LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-JRNL-804820

Performance and Accuracy Implications of Parallel--Split Physics--Dynamics Coupling in the Energy Exascale Earth System Atmosphere Model

A. S. Donahue, P. M. Caldwell

February 10, 2020

Journal of Advances in Modeling Earth Systems

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States government or Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Performance and Accuracy Implications of Parallel–Split Physics–Dynamics Coupling in the Energy Exascale Earth System Atmosphere Model

Aaron S. Donahue [1], Peter M. Caldwell [1]

[1]Lawrence Livermore National Laboratory, P.O. Box 808, Livermore, CA 94556

Key Points:

- Parallel coupling of physics and dynamics improves model scalability.
- Mass conservation and stability are major concerns for parallel splitting implementation.
- Parallel splitting may require a more restrictive model time step.

Corresponding author: Aaron Donahue, donahue5@llnl.gov

Abstract

Simultaneous calculation of atmospheric processes is faster than calculating processes one at a time. This type of parallelism is beneficial or perhaps even necessary to provide good performance on modern supercomputers, which achieve faster performance through increased processor count rather than improved clock speed. The scalability of the E3SM Atmosphere Model (EAM) is limited by the fluid dynamics which scales up to the number of mesh cells in the global mesh. In contrast, the suite of physics parameterizations in EAM are scalable up to the total number of physics columns, which is an order of magnitude greater than the number of mesh cells. A proposed solution to unlocking the greater potential performance from the physics suite is to solve the physics and dynamics in parallel. This work represents a first attempt at parallel–splitting of the grid–scale fluid dynamics model and the sub–grid scale physics parameterizations in a global atmosphere model. We will demonstrate that switching to parallel physics–dynamics coupling extends the scalability of the EAM to up to three times the previous peak scalability limit and is up to 20% faster than the sequentially split coupling at the highest cores counts and the same timestep. Decadal simulations of both coupling approaches show very little impact to the model climate. This improved performance does not come without drawbacks, however. Parallel splitting requires a shorter timestep and other modifications which largely offset performance gains. A mass fixer is required for conservation. Techniques for mitigating these issues are also discussed.

Plain Language Summary

In order to improve the computational performance of global atmosphere models the community must look at the impact of switching to more parallel splitting of individual processes, thus increasing the amount of work that can be distributed over a large number of computational units. This work looks at the impact to both accuracy and computational performance when the two most expensive components in the E3SM atmosphere model (EAM) are coupled in parallel instead of sequentially. The results are mixed: model climate does not appear to be impacted by the switch in coupling, however the performance gains were modest and fall short of the predicted improvements. This work discusses exactly why the performance gains fell short and points out the major infrastructure changes that would be required to extract the maximum performance gains. The results here will be helpful to any effort to adopt parallel-splitting in any multi-physics model.

1 Introduction

The next generation of high performance computing systems will be faster due to greater parallelism rather than faster chips. Another trend has been the shift towards hybrid systems that employ a combination of central processing units (CPUs) and graphics processing units (GPUs) [*Lapillonne and Fuhrer*, 2014]. GPUs show promise in that they provide an order of magnitude more processing for the same relative power consumption. Their downside is that they use simpler processors and porting code from CPUs to GPUs presents its own challenges [*Michalakes and Vachharajani*, 2008; *Bertagna et al.*, 2019]. In order to continue solving bigger and bigger problems the global climate modeling community must embrace these new systems, which often have hundreds of thousands of processing units.

In the past, when the next generation of high performance computing meant faster processors, a significant speed–up in the time–to–solution could be obtained simply moving to a new system. In particular, models had already been developed to scale up to the available number of computational cores. Next–generation machines, however, require more parallelism than current models can provide. This requires a revisiting of model implementation with a focus on how to produce code that can take advantage of more and more processing units. In other words, global climate models must explore ways to generate more parallelism.

Global atmosphere models solve a complex global system that spans a large range of temporal and spatial scales. To handle this, the atmosphere is separated into two parts: the resolved scale fluid dynamics and the sub–grid scale physical parameterizations. Physics and dynamics are typically coupled together using a sequential splitting (also known as time–splitting) approach. As described in more detail in Section 2, in sequential splitting (SS) each process is computed sequentially and its effect is felt by subsequent processes. An alternative approach which has not yet found wide–spread use but has recently attracted attention is parallel splitting (PS), whereby the tendency for each process is calculated independently and all process tendencies are combined at the end of each time step. Historically, evaluation of parallel splitting in climate models has focused on accuracy and climate impact rather than computational efficiency [*Beljaars*, 1991; *Williamson*, 2013; *Barrett et al.*, 2019]. One exception is *Balaji et al.* [2016], which showed that run-

ning radiation in parallel with other physics processes enabled more frequent radiation calls without a penalty in time–to–solution.

Here we focus on applying the physics suite in parallel with dynamics because –at least for the Energy Exascale Earth System Model (E3SM)– physics and dynamics are the two largest contributors to atmosphere runtime. Thus parallelizing them provides the biggest opportunity for potential speedup.

The paper is organized as follows. Section 2 will introduce the time–splitting techniques. Section 3 will discuss the E3SM Atmosphere Model (EAM) and our parallel–splitting implemention. Section 4 highlights a series of complications that arise when parallel–splitting physics and dynamics. Section 5 will present the results of the parallel–split simulations. Finally, Section 6 will present a summary of the results and potential future direction. Appendix A: illustrates the differences between parallel– and sequential–splitting by analytically examining a simple test problem and Appendix B: provides the processors counts for the simulations used in this study.

2 Time–splitting in global atmosphere models

As mentioned in the introduction, global atmosphere models must treat a complex global system with a range of temporal and spatial scales. The typical spatial resolution for a global climate simulation ranges from 25km [*Caldwell et al.*] to 100km [*Golaz et al.*] or more. Sub–grid scale processes are represented by a set of physics parameterizations. These provide a tendency which is then applied in the resolved scale fluid dynamics model, typically following an SS coupling implementation.

There are a handful of variations of the SS approach. *Sequential–update–splitting* (SUS) updates the model state after each process in the sequence. Thus, each process is passed an *updated* model state which reflects the impact of all processes that have preceded it. *Sequential–tendency–splitting* (STS) uses the same base state (typically the state at the beginning of the timestep) for each process and incorporates the *tendencies* from previous processes as a constant forcing term. This is particularly useful in cases where one process is sub–stepped at a shorter timestep than the other processes, in which case the tendencies can be applied partially at the substep timestep. Each of these approaches is used in EAM: SUS for the physics parameterization computations, and a variant of STS to incorporate the impact of physics in the dynamics calculation. In actuality physics/dynamics

coupling in E3SMv1 follows a *hybrid sequentially split* (HSS) approach which will be discussed in greater detail in Section 3. All of these approaches are formally first–order. Higher order *symmetric–update–splitting* methods can also be derived. These involve applying all processes at half a timestep in one order and then applying them again in the reverse order for another half timestep. These methods are not used in EAM and thus will not be discussed in this work. There are many references that explore these approaches with idealized settings and in more depth, see [*Sportisse*, 2000; *Dubal et al.*, 2004; *Staniforth et al.*, 2002a,b].

The parallel–split (PS) approach specifies that the individual tendencies for each process be calculated independently and then combined at the end of the timestep. The main distinction being that neither set of processes are aware of what the other process is doing until both have completed. This is currently the approach used in E3SM to handle sub–grid microphysical tendencies. In contrast to the SS approach, PS does not incur a splitting error for single step methods (e.g. forward Euler). A further advantage of this approach is that if more computational resources are available they can be fully utilized by handling each process simultaneously. Despite the promise of greater utilization of computer resources this coupling mechanism has not found wide–spread use in the weather and climate community. One reason for this is that until recently distributing work over the spatial degrees of freedom has provided more than enough work for even the biggest supercomputers. In addition Beljaars *et. al.* found that STS yielded more skillful forecasts [*Beljaars*, 1991; *Beljaars et al.*, 2004]. We find in this paper a third reason: parallel–splitting may require a shorter timestep for stability.

3 Study Methodology

3.1 Model Description

This study is based on the atmosphere component of version 1 of the Energy Exascale Earth System Model (E3SM) [*Golaz et al.*]. The E3SM Atmosphere Model (EAM) [*Xie et al.*, 2018; *Rasch et al.*, 2019] is broken into four main sections, which are run sequentially, see Figure 1 left panel; 1) the SE–dycore (Dynamics) which solves the equations of state [*Taylor et al.*, 2007; *Dennis et al.*, 2012], 2) a set of Before–Coupling Physics (BC Physics) parameterizations which handle cloud physics and radiation, 3) the surface model contributions which account for the other components in the E3SM suite (land,

ocean, sea ice) and 4) a set of After–Coupling Physics (AC Physics) parameterizations which handle the remaining physics parameterizations, as well as apply an energy fixer. The BC Physics set is the larger of the two and includes: 1) the Zhang–McFarlane deep convection scheme [*Zhang and McFarlane*, 1995; *Neale et al.*, 2008; *Richter and Rasch*, 2008], followed by 2) the Cloud Layers Unified By Binormals (CLUBB) [*Golaz et al.*, 2002; *Bogenschutz et al.*, 2013] which resolves shallow convection, turbulence and cloud macrophysics, 3) aerosol activation [*Abdul-Razzak and Ghan*, 2000], 4) the 2–moment Morrison–Gettelman cloud microphysics scheme [*Gettelman and Morrison*, 2015; *Gettelman et al.*, 2015], 5) aerosol wet deposition [*Wang et al.*, 2013; *Liu et al.*, 2016] and 6) Rapid Radiative Transfer Model for GCMS (RRTMG) radiation [*Iacono et al.*, 2008; *Mlawer et al.*, 1997]. The AC physics suite includes: 1) the dry deposition of aerosols [*Liu et al.*, 2012] and 2) gravity wave calculations [*Richter et al.*, 2010]. In both BC and AC physics the set of physics–parameterizations are solved using SUS splitting. As the names would suggest, BC Physics and AC Physics occur before and after EAM is coupled with the remaining surface components in the full E3SM suite.

