EGUsphere, referee comment RC1
https://doi.org/10.5194/egusphere-2022-943-RC1, 2022

Comment on egusphere-2022-943

Anonymous Referee #1

Referee comment on "Pace v0.2: A Python-based Performance-Portable Atmospheric Model" by Johann Dahm et al., EGUsphere,
https://doi.org/10.5194/egusphere-2022-943-RC1, 2022

Summary:
In this paper, Dahm et al. present a Python based implementation of the FV3 dynamical core that is compatible with both CPU and GPU hardware. Their Python implementation of the FV3 dycore is a step towards solving the "software productivity gap", aiming for a more productive development process through the use of a high-level programming language compatible with modern computing hardware. The authors demonstrate this by addressing their code porting strategy, describing optimization methods and showing scaling results, and by comparison with a baroclinic instability benchmark case. Limitations are addressed with a view towards future developments including parameterizations and CPU optimization. Overall, the manuscript makes a valuable contribution to the field, and is suitable for publication (provided the comments below are addressed to a reasonable extent) as a model development and technical paper.

Key to comments below:
LXX = Line XX
PXX = Page XX
FXX = Figure XX
TXX = Table XX
RC = Reviewer comment
[ ] -> [ ] = Suggested replacement

Major:
1. F1: Workflow pipeline …
RC: The pipeline figure is reasonable - however, please include a brief statement on the role of DaCe in this pipeline (either in the text associated with F1, or in the F1 caption). Figure 1 is referred to in line 98, however DaCe is not commented on until lines 125, and Section 5.

2. L12: Pace demonstrates how a high-level language can … facilitate the integration with new technologies such as machine learning.
RC: Are there currently available demonstrations of machine learning methods being used in Pace? If so, it would be useful to include them in the paper. Figure 13 and the text

in Section 6.2 address this to some extent. It is useful to include changes that "could be" (caption Fig 13) made to the model to highlight potential use-cases, however I believe a model development and technical paper must emphasize elements that are currently operational.

3. F3: RC: Is the model constrained to a specific floating point precision type? Have results been verified in all supported float-type options?

4. L205: … we occasionally needed to extend GT4Py …
RC: Is there clarity in when these extensions were/will be needed ? Are these discovered as the code is being ported, or did the design stage make it abundantly clear that certain extension patterns would be necessary? (I.e. do you anticipate additional scenarios where the DSL needs to be extended beyond those highlighted in section 4.1, especially when considering parameterizations?)

5. T1: RC: Exactly what is being timed ? Please be more precise in the table caption. (There is mention of the absolute runtime in L267 but the table needs a more appropriate caption).

6. F10: axis label [time per timestep]
RC: The units appear to be inconsistent with those in Table 1.

7. L314, L351: "certain conditions"
RC: Multiple references to possible optimization based on specific conditions - documentation of such conditions would improve the quality of the paper and enable readers to make decisions on porting strategies if a similar exercise were to be carried out in the future.

8. L333: We leverage DaCe …
RC: See comment (1).

9. L402: "We have preliminary results … competitive performance …"
RC: I believe including these (quantitative) results strengthens the paper by supporting the general statements made on the feasibility of incorporating parameterizations. Include quantitative measures of competitive performance here.

10. L433-435:
RC: Clarify - is the intent to represent Pace v0.1 as a complete atmosphere model with physical parameterizations or as the underlying dynamical core? The model development paper in review addresses the dynamical core component of a weather / climate model. As such I believe the statement "We have shown the advantages of a climate model written in a high-level language" should be rephrased to emphasize this. Following the code documentation, an alternative way to support the current statement in L434 could be to elaborate on the `fv3gfs-physics` component with supporting results (This is related to lines 401-404).

12. Plain Language Summary: Similar to (11) - "We re-wrote a Fortran code that simulates weather and climate into Python …" If you are suggesting here that, as submitted, Pace v0.1 is a Python weather and climate simulator then this must be supported by benchmarks in more complex scenarios (with the appropriate physical parameterizations included).

Minor:
1. L14: "Current weather and climate models are written in low-level compiled

languages…”
RC: Please support this with appropriate citations.

2. L19: “There [are] a handful of successful … “

3. F5: “… [figrue] …” -> “… [figure] …”

4. L47: RC: Please include an explicit citation for the FV3 dynamical core. I notice that this is done in line 131, but it is useful to include the citation at the first mention of this in the text.

6. L99: RC: Are there any pitfalls when handling “generated” Numpy code?

7. L119: [horizontal difference of the u and v contravariant velocity components respectively.]

8. L203: “… [lanches] …” -> “… [launches] …”

9. L284: [optimization].
RC: The text in this section in general is within the stated scope - the paper demonstrates portability but it does not demonstrate fully optimized CPU portability (authors state this is future work).

10. F11: RC: Is this simulated time per timestep? Please clarify the precise measure being timed in the figure caption.

11. L288: …[across simulation scale]
RC: … [across simulation scales] …

12. L408: “… as explained in [Section 5] …”

13. L414: Mixed tenses, please fix.

Other: Please include Zenodo DOI references to `Pace` and `gt4py` in the code availability section in the manuscript. (The code is otherwise accessible from the "assets" section on the preprint submission page.)