This is the second time I review this paper, and I must say the revised version is a significant improvement over the previous one. The methodology is now much clearer to me, and I truly believe this is a valuable paper that deserves to be published. However, I feel that further efforts to enhance clarity are essential. I still found myself spending an excessive amount of time trying to fully grasp the methods. This lack of clarity risks reducing the paper's impact, as readers unfamiliar with this type of model—like myself—may struggle to understand it. Many might simply give up, which would be unfortunate given the substantial effort and valuable insights the study provides.

To address this, I have provided detailed comments on sections I found unclear and suggested ways to enhance readability and accessibility. I apologize for the length of the review; my intention was to support the authors in maximizing the impact of their work by making it more approachable for a broader audience.

Specific comments

In the abstract, what do the authors mean by "adaptation measures"? Do they refer to actions aimed at enhancing adaptation to climate change? This point is not clear to me.

In the same vein, I find the use of "adaptation solutions" at L7 somewhat confusing. The term "adaptation" has different meanings depending on the field. In social science, it refers to human strategies and actions for adapting to climate change impacts. In evolutionary ecology, it describes the process by which populations evolve heritable traits through natural selection, increasing their fitness in response to environmental pressures. In the manuscript, the second definition is predominantly used (e.g., at lines 108 and 57). However, I believe this definition does not align with the meaning intended in the sentence at L7-8. Specifically, I do not see how promoting forest structure complexity can facilitate adaptation in the evolutionary sense, and the references cited do not appear to support this interpretation. It seems that "adaptation solutions" aligns more closely with the social science definition of adaptation. I recommend that the authors adopt a consistent definition of adaptation throughout the manuscript to avoid confusion. At L7, simply referring to "emerging solutions" may suffice, as adding "adaptation" here seems unnecessary and potentially misleading.

This comment also applies to the use of adaptation at L63.

L17. These are not trees that adapt to climate change, but rather tree populations (individual trees can only acclimate to climatic conditions over their lifetimes).

L18. I find the statement that questions about tree adaptation to climate change are focused on the potential for evolutionary rescue to be overly restrictive. Evolutionary rescue refers to a specific scenario where populations persist through rapid adaptation to severe environmental changes that would otherwise lead to extinction. However, in forest trees, much of the research focuses on populations that are not at risk of extinction but instead experience reduced productivity or increased mortality. These challenges, while significant, do not necessarily threaten the persistence of the population.

L19. I am sorry, but this is the same comment as in my previous review: it is important to provide the definitions of persistence and recovery. General definitions may be provided in the introduction. Then, in the M&M, I think it is important to specify the metrics and thresholds used to determine whether a population has persisted under climate change or recovered after climate change.

L23. “evolution and species diversity can interact in complex ways” Could the authors provide references to support this claim or include a specific example to illustrate the types of complex interactions they are referring to?

L53. I think it should be “heritable” instead of “hereditary”. Moreover, I think this sentence is somewhat reductive, as most traits have both a genetic (i.e., they are heritable) and an environmental (plastic) component. This aligns with the editor's comment that Sigma_P is not entirely defined by Sigma_G.

L78. Same comment as above, shouldn't be “heritability” instead of “heredity”?

L90. It seems a bit awkward to say that the site is "subject to RCP," as RCPs are future climate forecasts, not actual impacts. I would recommend rephrasing the sentence to reflect that the future climate at this site was projected based on RCP 4.5. Also, please write out "RCP" in plain text ("Representative Concentration Pathways") the first time it is mentioned.

L100: Why use “directional” instead of “assisted”? This is the only instance of “directional” in the manuscript, and I find it confusing, as I would say that “directional gene flow” typically refers to gene flow between nearby populations without human intervention. To maintain consistency and avoid confusion, I suggest using “assisted” here as well, in line with the terminology used throughout the manuscript.