The spectral element dycore solves the global set of equations using a spectral element cubed–sphere mesh [*Dennis et al.*, 2012]. Each cube face contains an $ne \times ne$ grid of spectral elements. The spectral element basis functions are defined as the set of Lagrangian interpolation functions at the Gauss–Lobatto–Legendre (GLL) points – typically referred to as Legendre–Gauss–Lobatto (LGL) point outside of the climate modeling community. There are $np \times np$ GLL points per element. Each GLL point is used as a column for physics calculations. There is no horizontal communication in the physics. For a domain with ne spectral elements per cube face and order np there are $6ne^2$ total elements and $6(np-1)^2ne^2+2$ unique GLL points (= number of physics columns), once overlapping GLL points between elements are accounted for. The default spectral resolution is $np = 4$, thus there are approximately 9 times more columns than elements in a given domain.

As mentioned in Section 2 and illustrated in Figure 2, EAM uses a hybrid–sequential–splitting (HSS) approach to handle the coupling of physics and dynamics. In HSS the physics provides an updated model *state* for conserved quantities (water vapor, liquid cloud mass, etc.) and model *tendencies* for non–conserved quantities (temperature and velocity) to the dynamics. The physics contribution to the conserved quantities such as water species and aerosol tracers is applied following a SUS approach, while the physics tendencies for the non–conserved quantities such as temperature and velocity are applied using

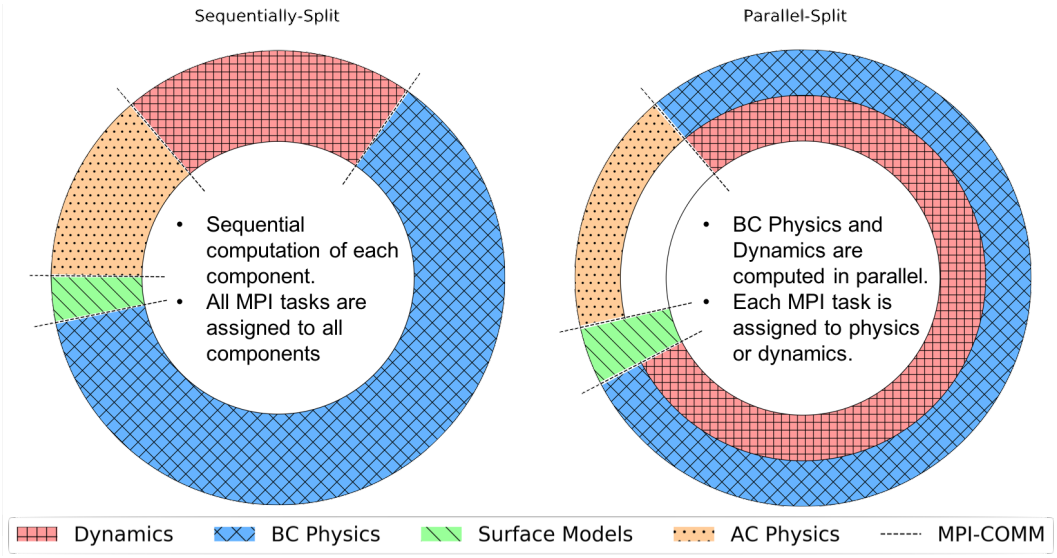

Figure 1. Flow diagram for the EAM, default sequential physics/dynamics coupling (left) and parallel physics/dynamics coupling (right).

an STS approach with their effect "dribbled in" over a set of *remapping* substeps. EAM adopted this HSS strategy to mitigate issues that arose when either SUS or sSTS alone were used to couple all physics with dynamics. The problem with STS is that conserved quantities became over–depleted and subsequently corrected in a way which violated energy conservation. SUS is not desirable because the rigid sudden adjustment to the model state every physics time step produced spurious gravity waves that degraded the solution fidelity. The application of HSS addressed these two concerns and produced a good solution. A detailed description of these problems and the solution are provided in *Zhang et al.* [2018]. All default model simulations employ this coupling mechanism.

3.1.1 Default model performance

As noted earlier the SE–dycore scales up to the total number of elements in the dynamics mesh while physics scale to the total number of columns (which is 9× larger). The default implementation allows the user to request as many cores as there are columns, but only as many cores as there are elements will be used for dynamics; all excess cores will sit idle during dynamics. This means that peak scalability is limited by the dynamics.

3.2 Implementation of parallel–splitting in EAM

The goal of this study is to separate the computation of the physics and the dynamics onto separate computational cores. This task is made difficult by the four components of the atmosphere model described in Section 3.1. In particular, there are a series of MPI communication barriers between these sections of code which limit when dynamics can be run in parallel with the physics. These barriers and our parallelization strategy are mapped out in Figure 1, right panel. The most critical use of MPI barriers is related to the surface coupling: at this point all computational cores must be reassigned to surface processes. With the exception of the ocean model, the surface models must be run sequentially with EAM. As a result it is not possible to run the atmospheric dynamics concurrently with the full suite of physics parameterizations without moving away from a 4–stage atmospheric software architecture — which would require massive infrastructure investment. Keeping the current software architecture, we could run dynamics concurrently with either the BC or AC physics parameterizations but not both. Given that BC physics is the most expensive of the two it was chosen. A complete restructuring of EAM would allow for more complete parallelization. In that sense, the results shown here provide a lower bound on the benefits of PS. That said, AC physics only accounts for $\approx 10 - 15\%$ of the overall physics run time so the potential for gains is fairly modest. We are currently rewriting the EAM in a way that will enable full physics dynamics parallelism as part of a larger long–term effort.

Mechanically the changes needed to run dynamics and physics in parallel were relatively minor and focused on two parts of the EAM infrastructure. First the initialization of physics and dynamics domain decompositions had to be adjusted so that a computational core could only be assigned to the dynamics or the physics but not both. Second, in the standard SS configuration the physics produces tendencies for the prognostic variables which are passed to dynamics. These tendencies are then applied in dynamics following one of the approaches described in Section 2. At the end of the dynamics step the model state is passed back to physics to initiate the next timestep. In the PS configuration the transfer and subsequent application of the physics tendencies was delayed until after both physics and dynamics had finished their respective calculations. As will be discussed in the following section, the final application of vertical remapping and horizontal hyperviscosity was also delayed until after the physics tendencies could be applied to the updated dynamics state, see Figure 3.

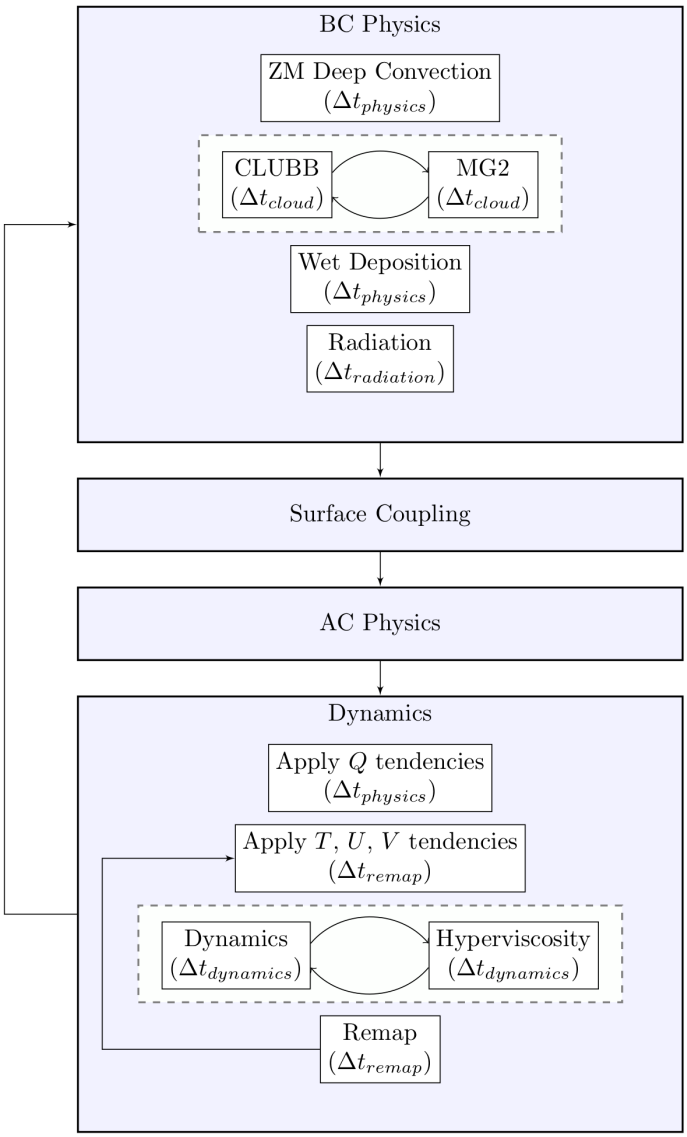

Figure 2. Flow diagram for the EAM using default sequential physics/dynamics coupling. All MPI ranks assigned the same set of tasks. Dashed boxes represent processes that are substepped.

