Journal of Open Aviation Science (2023), Vol.**1**
doi:10.59490/joas.2023.7325

JOAS

EDITORIAL

Reviews and responses for

AeroMAPS: a framework for performing multidisciplinary assessment of prospective scenarios for air transport

Authors: Thomas Planès, Scott Delbecq, and Antoine Salgas

Reviewers: Nicoletta Brazzola, Antoine Chevrot, Romain Sacchi,

Editor: Xavier Olive

1. Original paper

DOI for the original paper: https://doi.org/10.59490/joas.2023.7147

2. Review - round 1

2.1 Reviewer 1

Overall assessment: This article presents a new open-access software to perform a multidisciplinary assessment of future aviation. This contribution is very important and timely, as the aviation sector plays a key role, as a hard-to-abate sector, in future residual emissions and different tools to model it emerged in parallel, but most of these are private or hard to deploy/compare. While publishing such a tool is absolutely necessary, I believe that the article could benefit from some further clarifications on the main assumptions of the model, and by a more detailed discussion of what it can and what it cannot do. In the following, I make some suggestions, which I divide between major and minor.

Major comments:

In general, the study only summarily explains how exactly the different components and impacts of the aviation sector are modelled, and which key assumptions, couplings, and dynamics are implemented. For example, the fact that the cost module and the air traffic module are not coupled (and thus a substantial increase in aviation cost does not impact demand) is a key assumption that compromises the validity of the results, yet this is explicitly made clear only in the discussion of the study's weaknesses. How exactly the other model components are coupled is still very unclear (for example, how are the environmental impacts of alternative fuels calculated besides their impact on CO2/energy? Are non-CO2 effects included?). I think an expansion of the model description, including a clear overview of the key assumptions and simplification, is needed, as well as a more accurate description of what type of results the model delivers with a high level of confidence, and which results rely on simplistic assumptions.

To keep the article succinct and thus more impactful, I suggest that an expansion of the "Methods and model" section comes at the cost of the "Software development" chapter. The latter provides in fact more pragmatical descriptions of the model overview which are possibly already included in its documentation and of limited interest for a scientific publication.

Finally, the "Application" chapter shows an example of an application, but I think it could be substantially more useful if the validation described in the "Methods and model" section were expanded on here, and the results put into the context of existing studies and models. How do the results compare to other studies or models? What drives differences? What are the strengths and weaknesses compared to other models? Moreover, the application example lacks a clear overview of the key input variables and outputs. A Table with "main results" as well as some output figures are shown, but it is unclear to me whether the model produces a time series of simply a snapshot in time of costs, emissions etc., and what exact key variables can be output.

Minor comments:

Abstract – I think the abstract spends too many words on how the model is implemented (e.g., in Python and using different packages) while lacking a straight-to-the-point message on what exactly the model does and why it is useful and novel.

- l.19 – Only one citation when you refer to "numerous academic publications".
- l.24-28 – The motivational framing is somewhat weak. While it is clear that an advantage of the model is the fact that it is open source, it is not so clear what exactly this new framework is helping us with, and what previous challenges in the prospective modelling of the aviation sector can be tackled with it.
- l.62 – "similar analysis to those detailed previously" – be specific, what kind of analysis are we exactly talking about? You mentioned a lot of different ones earlier.
- l.116-120 –It would be good to have a systematic overview of the input data and key assumptions underlying the model, possibly as a Table and with explicit references to sources from which assumptions and data are taken.
- l.127-132 – Here too references to the specific assumptions and input data are missing (e.g., historical data of operational and load factor, public data on which energy carriers modelling is based, incorporation rates of energy carriers etc.)
- l.135 – How are emissions calculated for alternative energy carriers? Based on which data/assumptions?
- l.137-139 – Need to add an explicit calculation of how temperatures are derived from GWP*-based CO2-equivalent emissions. Are you using a reduced complexity climate model, or simply a function? Please also provide the key equations (e.g., how you calculate GWP*-based CO2 equivalent emissions.
- l.147 – Here information missing on how the carbon prices are implemented. Are these prices simply added to the total costs, without the modelling of their effects on demand? Explicitly state the assumptions and simplifications made.
- l.150-152 – It is unclear what types of allocation underlies the example provided in Figure 3b.
- l.162 – Why was the fleet renewal model based on logistic functions chosen if there are alternative methods that are being proposed? The text lacks a more in-depth justification.
- l.166-177 – It would be good to have a more systematic overview of the model validation. In lines 166-167 you state that you already validated the model against historical data. It would be great to show the performance of the model relative to historical data and other models instead of just describing it.
- l.173-188 – Here a systematic overview (e.g., table) of the input assumptions would be great to also give an example of what parameters can be easily changed to perform similar analyses.
- l.195-196 – What are exactly the outputs of the model? Are they time series or, for some values, only snapshots in 2050, as reported in Table 1?
- l.204-207 – Why has the fleet renewal had a larger effect than changing fuels? Is it because of the fact that, in the ReFuelEU targets, alternative fuels are only deployed later in time? Here it would be good to have a more thorough explanation of the drivers of these results.
- l.214 – After this paragraph, it would be beneficial to have a comparison of how these results relate

