

GMD Reviews and Authors' Response concerning the paper "A parallel implementation of the confined-unconfined aquifer system model for subglacial hydrology: design, verification, and performance analysis (CUAS-MPI v0.1.0)"

Yannic Fischler¹, Thomas Kleiner², Christian Bischof¹, Jeremie Schmiedel⁴, Roiy Sayag⁴, Raban Emunds^{1,2}, Lennart Frederik Oestreich^{1,2}, and Angelika Humbert^{2,3}

¹Department of Computer Science, Technical University Darmstadt, Darmstadt, Hesse, Germany

²Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Bremen, Germany

³Faculty of Geosciences, University of Bremen, Bremen, Germany

⁴Department of Environmental Physics, BIDR, Ben-Gurion University of the Negev, Sde Boker, Israel

Correspondence: Yannic Fischler (yannic.fischler@tu-darmstadt.de)

Copyright statement. ©2023 all rights reserved

1 General Comments

We are grateful for all suggestions concerning the structure of the manuscript that will help to improve the manuscript - we will incorporate the vast majority of the suggestions, including to move sections to the appendix. We are also happy to include more text in the introduction. We will extend the entire description of the state of the art of hydrology, the description of the hydrology of this model and the description of the modeling of the pumping tests.

We want to emphasize that CUAS-MPI is a code that has been developed from scratch. Its software architecture and numerics are different from former implementations of the same equations of the hydromechanical problem. The acronym CUAS stands for Confined-Unconfined Aquifer System, which both, the prototype code CUAS and CUAS-MPI are simulating. It is comparable to groundwater equations being solved by numerous different codes, ranging from very basic serial implementations to performance optimised codes for HPC.

We also want to lay out, that the model (and code) that the paper Fischler et al. (2022) was concerned with, is an entirely different model simulating another compartment of the Earth system, namely ice sheets. Consequently, none of the conclusions drawn in our manuscript can be drawn from Beyer et al. (2018) nor from Fischler et al. (2022).

2.1 General comments:

The manuscript by Fischler et al. discusses the development and performance of an MPI-aided parallel implementation of a pre-existing subglacial hydrology model (Beyer et al, 2018). The method and some specificities of the parallel implementation are briefly presented (including its reliance on external libraries such as PETSc). The model is then validated using several
20 test cases, some of which possess an analytical solution. Finally, runtime results for a series of bigger test cases (the size of the entire Greenland ice-sheet) are analyzed in order to derive tendencies concerning the scalability and throughput of the modeling tool, with the intent of comparing these to the performance of the well-known ice-sheet model ISSM (in view of a future coupling).

Overall I would not recommend this paper for publication in its present format. The development of an efficient subglacial
25 hydrology modeling tool capable of scaling on supercomputers is of interest to the community; so there is no doubt that CUAS-MPI has scientific relevance. However, the overall quality of the manuscript is poor and the scientific significance and originality of the results presented is only marginal. To be considered for publication, improvements should concern how the paper is structured –to be able to better distinguish what will be discussed in each Section; but there also should be a rethink of the general content. Indeed, my biggest concern is that the manuscript does not present any novel concepts or introduce any
30 new methodologies and that most conclusions drawn can already be found in previous work (see Beyer et al. 2018 and Fischler et al. 2022). The English quality also makes the manuscript difficult to read at times and many sentences/explanations would benefit from a rewrite. Finally I do not believe that the Introduction is complete with all relevant citations and references.

Due to the large number of major concerns, I have gathered all specific comments and technical corrections below, following
35 the current structure of the paper. For each Section, I start with a paragraph summarizing my concerns before I provide more technical, line by line, suggestions.

2.2 Specific comments:

On a general note and as suggested by reviewer #1, I would strongly advise making the titles of Sections more informative.

We will adapt the titles.

2.2.1 Abstract

40 – The abstract suffers from a lot of repetitions (CUAS-MPI).

– Acronyms (such as MPI) should be detailed at least once.

We will elaborate acronyms such as GCC, MPI and PETSc

– Beyer et al. 2018 should be cited.

We are citing Beyer et al. (2018) intensively in the entire manuscript, but we prefer to keep the abstract free from
45 citations.

- 1. 1 – 2: what is meant by “as well as marginal lakes and rivers” ? Rephrase this sentence

"The subglacial hydrological system affects ... ice marginal lakes and rivers". The subglacial water can feed into ice marginal lakes and rivers. Not only into the ocean. We will rephrase accordingly.

- 1. 2: “has been developed” makes it sound like CUAS is new, when this tool is pre-existing. Rephrase.

