

Geosci. Model Dev. Discuss., referee comment RC2  
<https://doi.org/10.5194/gmd-2021-239-RC2>, 2021  
© Author(s) 2021. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on gmd-2021-239

Anonymous Referee #2

---

Referee comment on "Enhancing the accessibility of unified modeling systems: GFDL System for High-resolution prediction on Earth-to-Local Domains (SHIELD) v2021b in a container" by Kai-Yuan Cheng et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-239-RC2>, 2021

---

This paper presents two approaches containerized solutions (over Docker and Singularity) to run GFDL SHIELD v2021b. In the last part of this work, the authors discuss and compare running the containerized model into three kinds of different hardware (Laptop, Desktop, and SuperComputer) and analyze the results and performance differences.

The quality of the paper is good and I don't have any major comment or concern about the discussed technical aspects but I think some minor aspects should be addressed:

- P2, l39: "A software container, or simply container, is a stand-alone, executable software ..." -> This is not correct, the container is a software artifact that can be instantiated (in an image) and the runtime can "run them". This is better explained later on in the paper (e.g. P3, l38). Could you please rephrase this?
- P3, l71: "As a result, a container is lightweight and fast." -> This is not necessarily true, containers can be big and problematic to build or run. In fact, your image is 576MB which is far from lightweight. I believe that if you want to mention one benefit of using containers (that is aligned with your work) it must be "portability". Could you please rephrase this?
- P3, l90: Recently, Docker has changed their licensing (<https://www.docker.com/pricing/faq>) and this could be problematic for reproducibility (which is one of the main topics of your work), I suggest mentioning this in the paper and some open alternatives (such as Podman, <https://podman.io/>).
- P10: I think you can find interesting (and very related to this section) the paper: "Montes D, Añel JA, Wallom DCH, Uhe P, Caderno PV, Pena TF. Cloud Computing for Climate Modelling: Evaluation, Challenges, and Benefits. Computers. 2020; 9(2):52. <https://doi.org/10.3390/computers9020052> "
- P12: Open question (that I believe can add more value to the conclusions): Can the model run in more than one container? If so, what would be the potential benefits of running it on an orchestration system such as Kubernetes or AWS ECS?
- Is there any public registry with an available SHIELD image that can be consumed? Building the image every time is a time-consuming task (>40 minutes for your case, on

a high-end machine) and could also lead to a potential issue with reproducibility (e.g. broken dependency). If so, could you please link it to the paper?