to other models and studies, and whether there are any significant differences or assumptions that undermine their validity.
- Figure 3b – It is unclear to me what the percentages refer to, e.g. in terms of impacts. While for the budget I understand that the percentage shows the share of budget used up by aviation, it is unclear to me what the baseline is for the calculation of the share of the impacts. Are you using a specific SSP-RCP scenario? Or are the impacts compared to current emissions, electricity use etc.?
- Figure 3c – It is unclear to me how are the different options calculated. Are there any constraints e.g., on biomass use that dictate the amount of CO2 abated, or are they simply derived by "top-down" targets for fuel substitution etc.?
- l.231-233 – This has quite substantial implications for the validity of the analysis – it requires to be highlighted earlier and here the implications need to be discussed more thoroughly. What kind of statements/insights does the model exactly allow with a high degree of confidence?
- l.237 – What perspective is taken on costs exactly? Are these the costs to airlines to operate flights? Or the cost to consumers?

2.2 Reviewer 2

AeroMAPS is a software written in Python to perform scenarios for air transport taking into account different aspects of the domain such as environmental or economic. It provides a framework to use different models in order to, for instance, evaluate the evolution of the Co2 emissions from public air transport.

I will separate the review in two. One for the software itself and the other for the paper.

Paper:

Overall, I feel the paper is lacking a lot of information. It is relatively short and does not delve deep enough into important aspects of the library. Here is a list of interrogations I had on my readings of the paper. They illustrate what I mean by "lack of information".

- l.127: historical data. This is referred several times in the paper. Where are they from? Are they freely available?
- l.127 again: logistic models. Which ones? How are they chosen?
- Data used in the example application shows a big drop in 2020 while the historical data stopped in 2019. Was the COVID crisis predicted by the scenario or is this drop due to something else?
- One of the critics made to another open-source solution (namely AIM2015) is said to be slow in its calculations. It would be interesting to have some comparisons in terms of time of execution between the author's approach and AIM2015 (or other models).
- The authors describe a modular code enabling the use of different models. How is this modularity implemented? Does this mean one could plug easily their own model?

Software:

The installation and setup (I used Python 3.9) of the Github repo was very easy thanks to poetry and seemed to work on the first try. I did not try the install through PiPy but I believe it works with no problems.

I enjoyed playing with the web app with the different knobs and their descriptions. This is a great addition to this library and will potentially be its main use.

Much of the information that could be lacking in my opinion in the paper (see above) can be found in the documentation of the library.

I regret the lack of some kind of API for future users to maybe plug their own models to perform comparisons. This could be added in future work as the modularity of AeroMAPS certainly could

allow it.

Minor caveat: maybe add the dependency to IPyKernel in poetry since it is supposed to be mostly used in notebooks?

To summarise, this work could be a major contribution for researchers and stake-holders to create scenarios and to identify the impact of different levers of actions on the evolution of air traffic. While the library itself seems quite complete and functional, the paper leads the reader to question some areas of the experimentation making them think there might be some gaps in the library itself while it is just a detail shortfall. I recommend the author add more information to their paper to ensure that future readers do not question the integrity of the library and to entice them to use it for their own research.

2.3 Reviewer 3

This paper provides a good and comprehensive presentation of the tool AeroMAPS.

However, the scope and system boundaries are not extremely clear when it comes to, for example, emissions accounting or energy. For example, are emissions from producing fuels or aircraft accounted for? Which types of emissions are considered? While I realize one cannot explain in detail every aspect of the model, it's probably worth spending a paragraph explaining this – and possibly adding to the limitations section if some relevant emissions are omitted.

