, (D) proportion of population that is sexually mature as a function of size at sexual maturity, and (E) per-time step per-capita reproductive rate as a function of size at sexual maturity, and (F) trade-off between the log of the density-independent rates with the log of the per-capita per-time step reproductive rate (the arrow represents the direction of evolution). The x-axis label is the combination of the density-independent life history traits. The dotted

green vertical lines in (B-E) represent the life history of minimum fitness. Fitness increases either side of the green line. Each point represents one of our 20 life histories. The redder the colour of a point, the fitter the life history strategy.”

Figure 4: What is  $a_s$ ? Is it age at maturity, and if so, why is it not  $a_m$  as before?

Well spotted! This was a typo and has been corrected.

What the polygons are is not explained clearly. Which axis shows which curve is not very clear. In general, I find figures with two y-axes very complicated and would avoid them if possible. I see why they are useful in this case, but everything should be very clearly explained.

We have extended the figure legend to read: “Figure 4. Summary of the age-structured life history dynamics of the model where reproduction is density-dependent. Blue lines represent the survivorship schedules (left y-axis) and red lines the fertility schedules (right y-axis). The polygons represent the product of the survivorship and fertility schedules that are described with the equations. The initial life history strategy is depicted in (A), the mutant strategy, with a delayed age at sexual maturity, in (B). The delay in age at sexual maturity results in a change in the development function that results in an elevation of the survivorship function (compare the solid blue line in (B) to the solid blue line in (A) which is also represented by the dotted blue line in (B)). Because the volume of the blue polygon in (A) and (B) must equal 1 (equations on plot), the reproduction function is depressed in (B) compared to (A) (compare the solid red lines in (A) and (B)). The grey and hashed blue polygons in (B) show how the polygon has changed shape between the two life histories. (C) Rectangular approximation of the life history function used to identify trade-offs. See Fig S3 for the rectangular approximation for each of our life histories.”

Figure 6: Maybe I am missing something, but why do blue lines end in the middle of x-axis? Does the carrying capacity stay the same after that point or does values were not looked at? If they were not looked at, why are they on the plots?

We have altered the legend to address these concerns. It is also now figure 5. It states: “Figure 5. Dynamical consequences of altering the intercept and slope of the body size-survival function in the density-dependent reproduction scenario. As the elevation of the intercept (rows and green numbers) and steepness of slope (columns and purple numbers) are altered, the change in the size-survival function alters selection on size at sexual maturity. The red lines represent the form of the size-survival function. The blue lines show how carrying capacity changes across the range of sizes of sexual maturity. The blue lines are analogous to Figure 2 for each parameterisation of the size-survival function but are drawn as continuous lines rather than as dots.”

Figure S1: The caption lacks what the green dot represents (the strategy with minimum fitness I assume)

This figure has been removed.

Table 2: It would be helpful to have the symbols of what these parameters are. In this table, do growth parameters represent 20 different life history scenarios? I am a little confused. And also, of the survival parameters, which ones are for juveniles and which are for adults? To my understanding, in scenario 2, only juvenile survival is density-dependent, whereas both juvenile and adult survival are a function of body size. I expect these to be reflected by non-zero body size and density slopes, but then this means, density slope should be zero for adults and non-zero for juveniles, right? Or am I missing something?

These are excellent suggestions. We have modified the table accordingly. It should now be clearer.

*Reviewed by Frédéric Guillaume, 21 Apr 2022 12:04*

This manuscript addresses the key question of how body size co-evolves with pace of life when some demographic rates are density-dependent. They show how density-dependence generates disruptive selection on body size by maximizing carrying capacity at equilibrium. The outcome of the model is thus the evolution of extreme body sizes and life-histories: small-fast or large-slow. One key interesting aspect of the model is the non-imposition of a trade-off between demographic rates. Instead, a linear trade-off between reproduction and offspring survivorship emerges from the model dynamics. The approach used is a size-structured model called IPM. The topic is complex and so is the methodological approach. I found the model and results descriptions hard to follow because rather abstract and technical. Certain aspects of the model must be clarified. The manuscript discusses in length some of the key model assumptions, namely fitness defined as carrying capacity and addresses theoretical and empirical implications.

