Responses to Reviewer #1

The aim of the manuscript is to demonstrate the benefits of having a SCM version of a global model to be able to test new parameterization in a more simplified environment than the full 3-D model. In the work of implementing a new parameterization suite, it is useful to be able to do technical tests in such a framework and the description of the workflow can be useful information to have documented. However, as the manuscript is written it is not clear that this is the main purpose of the paper as there are some discussion on results simulation results as well. It is very few testcases, three cloud cases and one idealized tropical storm. The cases and the discussion of the results is not very insightful and does not add to the scientific literature. Thus, I am quite puzzled with the purpose of the paper. Maybe it is worth publishing if the authors concentrate the storyline on the workflow and describe the choices and purpose of the testcases in a more general way. But as it is presented now, it is not sufficient on the methodology and the discussion of the results does not show insight to the processes that are studied and therefore I recommend reject.

Major response: We are grateful to this reviewer for the detailed and constructive comments (in blue text). Substantial revision has been made based on the comments, especially the SCM simulation and the conclusion sections. Regarding the model tests, the focus in the preprint is on the simulated rainfall and cloud when the physics package is coupled. Therefore, three SCM test cases including tropical convection, shallow convection, and stratocumulus are selected. In the revised manuscript, another two SCM cases, the GATEIII and the CGILS, have been added for further evaluation. The 3D simulations (short-term and long-term) are necessary to demonstrate that the transferred physics suite works reasonably in the new model system. Two coupling strategies for 3D simulations are studied via the idealized tropical cyclone test.

We have clarified the purpose of this paper more clearly. One main purpose is to demonstrate the usefulness of a single column model in coupling the CAM5 physics suite to a 3D model, not just to establish an SCM workflow for GRIST (although this is also very important for research and development of model physics). There are two major points that potentially make this work valuable, at least to the model development community.

First, to the authors' knowledge, transferring a physics suite to a new modeling system is nontrivial and should not be taken for granted (e.g., Ma et al. 2013; Giorgetta et al. 2018). But this procedure is generally lacking of a guiding principle. It becomes rather cumbersome if the target physics package does not have a "*highly-modular*" structure (i.e., all codes of a physics scheme are contained in one single-source file that does not have or only has minimum infrastructure dependency). As GRIST already provides a flexible and verified workflow to incorporate different physics packages and physics-dynamics coupling strategies, we wish to reduce the potential uncertainty for implementation of the physics suite as well. The full-physics model simulations are often directly compared in a 3D environment instead of a more-constrained configuration (i.e., SCM). As put by Giorgetta et al. (2018), "*details of the implementations can be quite intricate and nontrivial to resolve. It is neither clear that the physics has been implemented entirely correctly in the old model, nor can we know whether the implementation in the new model is entirely correct.*" In this paper, we demonstrate that an SCM and its experimental protocols can be a good standard to examine the performance of model physics in a more-constrained environment. The usefulness of SGRIST1.0 is demonstrated in two aspects. Firstly, it ensures the correctness of the import procedure of the CAM5 physics suite to a completely different modeling system via comparison between SGRIST1.0 and SCAM simulations. It isolates the impact of the physics package and assesses its behavior in the absence of 3D dynamics, thus separating the physical effect from the dynamical one. Secondly, the behavior of the CAM5 suite in the GRIST model can be evaluated with CAM in 3D test cases to further verify that the transfer is correct.

The second point is more specifically related to the CAM5 physics suite. The CAM5 suite has a sophisticated internal data dependency entangled with its infrastructure, making such a transfer not that easy. A very careful transfer approach should be used. There are technical code changes inside the physics suite for coupling with GRIST (i.e., uncertainty of implementation). Using an SCM has largely reduced the debugging and testing efforts compared with a 3D model, helping to extract this physics package more efficiently and effectively and understand its isolated behavior first in a rather cheap manner (this is actually what this reviewer has pointed out). We use the 3D simulations to further support the usefulness of using an SCM. The comparison of 3D modeling in both short-term deterministic and long-term statistical tests verifies that the transferred CAM5 suite in GRIST produces reasonable results. Of course, we will further tune this suite in a 3D model, but the associated efforts have been largely reduced thanks to the use of an SCM. We believe this work provides an example for people who are interested in developing a new modeling system, providing that they want to use a physics suite (as intricate as CAM5) from an established modeling system.

Specific comments:

Abstract, Line 21: “. . . of the simulated storm.” What storm? Why?