I only have a few In-text remarks:

- l.30: "generates significant calculation times". Reformulate.
- l.45: "In particular, strategy analysis should enable environmental sustainability assessments to be performed at the sectoral level": it's hard to understand this sentence. Consider reformulating.
- l.117-118: can you give an estimate of how much the commercial aviation traffic represents out of the total traffic, vs. military + private?
- l.159: "First, comparisons of methods from academic papers were performed. For instance, the Kaya identity for aviation given in Equation (1) is consistent with various proposals from academic papers [8, 9, 35, 36].": isn't always going to be pax-km*eficiency*carbon-intensity? If you found other definitions of the KAYA identity, an example would be nice, otherwise consider removing this sentence.
- L.164-166: these are very surprising values. Are you sure they are correct?
- l.172: besides the surprising RMS values, it is entirely unclear how this simulation was set up. This is all to back-test the model (presumably). So what were the input parameters? Did you take, for example, demand forecasts from 1990 and some static efficiency improvement rates? Be more explicit and detailed.
- l.171: this would be impressive if we knew more about the context of the simulation.
- l.180: "a hydrogen one": incorrect grammar.
- l.185: "different assumptions": which ones?
- l.197: "reduction is made possible by the improvement in emissions per passenger kilometre": what is this driven by? Should it not be an important part of the results to learn this?
- l.199: "and the lack of specific strategies for non-CO2 effects": maybe, as an example, you could consider the Google/AA contrail avoidance strategy. The data is very recent and rather promising.
- l.199-21: with a doubling of operational costs, a 3% annual growth rate seems unlikely then. Reconsider.
- l.232: "no coupling is included…": this should rather be mentioned in your example above, not at the end.
- Table 1: should not the unit be "CO2-eq." if you include GHG emissions as well as non-CO2 effects?

3. Response - round 1

4. Review - round 2

We have answered the various comments of the reviewers as follows. In particular, we detailed the model assumptions and limits based on the remarks of the reviewers. We also added a dedicated section (and an appendix) on the software validation. In the two cases, the modifications are due to the fact that all the reviewers commented on these aspects.

In order to facilitate the review, the modifications on the paper have been indicated with one colour per reviewer.

We sincerely thank the reviewers for their questions and comments and remain available for any additional information.

4.1 Response to Reviewer 1

The authors would like to thank you for the comments that helped improve the paper. The answers to the various remarks are given as follows.

> **Overall assessment** This article presents a new open-access software to perform a multidisciplinary assessment of future aviation. This contribution is very important and timely, as the aviation sector plays a key role, as a hard-to-abate sector, in future residual emissions and different tools to model it emerged in parallel, but most of these are private or hard to deploy/compare. While publishing such a tool is absolutely necessary, I believe that the article could benefit from some further clarifications on the main assumptions of the model, and by a more detailed discussion of what it can and what it cannot do. In the following, I make some suggestions, which I divide between major and minor.

[response] First of all, thank you for your many relevant remarks and suggestions, which we hope will improve the paper. We are going to answer each of your comments in the following. To help you understand our feedback below, we would like to remind you that the paper we are proposing is a software paper (and not a research paper). As mentioned on the JOAS website, a software paper is a "paper announcing a significant open-source tool or programming library for aviation research". As a consequence, the main objective is not to describe in detail the models used, but to present the software and its applications. Nevertheless, we tried to include your comments by explaining the positioning and different assumptions, as you will see in the following.

> **Major comments** 1. In general, the study only summarily explains how exactly the different components and impacts of the aviation sector are modelled, and which key assumptions, couplings, and dynamics are implemented. For example, the fact that the cost module and the air traffic module are not coupled (and thus a substantial increase in aviation cost does not impact demand) is a key assumption that compromises the validity of the results, yet this is explicitly made clear only in the discussion of the study's weaknesses. How exactly the other model components are coupled is still very unclear (for example, how are the environmental impacts of alternative fuels calculated besides their impact on CO2/energy? Are non-CO2 effects included?). I think an expansion of the model description, including a clear overview of the key assumptions and simplification, is needed, as well as a more accurate description of what type of results the model delivers with a high level of confidence, and which results rely on simplistic assumptions.

[response] As mentioned previously, due to the fact that this paper is a software paper, we have decided to keep an overview of the models, rather than a long full description, the latter being available in various dedicated articles and in the AeroMAPS documentation (a sentence has been added for better highlighting that fact in the introduction of the “Methods and models” section). However, we tried to integrate your main remarks on the models (see the following major and minor comments). In particular, we added some information concerning the limitations of the models which were not properly inserted in the paper. As suggested by another reviewer, the paragraph has been added at the beginning of the “Methods and models” section for introducing these limitations, choosing the cost-demand coupling and a comprehensive LCA as an example, including a reference to discussions in the "Conclusions" section for details. These discussions have been improved, with also a mention of the non-CO_2 modelling for alternative fuels.