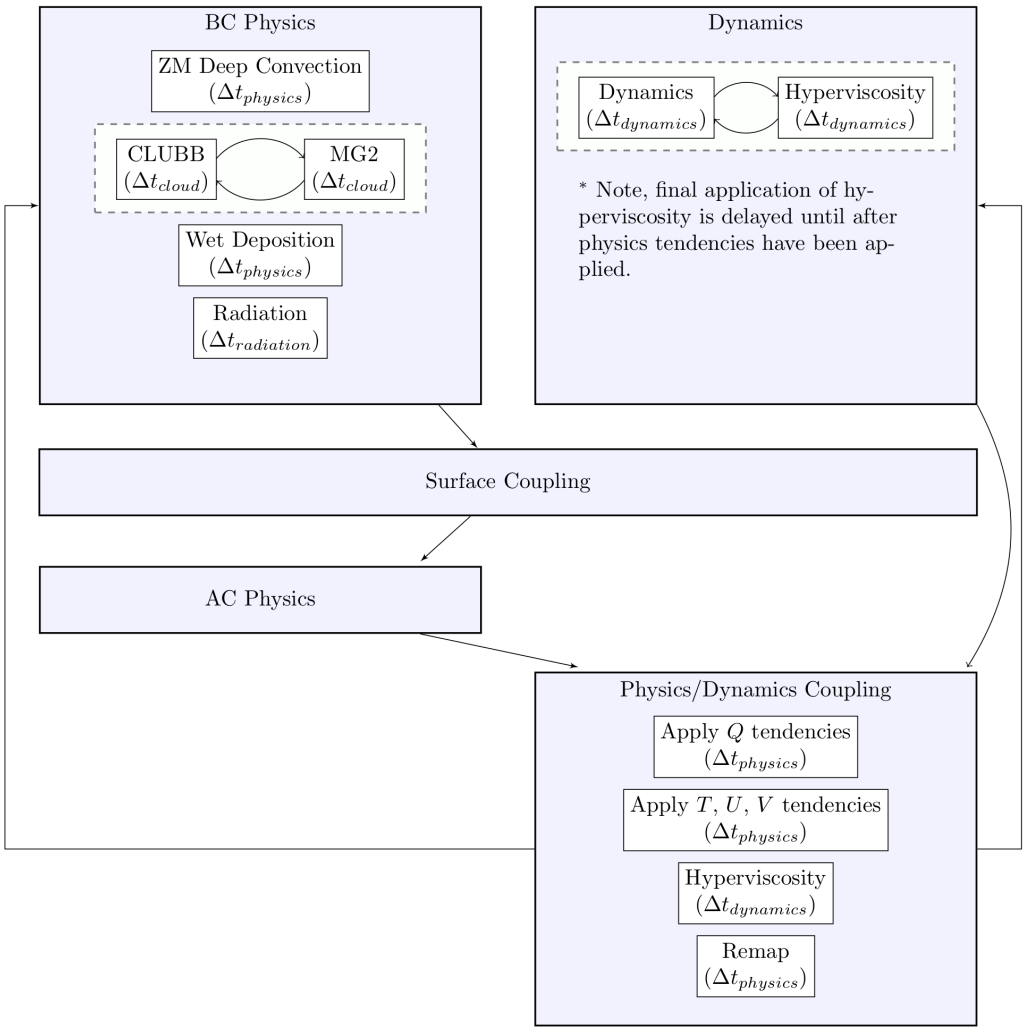

Figure 3. Flow diagram for the EAM using parallel physics/dynamics coupling. MPI ranks are separated into physics tasks (left) and dynamics tasks (right). Dashed boxes represent processes that are substepped.

Table 1. Spectral element domains and their respective number of elements and columns

Mesh Size	Avg. Nodal Distance	# of Elements	# of Columns
$ne4np4$	7.5°	96	866
$ne11np4$	2.7°	726	6536
$ne16np4$	1.9°	1536	13826
$ne30np4$	1.0°	5400	48602

3.3 Experimental Design

In order to assess the impact of shifting to a parallel-split physics/dynamics coupling paradigm, a series of simulations were conducted using the default sequentially–split version of E3SMv1 and the parallel–split version of E3SMv1. Parallel–split and sequentially–split simulations were conducted with a variety of total core counts to assess performance impact. All simulations were conducted with all forcings and sea surface temperatures (SST) set at anually–repeating year 2000 conditions. All simulations use 72 vertical levels in the atmosphere. Ten–year simulations at 1° resolution were used to evaluate the impact of parallel–split versus sequentially–split coupling on model climate. Computational performance was tested with 5 day simulations on the 7.5°, 2.7°, 1.9° and 1.0° degree meshes without writing any output, see Table 1 for mesh details. As explained further in Section 4.1, the parallel–split simulations required a reduced Δt for numeric stability, see Table 2. The simulations using the sequentially–split configuration were run with the default Δt as well as the reduced Δt. Scalability was tested on two computational platforms: a conventional CPU machine called Quartz at Lawrence Livermore National Laboratory (running 2.1 Ghz Intel Xeon E5-2695 chips with 36 cores/node) and a Knights Landing–based system called Cori–KNL at the National Energy Research Scientific Computing (NERSC) center (running 1.4 GHz Intel Xeon Phi Processor 7250 with 68 cores/node).

Table 2. Set of Δt values used for each simulation, as discussed in Section 4.1.

Mesh Size	Type	$\Delta t_{physics}$	Δt_{cloud}	$\Delta t_{radiation}$	Δt_{remap}	$\Delta t_{dynamics}$
7.5°	Default	120min	20min	120min	60min	30min
	Reduced	60min	20min	60min	60min	30min
2.7°	Default	120min	20min	120min	30min	15min
	Reduced	40min	20min	80min	40min	13.3min
1.9°	Default	30min	5min	60min	30min	10min
	Reduced	30min	5min	60min	30min	10min
1.0°	Default	30min	5min	60min	15min	5min
	Reduced	15min	5min	60min	15min	5min

4 Issues with parallel–splitting implementation

4.1 Stability due to timestep

The timestep limitations for SS can differ greatly from those for PS, see Appendix A.1 for greater detail. Thus many of the assumptions that control the timestep limitations for a SS model must be reconsidered when switching to a PS implementation.

As evident in Table 2, the timestep for a simulation in EAM is actually a collection of timesteps that are determined for each individual process in the model. The main timestep in the model can be referred to as the *physics* or *coupling* timestep, and represents the frequency of physics and dynamics coupling. Within both dynamics and physics there are a set of processes which are substepped at a smaller timestep. Dynamics is controlled by the CFL condition, which typically requires a shorter *dynamics* timestep than physics. Additionally, the vertical pressure level thickness is allowed to expand and contract due to the vertical Lagrangian transportation of mass. This requires a *remapping* timestep where the vertical pressure levels are remapped onto a standard set of pressure levels to avoid the complete collapse of any layers. See *Dennis et al.* [2012] for more details of the vertical pressure coordinate system. The remap timestep is typically shorter than the physics timestep but larger than the dynamics timestep. In physics, it was determined empirically that the CLUBB macrophysics and MG2 microphysics must be tightly coupled and run at a reduced *cloud* timestep. Conversely, the relatively slow process of

radiation can be run at a much larger timestep, and is thus typically super–stepped in the model. In the case of the default 1° EAM configuration the physics timestep is 30min. In dynamics the solution is remapped every 15min, or twice per full timestep, and the dynamics itself is run at a 5min timestep, or 3 times for every remap step and 6 substeps total. The cloud timestep was empirically determined to be optimal at 5min, so CLUBB and MG2 are substepped 6 times per physics step. Finally, radiation is run hourly and thus is only updated every other timestep. Similar configurations are set for other mesh resolutions.

Empirically we determined that stability for the PS implementation requires the physics timestep to be reduced to roughly the remapping timestep. For a typical 1° simulation this meant a reduction of physics timestep to 15min. This is a problem because it eats away at the efficiency gains we get by parallelizing physics and dynamics. This slowdown is not as detrimental as it sounds because the dynamics, macrophysics, microphysics and radiation timesteps remain unchanged. Thus, in order to maintain stability in the model, dynamics, macro– and micro– physics only need to be substepped 3 times instead of 6 and the radiation can still be super stepped hourly. As a result the reduction in timestep only produces about a 20% performance penalty in the cost of the atmosphere component. Note that for the 1.9° we found there is no need to reduce the timestep. The physics timestep is equal to the remapping timestep by default, and for this mesh simulations remained stable at the default timestep. The impact on the full model was not examined in this study, the increased coupling frequency between the atmosphere and surface is expected to incur an extra performance cost.

4.2 Stabilization mechanisms

Most global atmosphere models apply some form of diffusion to limit the propagation of high frequency modes. Formally the SE dycore has three steps, 1) application of the physics tendencies, 2) solve the conservation of momentum and mass from the equations of state and 3) the application of horizontal hyper-viscosity to remove high frequency modes in the solution, see Dennis *et al.* [*Dennis et al.*, 2012]. In the default SS configuration the hyper-viscosity is applied following every dynamics step. Because the physics tendencies are coupled before dynamics the hyper-viscosity smooths any high frequency noise introduced by the physics. In a PS configuration the hyper-viscosity becomes decoupled from the physics tendencies allowing high frequency waves to propagate through

the simulation. In order to ensure that hyper–viscosity is able to act on the solution after physics, the application of hyper-viscosity after the final dynamics step is delayed until after the physics and dynamics impacts have been combined.