Response: “The storm” refers to the simulated storm in the idealized tropical cyclone test case. To express more precisely, we have modified the sentence as “Two physics-dynamics coupling strategies are examined via an idealized tropical cyclone test case and found to have an impact on the intensity of the simulated storm.”

Line 45: “An AGCM coupled with a full physical . . .”. It is very unclear what it means that an AGCM can be evaluated using aqua planet experiments. Do you mean numerically or what? All AGCMs need parameterizations of physics.

Response: Yes, you are absolutely right that “*All AGCMs need parameterizations of physics*”, and we have no problem with this point. In the text, we are emphasizing on the hierarchical construction and evaluation of an AGCM. To avoid ambiguity, we have slightly modified that sentence.

The aqua planet experiment (APE) proposed by Neale and Hoskins (2000) is one component of the modeling test hierarchy that is of increasing complexity. The simplified short-term deterministic test cases (e.g., Jablonowski et al., 2008; Ullrich et al., 2012) can be only used to evaluate dry dynamical cores, while the real-world atmospheric simulations based on observation data, like AMIP, have been proven to be difficult to identify reasons for model errors since the inherent complexity of AGCMs and surface boundary interactions. The APEs that allow full physics-dynamics interaction in an AGCM under simplified surface boundary conditions provide an intermediate complexity configuration and drive models towards some expected behaviors, such as a stable and reasonable Hadley cell (Blackburn and Hoskins, 2013). Therefore, an AGCM can be evaluated with APEs to display its basic feature of simulating Earth climate and expose its characterizations on circulation, precipitation and so on (Tomita et al., 2005; Zarzycki et al., 2014; Medeiros et al., 2016; Zhao et al., 2016; Yang et al., 2017).

Line 50: “APE runs are usually evaluated by qualitative analysis” – what other ways are you suggesting since you use the word usually?

Response: Sorry for the confusion made by the wording. We delete the word “usually”. Neither analytical solution nor observation is available in the APE. The APE runs are evaluated by qualitative analysis or comparison with other model outputs.

Line 53: You are using the term “model error” in an unprecise way. There are several layers of model errors and several methods needs to be used to assure that the model is useful, it is not only dynamics and physics and their interaction that introduces errors.

Response: Thanks for your reminder. The sentence is modified as “Identification and interpretation of the model errors in the AMIP experiments depend on the dynamics, physics, the nonlinear interaction between them, the complex topography and boundary interactions, and some other factors (e.g., forcing data), all of which may introduce uncertainty that hinders the understanding of the model behaviors.”

Line 87: What is meant by a validated parameterization suite? That they technically work together? That they represent all physics?

Response: Yes, when transferring a physics suite across different modeling systems, it is very important to have a validated parameterization suite with minimum uncertainty. The physical parameterization suite is from CAM5 and has been evaluated and validated in CAM5 (Neale et al., 2010). In GRIST, the physical parameterization suite is evaluated with the SCM test cases before 3D simulation, since the implementation of the physics package code to GRIST might introduce uncertainties (see Major response). The “validated parameterization suite” in line 87 refers to what has been tested with the SCM cases in SGRIST1.0. To avoid confusion, the “validated parameterization suite” is modified as “parameterization suite”.

Line 122: This a very limited type of SCM where only part of the parameterization suite can be tested as you are not carrying any equations for momentum. The performance of the PBL scheme cannot be examined. What about the interaction with the surface?

Response: We are aware that there are more comprehensive SCMs that also include the momentum equation. But the focus here is to compare SGRIST with SCAM to demonstrate the correct implementation of CAM5 physics suite. SCAM also adopts such a formulation, so we did not add additional momentum equations for SGRIST, but this can be a point for future extension. Also, because of the lack of the large-scale forcing data for the momentum equations in the IOP dataset, only the thermodynamic budget is treated in the SCM (Hack and Pedretti, 2000; Zhang et al., 2013 and 2016; Gettleman et al., 2019). The momentum component can be evaluated in 3D simulations in this paper.

The boundary layer turbulence is necessary to all the SCM test cases. The thermodynamic process of the PBL parameterization is examined. We add 2 figures (Figure 2 and 5) to illustrate the water vapor budget of each physical parameterization, including the PBL scheme. In addition, sensible and latent heat fluxes are provided by the IOP data.

It is also quite strange that you put the subscript "obs" in the equations. Is it really observations you are relaxing to? Where do you get the large-scale terms from? What is the relaxation time scale?