2. To keep the article succinct and thus more impactful, I suggest that an expansion of the “Methods and models” section comes at the cost of the “Software development” chapter. The latter provides in fact more pragmatical descriptions of the model overview which are possibly already included in its documentation and of limited interest for a scientific publication.

[response] The "Software development" section is generally required for software papers. We therefore propose to keep it as it is, with additional information on model modularity and input and output data according to several comments (from you and other reviewers). Concerning the “Methods and models” section, we have expanded it to include your main comments (and those from other reviewers), while keeping it to a reasonable size of around 2 pages for this type of paper. We think that the total length of the paper is reasonable.

3. Finally, the “Application” chapter shows an example of an application, but I think it could be substantially more useful if the validation shortly described in the “Methods and models” section would be expanded on here, and the results put into the context of existing studies and models. How do the results compare to other studies or models? What drives differences? What are the strengths and weaknesses compared to other models? Moreover, the application example lacks a clear overview of the key input variables and outputs. A Table with “main results” as well as some output figures are shown, but it is unclear to me whether the model produces a time series of simply a snapshot in time of costs, emissions etc., and what exact key variables can be output.

[response] In order to highlight the validation of the software, we include a dedicated section as proposed, but in the "Software overview" section, because we wanted to preserve the "Application" section as an example. This new section has been enriched with new information, and in particular with an appendix for comparing the methods and models from AeroMAPS with other papers. However, the results have not been compared, because the papers deal with different scenarios (running all the models with the same scenario could be interesting, but the models are not easily available in most cases and it would require a long time whereas it is not the objective of the paper). Concerning the input and output data, a dedicated paragraph has been added in the "Software development" section for answering other reviewer comments, and a sentence has been modified at the beginning of the "Application" section to mention the fact that one can access to a Jupyter Notebook with all the data and results.

Minor comments Abstract – I think the abstract spends too many words on how the model is implemented (e.g., in Python and using different packages) while lacking a straight-to-the-point message on what exactly the model does and why it is useful and novel.

[response] Similarly to our previous comments, we propose to keep this sentence on Python and

the packages due to the fact that it is a software paper. However, we reformulated this part to better highlight the interest of the framework.

l.19 – Only one citation when you refer to “numerous academic publications”.

[response] The paper mentioned is a paper review. The sentence has been corrected to avoid misunderstandings.

l.24-28 – The motivational framing is somewhat weak. While it is clear that an advantage of the model is the fact that it is open source, it is not so clear what exactly this new framework is helping us with, and what previous challenges in the prospective modelling of the aviation sector can be tackled with it.

[response] The motivational framing has been separated from the rest of the paragraph (for better showcasing) and rewritten in order to integrate the interest in terms of prospective scenarios for aviation.

l.62 – “similar analysis to those detailed previously” – be specific, what kind of analysis are we exactly talking about? You mentioned a lot of different ones earlier.

[response] It concerns studies on the future climate impact of prospective scenarios for air transport. The sentence has been updated to be clearer.

l.116-120 –It would be good to have a systematic overview of the input data and key assumptions underlying the model, possibly as a Table and with explicit references to sources from which assumptions and data are taken.

[response] The input data description has been added in a dedicated paragraph in the "Software development" section, due to several comments from reviewers. We think that adding a Table is nevertheless not appropriate due to the numerous number of inputs used as detailed in the new paragraph. Concerning the references, we added the main ones in the paragraph.

l.127-132 – Here too references to the specific assumptions and input data are missing (e.g., historical data of operational and load factor, public data on which energy carriers modelling is based, incorporation rates of energy carriers etc.)

[response] As briefly mentioned in our answer to your major comment 2, we added a paragraph on input (and output) data. It includes in particular the main references used, which are all open access, but not the detailed ones (see next comment). Concerning assumptions such as incorporation rates, they depend on the scenario modelled, so we did not add them in the “Methods and models” section.

l.135 – How are emissions calculated for alternative energy carriers? Based on which data/assumptions?

[response] This sentence and a previous one have been enriched for specifying the parameters used. However, we think that it is not necessary to add the data references which are detailed in the reference papers on the models or in the documentation. As mentioned just before, we nevertheless added main references for data.

l.137-139 – Need to add an explicit calculation of how temperatures are derived from GWP*-based CO2-equivalent emissions. Are you using a reduced complexity climate model, or simply a function? Please also provide the key equations (e.g., how you calculate GWP*-based CO2

equivalent emissions.