L113. I believe the structure of the methods section could be improved, as some information related to the same parameters/processes is repeated in different parts. For example, information about the three evolving traits is given both in L136-147 and in the “Evolution module” section. I suggest merging this information into a single section (preferably the “Evolution module” section) to make it easier for readers to follow. Another example is the explanation of the slow growth state, which is first introduced at L124 and then explained in more detail at L240. A final example is the explanation of the regeneration process. I think the description of how the parents of the locally regenerated seedlings are chosen (L237-240) should be included alongside the regeneration information in the section about regeneration. So I recommend that the authors reconsider the structure of the methods section, as there is still room for improvement.

L129-130. “The sum of several patches simulates a forest” – However, in Figure 1, the sum of several patches appears to be referred to as a scene. Are these the same, or are they different? Please ensure consistent terminology to avoid confusion.

L140-141. “In order to be comprehensive with our investigations” Is this part really useful? I am not sure I understand what it means.

L147-149. “To facilitate further analysis, basal area (BA) (m²/ha) and its dynamic through time are output to describe the ecological state of the simulated forests.” Is it the total basal area for each species in each patch? Was it extracted based on an annual time step?

L150-151. Similar comment – Does “changes in the average values” refer to averages across patches? At which time step are these values extracted? I think this part could be clearer.

L152-175. I think the section on regeneration could still be clearer. If I understand correctly, the authors chose to adopt an approach previously used by other models that account for evolutionary processes (the first approach described in the paragraph). However, how is the number of new seedlings defined? And how is the mortality filter applied?

“*We incorporated a feedback mechanism from the current composition of the forest (based on the relative abundance of species in relation to their BA across all patches) on the identity of colonizing seedlings each year."* Does this essentially mean that the proportion of each species in the seedling pool was determined based on the proportion of the basal area of each species in the plot?

“Among the previous examples of models that account for evolutionary processes, most have adopted the first approach. However, the second approach renders regeneration independent of the current state of the simulated forest.” These two sentences are a bit confusing I think. I would suggest (if I have understood correctly): *“Most models that account for evolutionary processes have adopted the first approach because the second approach renders regeneration independent of the current state of the simulated forest and therefore does not allow evolution to occur.”*

L183-191. Would it be possible to briefly explain how these traits are obtained and the range of expected values? For example, what would be the expected values for a drought-tolerant species compared to a drought-intolerant one? Including the equation in the supplementary information or referencing relevant papers could help. This clarification is important because, as presented in Table 1, it is not possible to determine whether species are drought-tolerant, shade-tolerant, or exhibit high growth values without an idea of the range for these traits.

L185. The authors explained in their response to the reviewers: *“A new tree "appears" in the simulation at year n with a height greater than 1.3 meters, and is no longer considered a seedling at the year n+1. However, a tree can only reproduce once it reaches its species-specific age of maturity L185.”*

If I understand correctly, seedlings are simulated in each patch at a given year *n* based on local (and potentially migration) regeneration. These seedlings are directly assigned a height greater than 1.3m and, by the following year, are no longer considered seedlings. Am I correct?

I think this information should be included in the main text.

Then, it follows that seedling parameters are applied only for one year per individual. What is the justification for this? Is this step primarily aimed at removing seedlings that are not light-tolerant enough?

L197. I think the clarity of this section can still be improved.

L202. In which cases the initial state of a population is derived from previous simulations?

L212. I find the definition of VR unclear. When defining a fraction, it has to be specified what is being divided by what. In this case, I suggest clarifying that VR represents the fraction of intraspecific trait variability relative to interspecific trait variability observed in the species pools parameterized in ForCEEPS.
Relative to this, VR is referred to as “Trait variability” in some parts of the manuscript (e.g., L530 and in the figures), which I believe is not accurate.

L215. “This calculation requires us to assume that ΣG is proportional to ΣP which remain a major oversimplification of reality.”

First, “remain” should be written as “remains”.

Second, ΣE is also assumed to be proportional to ΣP, correct? This should be mentioned as well.