50

As outlined above, the CUAS acronym stands for “Confined-Unconfined Aquifer System” in general and the physics have been implemented in a new code called CUAS-MPI. We will rephrase this sentence here and elsewhere in the manuscript to make the distinction between “CUAS” and “CUAS-MPI” more obvious.

- 1. 7: repetition of throughput could be avoided

55

This is correct. We rephrased the sentence to “Our measurements show that CUAS-MPI reaches a throughput comparable to that of ice sheet simulations, e.g. the Ice-sheet and Sea-level System Model (ISSM).”

2.2.2 Introduction

My main issue with the introduction is that it lacks a proper state of the art. Context about the need for subglacial hydrology modeling should be provided, and different existing modeling approaches/proper citations for other similar modeling tools should also be acknowledged.

60

The place and goal of the present study are also not clearly stated and should be elaborated upon. Finally, a sentence underlining the main findings should be included.

We will elaborate upon the scientific goal(s) in the introduction, but prefer to state the main findings in the Abstract and Conclusion sections only.

- 1. 13: consider giving a ref to accompany the first sentence (especially for Antarctica)

65

This is a good point indeed. We have now gone this route: we provide a number of papers on observations of the impact, although this is not trivial as not many of such observations do exist. We cite now one for ice stream motion (Engelhardt and Kamb, 1997), one for tidal modulation (Bindschadler et al., 2003), one on seasonality of ice sheet motion due to meltwater input (Hoffman et al., 2011) and one on subglacial lakes (a review by Siegert (2000)). There is a by far larger number of papers on simulating ice stream flow and the effect different type of sliding has on the ice sheet motion, but we think it is useful in this very first sentence to stick with observations.

70

- 1. 15: use the same units (choose ma-1 or mmd-1)

We adopt the referee’s suggestion.

- 1. 19: “hundreds to thousands of metres thick ice” → “meters of thick ice”

We adopt the referee’s suggestion.

75

- The sentence starting at the end of 1. 20 should be rephrased, and the citations should be under (). I suggest transforming this sentence into a full paragraph with a proper state of the art of subglacial hydrology modeling.

Although the main topic of the manuscript is the performance of the new implementation CUAS-MPI of the published methods (Beyer et al., 2018), we are happy to add a paragraph about the state of the art of subglacial hydrology modeling, as suggested.

- 80 – The sentence starting at the end of l. 23 is too vague, consider quantifying “adequate spatial resolution” and “temporal resolution”

We omitted this sentence as we think the next, rephrased, paragraph motivates our work.

- l. 26: same issue, consider quantifying “the desired amount of timesteps”

We rephrased the paragraph to clarify the motivation for this work.

- 85 – l. 27: “ An example of relevance is the projection . . . until 2100. . . ” → why is that? (consider providing a reference)

Ice sheet dynamics is contributing significantly to sea level rise and projections show that this is going to increase in the next centuries. These projections are typically done up to 2100 as climate forcing for such simulations is available for this time period. As the subglacial hydrological system is having a big influence on the sliding of ice over the bed, projections of the contribution of ice sheets to sea level change will benefit if for the same time period simulations of the subglacial system are available. This is why we refer to this time period. However, indeed all ice sheet simulations will benefit from simulations of the subglacial system, paleo-simulations over millions of years similar to high resolution simulations of ice caps close to settlements. We are happy to provide a reference and extend the sentence in the manuscript to make this point more clear.

90

- l. 30: “This can only be achieved with a efficient codes”

95

We rephrased this section of the text.

- l. 31: again and unless I am missing something, CUAS is not new. The use of “developed” is misleading here.

We want to emphasize that CUAS-MPI is indeed an entirely new developed code. While the set of equations to be solved is similar to CUAS, the numerics is new, the software architecture, the parallelisation - as such the entire software is coded from scratch. Therefore, the use of ‘developed’ is indeed correct here.

100

I suggest merging this short sentence with the following one.

We adopt the referee’s suggestion.

- l. 33 “compute clusters” → supercomputers?

We adopt the referee’s suggestion.

- l. 34-35: provide more information about these “pumping tests”. A short description with a proper reference would do.

105

Indeed this is a very good suggestion. We will include a paragraph on the pumping test in the introduction with proper references.

- l. 35: “we employed CUAS-MPI then” → “we then employed CUAS-MPI”

We adopt the referee’s suggestion.