Response: The subscript "obs" stands for observed values of the Intensive Observation Period (IOP) data, as same as Hack and Pedretti (2000). The large-scale forcing terms are provided by the IOP data, which is derived from observation and 3D simulations of CAM. Only the CGILS simulations apply relaxation. The time scale is 3 hours. We have added this explanation in the revised manuscript.

Line 141: Here is it stated that the vertical velocity comes from observational data, so called IOP data. Is that really the case?

Response: Yes, the vertical velocity is provided by the IOP data.

Line 155: If you are not using prognostic aerosols, what are you doing? Neglecting or using prescribed values?

Response: Aerosols are currently neglected in the SCM test cases and the aqua planet configuration for both GRIST and CAM.

Line 161: How were the three cases selected and why?

Response: We care about the representation of cloud and precipitation after the physical parameterizations are coupled. The oceanic test cases are considered since the land surface model has not been used in the current SCM version. Deep and shallow convection, and stratocumulus are three common phenomena over tropical ocean. Hence, the tropical convective precipitation test case (TWP-ICE), the shallow tropical trade cumulus case (BOMEX), and the stratocumulus cloud case (DYCOMS-RF02) are selected to evaluate the coupled parameterization suite. Another tropical convection case (GATEIII) and a long-term integration test case (CGILS) have now been added in the revised manuscript.

Line 168: Is not standard SCAM using 31 levels? This is thus purely technical test (or possibly numerical on the time-step integration and relaxation) of the SGRIST as it is the same as in SCAM, or what am I missing?

Response: The standard (S)CAM5 has 30 full-levels (or called layers) and 31 interface levels. For comparison with SCAM, SGRIST1.0 are run at the same model levels to exclude possible influences of vertical resolution.

Line 192: If you see a deficiency in the midlevel humidity, would that not come from the large-scale fields, that you are relaxing to, and not the parameterisations? What parameterization would create that?

Response: Thanks for your reminder. Both SGRIST1.0 and SCAM simulate the midlevel humidity close to the IOP data in the convection active period where the large-scale forcing is relatively stronger. The large-scale forcing becomes much weaker in the suppressed period, and the physical parameterizations play a more important role. Hence, we speculated that the deficiency of midlevel humidity in the suppressed period is attributed to the physical parameterizations. Figure 2 in the revised manuscript shows that the evaporation in the convection scheme plays the most important role in the water vapor budget at 500 hPa, while the large-scale evaporation (condensation) is barely not active in the suppressed period. The lack of interaction between the dynamical forcing and the large-scale condensation might be another reason. We have deleted the figure of midlevel humidity in the revised manuscript, looking forward that future full tests may help to confirm the reasons.

Line 240: The conclusion that cloud variables are sensitive to the thermodynamical environment is trivial.

Response: Thanks for your reminder. The statement about the sensitivity of stratiform cloud parameterization is modified as: Figures 2 and 5 (in the revised manuscript) show that the large-scale condensation (evaporation) plays a more important role in the TWP-ICE and DYCOMS-RF02 cases than in the GATE III and BOMEX, and the sensitivity of water vapor budget to time step is more discernable. The PBL turbulence and the deep and shallow convection schemes are affected by the stratiform cloud scheme, showing an opposite time step sensitivity to balance the budget. A possible reason for the sensitivity of the stratiform cloud scheme in SCAM to time step lies in the macrophysics and microphysics being sequentially split. Macrophysics is the main source of cloud liquid water for microphysics. Gettelman et al. (2015) indicated that the forward-

Euler time integration scheme was used in the microphysics parameterization, and a longer time step facilitates the depletion of cloud liquid, resulting in more condensation by macrophysics for the next time step. This sensitivity to time steps is also affected by the time integration in the dynamical core, although the physics-dynamics interaction in a SCM is limited. SGRIST1.0 uses the third-order Runge-Kutta time integration scheme instead of the leapfrog scheme that used in SCAM. The higher order time integration in SGRIST1.0 decreases the sensitivity of stratiform cloud scheme to the time steps in comparison with SCAM.

Line 246: What exactly have you verified? That the code import was successful?

Response: Yes. As transferring the CAM5 physics suite to GRIST is the general background for this model development effort, it is very important to have a verified physics suite. As we have mentioned, such a transfer should not be taken for granted (as in the Major response). We use the SCM to ensure a safe and successful transfer of the CAM5 physics package to a new modeling system. We believe that the strategy we used for this transfer is useful for the model development community.

Line 258: Dribbling strategy is not clear and the reference to se_ftype0 does not make sense.