Third, “This calculation” refers to the preceding equation right? ΣG = h2 · ΣP (please number all equations). Then, I think this sentence should be included before the sentence “ΣP values depend on the initial intraspecific trait variability ratio (VR) [...] (Morin et al., 2021)” (which describes the following equation, equation 2).

L218. I would suggest mentioning that h² and VR vary between 0.1 and 0.4 depending on the simulations, and providing some references to support the biological realism of these values.

L220-221. ΣE and ΣP remain fixed across the simulations if I understand well. But what about ΣG? If ΣG changes, how? If ΣG changes, how does it change? And if it does not change, then h² cannot change either, as h² = ΣG / ΣP. I think it would be helpful to specify more clearly which parameters change across the simulations and, if so, how they change. I find it very hard to understand.

This comment follows the one of Reviewer 3 (“As h^2 and sigma_E are fixed, how do you control for sigma_G after local and migration regeneration?”), to which the authors' response is incomplete in my opinion. I could not understand from their reply how ΣG and h² vary throughout the simulation.

L226. Does this mean that the trait correlation matrix is assumed to be the same at both the intraspecific and interspecific levels? If so, I think this should be stated more clearly.

L234. Why are some equations numbered and others not? I would suggest numbering all equations. Regarding the equation below L234, I think the last component is $ShTol_{s}$, not $ShTol_{i,s}$, if I understood correctly.

L267. “ΣE variances are fixed” Why are only the variances mentioned? Are the covariances not also fixed?

L271. What do the authors mean by “central”? Does it mean that the site is located in the core of the beech and fir forest ecosystems?

L280-285. This section is somewhat confusing because the main text states that patches evolve over 2100 years, but in the figures, the timeline starts at year 1000 and ends at year 3000. Could this discrepancy be due to baseline simulations conducted prior to the evolutionary scenarios? If so, this should be explicitly clarified in the text.

For example, at L358: “ we computed the difference between the $\overline{DrTol_{s,p,3000}}$ and the

average $\overline{DrTol_{s,p,y}}$ over a reference historical period (year 1500 to 2000).” Does it mean that $\overline{DrTol_{s,p,3000}}$ is the final average trait value of the simulation? Additionally, isn’t the historical period supposed to be from 1900 to 2000 instead of 1500 to 2000?

L289. I would recommend explicitly clarifying what a replicate represents: is it 50 patches? If so, in Figure 2A, it would be helpful to clarify that the median basal area of the y-axis corresponds to the median basal area among the 50 patches.

L309. I do not understand the goal of using different VR values in the simulations. What hypothesis is being tested with this approach? Would it not be possible to use empirical values for VR? Is it particularly challenging to find values of intraspecific variation for the traits considered? I find this part unclear and think it would benefit from more explanation

L314. Which parameters are used to assess persistence and recovery? Which parameters are tracked to evaluate the "evolutionary trajectories of traits"? Are the trait values themselves used? I would suggest referring to the specific metrics rather than their interpretation.

Figure S.6. In the first frame, I would suggest writing "x50" next to each of the three treatments, rather than writing it only once. This would make it clearer that this refers to 3 x 50 simulations in my opinion.

For question 1 (L313), I understand that baseline simulations were run first to obtain the initial forest (shown in Figure S.6.). However, I do not understand what the starting points (i.e., the initial forests) are for questions 2 and 3.

Table S.4. I suggest mentioning in the caption that this table is used for Question 2 and that these trait values were obtained from simulations based on the historical climatic conditions in the Montagne Noire.

L330. Same comment as in my previous review: I believe more context and justification for the choice of these eight introduced species is needed here. In particular, it seems that they were chosen in part due to their higher drought tolerance, it has to be explicitly mentioned.

Moreover, the process of regeneration and migration remains unclear to me. For example, if M=0.1, does this mean that 100% of the local regeneration occurs (but how many seedlings are generated per patch? This is still unclear)? Then, in addition to the locally regenerated seedlings, the number of seedlings from migration (i.e., from other species) corresponds to 10% of the locally regenerated seedlings, correct? And how is the proportion of seedlings from the different introduced species determined?