Thank you – the description of what we did is accurate. The topic is indeed complex, and we have attempted to make the text more accessible. We hope we have succeeded.

The manuscript will gain by being shortened and streamlined, especially in the Results. Authors should strive to provide more biological intuitive understanding of the outcomes, especially when describing figures in the Results. My general feeling was that it is addressed to IPM specialists more than to a general audience.

On re-reading the manuscript we agree that the results are hard going. They have been completely rewritten.

Authors should improve the description of the simulation approach implemented, in relation to IPMs. In particular, it is unclear how simulations help in computing a life-history strategy's carrying capacity and whether any evolutionary dynamics are involved in the simulations. No description of the simulation procedure is provided. This should be improved.

We have added in a section describing how we conducted the simulations.

One key aspect of the approach not well delineated is whether a polymorphic population (ie containing multiple competing strategies) would actually evolve towards the strategy(ies) having the highest  $K$  identified in the IPM analysis and whether they may coexist.

Yes, it would, as long as competition between the strategies is symmetric. In other words, strategy carrying capacity will determine fitness if one individual of strategy A imposes a competitive pressure on an individual of strategy B that is equal to the competitive pressure that one individual of strategy B imposes on one individual of strategy A. We have previously demonstrated this in Kentie *et al.* 2020.

The manuscript misses a link between long term evolution as predicted from  $K$ -maximization principle à la Lande *et al.* 2009 and per-generation rate of change in average population trait values (eg. body size) provided by a selection gradient. Authors should clarify how such selection gradients, based on a  $K$ -definition of fitness can or cannot be derived. It would help link with more classical neo-Darwinian thinking about evolutionary dynamics.

Please see the comment above. This is a real can of worms! Selection differentials (and their scaled cousins, the selection gradients) are useful tools for describing changes in allele frequencies attributable to selection, or to phenotypic traits that have a simple, additive, genotype-phenotype map. Selection

differentials and gradients are less useful for describing change in characters that are the result of development trajectories. In other words, we could calculate selection differentials on the genetic variation that leads to variation in life history strategies, and they would explain evolution accurately. But the estimation of selection gradients or differentials on body size in such a population will not accurately predict evolution. This point has been made before in Childs *et al.* 2011.

In addition, it is assumed that carrying capacity is fitness and thus maximized by evolution but it is not demonstrated that it is indeed the case in the present model. A derivation of a selection gradient emerging from the model definition might clarify this point.

It is not an assumption. It has been previously proved analytically by Lande *et al.* as cited in the text and demonstrated through numerical simulation for this class of model by Kentie *et al.* We cite the relevant literature. Sadly, the derivation of a selection gradient is not possible for the reasons described above.

Discussion on carrying capacity as fitness is great but hard to follow. It would gain by being reduced, for instance by focusing on inter-specific interactions. As of now, a large part of that discussion is disconnected from the main subject of the paper. Moreover, a good part of the discussion consists in a verbal description of mathematical models from the literature, which is hard to follow without knowing them. Limiting the discussion to more intuitive arguments would ease the reading by a fair amount.

We have significantly simplified this section and have completely rewritten the discussion. A discussion of interpreting carrying capacity as fitness is now provided in the form of an appendix. We hope the reviewer likes the new version.

One key assumption leaves me unsure as whether the model is correct. On L167 (and Figure 1A) the distribution of the offspring size is independent of the adult phenotype. It is fine when considering the variance of the distribution but not when considering the mean. As stated, there is not inheritance of the parental traits in the offspring if the offspring size is invariant as shown in Figure 1A. Is this really the case? ( $h^2 = 0$ )? please clarify, provide a justification and explain how the model outcomes depend on that assumption. I would very much doubt that body size evolves in the model if offspring do not inherit their parental size.

An interesting comment. For body size to have a non-zero heritability *within* an IPM, the distribution of offspring body size needs to positively correlate with parental size. However, we are interested in evolution of the entire life history strategy – all the parameters used to define the life history strategy. The whole strategy – i.e. the parameter set – can be thought of as having a heritability of 1. This is essentially what is done in adaptive dynamic analyses. However, we do not need to conduct the invasion analyses because we know that carrying capacity is fitness for these models. The model is correct. For discussion on the heritabilities and IPMs please see Coulson *et al.* 2010, Janiero *et al.* 2017, Plard *et al.* 2021. To clarify this point in the manuscript, we have added the following sentence: “Assuming that body size is not heritable within an IPM is possible because, in these models, the entire life history strategy, i.e. the whole parameter set, can be thought of as having a heritability of 1.”