Response: We add references (Thatcher and Jablonowski, 2016; Zhang et al., 2018) which have a clear description of the dribbling strategy (se_ftype0) in the CAM model.

Ptend_f1_f1 for GRIST incrementally adds the physics tendencies to the dycore and the tracer transport, respectively. Within one model time step, the physics tendencies are applied as a constant source term in the dynamical sub-steps. The difference between the ptend_f1_f1 and the dribbling strategy (se_ftype0, Thatcher and Jablonowski, 2016; Zhang et al., 2018) used in CAM5-SE is that ptend_f1_f1 treats the dycore and passive tracer transport as two separate components, while the dribbling approach treats them as a whole. This is a subtle difference.

Line 263: "Evolution of the storm highly resembles the weather process . . ." This sentence gives a very strange impression, not sure what you mean, maybe reveals what your expectations of what a model is intended to do, or?

Response: We modify the sentence as "The simulation of the idealized storm with strong physics-dynamics interactions provides an assessment of the deterministic forecast behavior of GRIST."

Line 268: How do you motivate a time step for the physics to be longer than the dynamics?

Response: In general, time steps are selected to satisfy numerical stability (without degrading quality), and as large as possible for computational performance. As a horizontally-explicit model and for this CAM5 climate physics, it is overall efficient to use a longer physics time step (as CAM5 does). For high-resolution meso-to-cloud scale modeling, GRIST has a fast-slow physics separation due to other considerations (fast-physics may have a close-to-dynamics timestep to ensure accurate saturation adjustment), but it is still hoped that more schemes can use a larger step provided that the performance is not degraded.

Line 269: Why do you have one figure in supplement?

Response: Figure S1 is similar with Figure 6, both of which show the evolution of the idealized tropical cyclone. So, we put it in the supplement. The difference is that Figure 6 shows the simulation for ptend_f2_sudden configuration while Figure S1 the ptend_f1_f1.

Line 270: Why do you repeat the time steps here as you have them in the table?

Response: Thanks for your reminder. We have deleted the repeated statement.

Line 282: I would not like to have any conspicuous behavior in a model (at least not some that I know about!).

Response: We delete it in the revised manuscript.

Line 287: Normally one would expect a relationship between time step and model resolution, thus not all your combinations are worth testing and it is no surprise that they fail. However, one can possibly learn from the model if investigating why and where instabilities arise. Have you looked into that?

Response: We agree with your comment "one would expect a relationship between time step and model resolution", i.e., in the form of XX second/km for users. Here, we perturb the time steps to examine the model sensitivity. Due to the uncertainties of physics after its transfer across modeling systems, we believe that it is necessary to do such a time-step sensitivity test to understand some basic behaviors of model physics. When testing different physics-dynamics coupling strategies, we set these time step combinations to compare the ptend_f1_f1 and ptend_f2_sudden coupling strategies. Ptend_f2_sudden survives all the time step configurations, while ptend_f1_f1 cannot. Ptend_f2_sudden is better than ptend_f1_f1 in terms of longer-step stability, which is due to its pure operator splitting nature as also verified in previous simple-physics tests (Zhang et al., 2020).

We speculate that the instability of ptend_f1_f1 at long time steps comes from the nonlinear interaction between dynamics and physics. The physics increment of the ptend_f1_f1 is split into equal chunks and each chunk forces the sub-cycled advection in the dynamical core. The frequency of physics forcing applied in the dynamical core increases with the increasing time step. Firstly, the negative value of water vapor, caused by physical processes, will be fixed to zero. More frequent physics forcing on water vapor causes the fixer to be more active, and thus affects the water vapor budget. Secondly, the atmospheric state has been updated after the first sub-cycled dynamical core step that is adjusted with the physics increment, while the forcing of physical processes, especially those fast processes, remains the same as the previous physics time step. It may cause instability when using a longer time step.

Line 288: "This is consistent. . . "What is consistent and why?

Response: The conclusion about the stability of coupling strategies in full-physics simulations (i.e., ptend_f2_sudden has the largest long-time-step stability) is consistent with Zhang et al. (2020) that used simple-physics forcing.

Line 309-end: I do not think the concluding statements really contributes to any general understanding and they are not new. They are formulated for a specific model system and that limits even more the interest for them.

Response: We have modified the summary section of the manuscript. We believe that "*using an SCM as a guiding principle for transferring an established physics suite to a new modeling system*" is a general strategy, and can be adopted by the broad model development community.

Line 420: What do you mean by the sentence starting with "A way . . ."?