As water vapor mass is moved during the dynamics, pressure layers can collapse leading to numerical inconsistencies. To mitigate this issue the spectral element dycore employs a vertical remapper which projects the current solution onto a set of regular pressure based vertical coordinates. The remapper also has a slight smoothing effect on the vertical profiles of the solution. Remapping before physics tendencies are applied can introduce an inconsistency as the physics tendencies are defined on the set of pressure layers from the beginning of the timestep. To ensure consistency, the remapping step is similarly delayed until after physics and dynamics have been combined, see Figure 3.

This application of hyper–viscosity and vertical remapping occur once the physics is finished and are part of the non–parallel portion of the code. Since they are somewhat expensive this also limits the benefit of parallel–splitting. In our experiments the extra cost incurred by this step was equivalent to half the cost of the rest of dynamics. Models with more numerical diffusion and without a floating vertical grid may be able to reap more benefits from parallelization by not having to employ these steps.

4.3 Mass conservation

A major concern when applying the tendencies of two different processes in parallel is the risk of mass conservation violations. This issue arises from the fact that neither process is aware of the other, thus if both processes sufficiently deplete a resource, their combined tendency can cause that resource to become negative. This is already a risk for STS and is one of the main reasons that EAM switched to the HSS.

There are a number of approaches to mitigate mass conservation violations. The simplest solution is mass-clipping, by which any negative mass is "clipped" to zero. This has the impact of increasing the local mass to accommodate the strong negative tendencies. The risk of this approach is that globally you are increasing the mass which is non–physical and could lead to unrealistic feedbacks. Despite these drawbacks this is the approach employed by default in any case where mass violations might occur in EAM. More sophisticated methods "borrow" the mass from neighboring columns. This can be done locally over an element or a vertical column. Redistributing mass to account for mass con-

servation violations helps preserve global mass. However, this redistribution of mass in the system could lead to unrealistic states or tendencies. Additionally, there may not be a sufficient amount of mass in the neighboring columns or levels to account for the negative mass that must be fixed, so conservation may still not be entirely assured.

Global mass redistribution could potentially fix the issue by borrowing a very small amount of mass everywhere to fill in all holes. The two main benefits to this approach is that global mass will be conserved and that the amount of mass that must be redistributed is relatively small. Thus we may expect the non–physical effects of the movement of mass to be minor. This approach has not been widely adopted for two reasons, one is that the perturbations from the transportation of even a small amount of mass over long distances instantaneously could propagate errors that impact the fidelity of the long simulations typically run with GCMs. A second concern is that global operations require inter–processor communication, which can greatly slow down simulations.

We explored each of these approaches to address mass conservation violations in the PS version of EAM. There was little difference observed in the climate of the different simulations over our 10–year simulation horizon, so all simulations in this study use the mass–clipping approach. Figure 4 contrasts the total water mass (the sum of vertically-integrated vapor, liquid, and ice) and precipitation from the standard model?s HSS approach and from PS with mass clipping. For ease of comparison, the climatological annual cycle is removed from both panels. This graphic gives no indication that clipping negative water leads to growth in growth in total mass in the system. In fact, the least-squares regression for PS over this 10–year period (dashed red line near zero) has lower slope than for HSS. In addition to long-term drift, it is possible that clipping could affect individual events by changing water availability. We will show in Section 5.1 that such changes have no discernible impact on model climate over 10 years. We see in Figure 4 that there is a slight tendency (by ˜10%) for PS to have larger global-average anomalies than HSS. This may or may not be related to clipping. The main way that clipping might affect extremes is through increased precipitation, so we include in the lower panel the timeseries of global-average precipitation (with annual cycle removed). Precipitation does have non-zero trend over the 10 year simulation because the model hasn?t come into equilibrium yet. The long-term slope of the PS run is, however, lower than that for HSS. This lays to rest the concern that clipping is supplying extraneous precipitation to PS. PS standard deviation is ˜3% smaller than that of HSS. In short, we find no clear indication

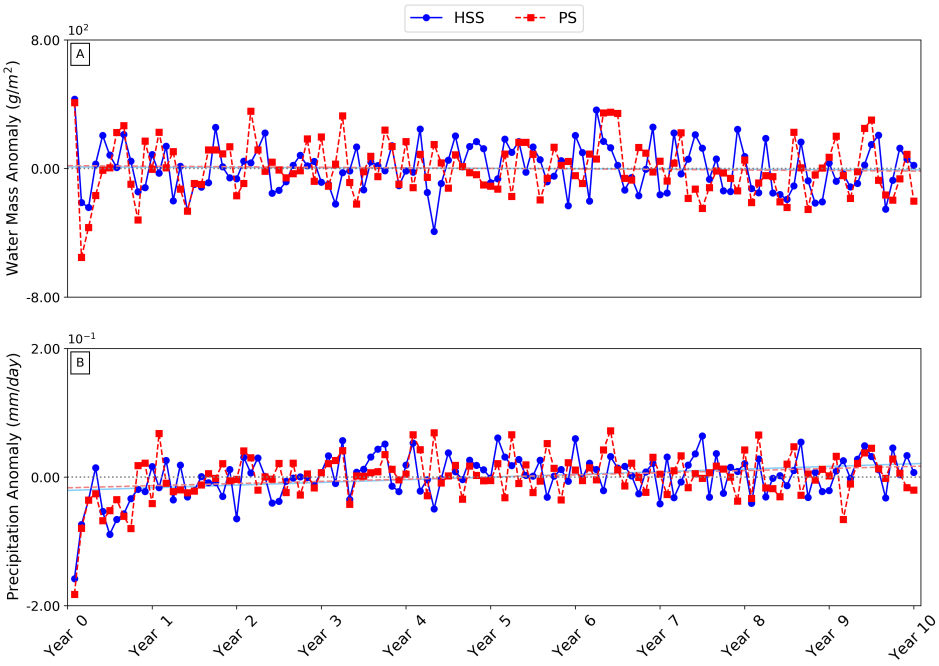

Figure 4. T

otal monthly global average anomaly, defined as the total average with the climatological annual cycle removed, for water mass g/m^2 (top panel) and precipitation mm/day (bottom panel) over ten years of simulation. Default HSS simulations shown in blue with a solid line and circles, PS results shown in red with a dashed line and squares. The overall trend lines for both HSS and PS are shown with light–blue and light–red lines respectively.

that clipping in PS is having an appreciable impact on model behavior. More testing (particularly of impact on local events and trends over centennial scales) would be necessary before PS was adopted operationally.

5 Results

5.1 Climatological Impact

Climatologies for a set of 10 year sequentially–split and a parallel–split simulations were evaluated using the E3SM–Diags toolkit [*Zhang et al.*, 2017], which is similar to the AMWG Diagnostics package (*http://www.cesm.ucar.edu/working_groups/Atmosphere/amwg-diagnostics-package*). The physics timestep for both simulations followed the reduced timestep criteria required for stability in the PS version of the model ($\Delta t = 15min$). See Table 2 for a full list of model timesteps used. Figure 5 shows a set of Taylor diagrams

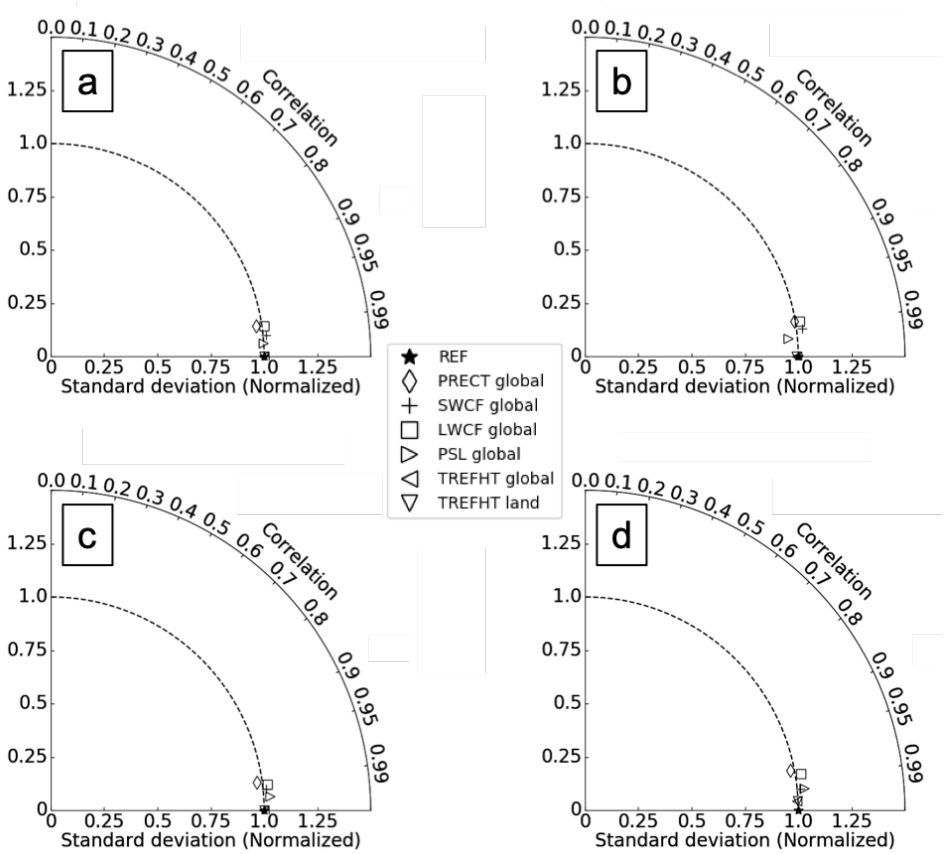

Figure 5. Taylor diagram for 10 year climatologies comparing the parallel–split solution of E3SM with the standard sequentially–split E3SM taken as "truth", a) Winter: DJF, b) Spring: MAM, c) Summer: JJA and d) Autumn: SON.