L89: clarify that the demonstration of "fitness is carrying capacity" is provided in Lande *et al.* 2009, but not in the other references.

It is provided in Kentie *et al.* via simulation. Please see SI of this paper.

L96/111: provide clarification and a definition of "asymptotic representation" and "asymptotic size".

We have replaced “asymptotic size” with “maximum body size” (introduction) and “asymptotic (i.e. maximum) body size” (methods).

L98-105: clarify if a demonstration exists or is this only a verbal argument? also clarify if it applies only to a monomorphic starting population, looking at invasion fitness or also applies to polymorphic populations?

This has been moved to an appendix. A demonstration has been provided in Kentie et al. 2020 and it is general to any case and any polymorphic starting populations. The key point is we can always avoid using invasion analyses if we can define fitness.

L219-220: please provide necessary details about calculations from model predictions.

Because these are standard calculations, and have been described in Coulson et al. 2010, we do not repeat them here as we have also been asked to reduce the length of the manuscript.

L325: "ever-earlier size" -> ever-smaller size?

In reworking the manuscript we removed the offending sentence.

L369: sentence needs correction. The reasons for this comment on linearization are unclear. Please provide more details on why it is necessary.

Reworded to: “Finally, to demonstrate that our results are not due to non-linearities in our model, we linearly approximated the model and explored outputs (Appendix). This revealed that the patterns we report are not a consequence of the linearities in our model functions.”

L411-413: It is unclear where such negative correlations can be directly observed. Please clarify.

In reworking the discussion we have removed the offending sentence.

table 2: please add symbols to relate to model definition (which are the rho's etc.)

Done.

*Reviewed by anonymous reviewer, 11 May 2022 13:03*

The authors investigate one of the problems of life-history evolution, the evolution of extreme body sizes. For this, the authors develop a size-structured integral projection model. Within this approach, the body size is associated with other life-history traits, such as survival, development, reproductive rate, and heredity of body size trait. What I find interesting in the paper is using a carrying capacity as a proxy for the fitness of a particular life-history strategy. Because the used model is density-dependent, the authors split the analysis into two scenarios. The first one, where density-dependence works on reproduction, and the second, where juvenile survival was density-dependant.

Carrying capacity is not a proxy for fitness, it is fitness in this scenario. This is a really important point we have stressed more clearly in the revised manuscript.

As a result, the authors showed conditions under which either body-size extremes evolve (small-bodied or large-bodied). For example, fast life-history evolves when delayed age at maturity leads to increased size at maturity, as well as an elevated mortality rate. In general, this paper adds an understanding of the conditions why a particular body size is selected. There are some assumptions and simplifications made, that the authors state clearly. At the same time, there are several comments I believe would help to improve the manuscript and make it easier to understand.

Thank you!

The authors can see my comments in the PDF file attached

comments

1. Consistency in notation of parameters. There is some inconsistency in using letters with subscripts throughout the manuscript. For example, on page 6 there are parameters with subscripts in italics ( $a_N$ ). The same parameter has no italics in the Table 1. Usually, it is up to the authors which symbols and letters to use for model parameters and variables. I would use no-italics subscripts everywhere and left italics to denote a parameter/variable that has a particular numerical value. However, as I said, it is up to the authors. The only think I would ask is to be consistent and stick to one chosen strategy.

Thank you. We have kept the italics, but we have checked for consistency. The mismatch between the Table and text was due to one author writing the text in LaTeX and another preparing the Table in Word. The conversion between Word and Latex using Pandoc hadn't worked particularly well. The table has now been written in LaTeX.

