Author's response

The comments in the community discussion and review have been very helpful to sharpen the manuscript. A summary of main changes:

- the verification section has been significantly reduced. Many cases have been moved to an appendix.
- the abstract is now more focused on the key points of the paper
- the context of the SERGHEI project and its outlook are now introduced earlier in the paper, and rounded up in the conclusions.
- we have better explained why there is a growing need for exascale-ready shallow water solvers, and how we think this will benefit the community.
- many minor corrections and clarifications

Herein, we include a point-to-point reply to the reviewer and community comments.

REVIEWER 1

The authors present a very timely and interesting topic introducing an open source modelling framework which in the perspective should be capable of simulating hydrodynamics, transport, morphodynamics, ecohydrology including feedback and interactions on large spatial and temporal scales using modern computational toolboxes and a broad range of high performance computing capabilities.The approach follows the tendency to also use hydrodynamic models for catchment scale cases where in the past hydrological models have been used which however cannot resolve the high resolution information on elevation etc. The paper can be published from my side after major revision.

1. Title: environmental hydraulics is mentioned; the reader may expect more than 2D hydraulics and rainfall runoff -> rethink

We understand the reviewer's point. However, we'd like to keep this as the broader idea of environmental hydraulics is where we envision the applicability of SERGHEI-SWE.

2. Abstract: it describes to a large extent the overall aim of the whole project where of course only parts are included in this paper -> reduce this a bit and therefore comment more on the results you obtained here

Done

3. There are many abbreviations in the text which are not all well known to the readers -> introduce a list of abbreviations

Indeed, this is particularly relevant considering the interdisciplinary target audience. We have included a list of abbreviations in the form of a glossary.

4. P. 4: as you are also addressing urban environments, I suggest to also comment on the linkage to sewer systems – although this might not be planned here

Indeed. It is currently not a high priority, but it is something that we aim for. We have included a comment in the text "...both natural and urban environments (for which coupling to sewer system models is a longer term objective)...".

5. 5: your framework is quite fundamental; why are viscous / turbulent terms not included, e.g. to also in the perspective include eg turbulence models (eg algebraic, k-epsilon); this also might be relevant in the view of environmental hydraulics (title); this might be extended later (include in chap. 7) ?

The reviewer brings up an important point. From a theoretical point of view, the addition of viscous / turbulent terms into the shallow water equations turns the hyperbolic PDE into a degenerate parabolic-hyperbolic PDE, which has some implications for the solution procedure. In general, the addition of such

terms in SERGHEI is straight-forward. We did not consider these terms, because we targeted rainfall-runoff simulations in this work. The term "environmental hydraulics" might have been used too broadly here. We used this term to refer to the use of computational hydraulics for environmental flow problems, in our case surface hydrology at the catchment scale. However, we are aware that the term often is used for river hydraulics—ecology relation investigations. We have written *The observant reader will note that in Equation 1, viscous and turbulent fluxes have been neglected. The focus here is on applications (rainfall-runoff, dam-breaks) where the influence of these can be safely neglected. Turbulent viscosity may become significant for ecohydraulic simulations of river flow, and turbulent fluxes play of course135 an important role in mixing in transport simulations. We will address these issues in future implementations of the transport solvers in SERGHEI.*

6. p. 5: there are also friction laws for small water depths; as rainfall runoff has an important role, should 1 or 2 be mentioned ? Ilhan knows

This is indeed a very relevant point. We have included some comments in the text for this. We do not wish to explore this in this manuscript, but it is certainly very relevant. We now state: *In addition, specialised formulations of the friction slope exist to consider the effect of microtopography and vegetation for small water depths, e.g., variable Manning's coefficients (Jain and Kothyari, 2004; Mügler et al, 2011} or generalised friction laws (Özgen et al., 2011). A recent systematic comparison and in-depth discussion of several friction models with a focus on rainfall-runoff simulations is given in Crompton et al (2020). Implementing additional friction models is of course possible, and relevant, especially to address the multiscale nature of runoff in catchments, but not essential to the points in this paper.*

7. p. 5: a comment that there are also many 2nd order schemes and unstructured grids