L348. This is the first time that populations are mentioned in the methods, what is a population in the simulations?

L353. “depending on its age at maturation (AMs)” To what does “its” refer? And what is the age of maturation? In Bone and Farres (2001), the evolutionary rates are calculated based on the number of generations instead of the age at maturation, so why not using the number of generations?

L354. “Exact H0 values are provided along with their 95% confidence intervals for all replicates for the h2 and VR = 0.3 evolutionary scenarios as this setting led to the most realistic evolutionary rates (Bone and Farres, 2001).”
What do the authors mean by “Exact”? Is this term necessary?
Then, I do not understand why the scenario with h2 and VR = 0.3 gives more realistic evolutionary rates - is this conclusion based on the results of the present study, or does it stem from prior findings, such as those in Bone and Farres (2001)? If it is the latter, it would be helpful to briefly explain why this is the case. And then, does this imply that the other evolutionary scenarios examined in the study are unrealistic? I am a bit confused here.

Why not write Equation 3 right after the text describing the equation ? (i.e., L352-353).

L356. A naive question: why can’t Haldanes be used to assess the long-term evolutionary response?

L364. The explanation of how the median minimal RBA and the median final RBA are calculated could be significantly improved, as it remains difficult to follow. Since these metrics are central to tracking the state of each species in the simulations, their explanation and interpretation would benefit from a dedicated section in the Materials & Methods. For example, sentences such as “The median minimal Relative Basal Area (RBA) (Fig. 2) can be interpreted as the population’s safety margin for persistence.” and “We used the final RBA to assess the population’s ability to recover its pre-warming basal area by the end of the 1000-year stabilization period in a warmed climate.” would be more appropriately placed in a section specifically focused on these metrics. It could be made clearer that median minimal RBA is used to assess the short-term responses of the populations (persistence under climate change), while the median final RBA is used to assess long-term response (recovery after climate change). This information is scattered in the first paragraph of the Results section and would benefit from being consolidated into a single, focused paragraph dedicated to these two metrics.

Then, in the equation provided in Figure 2A, what is exactly BA_i? Is it the median of the basal area of the 50 patches in the replicate i? I think it should be explicitly mentioned.

Figure 2A is really helpful to understand the equation but I think it should be explicitly mention below the equation that $Min(BA_{1500-3000,i})$ corresponds to the minimal of the median basal area (median across patches) shown on the figure, $Mean(BA_{1500-2000,i})$ corresponds to the mean of the median basal area across the reference period, and $BA_{3000,i}$ is the median basal area at the end of the simulation. So the authors may consider either adding $Min(BA_{1500-3000,i})$, $Mean(BA_{1500-2000,i})$ and $BA_{3000,i}$ on the figure, or adding some text below the equation (which requires writing the equation in the main text).

Is $Med(Min(BA_{1500-3000,i})/Mean(BA_{1500-2000,i}) * 100)$ the median calculated across replicates? If so, it has to be clearer in my opinion.

By the way, i was used to refer to individuals in previous equations so I think it would be better to use another letter to refer to the replicates.

There is an inconsistency in terminology: the main text refers to "median minimal RBA" (L364), while the figure uses "minimal median RBA". The correct term appears to be "median minimal RBA". Same for “Final median RBA” which should be median final RBA in my opinion.

It is unclear how the minimal median RBA can be interpreted as the population safety margin for persistence. Why couldn’t the population size decrease further and then recover? A justification for this interpretation is needed.