- 1. 36: perhaps expand on the “performance data” that you gather and provide a short description of the nature of Score-P
110 We expanded a half sentence to clarify the type of data collected and will add a paragraph on Score-P in the experimental section.
- 1. 37: “first we explain the underlying model ...” → “We explain the underlying model in Sect. 2”?
We adopt the referee’s suggestion.

2.2.3 Section 2

115 Section 2 appears a bit disorganized and could benefit from a complete makeover. The purpose of each subsection is not clear to me. Based on what I understand, I believe Appendix A should replace most of Section 2.1 (see below).

Thanks for the suggestion. We will move and extend Appendix A into Section 2. We expect this will improve the connection to the Validation section(pumping tests) and avoids duplications of text introducing certain quantities, symbols and units.

It also sounds like Section 2.2 could be a good place to describe your original work: how did you make the pre-existing
120 CUAS solver better, what did you change/improve?

Again, we want to emphasise, that CUAS-MPI is a new code, whereas CUAS was only a prototype or proof-of-concept code. But we do understand the intention of the reviewer and we will add more text to make clear what advantage a professional numerical code has. However, the description provided is too vague to be impactful (e. g. 1.74-90: “the distributed memory features of PETSc”, “a PETSc feature”, “list of available PETSc solvers”, “the iterative GMRES solver”, “an adapter which
125 provides a uniform interface to the features we require” etc.). I suggest rewriting this entire section with appropriate references to the external libraries (PETSc, NetCDF, HDF5) that you use, including a more specific description of the features that you select and why you chose to do so. Consider providing a more detailed layout of the code and its submodules in Fig 1 to accompany these developments

We have added a paragraph that outlines our design philosophy and the resulting choice of PETSc and HDF5: *The goal of
130 our code development was to arrive at a software artefact that is performant on current HPC systems but also maintainable in the light of future HPC system evolution. To that end, we based our development on the well-known PETSc parallel math library. PETSc is an open source software supported by a wide user base and, for example, part of the software stack supported by the U.S. exascale project <https://www.exascaleproject.org/research-project/petsc- tao/>. Similarly, to parallelize the NetCDF output, we used the HDF5 library, an open-source product maintained by the not-for-profit HDF group. By basing our
135 development work on these software infrastructures, we profit both from their maturity and performance on current systems, but also, in particular, from performance improvements that will be implemented by the community supporting these software libraries down the road.*

The purpose of Figure 1 is to show what physics is being solved (left side), and a high-level pseudo-code (right side) of how this is being done. The overall numerical approach is unchanged to that of Beyer (Beyer et al., 2018), except for the use of an
140 iterative solver for the linear systems of equation. As our paper focuses on the parallel implementation and its capabilities, we do not consider it appropriate to duplicate this information here. However, we will revise this section to provide more details on the libraries and their relevant features.

Finally, the point of Section 2.3 is also a little bit difficult for me to understand. I do not follow what is the relation between, and purpose of, the various scripts mentioned. Here also, I cannot differentiate what is original work when compared to the pre-existing CUAS workflow. I suggest rewriting this subsection, and accompanying the text of l. 109 - 118 (where “typical use cases” are employed as examples), perhaps sketching the workflow?

The workflow for the python implementation of CUAS and the new implementation, CUAS-MPI, is very similar, because it is mainly pre-processing available datasets and setting up the experiments. This part of the manuscript is intended as a suggestion and illustration of a workflow that worked for us and that could be starting point for new users. Nevertheless, this part seems to distract from the main line of thoughts and we move this to the Appendix section.

Section 2.1:

- l. 43: “different perspectives on the purpose of” → “different reasons for”? This sentence is awkward and could use a couple references to illustrate your point.

Changed to “different reasons for” as suggested. The initial (mostly descriptive) part of section 2.1 will be merged into a ‘state of the art’ paragraph in the introduction.

- End of l. 45, starting with “In the past . . .” is odd in this context. If a proper state of the art is placed in the Introduction, it should be removed altogether.

We agree and make changes accordingly.

- l. 47: “the code that we present” → “the code that we employ/use”

We do indeed present a new code here, this is why we do not agree to switch to employ/use.

- l. 50: “the void space is fully saturated” is confusing when it is said later in l. 55 that “it may happen that water supply is not sufficient to keep the water system fully saturated”. Please clarify.

This is exactly what the confined-unconfined scheme is about. If not sufficient water is available to have a fully saturated system over the entire thickness of the layer, the saturation thickness of the layer (Ψ) is reduced, while the system below that thickness remains fully saturated - this is the unconfined aquifer. To be more explicit: The saturation thickness is the depth from the water table to the base of the aquifer.