This is a good point. We have included the following: *It is relevant to acknowledge that second (and higher) order schemes for SWE are available (e.g. Buttinger-Kreuzhuber et al., 2019; Caviedes-Voullième et al., 2020b; Hou et al., 2015; Navas-Montilla and Murillo, 2018). However, first order schemes are still a pragmatic choice (Ayog et al., 2021), especially when dealing with very high resolution (as targeted with SERGHEI) which offsets their higher discretisation error and numerical diffusivity in comparison to higher order schemes. Similarly, robust schemes for unstructured triangular meshes are well established, together with their well-known advantages in reducing cell counts and numerical diffusion (Bomers et al., 2019; Caviedes-Voullième et al., 2012, 2020a). As these advantages are less relevant at very high resolution, we opt for Cartesian grids to avoid issues with memory mapping, coalescence and cache misses in GPUs (Lacasta et al., 2014) and additional memory footprint, while also making domain-decomposition simpler. Both higher order schemes and unstructured (and adaptive) meshes may be also implemented within SERGHEI.*

8. p. 12, Fig 5: add axis description specific discharge, why is discharge 0 ? or do you show the error or a relative error, unit ? in case it is a (relative) error, add description how it is computed in the text; supercritical flow: is this elevation or water depth ? the water depth goes down to 0 ? add a comment if it becomes very small and is no more visible

Done. The axis labels have been corrected.

9. p. 13, Tab 2: I suggest to explain the computation of the norms in the text add a, b, c, d also in Fig 5; I do not understand the numbers in Tab 2, they are far away from machine accuracy?

The norms in Tab. 2 are not normalised by the number of elements. Thus, in order to obtain an average error, values in the table have to be divided by the number of elements, in this case N = 1000. We added the equations to calculate the L-norms to clarify. There was also an error in the norms concerning the C-property. It is now zero, as expected.

10. chap. 4.3.1: a bit more description here

We have moved this case to the appendix and extended the description a bit.

11. p. 17, Fig. 9: the legends must be larger, add [m] for the axis

Done

12. p. 18, chap. 4.4: please comment a bit more on this rainfall; otherwise the results cannot be understood

We have now included a brief description of the rainfall: "a piecewise constant rainfall with two periods of alternating low and high intensities (50.8 and 101.6 mm/h) up until 2400 s."

13. p. 22, Fig 15: legend text a bit larger here

Done

14. p. 25, chap. 5.7: in the previous case 1cm, what resolution here ?

Same resolution. It is now explicitly stated and we have slightly reformulated the text. We meant to say that in principle, this case really demands this resolution, whereas in the previous case it is not really a requirement of the topology. We have also moved this to the appendix.

15. p. 26: add 1 or 2 sentences to the agreement of model and field results; Fig 20: explain these 3 terms in the text, otherwise the reader cannot understand

Thanks for pointing this out. We have now included a short description of the results, their visualisation and interpretation.

17. chap 4+5: there is a large number of test cases which of course is very impressive; I would include all these test cases in a validation document, however I suggest to reduce this number in a journal paper and possibly shift some cases in the appendix, for example: 4.1: 2 instead of 4 hydraulic cases; 4.2 show only 4.2.2; 4.3 show only 4.3.2; 5.3 show only 1 case; 5.6+5.7: show only 1; please rethink, this does not mean that all should be shifted to the appendix or removed from the paper

Indeed, it is our goal to be very thorough in the verification strategy. We do agree with the comments that the volume of tests distracts from the core message of the paper. Consequently, and since there is a nice consensus from the comments, we have moved many of the tests into the appendix. We still keep a strong verification content, as it is essential to demonstrate the validity and robustness of the implementation. It is also relevant to highlight that most of the case setups (except those which require data we are not at liberty to distribute) are available for the community to also test. We leave one test to show how SERGHEI-SWE conserves the C-property with a wet/dry front, two steady transcritical flumes, one analytical dam-break and one transient flow in a parabolic bowl. We leave one analytical rainfall-runoff case, one experimental dam-break flow in a flume, and one experimental steady flow in a flume. We also keep one experimental smooth transient flow over an island and one experimental rainfall-runoff case in a flume. Finally, we also keep a field runoff experiment and the Malpasset dam break. All others have been moved to the appendix.

18. chap. 6: from small scale to large scale: show 6.2 as 6.1 ?

