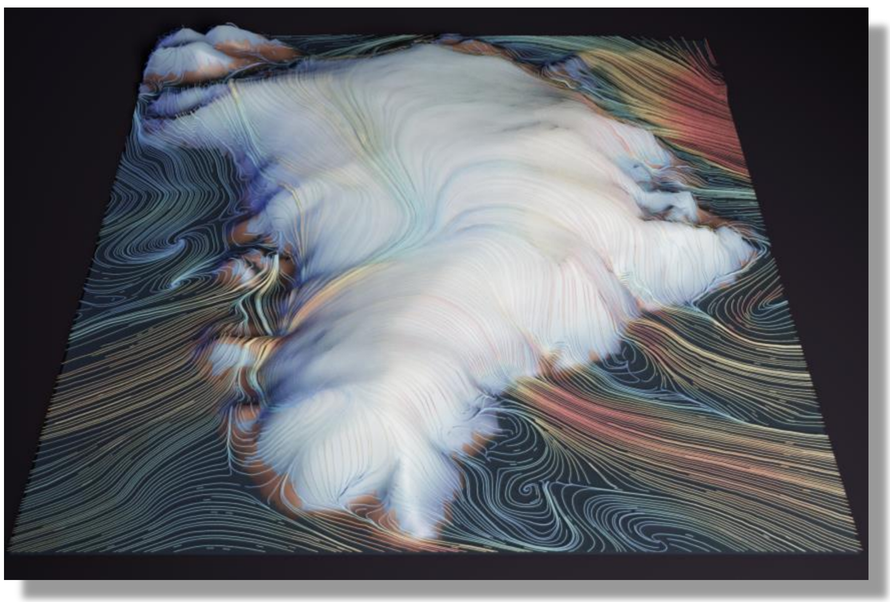
CESM

Community Earth System Model



Proposal for CSL Resources

PI: Gokhan Danabasoglu, CESM Chief Scientist

Period of Performance: 01 November 2020 – 31 October 2022

Total Request: 622.5 M Cheyenne (Equivalent) Core-Hours

Table of Contents

Cover image: Snapshot of the lowest model level streamlines, draped over the Greenland ice-sheet and colored by wind speed. Simulation was performed with a 1/8° refined grid over the island of Greenland using the variable-resolution configuration of the spectral-element atmospheric dynamical core in CESM2. Katabatic winds can be seen accelerating down the eastern slopes of the ice sheet. Visualization was developed by Matt Rehme (CISL) and Adam Herrington (CGD) of the National Center for Atmospheric Research, and was inspired by a visualization of winds over Antarctica by the Polar Meteorology Group at the Byrd Polar & Climate Research Center.

Acknowledgments

The CESM project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the NSF under Cooperative Agreement 1852977. Computing and data storage resources, including the Cheyenne supercomputer (doi:10.5065/D6RX99HX), are provided by the Computational and Information Systems Laboratory (CISL) at NCAR. We thank all the scientists, software engineers, and administrators who contribute to the CESM project.

Introduction

The Community Earth System Model (CESM) is a collaborative, community modeling effort between researchers at the National Center for Atmospheric Research (NCAR), universities, and other national and international research institutions. CESM is used for multiple purposes, including investigations of past and current climate, projections of future climate change, and subseasonal-to-decadal Earth system predictions. To keep CESM at the leading edge of climate and Earth system modeling and related science requires continuous developments, improvements, testing, and subsequent applications of the model to various problems. These efforts have been almost exclusively facilitated through access to the NCAR Computational and Information Systems Laboratory (CISL) Climate Simulation Laboratory (CSL) computational resources.