- l. 54: “inefficient water transport”, but the entire sentence should be rephrased to be more impactful and precise.

Very good suggestion. We rephrased this to: “The ice sheet is acting as a confining layer of the aquifer. It may happen that water supply is not sufficient to keep the water system fully saturated. To this end, an unconfined layer is incorporated, capturing the dynamics if the head is falling below the layer thickness allowing for further water drainage.”

- l. 58 - 63: As suggested by reviewer 1 I would strongly suggest moving Appendix A here, along with providing proper citations (to Beyer et al 2018 and de Fleurian et al 2014). Describing equations is always a difficult task, and given the length of Appendix A and the fact that the equations are referenced in the body of the text I do not believe it would make the paper unnecessarily longer (especially if the rest of Section 2.1 is rewritten in a more concise manner). Consider

175 using an image to describe the model? Or at least a reference to Fig. 1 of Beyer et al 2018.

We agree.

Section 2.2:

– l. 66: “its performance was too low for larger setups such as Greenland” is too vague. Please clarify

We will explain this from the viewpoint of a domain scientist.

180 – l. 68: please provide a description of the “physics kernels”

Many thanks for this feedback. We originally thought that Figure 1 is giving this overview, but we learned that this is not detailed enough. In the revised version we will add a paragraph to elaborate more on the physics kernel.

– l. 74: provide a reference to PETSc

We will add an additional reference to PETSc - please note the PETSc has been cited already.

185 – I feel like Table 1 and the associated discussion in Section 2.2 (l. 87-90) is unnecessary

In the usage scenarios we encountered so far, the output and output options were relevant.

Section 2.3:

We moved this section to appendix.

– l. 96 – 97: “a setup script”, “the mask”, “any environment”... please be more precise

190 We are not sure we understand what exactly the reviewer means. In the sentences before we detail what the setup script is actually for: ‘A typical workflow to run a simulation is described by a setup script. In that script the model domain and the grid are defined 95 and used to create the mask including boundary conditions. This mask is one of the input fields.’ And then we explain in which environment you can perform this - on any environment, a local cluster, a desktop computer, a laptop or a HPC cluster. We are happy to improve our text once we understand what exactly the reviewer is referring to.

195 – Usefulness of l.100 – 104 is questionable (restarting a run is common practice)

Indeed, restarting a model run is common practice. Here we try to explain, that the restart capability of CUAS-MPI, and also the former python implementation, can be used in alternative ways. E.g., restarting from a coarse-resolution spinup after remapping onto a high-resolution mesh.

200 – l. 103: “has be remapped” → “has been remapped”

We adopt the referee’s suggestion.

– l. 111 – 112: “Here the seasonal . . . and others with seasonal forcing” should be rephrased

205 Weill be rephrased. “At this point the actual simulation with seasonal forcing starts using the refined mesh. It is highly recommended to also run a control simulation on the refined mesh by further using the stead-state forcing. This allows to account for model transients that may still be contained in the model after the spin-up.”

- I. 113: can you elaborate on this “nesting approach” mentioned?

As this part of the manuscript (workflow) is of only minor relevance for the main topic (performance) we will only briefly elaborate on this topic. A regional subdomain nesting for a particular area of interest can be used to reduce the computational costs of a model run. If a drainage basin can only covered partially it is very beneficial to embed this area into a large-scale model run that provides boundary conditions along the subdomain margin.

210

- I. 120: what is the meaning of that last sentence? (what are “these numbers”?)

After restructuring Section 2 and moving the ‘workflow’ section into the appendix, the last two sentences of this section are no longer required.

2.2.4 Section 3

215 This Section is interesting and the work conducted is solid but my impression is that the CUAS-MPI equations and numerical methods employed to solve them have not changed since what is described in the paper by Beyer et al., 2018. This makes the relevance of the proposed validations questionable.

Indeed, the equations have not changed in the CUAS-MPI implementation since Beyer et al. (2018), but the equations have also not been tested with pumping tests before. Nevertheless, there is a need to show that the new implementation leads to expected results. We think, that pumping tests are very good test cases, because they are very well established in the field of hydrology. Specifically, analytical and semi-analytical solutions exist, which CUAS-MPI does compute correctly. With the pumping tests, one can verify the implementation of the governing equation (Eq. A1), as well as the implementation of Neumann and Dirichlet boundary conditions. They are further relatively easy to set-up and conduct.

220

I would suggest either to elaborate on the relevance of these test cases, describing them in more depth, or moving them to an Appendix.