This comment aligns with Reviewer 4’s concerns from the previous review regarding the use of the median minimum RBA as a metric for evolutionary rescue. Reviewer 4 suggested examining tree density and diameter to provide a more comprehensive view of the forest's state. While I acknowledge that the paper and figures are already dense, I believe the justification for using median minimum RBA still needs to be strengthened. To address this, the authors could consider adding graphs in the supplementary information that depict tree density and diameter as functions of the median minimum basal area. These visualizations would aid readers in understanding how patches vary across low and high densities. Furthermore, the authors might include their justification to Reviewer 4 directly in the main text, ideally supported by some references.

Also, for the discussion on effective population sizes at the end of the manuscript (L510), it would be helpful to know what the tree densities are at median minimal RBA and median final RBA for each species. It would provide an idea of the minimal effective population sizes that the species go through during the simulations.

L366. I do not see how variation in outputs due to variation in VR provides insight into evolutionary processes. Higher VR does not necessarily indicate higher genetic variation,

right? If h2 is fixed, then higher VR would lead to higher environmental variation, correct? Therefore, I am unsure how this can be interpreted as shedding light on evolutionary processes. Could the authors clarify this? Additionally, for h2, I believe it would be more accurate to discuss the relevance of accounting for "genetic parameters" rather than "evolutionary processes," as h2 does not describe a process, but rather a genetic quantity.

L368-369. “cases with lower h2 and VR values” Which values are considered high and which are considered low? Please be more precise.

L371. “for minimum h2 and VR of 0.3” I do not understand this. Does it mean that the median minimal RBA was lower than or equal to 5% when both h2 and VR were lower than or equal to 0.3?

L384. Shouldn't it be 118% instead of 120%?

L407. “In beech-fir mixtures, evolution favored the dominant species” is confusing in my opinion; I initially understood that beech was favored simply because it was dominant. I would suggest rephrasing it with something like: “Beech better recovered than fir in the beech-fir mixtures […]”.

Maybe I missed the information, but is there a comparison between the initial proportion of beech and fir versus the final proportion? Did the proportion change over the course of the simulations? Figure 2C shows this, but it would be useful to provide the initial and final proportions. Additionally, it seems that in Figure 2C, the proportion of fir is slowly increasing after the warm period, and the simulation stops before any stabilization occurs. Could fir have recovered its initial proportion if the simulation had continued?

L413. Looking at Figure 5B, it seems that both species evolved towards greater drought tolerance from the beginning of the simulations, even before the warm period. It appears that beech benefited from being more drought-tolerant at the start, as its DrTol values remain higher than fir's over the course of the simulation. Perhaps the authors could elaborate on this in the discussion, as this seems to be an interesting point.

L423. I think this is misleading to refer to h2 and VR as evolutionary rates as the latter are estimated with H_0 and the difference in trait values between the reference period and the end of the simulations.

L426. But recovery was generally higher for higher h2 and VR (even though the maximal final RBA was not for VR = h2 = 0.4).

L427-428. Can we really say that $\bar{g}$ values were lower in the assisted gene flow treatment? It seems that their mean value is very close to the one of the beech-fir mixture treatment, and that quantile intervals are very wide.

L432. Could the authors elaborate further on this interpretation? Does this imply that the gene flow from the beech population of the Montagne Noire was better pre-adapted to a

warmer climate compared to the fir population? Would it be possible to compare the DrTol values of the fir and beech populations from the Montagne Noire to support this point?

L435. The effect of trade-offs among traits on the final RBA of each species does not seem entirely convincing. Could the authors provide more robust arguments, supported by quantitative parameters (e.g., is there really a meaningful difference in $\bar{g}$ values between the beech-fir mixture and assisted gene flow treatments?), to support their claim that these trade-offs significantly impact the final RBA?

L442-446. So I am not sure to understand; did the proportion of basal area of fir vs beech change between the reference period and the end of the simulation? Maybe it would be useful in this treatment to compare the absolute basal area of the two species before and after the warming period, as the relative basal area does not seem very useful for the interpretation in this treatment.

L478-479. This is not clear to me why the authors consider that the “choice of stabilizing the climate after 100 years of warming [...] limits conclusions on any long-term effects.”. I don't believe this is the case, as this study effectively assesses the long-term effects of a warming climate period on the proportion of each species.