2. Recruitment vs Reproduction. Through the entire manuscript, the authors use the term "reproduction". However, in the Table 1, there is another term used – "recruitment" – describing the same notations ( $r$  and  $R$ ). I believe that both terms might describe similar processes, although not the same. For me, "reproduction rate" could be either number of offspring produced or number of copulations (both per-capita). "Recruitment" would describe number of juveniles reaching reproductive state. While there might be some discussion what either term means, I would recommend to use only one of them, stating clearly in the text what the authors want to say with it. Using both terms confused me a little when I was reading the paper and referring to the Table with the list of parameters.

Thank you. We now use reproduction throughout. We do sometimes use reproductive rate where appropriate – i.e. we use reproduction to describe the function and reproductive rate to describe the rate it predicts.

Minor comments

- Line 34. It feels like a citation is needed after "There are clearly physiological limits..."

Added.

- Line 35. "... but the factors that push organisms..." Did you mean evolutionary factors here? Or factors, acting on the evolutionary timescale?

Replaced with 'selective forces'.

- Line 39. “In contrast, as size increases, the values ...” It is a nice statement. However, it might be good to provide a citation here, because it is not entirely clear whether you still refer to sources mentioned earlier in the text.

It was the same sources as before, but we have cited them again for clarity.

- Lines 56-59. You listed here a set of factors able to generate trade-offs. Can we consider here time as a limiting factor? Individuals may have enough resources but limited time to gain, incorporate or use resources.

Yes, time is definitely limiting, and in particular time required to develop. We have not added it in the list, however, as it is not intended to be exhaustive.

- Line 59. “Trade-offs consequently arise...”. Here I would mention that trade-offs arise on an individual level (a particular individual meets resource/time/space constrains), that result in population growth limits

We now state this explicitly.

- Line 89. “...fitness is carrying capacity...”. This is, probably, the most subjective comment. In most cases, I prefer using the phrase “X is a proxy for fitness”. Considering the variety of modelling approaches, there are several ways how one can define fitness. But in all the approaches, there is always a proxy for fitness.

This is an interesting comment, and one we not completely agree with, although we do see where the referee is coming from. In these models carrying capacity is fitness in the sense that it perfectly predicts evolutionary outcomes. We are concerned that if we state it is a proxy for fitness then readers won't realise that it does this. They may think we have arbitrarily defined fitness to be carrying capacity, and that is not the case.

- Line 101. “... would be less than unity.” Unity = one?

Yes. We have replaced unity with one throughout.

- Line 106. “Since we know what...” I would recall here carrying capacity.

The offending sentence has been removed as we have reworked the manuscript.

- Lines 112-124. This part seems somehow not in the right place. You describe well two scenarios how density-dependence acts later in the Methods. Also, you explain nice results for each of the scenarios. With this, your discussion and conclusion look logic. In contrast, mentioned part here is perceived a bit dry. I would either add some examples (for instance, after the third paragraph), or move this part of the text entirely at the end of the manuscript.

We have reworded the text here. We acknowledge this is a matter of style, and there is no right or wrong solution. The lead author likes to state what the paper reports at this point in the paper, even if not all his co-authors agree.

- Line 132. Extra closing bracket in the denominator of the equation.

Corrected.

- Line 145. One can deduce what each capital letter (but S) from the equation means. However, recalling meaning of letters would increase understanding.

Good idea – we have explained the meaning of the emboldened letters.

- Lines 147-150. This is one of the most crucial parts of the Methods, where you describe two scenarios. It will not hurt if you write here sentences in a more pro-active way. For example, one could phrase “In the first scenario, density dependence acts on reproduction, limiting number of offspring produced. We modelled this assuming  $r < 0$ . In the second scenario, population size is controlled through juvenile survival...”. By no means I want to force you rewriting this part of the methods. I simply believe that pro-active phrasing makes the text much more attractive, especially when you describe your model assumptions.

We have made the requested change.

- Line 151. Unity=one?

Altered.

- Line 156-159. Just to clarify. Here you mentioned 20 models for each of the scenarios. Models represent different life histories. Later, you say that all but two functions (development and size at maturity) are identical. Does this mean that difference in life histories is based on these two functions?

Correct – and we have modified this text to make it explicit.

- Lines 197-199. There are several terms provided (i.e.  $rNK$ ). It will increase understanding if you could provide or recall a biological meaning of all these terms.

We have modified this sentence in an attempt to make it clearer.