[response] A dedicated equation for the calculation of the temperature-induced and the warming-equivalent emissions for non-CO_2 has been added. The estimation of equivalent emissions is based on a function including annual forcing and forcing variation for each non-CO_2 effect. The temperature estimation relies on cumulative emissions and the use of the TCRE coefficient.

l.147 – Here information missing on how the carbon prices are implemented. Are these prices simply added to the total costs, without the modelling of their effects on demand? Explicitly state the assumptions and simplifications made.

[response] These prices are simply added to the total costs. As mentioned previously, a sentence has been added at the beginning of the “Methods and models” section concerning the limitations of the framework, taking as an example the cost-demand coupling.

l.150-152 – It is unclear what types of allocation underlie the example provided in Figure 3b.

[response] Concerning the allocation rules in the application, we mentioned "Finally, concerning allocation rules, a grandfathering approach is assumed for climate issues, which means that 2.6% (i.e. aviation’s current share of CO_2 emissions) of the world carbon budget is allocated for aviation. For energy resources, an illustrative allocation of 5% is assumed.". For climate issues, the methodology used is mentioned (grandfathering). For energy issues, we added the term "arbitrary" to indicate that no specific methodology was applied.

l.162 – Why was the fleet renewal model based on logistic functions chosen if there are alternative methods that are being proposed? The text lacks a more in-depth justification.

[response] The alternative method is suitable for air route modelling, taking into account aircraft TLARs (Top Level Aircraft Requirements) for the replacement of aircraft on different routes. They are too detailed and difficult to use for our approach, at least in the short term. We added a specification concerning the route modelling in the text.

l.166-177 – It would be good to have a more systematic overview of the model validation. In lines 166-167 you state that you already validated the model against historical data. It would be great to show the performance of the model relative to historical data and other models instead of just describing it.

[response] As mentioned in major comment 3, we included a dedicated section on software validation in order to have an overview of the models. Other comparisons have been added to show the performance. For information, some reformulations have been performed and the performance has been expressed via R^2 for answering another reviewer’s comment.

l.173-188 – Here a systematic overview (e.g., table) of the input assumptions would be great to also give an example of what parameters can be easily changed to perform similar analyses.

[response] The number of inputs is important (180 including cost parameters). The objective for us was to describe the "philosophy" of the scenario and the main assumptions in the paragraph. As mentioned previously, we clearly added a reference to the Jupyter Notebook which includes all the inputs. One can easily change the different settings for performing another study directly with AeroMAPS.

l.195-196 – What are exactly the outputs of the model? Are they time series or, for some values, only snapshots in 2050, as reported in Table 1?

[response] The outputs are time series mostly, but also some float outputs. This information has been added in the "Software development" section as mentioned in the major comment 2.

l. 204-207 – Why has the fleet renewal had a larger effect than changing fuels? Is it because of the fact that, in the ReFuelEU targets, alternative fuels are only deployed later in time? Here it would be good to have a more thorough explanation of the drivers of these results.

[response] Indeed, the fleet renewal is currently in progress, whereas significant alternative fuel deployment comes later in this scenario. The explanation has been added.

l.214 – After this paragraph, it would be beneficial to have a comparison of how these results relate to other models and studies, and whether there are any significant differences or assumptions that undermine their validity.

[response] As mentioned in the major comment 3, we included a dedicated section on software validation, concerning in particular the comparison of methods and models with other papers.

Figure 3b – It is unclear to me what the percentages refer to, e.g. in terms of impacts. While for the budget I understand that the percentage shows the share of budget used up by aviation, it is unclear to me what the baseline is for the calculation of the share of the impacts. Are you using a specific SSP-RCP scenario? Or are the impacts compared to current emissions, electricity use etc.?

[response] On the one hand, the orange part is the "real impact" due to the scenario (for instance, the consumption of around 14 EJ of electricity here), expressed as a percentage of a world target (for instance, the electricity available in the world in 2050, assumed here at 200 EJ, see notebook): we obtain the value by the division of the two parameters (for instance, around 7% of biomass here). This estimation is based on the studied scenario (whose assumptions are described above), not on the SSP-RCP scenario. On the other hand, the green part is the "allocated budget" or the "sectoral objective for aviation", also expressed as a percentage of a world target: the value has been given in the assumptions (for instance, 5% of the available biomass, which was an arbitrary choice in this case, see previous comment on allocation). In the paper, we added a sentence to be clearer concerning targets for aviation, and we also reformulated a sentence for Figure 3b.