[*Taylor*, 2001] comparing the parallel–split simulation with the sequentially–split simulation for the four seasons. It is clear from the figure that the two simulations are incredibly similar. The worst case is over the northern hemisphere winter months of December, January and February. But even here, the spatial correlation for all variables is greater than 0.97. The main takeaway from these results is that switching from sequentially–split to parallel–split has very little climatological impact on the solution. Running longer and including an interactive ocean may change this result, but it is outside the scope of this initial performance–focused exploratory study.

5.2 Performance Impact

Our main motivation behind adopting a parallel–split paradigm for coupling physics and dynamics is to improve the computational efficiency of the model at high core counts. In principal both of these components are perfectly scalable. As noted earlier, the dycore scales to the total number of elements and the physics scales to the total number of columns which is $9\times$ larger for typical E3SM configurations. In the default SS configuration, the overall scalability of the model is limited to the least scalable of these two components, in this case the total number of mesh cells. A PS configuration scales to roughly the sum of maximum core counts for both physics and dynamics. Thus, there is the potential to extend the scalability of the model by a full order of magnitude using parallel–split physics/dynamics coupling.

In reality, the time to solve the dynamics on one mesh cell is larger than the time to solve one physics column. Thus in PS with a dynamics core for each element and one column per physics core, the cores assigned to physics can sit idle for some time. Optimum performance can be gained by a proper load balancing of physics and dynamics cores such that the approximate time to complete one dynamics step and one physics step are equal. For all simulations here a reasonable effort was made to optimize the physics/dynamics load–balancing on a case by case basis to provide the best estimates of PS performance, see Appendix B: for details. The ratio of computation time between a mesh cell and a physics column varied by mesh resolution. At the lowest resolution, 7.5°, an individual physics column was on average more computationally expensive than a mesh cell, with an average element:column cost ratio of 0.53. This is most likely related to the increased frequency of radiation, which has the same timestep as the physics. For the simulations at higher resolution, the average cost of a mesh cell varied depending on if radiation was called during the timestep. For timesteps when radiation was not active the average cost of a mesh cell was greater than a physics column with ratios of 2.17, 1.52 and 2.51 for resolutions of 2.7°, 1.9° and 1° respectively. These ratios dropped to 1.34, 1.11 and 1.51 when considering timesteps where radiation was active. A more comprehensive study focusing on load balancing could potentially increase the performance gains obtained in this study.

The timing results are shown in Figure 6. On the positive side, the parallel–split simulations scale to roughly triple the number of cores that the sequentially–split simu-

lations do. Unfortunately, the parallel–split simulations are much slower overall at most core counts. It isn't until we reach the highest possible core counts of 1 or 2 columns per core that the parallel–split runs surpass the performance of the sequentially–split solution. At the highest core counts, the parallel–split solution is roughly 15–20% faster than the sequentially split solution with the same timestep and only slightly faster than the sequentially–split solution at the default timestep for the 1.9° and 1° resolutions. This pattern is observed on both sets of computational architectures, suggesting that the scalability and performance improvements of the parallel–split solution is independent of computational platform. Given that we are -in theory- unlocking $\approx 10\times$ more parallelism and running the two most expensive pieces of the atmosphere simultaneously, this result is a bit disappointing. Better results may be possible with other models or more massive infrastructure changes.

Figure 7 examines the timing results in terms of computational cost vs. time–to–solution. Here we see that PS is consistently more expensive in terms of overall computational hours for similar time to solution when compared with HSS at both the default and reduced timesteps. It is only for the fastest time to solutions that PS is cheaper than HSS, and only marginally so, and only for the same timestep. This is a reflection of the relatively high base computational cost for PS, as measured by the static computational cost at the slowest simulations. This floor in computational cost for PS is ~56% higher than the HSS ($\Delta t_{reduced}$) simulations and ~84% higher than the default HSS simulations. The general trend of PS accomplishing a shorter time–to–solution than HSS for the same timestep demonstrates that there is room for improvement. As long as the base computational cost for PS could be significantly reduced.

5.3 Impact of stability fix on performance

The most striking observation from the performance results is that the parallel–split model underperforms at low core counts. This can be explained by considering the steps taken to stabilize the model after the tendencies from both physics and dynamics are combined at the end of the timestep. The hyperviscosity and vertical remapping steps are expensive. As a result, there is significant speedup in terms of the actual time–to–solution for the individual dynamics and physics solvers, but a tremendous sink in time spent in the hyperviscosity and remapping steps as the last action in a timestep. The hyperviscosity and remapping infrastructure is contained inside the dynamics solver, thus only the cores

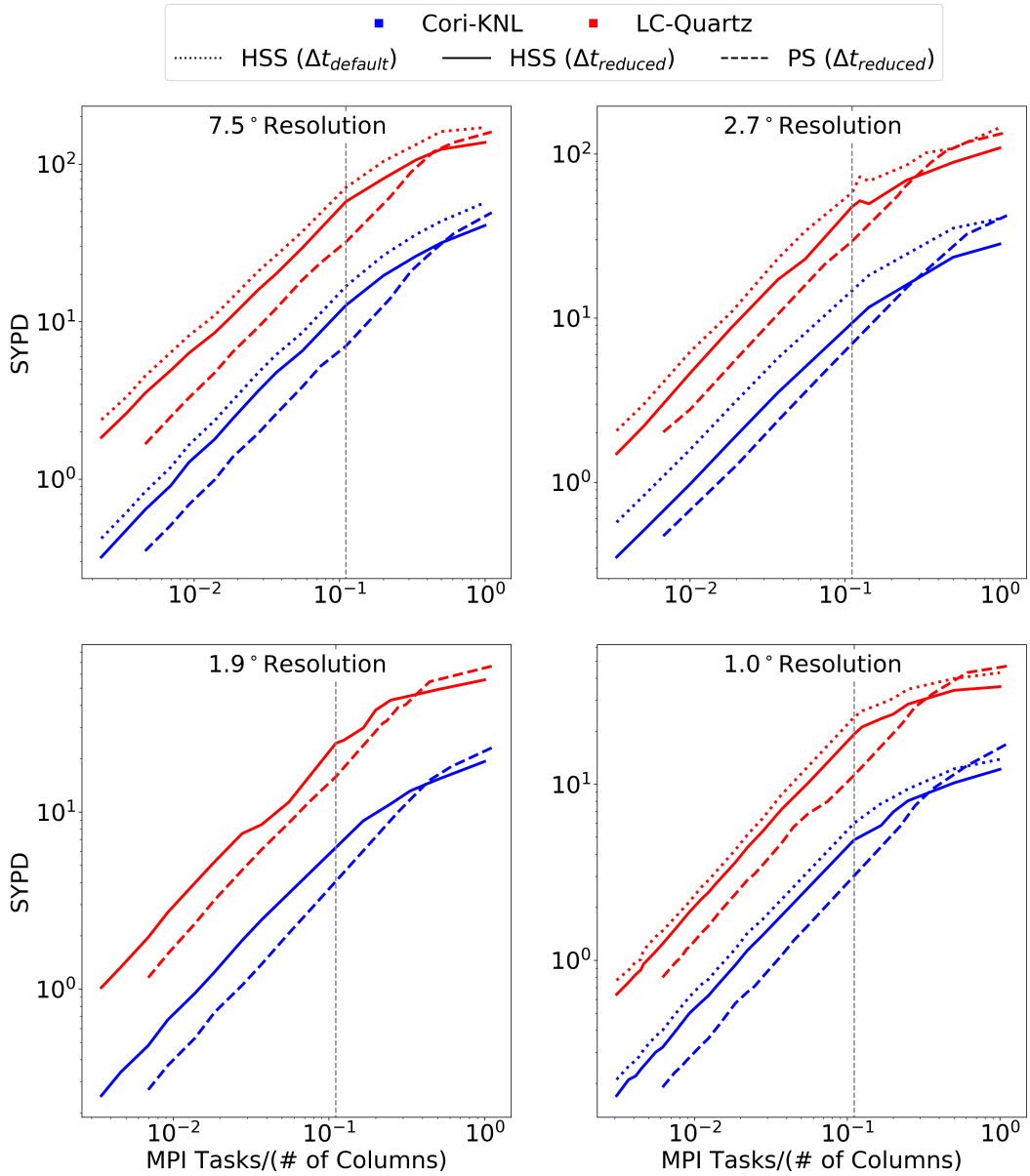

Figure 6. EAM performance in simulated years per compute day (SYPD) for four meshes at multiple core counts for sequential–split (solid lines) and parallel–split (dashed lines) at the reduced Δt and at the default Δt (dotted line). The vertical dashed lines show the peak scalability for sequentially–split. Recall that for the 1.9° resolution there is no reduced Δt simulation. Cori–KNL results shown in blue, Quartz results shown in red.