Yes, this is a good suggestion. We have changed the order.

19. 4, 5 + 6: I understand and it is good that the cases descriptions are very short as most cases have been widely used in the literature; however not every reader know every cases; to avoid that the reader has to search for several other papers to understand the systems and results in the test cases I suggest to add core information to each test case in an appendix, eg: figure of system, initial and boundary condition, core parameters; it is not a must, but would be helpful for the reader

We have extended some of the descriptions, to ensure that all the critical information is available. However, given that there are plenty of reproductions of the setups in the literature, we believe it is not particularly critical to contribute with yet-one-more detailed description or sketch of the cases. We also highlight that most of the cases are openly available to download and run with SERGHEI, where all the details of the setups are necessarily included.

20. chap. 7.2+3: you had that many cases before, therefore I am wondering why not one of the previous cases was chosen or the cases of 7.2+3 should be included earlier -> rethink

This test is particularly interesting for scalability analysis. We have now written out this explicitly and explained why:

"This is a simple analytical verification test in the shallow water literature, which generalises the 1D dam-break solution. We purposely select this case (instead of one of the many verification problems) for its convenience for scaling studies. Firstly, resolution can be increased as will. Additionally, the square domain allows for trivial domain decomposition, which together with the fully wet domain and the radially-symmetric flow field minimises load balancing issues. Essentially, it allows for a very clean scalability test with minimal interference from the problem topology, which facilitates scalability and performance analysis (in contrast to the limitations of the Malpasset domain discussed in section 7.1)."

For verification, this test overlaps with the dam break over a wet bed, which is why we choose not to include it.

21. p. 32, close to end: add references to Ecosys, EcH20

Done

22. chap. 8: is further (substantial) performance improvement possible using domain decomposition with MPI ?

Our multi-GPU and multi-node simulations already rely on domain decomposition and MPI. We have made this a bit more clear in the HPC implementation section for which we have written a new short introductory paragraph explaining the different layers of parallelism:

"In this section we describe the key ingredients of the HPC implementation of SERGHEI. Conceptually, this requires, firstly, handling parallelism inside a computational device (multicore CPU or GPU) with shared memory, and the related portability and corresponding backends (i.e., OpenMP, CUDA, HIP, etc.). On a higher level of parallelism, distributing computations across many devices requires domain decomposition and a distributed memory problem, implemented via MPI. The complete implementation of SERGHEI encompasses both, distributing parallel computations into many subdomains, each of which is mapped onto a computational device. Here we start the discussion from the higher level of domain decomposition."

23. Further minor comments are in the commented pdf. We have addressed the typos and suggestions.

24. Experimental idealised urban dam-break: comment on model results and measurements, does a refined grid improve the agreement. similar deviations found in the literature ? We included some qualitative comments on the quality of the solution. The results are comparable to those in the literature. We do not wish to go deepen the analysis in terms of grid convergence. Proper quantitative analysis of the results, especially for different grids is rather complex for this case, as there are many dynamics and many different factors playing a role. This would require a dedicated effort which is not the point of this paper.

REVIEWER 2

Comments: https://doi.org/10.5194/gmd-2022-208-RC2

1. The paper deals with the development of, what the authors call, Simulation Environment for Geomorphology, Hydrodynamics and Ecohydrology (SERGHEI) - Shallow Water Equation (SWE) for hydrodynamics, ecohydrology, morphodynamics simulations. Although there are many shallow water equation models in the world, I think the SERGHEI-SWE is new in that it can handle even exa-scale problem such as with 122000000 computational cells (dx=0.5m) for the rainfall-runoff processes at lower Triangle region in the East River Watershed with fast parallel computing.

This is a fair summary. We have decomposed the reviewer's comments to address them in order.

2. My impression of the overall contents of the paper was that actually the contents is basically for the evaluation of shallow water model, however, the authors are emphasizing more on the broad view of SERGHEI that it can handle hydrology and environmental hydraulics problems.

We intend to convey both aspects. We specifically describe the SERGHEI-SWE module, and evaluate in terms of accuracy, applicability and performance. Nonetheless, since SERGHEI-SWE is the core hydrodynamic

module of the broader SERGHEI, and because the modularity and broader project shape both the motivation and the design of SERGHEI-SWE, it is relevant and interesting to discuss in this paper as well. We intend to convey that there is a clear outlook in proposing SERGHEI-SWE.