Figure 3c – It is unclear to me how are the different options calculated. Are there any constraints e.g., on biomass use that dictate the amount of CO2 abated, or are they simply derived by “top-down” targets for fuel substitution etc.?

[response] The availability of the different energy resources is based on the allocation chosen in this plot. Here, we consider that 5% of the biomass and electricity available in the world is allocated to aviation. It allows estimating the potential CO_2 abetted for each energy carrier. The paragraph has been rewritten to make it easier to understand the plot. Note that this plot is used for ranking decarbonisation solutions depending on feedstock availability, and does not directly depend on the incorporation rate of the scenario (this plot can help to establish them for instance, here we considered blending mandates from ReFuelEU).

l. 231-233 – This has quite substantial implications for the validity of the analysis – it requires

> to be highlighted earlier and here the implications need to be discussed more thoroughly. What kind of statements/insights does the model exactly allow with a high degree of confidence?

[response] As mentioned in your major comment 1, we added a sentence referring discussion of model limits to the "Conclusions" section, and we also expanded the discussion on models in the new "Software validation" section, including the analyses that this version of AeroMAPS enables performing. Moreover, we modified the formulation in the "Conclusions" section to distinguish missing independent models from couplings.

> l. 237 – What perspective is taken on costs exactly? Are these the costs to airlines to operate flights? Or the cost to consumers?

[response] We have given examples of costs currently covered. For information, we could for instance add models on airfare in the future.

4.2 Response to Reviewer 2

The authors would like to thank you for the comments that helped improve the paper. The answers to the various remarks are given as follows.

> **Paper** Overall, I feel the paper is lacking a lot of information. It is relatively short and does not delve deep enough into important aspects of the library. Here is a list of interrogations I had on my readings of the paper. They illustrate what I mean by "lack of information".

[response] We tried to complete the paper concerning the information missing, based on your comments and the ones from other reviewers: the software characteristics, the data used and some methods and models have been for instance detailed. We describe in particular in the following our answers concerning your specific questions. Don't hesitate to consult the other reviews to check that they contain the main missing information.

> - l 127: historical data. This is referred several times in the paper. Where are they from? Are they freely available?

[response] Based on your comment and the ones from other reviewers, we added a dedicated paragraph on input (and output) data in the "Software development" section. It includes in particular the type and main references. Only open-access data have been used.

> - l 127 again: logistic models. Which ones? How are they chosen?

[response] We added a description of the logistic functions in the case of fleet renewal, with in particular the corresponding equation and parameters. The setting of the parameters depends on the user assumptions (entry-into-service year, duration for replacing aircraft...). A sentence has also been added concerning the calibration of the historical fleet.

> - Data used in the example application shows a big drop in 2020 while the historical data stopped in 2019. Was the COVID crisis predicted by the scenario or is this drop due to something else?

[response] The drop is indeed due to COVID-19. We added a sentence indicating that its impact is taken into account in the application. For your information, we included the impact of COVID-19 in the AeroMAPS parameters by default, based on partial 2020 data and assumptions (see documentation, in Documentation/Use of the graphical user interface, available here https://aeromaps.github.

io/AeroMAPS/books/documentation_gui.html, in section Reference settings/Discovery mode/Air traffic), which correspond to the ones used in this application.

> - One of the critics made of another open-source solution (namely AIM2015) is said to be slow in its calculations. It would be interesting to have some comparisons in terms of time of execution between the author's approach and AIM2015 (or other models).

[response] We added this information (around 1 s for AeroMAPS, see JOAS notebook that has been updated for exact value, and 40min-2h for AIM2015, see sentence added in the statement of need) in a new dedicated section, including also the comparison of the framework with other ones (with an appendix) and the software validation that was already present in the initial paper. However, even if we provided the information on calculation time for the two software, we did not directly perform a duration comparison due to the fact that AIM2015 provides much more detailed modelling and results.

> - The authors describe a modular code enabling the use of different models. How is this modularity implemented? Does this mean one could plug easily their own model?

[response] The "Software development" section has been completed to detail the model management. In terms of model plugging, the user just has to create a new model, with respect to the variable names, and add it to the list of the model. GEMSEO allows an automatic connection with the other models. The user just has to check that the model does not replace another one (in particular to avoid that an output has two ways of calculation).