The CESM and its predecessor, the Community Climate System Model (CCSM), have been at the forefront of both national and international efforts to understand and predict the behavior of Earth's climate. Output from numerous simulations using CCSM and CESM are routinely used in thousands of peer-reviewed studies to better understand the processes and mechanisms responsible for climate variability and change. Specifically, the overwhelming majority of these studies make use of CCSM's and CESM's contributions to the Coupled Model Intercomparison Project phase 3 (CMIP3), phase 5 (CMIP5), and, most recently, phase 6 (CMIP6). Evaluation of the first generation CESM, CESM1, has identified it as a model among the most realistic climate models in the world based on some metrics that compare the model outputs against present-day observationally-based data sets (Knutti et al. 2013). This tradition continues with the latest model version, CESM2, as initial analysis – similarly based on comparisons of a large set of model fields with available observationally-based data – indicates that CESM2 simulations also rank among the most realistic coupled models in the CMIP6 archive with all CESM2 simulations being in the top ten (Fasullo 2020). As a further testament to this point, the two primary manuscripts, Gent et al. (2011) and Hurrell et al. (2013), that introduced and described the previous two versions of the model, CCSM4 and CESM1, have been cited > 2100 and > 1200 times, respectively, since their publications. Significant CSL-supported efforts such as the CESM1 Large Ensemble (CESM1-LENS; Kay et al. 2015), its sibling CESM1 Decadal Prediction Large Ensemble (CESM1-DPLE; Yeager et al. 2018), and the CESM1 Last Millennium Ensemble (CESM1-LME; Otto-Bliesner et al. 2016) have been key in advancing our understanding of the climate system and its variability and predictability, supplementing CESM's contributions to the CMIPs with community-driven science efforts. CESM source code and simulation output are made freely available to the broad scientific community. Additionally, CCSM and CESM simulations provide the quantitative modeling basis for both national and international assessments of climate, including those of the Intergovernmental Panel on Climate Change (IPCC) and the U.S. Global Change Research Program (USGCRP). Furthermore, CESM provides the National Science Foundation (NSF), its primary sponsor, and the national and international research communities with a well-supported Earth system modeling framework.

More than one half of our 2018-2020 CESM CSL allocation has been used to perform simulations in support of CESM's participation in CMIP6 (Eyring et al. 2016). The majority of the simulations use the nominal 1° horizontal resolution configuration with both the low-top (with limited chemistry) and high-top with comprehensive chemistry versions of the atmospheric model. The core simulations are the so-called Diagnostic, Evaluation, and Characterization of Klima (DECK) experiments that consist of a long preindustrial (PI) control simulation; an abrupt quadrupling of CO_2 concentration simulation; a 1% per year CO_2 concentration increase simulation; and an AMIP (Atmospheric Model Intercomparison Project) simulation forced with prescribed observed sea surface temperatures and sea-ice concentrations. These are complemented by multiple simulations of the historical (1850-2014) period. In addition, CESM2 has been participating in about 20 Model Intercomparison Projects (MIPs). To increase the usefulness of simulations for a broad range of applications, for the low-top version, we have extended the PI control simulation to 2000 years (from the requested 500 years) and increased the ensemble size to 11 (from the requested 3) for the subsequent historical simulations. In addition to the DECK and MIP Tier 1 experiments, many Tier 2 and higher simulations have been completed by the related Working Groups (WGs) using their WGs' allocation (see the Accomplishments Report). To provide a computationally more economical model for long time scale, e.g., paleoclimate, applications, several DECK and MIP simulations were also conducted with a version that uses a 2° horizontal resolution in its atmospheric component only.

CESM's participation in CMIP6 DECK and MIPs simulations is its broadest community project. These simulations reach a vast national and international group of scientists and researchers who rely on NCAR, and CESM in particular, to perform these simulations. Among the modeling groups which contribute to CMIP6, CESM is very unique in its community involvement and the level of transparency with which it approaches model development. Moreover, many national and international researchers make use of CESM in their research proposals, e.g., submitted to the NSF, counting on the CESM's contributions to the CMIPs. Additionally, NCAR and CESM assist the community with the analysis of CMIP simulations by providing a CMIP Analysis Platform (https://www2.cisl.ucar.edu/resources/cmip-analysis-platform) as well as by making many diagnostics tools available for use of the community such as the Climate Variability and Diagnostics Package (CVDP). Of course, high level of scrutiny and analysis of the CESM simulations, in turn, feeds back to CESM, promoting further model development as well as enhancing collaborations.

The data sets from CESM2 CMIP6 simulations are available on the Earth System Grid Federation (ESGF; https://esgf-node.llnl.gov/search/cmip6). To date, about 1000 CMIP6 experiments / cases have been run. About 1.7 PB of lossless-compressed time series files have been generated. 600 TB of these time series files, corresponding to over 830,000 files, have been published on the ESGF. This volume of data is roughly 7 times larger than CESM1's contributions to CMIP5.