225

We probably cut this part of the manuscript too short in the submitted version. We think pumping tests are very important and extend this part of the manuscript.

If kept, consider providing a Table summarizing the various test cases performed (from the body of the text it is unclear how many tests are actually designed), and clearly identifying what code feature they are testing.

230 The experiments conducted are shown in Fig. 2 and we don’t think a table is needed with the extended version of this section.

Finally, I cannot emphasize enough how much a sketch or a picture would help the reader understand what these “pumping tests” are.

This was also recommended by RC1 and we are happy to provide a new figure illustrating the set-up including boundary conditions, distances and image-well types and locations.

235 There is a lot of math involved here that is not trivial, and a proper introduction of the terminology would be helpful.

We moved the equations from Appendix A to the main text in section 2, as suggested by RC1 and RC2. Now, most of the quantities and units are already mentioned prior to this section.

– I. 127: consider providing a more detailed description of these “pumping tests”.

We agree and will provide a more detailed description in the revised version.

240 – I. 134: I am unfamiliar with the method of images in this context, I believe a short justification or a reference would be helpful

The justification is that the partial differential equation for the confined case is linear, thus allowing the use of a superposition (known as the method of images), the reference is Ferris et al. (1962), mentioned in I. 135.

– I. 138: please provide a brief description of what is an “image wells”

245 An image well (or an imaginary well) represents the numerical boundary in the mathematical terminology. Thus, the two boundary conditions can be described with the method of images using wells projected perpendicular and equidistant to the boundaries as the "real" pump (Ferris et al., 1962). We hope to clarify this with our new figure illustrating the set-up including image-well types and locations.

– I. 140: define “drawdown”

250 The drawdown is already defined in I. 140 with $s = h(x, y, 0) - h(x, y, t)$. Alternatively: drawdown, the amount the hydraulic head has changed since its initial state $s = h(x, y, 0) - h(x, y, t)$.

– I. 145: what is “the non-linear unconfined case” ? (again a Table summarizing the test cases and giving them names would help)

255 We appreciate the reviewer suggestion. The simulations conducted are shown in Fig. 2 and we think with the extended version of this section we don’t need a table. We will mention the abbreviations of Fig. 2 in the text directly.

– I. 153: what is “the specific yield” (another case for describing the model’s equation in a previous section)?

The specific yield is the ratio of water volume per unit volume which gets released once the aquifer drains. This is only relevant for the unconfined case due to a partially filled aquifer $\Psi < b$ (Eq. A2). We moved the equations from Appendix A to the main text in section 2, as suggested by RC1 and RC2 and changed the text accordingly.

260 2.2.5 Section 4

I feel like this section should be a subsection of Section 5.

We changed the text accordingly.

It could be useful to see the various grid resolutions, perhaps as a Supplementary Material?

– I. 162: the first sentence should be rephrased.

265 Rephrased and split into two sentences.

– Fig 3: the bed topography is not visible, consider changing the color scheme or the transparency. Also, there should be a scale for it.

The shading was for three-dimensional appearance of the figure only. There is no scientific benefit on showing the shading with a proper scale so we update the figure without shading.

270 – l. 165: please elaborate on the “Dirichlet boundary condition for the head”. A pressure is not homogeneous to the hydraulic head.

This reviewer is correct. The hydraulic head is zero at the ocean boundary and is imposed as a Dirichlet condition. This is equivalent to assuming, that the water pressure in the subglacial aquifer is equivalent to the pressure in the ocean at the same depth. We ignore the slight density difference between sea water (about 1028 kg m^{-3}) and fresh water (1000 kg m^{-3}). We correct this in the text.

275

– l. 168: the citation Christmann et al. 2021 is odd. What are you trying to back up with it?

We are referring to a version of BedMachine that has been created with new airborne data from AWI for the publication Christmann et al. (2021). We have used this new BedMachine dataset for our simulations here and there is no other way to reference it, because it is not one of the major BedMachine releases.

280 – l. 168: “To summarise it”, what is “it” ? Also, I am not sure what you mean by this sentence.

The margins along the ice sheet consist of marine terminating ice masses and land terminating ice masses. The marine terminating ice masses have a different boundary condition (head boundary condition) then the land terminating ones (no flux). This way the water moves from the inland into the fjords (marine terminating glaciers) only.

– l. 171: “Ice sheet basal melt” can you clarify what you mean here, where are you using this information into your equation system (what are the units) ? This sentence seems redundant with l. 182

285

Ice sheet basal melt (unit: m a^{-1} water equivalent) is the freshwater source Q for the model (Eq. A1). This will be clear in the revised version with Appendix A moved into section 2. We drop the last part of the sentence in l. 182, because of this redundancy.