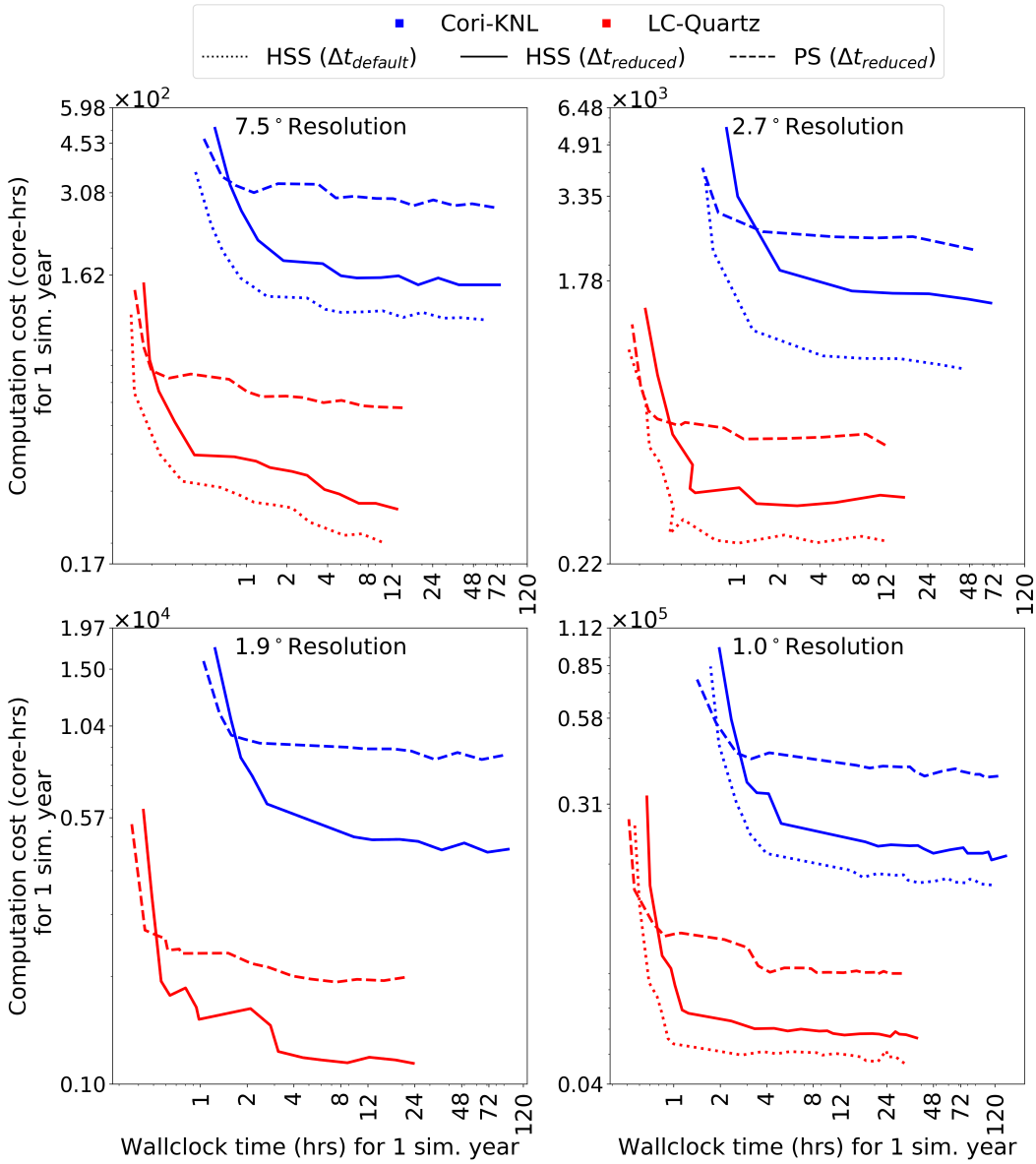

Figure 7. EAM simulation cost in computational core hours per simulated year for four meshes at multiple core counts for sequential–split (solid lines) and parallel–split (dashed lines) at the reduced Δt and sequential–split (dotted line) at the default Δt. Recall that for the 1.9° resolution there is no reduced Δt simulation. Cori–KNL results shown in blue, Quartz results shown in red.

which have been assigned a section of the dynamics domain are able to contribute to the hyperviscosity and remapping steps. As a result, the cores assigned to physics must all sit idle waiting for the final stabilization step to be applied. In a simulation with optimal load balancing, the number of cores assigned to physics is significantly greater than the number assigned to dynamics, thus having physics cores sit idle during hyperviscosity and remapping is inefficient. Recall that the final application of hyperviscosity and vertical remapping has a cost equivalent to 50% of the dynamics step. In a case where physics and dynamics are properly load–balanced an improvement in the PS efficiency by roughly 50% would be possible without this extra stabilization step. Instead, it is necessary to maintain significant resources for dynamics to minimize the cost of this step. We were able to run the SE dycore stably without moving the final hyperviscosity step, but doing so required a significantly reduced model timestep of Δt = 5min.

Recent work by Herrington *et al.* [*Herrington et al.*, 2019], which explores running the physics parameterizations on a coarser finite volume mesh, may obviate the need to apply hyperviscosity after physics. Preliminary parallel–split results using a PG2 physics domain have shown that smoothing from mapping between SE and FV meshes is enough to retain stability at the reduced timestep.

6 Conclusions

A parallel-split algorithm for coupling physics and dynamics in an atmosphere model has been successfully implemented. Parallel splitting did not show a significant change to model climate. There was a marginal improvement to model performance at the highest core counts. It was necessary to reduce the timestep of the simulation to maintain stability, in practice the timestep matched the remapping timestep. Additionally, the internal diffusion mechanisms of hyperviscosity and vertical remapping needed to be applied after the coupling of physics and dynamics tendencies to keep the model stable. These two requirements incurred an extra computational cost on the parallel-split solution that largely countered the performance gains from running dynamics and physics in parallel on separate cores. Note, however that using a shorter physics/dynamics coupling timestep has benefit in itself because it reduces time truncation error. PS may be viewed as an efficient way to accomplish this. At the highest possible core counts a performance improvement of 10 – 20% over the default model was observed.

It is important to consider that this study represents in some sense a lower bound on the performance improvements that can be gained from parallel–splitting. Software infrastructure issues in the E3SM model limited the degree of parallelism we could obtain without a significant overhaul of the code. A more comprehensive rewrite of the atmosphere model could unlock greater performance gains from parallel–splitting, as evidenced by the improved scalability of the parallel–split simulations. Most of the performance loss is due to the extra stabilization step that is taken after the physics and dynamics tendencies are coupled together. If hyperviscosity and vertical remapping could be a) made computationally more efficient or b) be distributed over more cores this would decrease their performance impact. Similarly, in models where the physics calculations are considerably more expensive the benefit of being able to apply more cores to physics through parallel–splitting will improve the relative performance. Examples of these sorts of models would be bin microphysics or cloud resolving models (CRMs).

Based on these results, parallel–splitting of physics and dynamics does offer some potential for improved performance, especially at the highest core counts. But a great deal more work would need to be done in terms of the code infrastructure and efficient stabilization mechanisms before it would be a valid approach for the domains and machines that currently run global atmosphere models to adopt.

A: Canonical problem

To illustrate the properties of parallel– and sequential– splitting approaches we consider the simple problem used in Caya *et al.* [*Caya et al.*, 1998] which contains the "minimum ingredients",

$$\frac{dF(t)}{dt} + \beta F(t) = G, \tag{A.1}$$

where β and G are both constant and represent fast and slow processes respectively. The term G can be thought of as a sub–grid scale forcing term which is constant over the timestep, while $\beta F = (a + bi)F$ may represent some combination of damping (real part) and oscillation (imaginary part). We will use time–splitting to apply different timesteps for each process in the sequence. Suppose $\frac{dF}{dt} = -\beta F$ is sub–stepped n times at a timestep of Δt_β, then $\frac{dF}{dt} = G$ is integrated with a long timestep of $\Delta t_G = n\Delta t_\beta$. If we consider the fully–explicit sequentially–update–splitting (SUS) solution where the slow process is

applied first in the sequence we get,

Step 1:

$$F^* = F(t) + n\Delta t_\beta G, \tag{A.2}$$

Step 2:

$$F(t + n\Delta t_\beta) = \left(1 - \beta\Delta t_\beta\right)^n F^*. \tag{A.3}$$

Note, this has an advantage in global atmosphere models where the physics parameterization can use much larger time steps than the dynamics.

If we consider the parallel–split (PS) solution we first must determine the tendency for each process in the absence of the other and combine the tendencies to advance one timestep, as follows,

Step 1:

$$\delta_G = G$$
$$\delta_\beta = \left(\frac{(1 - \beta\Delta t_\beta)^n - 1}{n\Delta t_\beta}\right) F(t)$$

Step 2:

$$\frac{F(t + n\Delta t_\beta) - F(t)}{n\Delta t_\beta} = \delta_G + \delta_\beta \tag{A.4}$$

Simplifying the terms in equation (A.4) we get,

$$F(t + n\Delta t_\beta) = (1 - \beta\Delta t_\beta)^n F(t) + n\Delta t_\beta G \tag{A.5}$$

A.1 Timestep constraints

To illustrate the timestep constraint in SUS we apply the following condition to Equation A.3,

$$|F(t + \Delta n t_\beta)/F(t)| \leq 1 \tag{A.6}$$

Which yields,

$$\left|(1 - \beta\Delta t_\beta)^n \left(1 + n\Delta t_\beta \frac{G}{F}\right)\right| \leq 1 \tag{A.7}$$

which implies the recursive formula for the upper bound,

$$n\Delta t_\beta \leq \frac{\left(1 \mp \left|(1 - \beta\Delta t_\beta)^n\right|\right)}{\left|(1 - \beta\Delta t_\beta)^n\right|} \left|\frac{F}{G}\right| \tag{A.8}$$

where the $\mp$ is determined by the sign of the ratio F/G, i.e. if $F/G > 0$ then $(-)$ and if $F/G < 0$ then $(+)$. An obvious difficulty with this analysis is that we must recursively find

the largest n such that this relationship still holds. Another issue is the dependence on F for determining Δt_β and n. Although some knowledge regarding the behavior of G and/or limits on F can mitigate some of these issues.