CESM2 was released to the community in June 2018. To expedite the use of CESM2 by the community primarily for CMIP6-related science and simulations, three *incremental* releases of CESM2 with the same base code were made available in December 2018 (CESM2.1.0), June 2019 (CESM2.1.1), and February 2020 (CESM2.1.2) in which many of these simulations can be run as out-of-the-box configurations. All model versions are available at www.cesm.ucar.edu:/models/cesm2/. While the CESM2.1.x releases have been non-climate-changing, the upcoming CESM2.2 release scheduled for mid-September 2020 will contain many additional, answer-changing features that have been in development since mid-2018.

Manuscripts describing and analyzing these CESM2 CMIP6 experiments in detail are collected in *the AGU CESM2 Virtual Special Issue*. As CESM2 simulations are used in many studies, it is important that its main characteristics are thoroughly analyzed and documented. The Special Issue is spread across several AGU journals that include *Journal of Advances in Modeling Earth Systems*, *Global Biogeochemical Cycles*, *Journal of Geophysical Research – Atmospheres*, *Journal of Geophysical Research – Earth Surface*, *Journal of Geophysical Research – Oceans*, and *Geophysical Research Letters*. To date, over 40 published or submitted manuscripts have been contributed by members of the broad Earth system modeling community. For timely dissemination of the results, we have made all the to-date published and submitted manuscripts available at http://www.cesm.ucar.edu/publications/. An introduction to CESM2 is provided in Danabasoglu et al. (2020), summarizing many new scientific and technical advances in CESM2 compared to its previous version, CESM1.

With the advancements brought about through CESM2, community involvement in CESM development and its applications has continued to expand further. Indeed, an exciting new effort has been recently established under the Earth System Prediction Working Group (ESPWG). This newly formed WG focuses on understanding the processes responsible for predictability on scales from subseasonal to decadal, filling a very much needed niche of providing the community a framework for performing and analyzing initialized predictions of the Earth system as well as serving as a community nexus for related research.

The priorities and goals outlined in this proposal emanate directly from the community of scientists who participate in the CESM project through the CESM WGs and the CESM Scientific Steering Committee (SSC, whose membership consists of not only NCAR scientists but also scientists from universities and government laboratories). Specifically, to prepare this proposal, each WG reached out to their constituents – beginning in May 2020, with widely distributed emails to discuss model development goals towards the next generation model version, CESM3, and production simulations (mostly with the CESM2 versions) required to address high priority scientific questions, especially those that benefit from analysis and interpretation by the broader community. During drafting of the WG plans, there were substantial discussions and planning both within and across the WGs. The resulting drafts were further reviewed, revised, refined, and prioritized by the CESM Chief Scientist in frequent communication with the WG cochairs and the SSC with the goal of producing a coherent and coordinated plan for the use of the CSL resources over the upcoming period of performance. These reviewed plans and resource requests of the individual WGs and community projects, which are included in the Supplementary Material, then served as the draft source material for further review by the SSC and WG cochairs. The responsibility of the SSC in this proposal was to review the overarching development and production simulations, including priorities, as well as the *Community Projects*, and to provide input on the main development and production activities and the required computing and data resources. We note that the CESM SSC has been continuously providing oversight and guidance on CESM objectives and priorities, including development activities towards CESM3. As such, proposed efforts in this proposal directly reflect the inputs from the CESM SSC. A similar process was followed during the last several CESM CSL proposals and, we believe, has resulted in a coherent overview of the testing, development, and application needs of the broad CESM project.

The guidance from CISL indicated that the first 18 months of our 2-year request would be on Cheyenne while the last 6 months would be on the new Supercomputer designated as NWSC-3. The guidance also indicated that our target allocation on Cheyenne would be 230 M core-hours per year and would be scaled according to the following formula for NWSC-3 to obtain Cheyenne Equivalent (CE) core-hours: 3 (computational power increase) x 230 M core-hours / year (CESM allocation on Cheyenne) x 0.5 years x 0.8 (CPU fraction of NWSC-3) = 276 M CE core-hours. So, our second year target allocation is calculated as 115 M (on Cheyenne) + 276 M (on NWSC-3) = 391 M CE core-hours. We certainly understand the need to transition to the new NWSC-3 Supercomputer in the middle of our second year allocation. Although we do not anticipate major disruptions in our simulations during this transition period because NWSC-3 will likely be very similar to Cheyenne in many aspects, the CESM Software Engineering Working Group (SEWG) will dedicate resources (in collaboration with CISL) to ensure a smooth transition.

Overarching Priorities, Goals, and Summary Plan

During the 2-year period of performance of this CESM CSL allocation request, we have two primary overarching priorities:

- Creation of the next generation CESM version, CESM3; and
- Performing and providing simulations with CESM2 to enable and support community-driven science to answer fundamental Earth system related questions.