- Line 254. Can you recall briefly, what the equation here stands for?

We have completely rewritten and simplified this section of the discussion, and the offending sentence has been removed.

- Line 286. “... was highly non-linear...”. Can non-linearity be of a particular level? Is it measured somehow?

We have removed the offending sentence in the rewrite.

- Line 306. It might be helpful to explain briefly what the threshold means. You refer to the figure, where one can understanding what a threshold is. But, the description is lacking here.

We have explained this is the threshold of minimum fitness.

• Line 323. I am confused with phrasing “Below the threshold of minimum fitness...”. Does it mean that every life-history strategy that is below this threshold is not selected at all? It would be great if you could clarify here. The same comment relates to Line 306

We have modified the text to read: “In both the density-dependent reproduction and density-dependent juvenile survival scenarios we observe disruptive selection on body size (Figure 2(A,B)). Below a threshold size at sexual maturity where lowest carrying capacity is observed there is directional selection for small size at sexual maturity and a fast life history. Above the threshold, evolution of gigantism is observed. Why do we observe these patterns?”

• Line 348. What is the biological meaning of these equations?

We have attempted to explain earlier in the text that it is the quantity that evolution acts to minimize.

Table 1

1. It might be useful for readers to have some information about ranges of parameters and variables (where possible) used in the model. This can be done by adding an additional column. There is table 2 with (presumably) values for most of the parameters. If so, it might be good to refer table 2 here

In Table 1’s caption we now refer the reader to table 2 for parameter values rather than complicating table 1 further.

2. Recruitment vs Reproduction comment (see above). There are recruitment terms in the table, while reproduction in the text

Altered – thanks.

Table 2

1. I would add a few more lines describing this table. For example, mention that all these parameter values are the same across all the models you constructed (according to what you wrote in lines 156-157). In addition, it would help readers to see connection if you provide parameter notations from Table 1.

We have altered the table significantly. Now each parameter is referred to with its appropriate notation, and the two scenarios are clearly described, as are the parameters that vary across life histories within the scenarios.

2. There are no rows indicating values for the development function and the size at sexual maturity, which are unique for every model (according to lines 156-157)

Correct. Size at sexual maturity is an emergent property of the growth function, being 80% of asymptotic size. We do not wish to further complicate the table, so have not added in any further information to this table.

3. What the empty space and tree dots under “Growth parameter” stand for?

Hopefully this is now clear. They represent life histories 4 through 19 of the 20 we use.

## Figure 1

1. Fig1A – What do black lines mean? If they are necessary and bring some useful information, can you describe them?

They are the contours of the probability distribution of offspring size. This is now explained in the legend.

## Figure 2

1- “Vertical lines separate juveniles from adults.” – I would assume that juveniles are on the left side of the graphs (C) and (D). However, there is inconsistency in color use for slow life history strategy (shades of red). On the (C) graph yellowish color indicates (supposedly) adults (right side from the vertical line). While on the (D) graph the same shade indicates the left part of the area under the curve.

We have removed figures (C) and (D) as they were too hard to interpret.

2- (D). “However, because juvenile and adult size...”. I am confused by this sentence. I would assume you meant that survival rates are higher than for the faster life history, considering what we can observe on the Fig 2D. But I may be wrong. Can you clarify this part?

We have removed figures (C) and (D) as they were too hard to interpret.

3- Survival rates are indicated with pastel dots on Fig 2C, while with bright dots on Fig 2D. It might be good to be consistent how you picture the same parameter.

We have removed figures (C) and (D) as they were too hard to interpret.

## Figure 3

1. Vertical green line. From these plots and your description, it is clear that this line shows indeed the least fit life history. However, from the text (see my comments above – line 306 and line 323), one could assume that this is a threshold separating the fittest life histories from less fit.

We have modified the legend to explain that fitness increases either side of the green line.

2. What does the x axis label at Fig 3F mean? Could you add a brief explanation in the figure capture?

Done.

## Figure 4

1. Fig 4A and B. It is a nice idea to add a dashed line to Fig 4B for better comparison. In such a case, I would not mention Fig 4A at all, because it might be hard to compare lines on separate plots. Even if there is some difference, it can be unnoticeable.