3. I had a feeling that you could simply describe the importance of SERGHEI for the future of pure shallow water equation modelling in the paper, but you didn't.

This is a fair point. As the reviewer points out several times, there is some opportunity in the manuscript to extend on the outlook of HPC (and exascale) enabled shallow water modelling. We have written some remarks on this.

4. Likewise, you do not have to say fluvial or urban flood modelling as classical engineering problem. If you model fluvial or urban flood modelling with e.g. 100 m resolution, you may be justified to say that it is established engineering problem, but I think it is still new if you model them with one-digit resolution and large area with scientific scope such as considering sub, super-critical flow distinction.

From a historical point of view, shallow water solvers received wide attention for fluvial and urban flooding applications, and a large part of the numerical developments, as well as software implementations were tailored to these applications. However, this trend has been systematically changing, opening up shallow water models for broader hydrological applications. This difference is what we mean to highlight here. By "classical", we meant that it is the more established application, but we did not intend to convey that it is one for which everything is solved and that there are no scientific nor technical questions to address. We do agree that flood modelling applications with large domains and very high resolution are NOT an established thing, especially with complex transcritical flow conditions for which simplified models are not well suited. Such applications can of course benefit from this technology. We have included some remarks to clarify this, and better frame the applicability and interest of SERGHEI across different use cases.

5. In this sense, turbulent modelling may be interesting to add more scientific essence in the analysis. Yes, this is of course an interesting point, and it will come up as more relevant as we move towards additional applications, especially those concerning transport. We intend to address this in the future, when we have the transport modules further developed. This is now also mentioned in the text, coinciding with a point from another reviewer.

6. Anyway, I would appreciate it very much if you could explain more why the exa-scale handling is necessary for the shallow water modelling. Lower Triangle region is 14.82km2 which is small. Why do you need to model such catchment with 0.5m resolution? Concerning the first part of this question, please see the reply to point 7 below. There are two parts to the second question, of why do we need to model the Lower Triangle at 0.5m resolution:

a. for the purposes of this paper, there is no need to do this. We chose the Lower Triangle because a high-resolution DEM was available, and because it has a complex topography which should be general enough to test SERGHEI's performance under real conditions. It creates a large computational problem in space, and in time due to high velocities thanks to the high slopes. In this sense, we stress that this is simply a computational performance test. We set out to test the scalability of SERGHEI on something that was close to reality, or a plausible application. We note that the value of such high-resolution data has been recognised by the shallow water community and the Lower Triangle case has been used in a similar manner as a benchmark in Shaw et al. (2021) (https://doi.org/10.5194/gmd-14-3577-2021) to assess the parallel performance of LISFLOOD-FP 8.0.

b. From a broader perspective, and to frame the context of why the Lower Triangle is of interest. The Lower Triangle catchment is part of the U.S. Department of Energy's Watershed Scientific Focus Area (Watershed SFA) project and is intensively monitored, see Hubbard et al. (2018) (doi: 10.2136/vzj2018.03.0061). The catchment is strategically important because it collects snowmelt-based headwater from upstream catchments throughout the year, and thus, the discharge at the outlet of this catchment is an integrator of many hydrological signatures. From a water resources management point, the catchment is important because it contributes significantly to the water supply of the midwest of the United States. Locally, a pump house at the outlet of the catchment provides water for the municipality of Mt. Crested Butte, Colorado, see Maavara et al. (2021) (doi:

10.1371/journal.pone.0247907). One of the aims of Watershed SFA is to link hydrological signatures to geochemical processes. It is well-known that hydrological fluxes create hot spots and hot moments in watershed geochemistry, see, for example Arora et al. (2022) (https://doi.org/10.1007/978-3-030-95921-0_2), Zhao et al. (2021) (https://doi.org/10.1016/j.scitotenv.2020.144168), and Krause et al. (2017) (https://doi.org/10.1002/2016WR019516). These hydrological fluxes are dependent on small-scale watershed features that can only be captured at these high resolutions, for example, infiltration and exfiltration patterns are strongly controlled by microtopography. Thus, in order to properly detect potential geochemical hot spots and study how they scale up to the watershed scale, this high resolution information is required. In the future, we will continue to use this area to test additional modules oriented towards these purposes.