L498-499. I may be mistaken, but such opposing selective forces seem to accurately capture what happens in real forests, which I believe is a strength of the study.

L513. “Another reason to improve the integration of migration into models is to explore its demographic impact on populations. In our case, we have deliberately limited these impacts to highlight the genetic consequences of the arrival of pre-adapted individuals.” Why not recording also the number of individuals over the course of the simulations as suggested by Reviewer #4? How did the authors limit the demography? I do not follow the rationale of those sentences.

L517-518. Since the regeneration process remains highly unclear to me, I do not understand what proportion of seedlings from other species is added in each patch and across the entire forest. Are these numbers realistic?

L521. “We therefore think it is urgent to explore the case where intra- and interspecific interactions are at the origin of selection pressures that accompany environmental change.” I’m not sure I fully understand. Do the authors mean that it is important to better understand the interactions between selection pressures from the environment and competition from other trees?

Figures

Figure 1. “The initial state of any evolution simulation is a forest inventory (see methods).” Maybe mention here that the initial state is based on *real* forest inventory data.

Figure 2. In the figure legend, panels are inverted.

Then, for figure 2B, it is written: “numerical values of median relative basal area are indicated for maximum and for minimum value above 5%”. But I am unclear about what these numbers represent. Aren’t they the minimum and maximum values for the median minimal RBA in each treatment? If so, why are the numbers in the third row (whole forest) not approximately equal to the sum of the numbers in the first and second rows (species-specific BA)? Please clarify all of this.

Additionally, I think it is important to consistently use the same terms when referring to each treatment. For instance, the assisted migration treatment is called “beech-fir mixture + assisted migration” in Figure 2B, “species enrichment” in Figure 2C, and “assisted migration” in Figure S7. Please use consistent terminology throughout the manuscript, as this is confusing.

Also, what is the “median percentage of reference basal area” (title of the color scale)? Does it refer to median minimum and final RBA?

How can the final RBA exceed the RBA of the reference period ? I would expect that trees are optimally adapted to their environment, with maximal basal area during the reference period. This outcome is understandable when other species are introduced or in scenarios with gene flow, but it is unclear how this could occur, for example, in a monospecific patch of fir

Figure 3. This figure is much improved compared to the previous version!
“Median *across* 25 to 50 *replicates* of the mean relative light availability perceived by individuals of beech and fir (model output) *across replicates*”. This is unclear. Are the median and quantiles across replicates being represented? If so, what exactly is meant by "mean relative light availability"? Is the mean calculated over the patches, or over some other quantity?
The legend in Figure S7 is also unclear, where it states “mean individual relative light availability over time.” This doesn’t help clarify how the mean is calculated. Please clarify the legends of those two figures.

By the way, how is the light availability calculated exactly? It is stated to be the “light availability perceived by individuals,” but since the model is not spatially explicit ("individuals are generally not spatialized within a plot" L127), there is no direct information on the light perceived by each tree, right? Could the authors explain how this was calculated? It would be sufficient to include this explanation in the supplementary information.

Figure 5B. Why looking at negative values for $ShTol$?

Table 1. Full names and units of all parameters should be provided; otherwise, the table is not useful as it cannot be understood. Since there are many parameters, perhaps only those of particular interest for this study could be included here, with the others provided in the Supplementary Information.

Supplementary Information

Natural selection in mixture validation

“Demography was assessed using basal area (BA)” – The median basal area across simulations? Please be more precise.

How do the authors interpret the fluctuations in basal area (BA) observed in the scenario with +3°C?

I really appreciated the simulations in the Supplementary Information, as they help to understand how the model behaves and build confidence in its predictions. The authors effectively demonstrate that the model behaves as expected for traits under drift or natural selection.

Fig. S.7. Why does the blue line corresponding to the assisted migration treatment stop before the end of the simulation in some scenarios?