We don't understand this comment. Sorry!

## Figure S1

We have removed this figure from the manuscript.

1. From the text, one can get that these plots depicts scenario 1 (am I right?). But there is no notation in the figure capture.
2. What does a green dot indicate?
3. I am confused with the description of “fast vs slow” life histories. In the text, you wrote about the positive covariance between parameters determining the evolution of fast life histories (lines 279-280). Do these histories evolve from green dots? Evolving from green dots, are these histories stay “fast-life”? Based on your figure legend, one may assume that not – “Blue dots represent fast life histories, red dots slower ones”.

Figure S2

The figure has been removed from the manuscript.

1. Do you need graphs E and H? I would say that they are rather not necessary.
2. Same comment on the green dots and mixing colors to picture both “slowvs-fast” life histories and fitness of strategies

#### References

- Bakewell, A. T., Davis, K. E., Freckleton, R. P., Isaac, N. J., & Mayhew, P. J. (2020). Comparing life histories across taxonomic groups in multiple dimensions: how mammal-like are insects?. *The American Naturalist*, 195(1), 70-81.
- Bielby, J., Mace, G. M., Bininda-Emonds, O. R., Cardillo, M., Gittleman, J. L., Jones, K. E., ... & Purvis, A. (2007). The fast-slow continuum in mammalian life history: an empirical reevaluation. *The American Naturalist*, 169(6), 748-757.
- Childs, D. Z., et al. "Predicting trait values and measuring selection in complex life histories: reproductive allocation decisions in Soay sheep." *Ecology Letters* 14.10 (2011): 985-992.
- Childs, Dylan Z., Ben C. Sheldon, and Mark Rees. "The evolution of labile traits in sex-and age-structured populations." *Journal of Animal Ecology* 85.2 (2016): 329-342.
- Coulson, Tim, Shripad Tuljapurkar, and Dylan Z. Childs. "Using evolutionary demography to link life history theory, quantitative genetics and population ecology." *Journal of Animal Ecology* 79.6 (2010): 1226-1240.
- Del Giudice, M. (2020). Rethinking the fast-slow continuum of individual differences. *Evolution and Human Behavior*, 41(6), 536-549.
- Gaillard, J. M., Pontier, D., Allaine, D., Lebreton, J. D., Trouvilliez, J., & Clobert, J. (1989). An analysis of demographic tactics in birds and mammals. *Oikos*, 59-76.
- Janeiro, Maria Joao, et al. "Towards robust evolutionary inference with integral projection models." *Journal of Evolutionary Biology* 30.2 (2017): 270-288.

Kentie, Rosemarie, et al. "Life-history strategy varies with the strength of competition in a food-limited ungulate population." *Ecology Letters* 23.5 (2020): 811-820.

Plard, Floriane, et al. "Demographic determinants of the phenotypic mother–offspring correlation." *Ecological Monographs* 91.4 (2021): e01479.

Promislow, D. E., & Harvey, P. H. (1990). Living fast and dying young: A comparative analysis of life-history variation among mammals. *Journal of Zoology*, 220(3), 417-437.

Snyder, Robin E., and Stephen P. Ellner. "Pluck or luck: does trait variation or chance drive variation in lifetime reproductive success?." *The American Naturalist* 191.4 (2018): E90-E107.

Stearns, S. C. (1983). "The influence of size and phylogeny on patterns of covariation among life-history traits in the mammals". *Oikos*, 173-187.

Steiner, Ulrich K., Shripad Tuljapurkar, and Steven Hecht Orzack. "Dynamic heterogeneity and life history variability in the kittiwake." *Journal of Animal Ecology* 79.2 (2010): 436-444.

Swanson, E. M., & Dantzer, B. (2014). Insulin-like growth factor-1 is associated with life-history variation across Mammalia. *Proceedings of the Royal Society B: Biological Sciences*, 281(1782), 20132458.

Tuljapurkar, Shripad, et al. "Distributions of LRS in varying environments." *Ecology letters* 24.7 (2021): 1328-1340.

Tuljapurkar, Shripad, et al. "Skewed distributions of lifetime reproductive success: beyond mean and variance." *Ecology letters* 23.4 (2020): 748-756.