Overall, in the research outlook of the SERGHEI team, sites such as the Lower Triangle are interesting to simulate at very high resolution to study how small scale processes produce large scale hydrological signatures in general. Please see next point for further discussion.

7. What is your vision for the exa-scale modelling using a shallow water equation?

We interpret this point as a request to better explain why there is a need for exascale-ready shallow water modelling. We appreciate this very much, as this is indeed at the core of our proposition with this model.

There is the opportunity to exploit increasingly better and highly resolved geospatial information (DEMs, land use, vector data of structures) which prompts the need for high resolution solvers. At the same time, the push towards the study of multiscale systems and integrated management warrants for increasingly larger domains. Together, these trends result in larger computational problems, and consequently HPC shallow water solvers.

Although it is likely that for purposes such as generating hydrographs (which would be the interest for flood modelling) such high resolution is unnecessary, to capture the spatial dynamics, pathways and transit times (all of which are of ecohydrological interest), it is most likely beneficial. High resolution is necessary for such purposes to avoid some type of a-priori averaging of properties/parameters/processes. We are interested in ecohydrological processes, which will arguably benefit from this larger resolution.

We have included these arguments in the "Conclusions and outlook" section.

8. In addition, it is little hard to follow all the benchmark cases one by one. The total number of page is 44. You can remain for example only the essential benchmark for readers' sake. If you say, that every benchmark is necessary for the true evaluation of the shallow water model by SERGHEI, then you can remain all of them, but you can as well keep them in the appendix for example. Please consider them. Please sharpen the contents for what you really want to convey.

There is a strong consensus from the review comments on this point, and we actually agree that the verification section is far too long. We do keep all of them, but around half of them have been moved to the appendix. The ones we keep intend to show the robustness and correctness of the SERGHEI-SWE implementation for a variety of problems which can be tackled with SWE solvers.

9. Overall, I think there are shallow water models which can simulate benchmarks with the same level accuracy with SERGHEI-SWE except that those are slower than SERGHEI-SWE.

This indeed is likely the case. We do not make any comparative arguments in terms of the accuracy of SERGHEI-SWE. The main reason is that we implement well-established numerical schemes, and we only wish to verify that the implementation is robust and accurate with our verification exercise. Whether other solvers are "faster/slower" is a matter of debate. Our argument is not really one of speed, but of scalability and portability. That is, a very lightweight, well-implemented solver, running on a GPU and using potentially a cheaper numerical scheme may do the job very fast for any given application. However, it would be limited to off-the-shelf consumer hardware. With SERGHEI, what we are proposing is that the community knowledge, developed and gathered in the past decade, can be harnessed into state-of-the-art computational technology, and it can be scaled up, in the same way that many communities in the geosciences and engineering do. Since HPC implementations require additional attention to implementation details, and because we explicitly set out for it, SERGHEI is likely to be a comparatively efficient implementation — although we would not

outright claim this, since we do not wish to engage in an intermodel performance comparison. However, how "fast" it is, is a function of the hardware, reason for which portability is relevant. The point is really, that with SERGHEI, we offer the community a scalable solver, which, sure, allows to tackle "normal" problems faster, but most importantly, enables larger problems and sets a tool that can harness upcoming computing technology.

10. But other shallow water models could be more user-friendly. Likewise, the users of other shallow water models may consider that they do not need exa-scale computation for their purpose. The authors need to explain the vision how the SERGHEI-SWE needs to be used in the near future.

In general terms we agree, although this depends on the definition of user-friendliness. In the context of engineering-oriented shallow water solvers "user friendly" has often meant a sophisticated Graphical User Interface (usually topped with Windows support). This is something mostly out-of-scope for our purposes. We envision SERGHEI mostly as a scientific tool, and assume that users are capable of interacting with the model via a small set of input files. We also provide some R-based interfaces for input/output, which aside from being our tools to pre/post process SERGHEI input/output, also facilitate its usage. Our contribution in terms of user-friendliness is that SERGHEI is easily deployed on a local GPU-enabled workstation. Although we target large HPC facilities, the required dependencies and underlying software stack is quite lightweight. Therefore, setting up and building SERGHEI on a workstation is rather straightforward. Most importantly, our vision of SERGHEI as a community model has required for the framework to be simultaneously performance-portable and to facilitate development by the community (i.e., the code is quite readable, and not obfuscated by hardware specific code or complicated software abstractions). In this sense, it is a different type of user friendliness. We of course don't argue that everyone needs exascale. But most likely every user of shallow water models will benefit from performance, and we hope, from a transparent, open-source community model.