> **Software** The installation and setup (I used Python 3.9) of the Github repo was very easy thanks to poetry and seemed to work on the first try. I did not try the install through PiPy but I believe it works with no problems. I enjoyed playing with the web app with the different knobs and their descriptions. This is a great addition to this library and will potentially be its main use. Much of the information that could be lacking in my opinion in the paper (see above) can be found in the documentation of the library. I regret the lack of some kind of API for future users to maybe plug their own models to perform comparisons. This could be added in future work as the modularity of AeroMAPS certainly could allow it. Minor caveat: maybe add the dependency to IPyKernel in poetry since it is supposed to be mostly used in notebooks?

[response] We are satisfied that the installation instructions are sufficiently clear, that the various modes of use are interesting, and that the documentation includes sufficient information, particularly that required for the paper. Concerning your software comment, we are currently developing a reference API which will be available in the coming weeks, and which will be integrated directly into the documentation. Finally, ipykernel is included in jupyterlab package dependencies (present in the AeroMAPS pyproject.toml file) which is useful for manipulating notebooks (see https://github.com/jupyterlab/jupyterlab/blob/main/pyproject.toml). For instance, using the installation of aeromaps from PyPI, we obtain ipykernel in the poetry.lock file (version 6.23.1).

4.3 Response to Reviewer 3

The authors would like to thank you for the comments that helped improve the paper. The answers to the various remarks are given as follows.

> This paper provides a good and comprehensive presentation of the tool AeroMAPS. However, the scope and system boundaries are not extremely clear when it comes to, for example, emis-

sions accounting or energy. For example, are emissions from producing fuels or aircraft accounted for? Which types of emissions are considered? While I realize one cannot explain in detail every aspect of the model, it's probably worth spending a paragraph explaining this – and possibly adding to the limitations section if some relevant emissions are omitted.

[response] We agree that certain assumptions and limitations of the model were not sufficiently highlighted. We have therefore completed the "Methods and models" section in response to your comments and those of the various reviewers. In particular, with regard to the limitations, we have added a sentence indicating the limits (with the example of the absence of cost-demand coupling and a comprehensive LCA) and included a link to a discussion in the "Conclusions" section which has been completed as suggested by another reviewer. More specifically, concerning your questions, the emissions from fuel production and combustion are included, but not the emissions for the rest of the life cycle (aircraft production and end-of-life, airports...). We only consider emissions that have an impact on climate change (CO_2, basic GHG for fuel production, and non-CO_2 effects from aviation). The information has been added in the dedicated paragraph.

l.30: "generates significant calculation times". Reformulate.

[response] Corrected.

l.45: "In particular, strategy analysis should enable environmental sustainability assessments to be performed at the sectoral level": it's hard to understand this sentence. Consider reformulating.

[response] We reformulated this sentence in a simple way, the words "strategy analysis" were not appropriate.

l.117-118: can you give an estimate of how much the commercial aviation traffic represents out of the total traffic, vs. military + private?

[response] We used Gossling and Humpe data (see paper: https://doi.org/10.1016/j.gloenvcha.2020.102194). According to this reference, commercial aviation represents 88% of aviation fuel consumption. We added the reference in the paper.

l.159: "First, comparisons of methods from academic papers were performed. For instance, the Kaya identity for aviation given in Equation (1) is consistent with various proposals from academic papers [8, 9, 35, 36].": isn't always going to be pax-km*eficiency*carbon-intensity? If you found other definitions of the KAYA identity, an example would be nice, otherwise consider removing this sentence.

[response] Not necessary, [36] for instance includes economic factors, and [35] provides an adaptation based on the Breguet-Leduc equation (be careful, the reference numbers have been updated in the correcter paper). The sentence has been reformulated to be clearer, separating similar and different formulations. More generally concerning the validation of the software (discussed in this and subsequent comments), we created a dedicated section which has been expanded from the paragraph in the original version.

L.164-166: these are very surprising values. Are you sure they are correct?

[response] Thank you for the comment, we indeed made a mistake reporting the values... Moreover, RMSE is not easily usable for comparisons between models (another reviewer asked us to add another

example of modelling in this paragraph). As a consequence, we replaced them with R^2 (with values checked!).

l.172: besides the surprising RMS values, it is entirely unclear how this simulation was set up. This is all to back-test the model (presumably). So what were the input parameters? Did you take, for example, demand forecasts from 1990 and some static efficiency improvement rates? Be more explicit and detailed.

[response] The paragraph has been rewritten in order to explain clearly the type of models considered, which were not explicit and too short in the initial paper. Another example has also been added as suggested by another reviewer. The inputs are historical data which are used to calibrate the models. More details can be found in the original paper (Planès *et al.*, 2021).

l. 171: this would be impressive if we knew more about the context of the simulation.