If we conduct a similar analysis for the PS approach we get.

$$\left|(1-\beta\Delta t_\beta)^n + n\Delta t_\beta \frac{G}{F}\right| \leq 1 \tag{A.9}$$

Which leads to an upper stability bound of,

$$n\Delta t_\beta \leq \left(1 \mp (1-\beta\Delta t_\beta)^n\right)\left|\frac{F}{G}\right| \tag{A.10}$$

where again the sign chosen for $\mp$ is determined by the sign of the ratio F/G, as with SUS.

Given equations (A.8) and (A.10) and taking $\Delta t_\beta \rightarrow \frac{1}{b} = \frac{1}{Re(\beta)}$, which is the largest timestep that won't change the sign of F, we can clearly see the differences in both methods. For the SUS approach there is no bound on n. In the PS approach the timestep is bound by the ratio of F and G. In fact, when F and G are the same sign SUS will always have a larger upper bound on n. When the signs of F and G are opposite the PS approach only has a higher upper bound on n when $\Delta t_\beta > 1$. Thus we can conclude that SUS and PS have different timestep criteria, and it is only SUS that allows for an unbounded n.

B: Case setup for simulations

Set of tables describing the number of processors assigned to dynamics computations and physics computations for performance study comparing hybrid–sequentially–split (HSS) and parallel–split (PS) simulations.

Acknowledgments

This work was part of the CMDV Software Modernization program, funded by the U.S. Department of Energy (DOE) Office of Biological and Environmental Research. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 IM Release Number LLNL-JRNL-804820. Data for this work is available at https://dabdceba-6d04-11e5-ba46-22000b92c6ec.e.globus.org/publications/ASDonahue_JAMES_2019.tar.gz. The E3SM code is open access and can be found at https://e3sm.org/.

Table B.1. EAM configuration for the sequentially split 7.5° simulations

HSS			PS		
Total	Dynamics	Physics	Total	Dynamics	Physics
2	2	2	4	2	2
3	3	3	6	3	3
4	4	4	8	4	4
6	6	6	12	6	6
8	8	8	16	8	8
12	12	12	24	12	12
16	16	16	32	16	16
24	24	24	48	24	24
32	32	32	64	32	32
48	48	48	96	48	48
96	96	96	192	96	96
174	96	174	270	96	174
290	96	290	386	96	290
433	96	433	529	96	433
866	96	866	962	96	866

References

Abdul-Razzak, H., and S. J. Ghan (2000), A parameterization of aerosol activation: 2. multiple aerosol types, *Journal of Geophysical Research: Atmospheres*, *105*(D5), 6837–6844, doi:10.1029/1999JD901161.

Balaji, V., R. Benson, B. Wyman, and I. Held (2016), Coarse-grained component concurrency in earth system modeling: parallelizing atmospheric radiative transfer in the gfdl am3 model using the flexible modeling system coupling framework, *Geoscientific Model Development*, *9*(10), 3605–3616, doi:10.5194/gmd-9-3605-2016.

Table B.2. EAM configuration for the sequentially split 2.7° simulations

HSS			PS		
Total	Dynamics	Physics	Total	Dynamics	Physics
22	22	22	44	22	22
33	33	33	66	33	33
66	66	66	132	66	66
121	121	121	242	121	121
242	242	242	484	242	242
363	363	363	726	363	363
726	726	726	1452	726	726
817	726	817	1543	726	817
934	726	934	1660	726	934
1634	726	1634	2360	726	1634
2179	726	2179	2905	726	2179
3268	726	3268	3994	726	3268
6536	726	6536	7262	726	6536

Barrett, A. I., C. Wellmann, A. Seifert, C. Hoose, B. Vogel, and M. Kunz (2019), One step at a time: How model time step significantly affects convection-permitting simulations, *Journal of Advances in Modeling Earth Systems*, *11*(3), 641–658, doi: 10.1029/2018MS001418.

Beljaars, A. (1991), Numerical schemes for parametrizations, in *Proc. ECMWF Seminar on Numerical Methods in Atmospheric Models*, vol. 2, pp. 1–42.

Beljaars, A., P. Bechtold, M. Köhler, J.-J. Morcrette, A. Tompkins, P. Viterbo, and N. Wedi (2004), The numerics of physical parametrization, in *Proc. of ECMWF Seminar on Recent Developments in Numerical Methods for Atmosphere and Ocean Modelling, ECMWF, Reading, UK*.

Table B.3. EAM configuration for the sequentially split 1.9° simulations

HSS			PS		
Total	Dynamics	Physics	Total	Dynamics	Physics
48	48	48	96	48	48
64	64	64	128	64	64
96	96	96	192	96	96
128	128	128	256	128	128
192	192	192	384	192	192
256	256	256	512	256	256
384	384	384	768	384	384
512	512	512	1024	512	512
768	768	768	1536	768	768
1536	1536	1536	3072	1536	1536
1729	1536	1729	3265	1536	1729
2305	1536	2305	3841	1536	2305
2766	1536	2766	4302	1536	2766
3457	1536	3457	4993	1536	3457
4609	1536	4609	6145	1536	4609
6913	1536	6913	8449	1536	6913
13826	1536	13826	15362	1536	13826

Bertagna, L., M. Deakin, O. Guba, D. Sunderland, A. M. Bradley, I. K. Tezaur, M. A. Taylor, and A. G. Salinger (2019), Hommexx 1.0: a performance-portable atmospheric dynamical core for the energy exascale earth system model, *Geoscientific Model Development*, *12*(4), 1423–1441, doi:10.5194/gmd-12-1423-2019.

Bogenschutz, P. A., A. Gettelman, H. Morrison, V. E. Larson, C. Craig, and D. P. Schanen (2013), Higher-order turbulence closure and its impact on climate simulations

in the community atmosphere model, *Journal of Climate*, *26*(23), 9655–9676, doi: 10.1175/JCLI-D-13-00075.1.

Caldwell, P. M., A. Mametjanov, Q. Tang, L. P. Van Roekel, J.-C. Golaz, W. Lin, D. C. Bader, N. D. Keen, Y. Feng, R. Jacob, M. E. Maltrud, A. F. Roberts, M. A. Taylor, M. Veneziani, H. Wang, J. D. Wolfe, K. Balaguru, P. Cameron-Smith, L. Dong, S. A. Klein, L. R. Leung, H.-Y. Li, Q. Li, X. Liu, R. B. Neale, M. Pinheiro, Y. Qian, P. A. Ullrich, S. Xie, Y. Yang, Y. Zhang, K. Zhang, and T. Zhou (), The doe e3sm coupled model version 1: Description and results at high resolution, *Journal of Advances in Modeling Earth Systems*, *n/a*(n/a), doi:10.1029/2019MS001870.

Caya, A., R. Laprise, and P. Zwack (1998), Consequences of using the splitting method for implementing physical forcings in a semi-implicit semi-lagrangian model, *Monthly Weather Review*, *126*(6), 1707–1713, doi:10.1175/1520-0493(1998)126<1707:COUTSM>2.0.CO;2.

Dennis, J. M., J. Edwards, K. J. Evans, O. Guba, P. H. Lauritzen, A. A. Mirin, A. St-Cyr, M. A. Taylor, and P. H. Worley (2012), Cam-se: A scalable spectral element dynamical core for the community atmosphere model, *The International Journal of High Performance Computing Applications*, *26*(1), 74–89, doi:10.1177/1094342011428142.

Dubal, M., N. Wood, and A. Staniforth (2004), Analysis of parallel versus sequential splittings for time-stepping physical parameterizations, *Monthly Weather Review*, *132*(1), 121–132, doi:10.1175/1520-0493(2004)131<0121:AOPVSS>2.0.CO;2.

Gettelman, A., and H. Morrison (2015), Advanced two-moment bulk microphysics for global models. part I: Off-line tests and comparison with other schemes, *Journal of Climate*, *28*(3), 1268–1287, doi:10.1175/JCLI-D-14-00102.1.

Gettelman, A., H. Morrison, S. Santos, P. Bogenschutz, and P. M. Caldwell (2015), Advanced two-moment bulk microphysics for global models. part II: Global model solutions and aerosol cloud interactions, *Journal of Climate*, *28*(3), 1288–1307, doi: 10.1175/JCLI-D-14-00103.1.