– l. 176: remove the comma

290

We adopt the referee’s suggestion.

– l. 184: you defined a symbol for the specific yield earlier, use it

We adopt the referee’s suggestion.

– l. 191: remove the “residuum”

We adopt the referee’s suggestion.

295 2.2.6 Section 5

Section 5.1 is actually well written, and Fig. 4 is a good illustration, but the conclusions lack generality: do they hold for other node architectures? Can we identify directions for potential code improvement? It would help to more thoroughly quantify the memory requirements and their nature to answer these questions. It sounds like the point that is made is simply to provide users

with recommendations when running this specific code on this specific supercomputer, which is not extremely useful.

300 The Lichtenberg Architecture is actually very common as far as HPC systems are concerned: Intel- or Intel-compatible chips
connected with a high-speed network. We have made that point in the text. So our findings are highly relevant for HPC
installations other than TU Darmstadt. In the section on software design, we added a paragraph on why we chose PETSc and
HDF5 as basis of our parallel implementation, as they are mature software infrastructures that are being developed further and
where we will be able to profit from this work just by linking new versions of these libraries. So, in fact, our software design is
305 designed to run well across a wide variety of platforms, now and in the future.

In Section 5.2 the contributions from various pieces of the solver to the total runtime are quantified. The first comment that
can be made here is that if outputting destroys the analysis that much, and you choose not to discuss this phenomenon, why
include it in the curves? It is very distracting, particularly for the lower resolutions since “NetCDF output” is seen to dominate
runtime.

310 Depending in the application scenario, the state of the system will be saved more or less often. So output is an important
component of the overall CUAS-MPI software. Even parallel output is, however, an expensive operation, even with a state-of-
the-art implementation such as HDF5, and that is particularly the case in application scenarios where there is little computation,
i.e., low resolution runs. Figure 5 shows the regime where one output is written daily. In Figure 7a, performance results with
daily and annual output are shown, and in the annual case the cost of output is negligible.

315 More importantly, the conclusion here seems to be that there is a tipping point where scalability breaks, which depends
on the relative costs of the various “subcategories” and the general amount of work per MPI process (G600 is experiencing
scalability issues faster than G150 because it has less elements per threads). Not surprisingly, the curves presented in Fig.
5 and the explanation behind the observed tendencies are similar to what is discussed in depth in Fischler et al., 2022. The
culprit for the inflexion is the same: both system matrix creation and PETSc communications experience scalability issues due
320 to the overhead costs of MPI communications (citation from the 2022 paper: “with an increasing number of MPI processes,
the communication overhead of the assembly starts dominating at some point.”). Unfortunately, the same can be said of Sect.
5.3: the two conclusions starting l. 279 “In general we see that for smaller grids there is no sense in using a large number of
processes, as there is not enough work to be done for an efficient parallelization” and l. 285 “...more MPI processes, in general,
cause more communication and synchronization overhead and more computation is needed to offset this” are similar to those
325 drawn by Fischler et al. 2022, limiting the impact and importance of this study. Is there a way to differentiate this work from
the 2022 paper?

Our manuscript discusses an entirely different model than Fischler et al., 2022. Fischler et al. (2022) was about a model about
ice dynamics, comprising an approximation of the Stokes equation, enthalpy balance, evolution equations for the geometry
with boundary conditions. The equations are discretized using finite elements. CUAS-MPI is a hydromechanical model solving
330 basically a groundwater equation. Both models are concerned with compartments of the Earth system that interact with each
other, but this is all they have in common. Both models are in the same environment, on HPC platforms, and if they are run in
a coupled mode, we find it worth to show the scaling of both applications. Perhaps perspective on code improvements could
be discussed?