[response] The comparisons were performed on historical data, so the results do not depend on prospective assumptions. We added additional comparisons (CO_2 emissions for other years, historical ERF over the period 2000–2018) to better highlight the AeroMAPS performance.

l. 180: "a hydrogen one": incorrect grammar.

[response] The sentence has been corrected.

l. 185: "different assumptions": which ones?

[response] The assumptions have been detailed for the main ones (carbon tax, electricity, fuel MFSPs).

l.197: "reduction is made possible by the improvement in emissions per passenger kilometer": what is this driven by? Should it not be an important part of the results to learn this?

[response] The information has been added, with a reduction of more than 60% due to the use of alternative fuels. We also discussed this aspect using Figure 3a.

l. 199: "and the lack of specific strategies for non-CO2 effects": maybe, as an example, you could consider the Google/AA contrail avoidance strategy. The data is very recent and rather promising.

[response] In this illustrative scenario, we wanted to focus on CO_2 strategies (in particular because the effect of SAF on contrails has not yet been considered as indicated). Nevertheless, we added a sentence for explicitly mentioning the non-use of contrail avoidance strategies, using the reference of *Teoh et al.* (https://doi.org/10.1021/acs.est.9b05608), more easily quotable than the reference you mentioned (https://blog.google/technology/ai/ai-airlines-contrails-climate-change/), which is also an really interesting web page.

l. 199-21: with a doubling of operational costs, a 3% annual growth rate seems unlikely then. Reconsider.

[response] We considered in this arbitrary example a lower annual growth rate of 2% (and not 3%). Moreover, we added information on cost effects on demand before and after the "Application" section (see next comment).

l. 232: “no coupling is included…”: this should rather be mentioned in your example above, not at the end.

[response] As mentioned previously for your main comment, we added this information in the “Methods and models” section. Moreover, for information, we also reformulated this paragraph in the "Conclusions" section for answers to another reviewer’s comment.

Table 1: should not the unit be “CO2-eq.” if you include GHG emissions as well as non-CO2 effects?

[response] We do not include non-CO_2 effects in the two first lines of the table. Indeed, it requires choosing a climate metric (GWP, GTP) and a time horizon, which can be complex to interpret due to the very short-term effect of aviation non-CO_2 effects. However, it is true that some GHG (methane and N_2O in particular) are included in the biofuel and electricity LCA. Strictly speaking, the notation should be CO_2-eq. with GHG but without aviation non-CO_2 effects. We prefer using the basic notation (because it could be misleading and CO_2 represents most emissions), but we added a footnote to the table to specify the perimeter.

5. Review - round 2

5.1 Reviewer 1

Thanks for your response and congratulations on your new manuscript version, which I think is now more comprehensive in terms of information and descriptions.

I understand the focus of a software paper but believe essential details, like key assumptions and result reliability, are needed for a comprehensive understanding, as highlighted as well by other reviewers. While I believe that the expansion of the discussion on limitations and the aim and scope of the model are now sufficient, I still suggest referring to your input data, model structure, and assumptions more explicitly. As it is now, in fact, the references to the articles where you describe the model are all concentrated in the first paragraph of the "Previous works" section (l.19 and following), but are not explicitly referenced later on. You moreover suggest that many of the input data and assumptions can be found in the documentation/Jupyter notebook, but I still believe that it would be more useful to the reader to have a brief overview of key elements in the current paper to make it a standalone resource (especially since there is not a single, up-to-date model description paper, but different additions are described in different articles or conference proceedings, and not all your readers may be literate enough in Python to understand the key assumptions by quickly having a look at your Jupyter notebook). To not overwhelm readers with a detailed description of the model, it would be sufficient to add explicit references to the detailed description at each point where you did not go into the details of the model, especially since the different information is spread across different articles, conference proceedings, the documentation, and the Jupyter notebooks. An example of where explicit references are still missing are:

118-119: Add explicit references to the articles where you give a full description of the model

105-107: Add a reference to where to find a full list/description of input data used

5.2 Reviewer 2

One of the main issues I had with the original submission was the lack of information and not delving enough into different aspects of the library. From what I understood, the other reviewers felt the same and the authors made many changes accordingly to add more context to their work.

It now feels like a more complete software submission worth being presented in the JOAS. I do not have any further demands for the authors concerning the paper or the software.

Minor edit: l203 - models' accuracy

5.3 Reviewer 3

The authors have adequately addressed my concerns.