Golaz, J.-C., P. M. Caldwell, L. P. Van Roekel, M. R. Petersen, Q. Tang, J. D. Wolfe, G. Abeshu, V. Anantharaj, X. S. Asay-Davis, D. C. Bader, S. A. Baldwin, G. Bisht, P. A. Bogenschutz, M. Branstetter, M. A. Brunke, S. R. Brus, S. M. Burrows, P. J. Cameron-Smith, A. S. Donahue, M. Deakin, R. C. Easter, K. J. Evans, Y. Feng, M. Flanner, J. G. Foucar, J. G. Fyke, B. M. Griffin, C. Hannay, B. E. Harrop, E. C. Hunke, R. L. Jacob, D. W. Jacobsen, N. Jeffery, P. W. Jones, N. D. Keen, S. A. Klein,

V. E. Larson, L. R. Leung, H.-Y. Li, W. Lin, W. H. Lipscomb, P.-L. Ma, S. Mahajan, M. E. Maltrud, A. Mametjanov, J. L. McClean, R. B. McCoy, R. B. Neale, S. F. Price, Y. Qian, P. J. Rasch, J. J. Reeves Eyre, W. J. Riley, T. D. Ringler, A. F. Roberts, E. L. Roesler, A. G. Salinger, Z. Shaheen, X. Shi, B. Singh, J. Tang, M. A. Taylor, P. E. Thornton, A. K. Turner, M. Veneziani, H. Wan, H. Wang, S. Wang, D. N. Williams, P. J. Wolfram, P. H. Worley, S. Xie, Y. Yang, J.-H. Yoon, M. D. Zelinka, C. S. Zender, X. Zeng, C. Zhang, K. Zhang, Y. Zhang, X. Zheng, T. Zhou, and Q. Zhu (), The doe e3sm coupled model version 1: Overview and evaluation at standard resolution, *Journal of Advances in Modeling Earth Systems*, *0*(ja), doi:10.1029/2018MS001603.

Golaz, J.-C., V. E. Larson, and W. R. Cotton (2002), A pdf-based model for boundary layer clouds. part I: Method and model description, *Journal of the Atmospheric Sciences*, *59*(24), 3540–3551, doi:10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2.

Herrington, A. R., P. H. Lauritzen, M. A. Taylor, S. Goldhaber, B. E. Eaton, J. T. Bacmeister, K. A. Reed, and P. A. Ullrich (2019), PhysicsâĂŞdynamics coupling with element-based high-order galerkin methods: Quasi-equal-area physics grid, *Monthly Weather Review*, *147*(1), 69–84, doi:10.1175/MWR-D-18-0136.1.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins (2008), Radiative forcing by long-lived greenhouse gases: Calculations with the aer radiative transfer models, *Journal of Geophysical Research: Atmospheres*, *113*(D13), n/a–n/a, doi:10.1029/2008JD009944, d13103.

Lapillonne, X., and O. Fuhrer (2014), Using compiler directives to port large scientific applications to gpus: An example from atmospheric science, *Parallel Processing Letters*, *24*(01), 1450,003.

Liu, X., R. C. Easter, S. J. Ghan, R. Zaveri, P. Rasch, X. Shi, J.-F. Lamarque, A. Gettelman, H. Morrison, F. Vitt, A. Conley, S. Park, R. Neale, C. Hannay, A. M. L. Ekman, P. Hess, N. Mahowald, W. Collins, M. J. Iacono, C. S. Bretherton, M. G. Flanner, and D. Mitchell (2012), Toward a minimal representation of aerosols in climate models: description and evaluation in the community atmosphere model CAM5, *Geoscientific Model Development*, *5*(3), 709–739, doi:10.5194/gmd-5-709-2012.

Liu, X., P.-L. Ma, H. Wang, S. Tilmes, B. Singh, R. C. Easter, S. J. Ghan, and P. J. Rasch (2016), Description and evaluation of a new four-mode version of the modal aerosol module (mam4) within version 5.3 of the community atmosphere model, *Geoscientific Model Development*, *9*(2), 505–522, doi:10.5194/gmd-9-505-2016.

Michalakes, J., and M. Vachharajani (2008), Gpu acceleration of numerical weather prediction, *Parallel Processing Letters*, *18*(04), 531–548.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough (1997), Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, *Journal of Geophysical Research: Atmospheres*, *102*(D14), 16,663–16,682, doi:10.1029/97JD00237.

Neale, R. B., J. H. Richter, and M. Jochum (2008), The impact of convection on ENSO: From a delayed oscillator to a series of events, *Journal of Climate*, *21*(22), 5904–5924, doi:10.1175/2008JCLI2244.1.

Rasch, P. J., S. Xie, P.-L. Ma, W. Lin, H. Wang, Q. Tang, S. M. Burrows, P. Caldwell, K. Zhang, R. C. Easter, P. Cameron-Smith, B. Singh, H. Wan, J.-C. Golaz, B. E. Harrop, E. Roesler, J. Bacmeister, V. E. Larson, K. J. Evans, Y. Qian, M. Taylor, L. R. Leung, Y. Zhang, L. Brent, M. Branstetter, C. Hannay, S. Mahajan, A. Mametjanov, R. Neale, J. H. Richter, J.-H. Yoon, C. S. Zender, D. Bader, M. Flanner, J. G. Foucar, R. Jacob, N. Keen, S. A. Klein, X. Liu, A. Salinger, M. Shrivastava, and Y. Yang (2019), An overview of the atmospheric component of the energy exascale earth system model, *Journal of Advances in Modeling Earth Systems*, *11*(8), 2377–2411, doi:10.1029/2019MS001629.

Richter, J. H., and P. J. Rasch (2008), Effects of convective momentum transport on the atmospheric circulation in the community atmosphere model, version 3, *Journal of Climate*, *21*(7), 1487–1499, doi:10.1175/2007JCLI1789.1.

Richter, J. H., F. Sassi, and R. R. Garcia (2010), Toward a physically based gravity wave source parameterization in a general circulation model, *Journal of the Atmospheric Sciences*, *67*(1), 136–156, doi:10.1175/2009JAS3112.1.

Sportisse, B. (2000), An analysis of operator splitting techniques in the stiff case, *Journal of Computational Physics*, *161*(1), 140 – 168, doi:http://dx.doi.org/10.1006/jcph.2000.6495.

Staniforth, A., N. Wood, and J. Coté (2002a), A simple comparison of four physics dynamics coupling schemes, *Monthly Weather Review*, *130*(12), 3129–3135, doi:10.1175/1520-0493(2002)130<3129:ASCOFP>2.0.CO;2.

Staniforth, A., N. Wood, and J. Coté (2002b), Analysis of the numerics of physics dynamics coupling, *Quarterly Journal of the Royal Meteorological Society*, *128*(586), 2779–2799, doi:10.1256/qj.02.25.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, *106*(D7), 7183–7192, doi: 10.1029/2000JD900719.

Taylor, M. A., J. Edwards, S. Thomas, and R. Nair (2007), A mass and energy conserving spectral element atmospheric dynamical core on the cubed-sphere grid, *Journal of Physics: Conference Series*, *78*(1), 012,074.

Wang, H., R. C. Easter, P. J. Rasch, M. Wang, X. Liu, S. J. Ghan, Y. Qian, J.-H. Yoon, P.-L. Ma, and V. Vinoj (2013), Sensitivity of remote aerosol distributions to representation of cloud aerosol interactions in a global climate model, *Geoscientific Model Development*, *6*(3), 765–782, doi:10.5194/gmd-6-765-2013.

Williamson, D. L. (2013), The effect of time steps and time-scales on parametrization suites, *Quarterly Journal of the Royal Meteorological Society*, *139*(671), 548–560, doi: 10.1002/qj.1992.

Xie, S., W. Lin, P. J. Rasch, P.-L. Ma, R. Neale, V. E. Larson, Y. Qian, P. A. Bogenschutz, P. Caldwell, P. Cameron-Smith, J.-C. Golaz, S. Mahajan, B. Singh, Q. Tang, H. Wang, J.-H. Yoon, K. Zhang, and Y. Zhang (2018), Understanding cloud and convective characteristics in version 1 of the e3sm atmosphere model, *Journal of Advances in Modeling Earth Systems*, *10*(10), 2618–2644, doi:10.1029/2018MS001350.

Zhang, C., Z. Shaheen, J.-C. C. Golaz, and G. J. Potter (2017), Acme diagnostics, doi: 10.5281/zenodo.1009157.

Zhang, G., and N. A. McFarlane (1995), Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate centre general circulation model, *Atmosphere-Ocean*, *33*(3), 407–446, doi:10.1080/07055900.1995.9649539.

Zhang, K., P. J. Rasch, M. A. Taylor, H. Wan, R. Leung, P.-L. Ma, J.-C. Golaz, J. Wolfe, W. Lin, B. Singh, S. Burrows, J.-H. Yoon, H. Wang, Y. Qian, Q. Tang, P. Caldwell, and S. Xie (2018), Impact of numerical choices on water conservation in the e3sm atmosphere model version 1 (eamv1), *Geoscientific Model Development (Online)*, *11*(5), doi: 10.5194/gmd-11-1971-2018.

Table B.4. EAM configuration for the sequentially split 1.0° simulations

HSS			PS		
Total	Dynamics	Physics	Total	Dynamics	Physics
150	150	150	300	150	150
180	180	180	360	180	180
200	200	200	400	200	200
216	216	216	432	216	216
225	225	225	450	225	225
270	270	270	540	270	270
300	300	300	600	300	300
360	360	360	720	360	360
450	450	450	900	450	450
540	540	540	1080	540	540
600	600	600	1200	600	600
675	675	675	1350	675	675
900	900	900	1800	900	900
1080	1080	1080	2160	1080	1080
1350	1350	1350	2700	1350	1350
1800	1800	1800	3600	1800	1800
2700	2700	2700	5400	2700	2700
5400	5400	5400	10800	5400	5400
6076	5400	6076	11476	6076	5400
8101	5400	8101	13501	8101	5400
9721	5400	9721	15121	9721	5400
12151	5400	12151	17551	12151	5400
24301	5400	24301	29701	24301	5400
48602	5400	48602	54002	48602	5400