- As we present a new implementation of a model, code improvements are not the scope of this paper, but may well be a topic in the next years.
- The first paragraph starting l. 193 should be rephrased
 - l. 195: “()” should be removed around Section 4
We adopt the referee’s suggestion.
 - l. 201: is there a link to learn more about that particular supercomputer?
A link to the system web site was added: <https://www.hrz.tu-darmstadt.de/hlr/hochleistungsrechnen/index.en.jsp> l. 203: what is “turbo-mode”?
Explanation added. "Turbo-mode allows the CPU to exceed its base frequency for short time, but also considerably increases power consumption."
 - l. 205: “relative standard deviation” of . . . ? (the time)
Yes, indeed the time, we will be more precise at this point. Many thanks for pointing this out.
 - I agree with reviewer #1 that the vertical coordinate of Fig. 4 is not consistent with the runtimes reported in the body of the text.
We will check and fix it and apologize for this unwariness.
 - l. 234 - 242: I would refrain from using too many “”, as they clutter the text.
We agree and will use italics instead.
 - The terminology is awkward at times: these so-called functional categories should be put in the context of a workflow akin to Fig. 1 in Fischler et al 2022 (references to Fig 1 are not helping until the layout in Fig 1 is made more descriptive). . .
See our comment on Figure 1 in the comments on section 2.
 - but also, it turns out that these functional categories are not directly the categories used in Fig. 5, so why bother? Explain what is shown in Fig 5 without the confusion of the “functional categories”. I also have no idea what the “NETCDF output” is part of: is it the I/O interface or is it the post-processing of the solution vector which ultimately would be a sub-category of the “CUAS-MPI solver “?
There is, in fact an inconsistency between the labeling of Figure 1 w.r.t. the use of "The CUAS-MPI Solver". We will revise Figure 1 and be more precise which lines from the pseudocode in Figure 1 correspond to the functional categories we measured.
 - l. 248: what does the “They” refer to here?
We reformulated it.

- 365 – l. 267: consider rephrasing that first sentence (why using the past tense here? Maybe I am missing the point).
We will check.
- l. 272: the choice to output a solution here is odd, since this cost is obviously going to dominate the runtime for all coarser grids, based on the results from G600 of Fig 5.
See our previous comment on output at the start of this section.
- 370 – l. 276: I wouldn't say "profitably"
We think "efficiently" fits best in this context.
- Fig 7a: why no more log scale for the x axis of Fig 7a?
Good point — we will prepare Fig7a with a log scale for the revised version.
- Fig 7b and c are difficult to read and the accompanying text is not very easy to understand. Are we supposed to "read" them as going from right to left = increasing the number of MPI processes?
375 Indeed reading them from left to right would be a good way to draw conclusions from them, but similarly looking for minima of the individual curves is an approach. We understand that the current text is not sufficient for the reader and will improve the main text. This figure is actually most useful for planning simulations and as thus potential users do indeed need to be given a good explanation of this figure. Many thanks for pointing this out! Also why is the inflexion (breaking of scalability) starting for a higher number of cells per MPI processes for finer resolutions?
380 We will be more descriptive for this: finer grids have more cells in total than coarser grids. Thereby more processes are necessary to reach a lower number of cells per process. More processes cause more computational overhead, resulting in lower maximum throughput for finer grids.
- l. 279: "Smaller grids"= coarser resolution? (consider changing "larger grids" too)
We adopt the referee's suggestion.
- 385 – l. 288: what is the "sweet spot"?
In this context, we assume that the sweet spot is the minimum runtime and will clarify this in the text.
- l. 292: that last sentence needs rephrasing, not sure what is said here (what is "the spread"?)
We rephrased the sentence.

2.2.7 Section 6

- 390 I suggest merging part of this section with the Conclusions, and including the rest (the discussion about runtime and throughput) in the previous section.
We appreciate the suggestion, but in our view it is appropriate to differentiate between a more expensive discussion and precise formulation of the conclusion.

- I. 296: First sentence should be rewritten. Give specifics -what is high, highest?
395 We have included numbers for the resolution and have rephrased the sentence to 'The throughput shown by CUAS-MPI enables ice sheet wide simulations in high (600 m), but potentially not highest resolution (150 m) due to the computational costs.'
- What does “the code performs well” means ?
400 Yes, we agree 'the code performs well' is too vague as a statement. The intention here was to make the statement that one can compute longer time periods without having simulations that take months' to complete, seasonal simulations, which need hourly time steps and daily output, are taking too long in highest resolution. Highest resolution refers to the resolution of the bedrock topography, which is (currently!) 150 m. We are rephrasing the sentence in the revised version.
- What are the number of time steps required for a seasonal cycle?
405 We compute hourly time steps for the seasonal cycle but we write only daily output. This is of course a choice of the modeller, balancing available forcing data and affordability. It is worth to note, that seasonality does not end once the seasonal melt water availability vanishes, as the system still evolves beyond the end of the melt season.
- The requirements of 56 days for a 90 years run with 600m temporal res seems off, based on the results from Fig 7a. Please clarify. Also this is very much supercomputer-dependent and this should be specified.
410 This is indeed correct. We are into investigating this and will provide updated numbers in the sentence in the revised version. Many thanks for pointing this out!
- I. 301: what are “panGreenland simulations?”
“panGreenland” → “Greenland-wide” - we will change this in the revised version.
- I. 302: what do you mean by that last sentence “The costs are also ...”?
415 When we aim to run coupled simulations for projections, they are typically simulating 100 to 300 years. If we include a seasonal cycle of meltwater input, we end up with a daily coupling between ice sheet and hydrology models to capture the seasonal effect on the ice sheet, too. This is indeed costly, as Fig. 7 shows. This is what is meant here.
- What is the paragraph starting I. 307 doing here? It sounds like it would be better placed in the Introduction
Very well spotted - many thanks! We will move this to the introduction
- I. 313: “is” at the end of the sentence should be removed
420 We adopt the referee’s suggestion.
- I. 315 “In out work, the next step ...” → “The next step ...”
We adopt the referee’s suggestion.