11. The authors do not necessary needs to write the answers to the above questions in the paper itself but I would like to know the vision of exa-scale computing.

As the reviewer can tell, we have addressed these points in an extended discussion. Additionally, some of these points are also succinctly treated in the manuscript, as these issues are relevant to the overall presentation of SERGHEI and its expected use and place in the community.

Georges Kesserwani

Comments: https://doi.org/10.5194/gmd-2022-208-CC2 This is an interesting contribution that I think it would be a nice read if made shorter and more focused, by mainly:

1. reorganising the methodology to ensure that the interesting parts appear first, or are at least introduced. For example, it would help to introduce the concept of SERGHEI here, and narrow down the scope to the SWE module.

Thanks for this comment. That was the original organisation of our manuscript, which we changed as it evolved. However, it does seem to work better by including the overall framework in the introduction.

2. cutting the verification and validation section, significantly, to only focus on the analysis of performance and scaling for selected test cases. The 2D dam-break is a keeper because its symmetricity is important to ensure that the code re-implementation is valid (i.e., illustrating the 1D radial profile for the 2D dam-break test will provide stronger evidence of code validity). Many the test cases can either removed or put in an Appendix.

We have significantly reduced the section on verification and validation, essentially leaving only a single test of each type. We do think this is necessary to establish the accuracy, robustness and applicability of SERGHEI to multiple use cases.

3. The abstract: It can be made more personalised to just reflect on the content of this contribution. For example, the first 10 lines can be cut to a sentence or two to just introduce the concept of SERGHEI and its relevance to portability.

Done

4. The Introduction: What does ESM stand for? Earth System Modelling?

Yes, ESM stands for Earth System Modelling. This was defined already here, but we now include a list of abbreviations for clarity.

5. Line 89-90. The trend to address "rainfall-runoff problems in both natural and urban environments, fluvial problems, and other flows" seems to be the case with many of existing models in Table 1. That may not be the novelty here, as also shown later by the type of test cases reported.

This is a fair point. We have reformulated to express which intentions and aims have shaped our design and implementation. More than a novelty, the point is that SERGHEI is conceived already as a framework for wider usage. This is now better conveyed.

6. The organisation of the methodology: Section 2 and Section 3.1 seem to be overviewed from past published works. They are useful for completeness, but their presence at the start is confusing without any introductory text saying about what is where, and why, with reference to the novel aim and scope of this work. Could Section 3, or the methodology, be titled more specifically? For example, SERGHEI-SWE?

Some context was required at the beginning of both sections. We also take in the idea of providing somewhat more informative section titles. For section 2 we now explain that there is indeed no real novelty concerning the mathematical and numerical formulations. For section 3, we provide an introductory paragraph better explaining the levels of parallelism, and the technical novelties.

7. Verification and validation: The mathematical and the numerical formulations are already valid on existing platform. From this stand, verification should be brief to just confirm that the new implementation, using kokkos, remains valid too. Therefore, there is no need to have the main body of the results for the detailed verification of accuracy and error convergence. This could go into an appendix for the basic test cases. This will keep room for the interesting part to be the center of the main discussion: that is, Section 7 of the performance and scaling.

The extensive verification strategy is important to establish the correctness and robustness of the implementation. As you point out, we are not really intending to verify the underlying numerical schemes. Therefore, and following your suggestion and that of the reviewers, we have significantly shortened the verification section, to keep basically one example of each type of problem. The rest has been moved to the appendix.

8. Vision and future work: Couldn't this part be repositioned, re-edited higher up to be integrated at start the methodology to better scope this paper and set the expectation earlier/clearer on the novelty? Yes, we agree. This actually works better at the end of the introduction. We leave only the reference to potential future innovations as "Outlook" in the conclusions section (as they are not really part of the envisioned framework).