Response: Per the current policy on code sharing at Chinese Academy of Meteorological Sciences, the GRIST model code can be accessed via a registration-like procedure (but not freely open yet). We have provided to the topical editor an assessable Zenodo link of a frozen code version for manuscript review before the preprint was posted. The version of the CAM5 physics package and the associated input and output data can be freely accessed on Zenodo.

References

Gettelman, A. and Morrison, H.: Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes, J. Climate, 28(3), 1268–1287, 2015.

Gettelman, A., Truesdale, J. E., Bacmeister, J. T., et al.: The Single Column Atmosphere Model version 6 (SCAM6): Not a scam but a tool for model evaluation and development, J. Adv. Model. Earth Syst., 11, 1381–1401, 2019.

Giorgetta, M. A., Brokopf, R., Crueger, T., et al: ICON-A, the atmosphere component of the ICON Earth system model: I. Model description. J. Adv. Model. Earth Syst., 10, 1613–1637, 2018.

Hack, J. J. and Pedretti, J. A.: Assessment of solution uncertainties in single-column modeling frameworks, J. Climate, 13(2), 352-365, 2000.

Jablonowski, C., Lauritzen, P. H., Nair, R. D., and Taylor, M.: Idealized test cases for the dynamical cores of Atmospheric General Circulation Models: A proposal for the NCAR ASP 2008 summer colloquium, available online: http://www-personal.umich.edu/~cja-blono/DCMIP-2008_TestCaseDocument_29May2008.pdf., 2008.

Ma, P. -L, Rasch, P. J., Fast, J. D., et al.: Assessing the CAM5 physics suite in the WRF-Chem model: implementation, resolution sensitivity, and a first evaluation for a regional case study. Geosci. Model Dev., 7, 755–778, 2014.

Medeiros, B., Williamson, D. L., and Olson, J. G.: Reference aquaplanet climate in the Community Atmosphere Model, Version 5, J. Adv. Model. Earth Syst., 8, 406–424, 2016.

Neale, R. B., Chen, C. C., Getteman, A., et al: Description of the NCAR Community Atmosphere Model (CAM 5.0). NCAR Technical Note. NCAR/TN-486+STR, pp 108–117, 2010.

Thatcher, D. R. and Jablonowski, C: A moist aquaplanet variant of the Held-Suarez test for atmospheric model dynamical cores. Geosci. Model Dev., 9, 1263–1292, 2016.

Tomita, H., Miura, H., Iga, S., Nasuno, T., and Satoh, M.: A global cloud-resolving simulation: Preliminary results from an aqua planet experiment, Geophys. Res. Lett., 32, L08805, 2005.

Ullrich, P. A., Jablonowski, C., Kent, J., Lauritzen, P. H., Nair, R. D., and Taylor, M. A.: Dynamical core model intercomparison project (DCMIP) test case document, Tech. rep., 2012.

Yang, Q., Leung, L. R., Lu, J., Lin, Y.-L., et al: Exploring the effects of a nonhydrostatic dynamical core in high-resolution aquaplanet simulations, J. Geophys. Res. Atmos., 122, 3245–3265, 2017.

Zarzycki, C. M., Michael, N. L., Jablonowski, C., et al.: Aquaplanet Experiments Using CAM's Variable-Resolution Dynamical Core, J. Climate, 27, 5481-5503, 2014.

Zhang, M., Bretherton, C. S., Blossey, P. N., et al.: CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models, J. Adv. Model. Earth Syst., 5, 826–842. 2013.

Zhang, M., Somerville, R. C. J., and Xie, S. C.: The SCM concept and creation of ARM forcing datasets. Meteorological Monographs, 57, 24.1–24.12, 2016.

Zhang, K, Rasch, P. J., Taylor, M. A. et al: Impact of numerical choices on water conservation in the E3SM Atmosphere Model version 1 (EAMv1), Geosci. Model Dev., 11, 1971–1988, 2018.

Zhang, Y., Li, J., Yu, R. C., et al: A Multiscale Dynamical Model in a Dry-Mass Coordinate for Weather and Climate Modeling: Moist Dynamics and Its Coupling to Physics, Mon. Wea. Rev., 148, 2671-2699, 2020.

Zhao, C., Leung, L. R., Park, S., et al.: Exploring the impacts of physics and resolution on aqua-planet simulations from a nonhydrostatic global variable-resolution modeling framework, J. Adv. Model. Earth Syst., 8, 1751–1768, 2016.