- l. 316: the ISSM citation is odd, the website should be a footnote or a real citation altogether?

425

We will apply according to the GMD guidelines

- l. 320: odd citation for preCICE too

We will apply according to the GMD guidelines

- Paragraph starting l. 315 can be a perspective for, say, the Conclusions section but its placement here is odd. You could merge with the next paragraph but emphasis should be on the test cases results

430

We agree on expanding the conclusion, but we prefer not to shorten the discussion.

- l. 326-333: what is said here ? Is it an extension of the paragraph starting l. 315?

we will rephrase this

2.2.8 Conclusions

This section is too incomplete. It should either be augmented with the suggested paragraphs of Sect. 6, or merged with the discussion section altogether.

435

We believe it to be valuable to separate a discussion of the results from a concise description of our conclusions. However, our conclusion is, in fact, currently very short and we will expand it in the revised version.

- l. 349: CUAS-MPI is not a new code – or if it is it has not been demonstrated properly

440

As we laid out above, CUAS-MPI is written from scratch and we will make this more clear in the revised version of the manuscript.

- l. 355-356: too vague - consider describing what these enhancements and model adaptation should entail. The last sentence should be impactful.

Yes, we admit the last part of the conclusion is neither inspiring nor impactful. We will rephrase the conclusion for the revised version.

445 2.2.9 References

- The doi for the paper by Young, T. J., Christoffersen, P., Bougamont, M., and Stewart, C. L. should read <https://doi.org/10.1073/pnas.2116036119>

We adopt the referee's suggestion.

References

- 450 Beyer, S., Kleiner, T., Aizinger, V., Rückamp, M., and Humbert, A.: A confined–unconfined aquifer model for subglacial hydrology and its application to the Northeast Greenland Ice Stream, *The Cryosphere*, 12, 3931–3947, <https://doi.org/10.5194/tc-12-3931-2018>, 2018.
- Bindschadler, R. A., Vornberger, P. L., King, M. A., and Padman, L.: Tidally driven stick–slip motion in the mouth of Whillans Ice Stream, Antarctica, *Annals of Glaciology*, 36, 263–272, <https://doi.org/10.3189/172756403781816284>, 2003.
- Engelhardt, H. and Kamb, B.: Basal sliding of Ice Stream B, West Antarctica, *Journal of Glaciology*, 44, 223–230, 455 <https://doi.org/10.3189/S0022143000002562>, 1997.
- Ferris, J. G., Knowles, D. B., Brown, R. H., and Stallman, R. W.: *Theory of Aquifer Tests*, Tech. rep., U.S. Government Print. Office, <https://doi.org/10.3133/wsp1536E>, 1962.
- Fischler, Y., Rückamp, M., Bischof, C., Aizinger, V., Morlighem, M., and Humbert, A.: A scalability study of the Ice-sheet and Sea-level System Model (ISSM, version 4.18), *Geoscientific Model Development*, 15, 3753–3771, <https://doi.org/10.5194/gmd-15-3753-2022>, 2022.
- 460 Hoffman, M. J., Catania, G. A., Neumann, T. A., Andrews, L. C., and Rumrill, J. A.: Links between acceleration, melting, and supraglacial lake drainage of the western Greenland Ice Sheet, *Journal of Geophysical Research: Earth Surface*, 116, <https://doi.org/10.1029/2010JF001934>, 2011.
- Siegert, M. J.: Antarctic subglacial lakes, *Earth-Science Reviews*, 50, 29–50, [https://doi.org/10.1016/S0012-8252\(99\)00068-9](https://doi.org/10.1016/S0012-8252(99)00068-9), 2000.