With the completion of the CESM2 CMIP6 simulations and release of both the data sets and model code, we have been turning our attention to our next model version, CESM3. The majority of the proposed development activities by the core modeling WGs are devoted to CESM3 development efforts. We expect that a vast majority of the proposed developments will be completed and preliminary CESM3 versions will be available in Spring-Summer 2022.

The path towards CESM3 had already started in earnest a few years ago with the move to the Modular Ocean Model version 6 (MOM6) as the new CESM ocean component. Although during the current allocation cycle, substantial progress was made toward developing a workhorse version of the ocean component based on MOM6, much work remains to be completed in developing, implementing, and tuning parameterizations for the

MOM6 framework as we progress towards CESM3. (A functional version of the workhorse CESM2(MOM6) model has been released in CESM2.2.) The development efforts also include bringing Marine Biogeochemistry Library (MARBL) into MOM6.

On the atmospheric side, there are two major efforts. The first concerns the choice of an atmospheric dynamical core (dycore) for the workhorse CESM version. Three new dycores defined on quasi-isotropic grids (2 cubed-sphere and 1 Voronoi) are being integrated into the CESM: Finite Volume (FV3, a non-hydrostatic cubed-sphere version of FV), Model for Prediction Across Scales (MPAS, a global version of Weather Research and Forecasting, WRF, model discretized on a Voronoi grid), and Spectral Element - Conservative Semi-Lagrangian Multi-tracer dynamical core with finite-volume transport (SE-CSLAM). The second key effort is to deliver a scientifically validated atmospheric model with a higher top as well as a well-resolved stratosphere and boundary layer. An outcome of this effort is creation of a single workhorse atmospheric model that can be used for both low- and high-top applications with high enough resolution to adequately represent processes both within the boundary layer and stratosphere. It would also simplify the preparation for coupled experiments, an effort which currently requires ancillary simulations with the high-top model first to generate the necessary chemical forcings for climate simulations. This configuration is expected to have around 85 levels. This effort is long overdue as our workhorse Community Atmosphere Model's vertical resolution is well below those of most other major global coupled models. In addition to this configuration, a *half-top* version with 40-45 levels for more efficient tropospheric physics development will be created. Similar to the ocean model selection, a process will be established for both the atmospheric dycore and vertical grid / model top evaluations that will involve the CESM SSC and the broader community.

On the land modeling front, an ambitious program of model development is planned. In particular, there are several large development projects that are underway, including a multi-layer canopy scheme, a representative hillslope hydrology model, and the Functionally Assembled Terrestrial Ecosystem Simulator (FATES) configuration of the Community Land Model (CLM). Another major activity is work towards unifying land modeling activities across NCAR within the Community Terrestrial Systems Model (CTSM).

Major land-ice development goals towards CESM3 include subglacial hydrology; improved treatments of ice fracture and calving; and new parameterizations and reduced-order ocean models to compute sub-ice-shelf melt rates. The sea-ice model developments include incorporation of a new sea-ice model version; numerous new parameterizations; biogeochemistry coupling; and addressing issues with use of different ocean and sea-ice grids.

Although many of the developments will be initially tested and evaluated in component-only configurations, WGs have also requested time to perform fully-coupled simulations to evaluate the new developments within the coupled climate context before finalizing their component models. Due to the natural uncertainties in timelines in model development, such coupled simulations will use intermediate versions of the other components also in development. Therefore, these simulations will need to be coordinated among the WGs to optimize our resources. If additional computational resources would be necessary, they would be provided either from our Emerging Science allocation (see Community Projects C6) or via reallocations among the WGs.

In parallel with the development efforts towards CESM3, WGs plan to perform simulations to advance Earth system science using various CESM2 configurations. The scope of these experiments is rather broad, reflecting the diversity of the CESM community. Just to provide a few examples, they include: continued contributions to national and international model intercomparison projects; simulations to study how wildfires change in a warming climate and the impacts of fire variability and change on the climate system; climate sensitivity experiments to further our investigation of higher climate sensitivity in CESM2; fully-coupled carbon cycle experiments; many targeted simulations to address specific topics in climate variability and change; simulations on understanding predictability on seasonal to multi-year timescales as well as a detailed assessment of sources of predictability and the role of stochastic physics on the subseasonal timescale; multi-century simulations of future Greenland deglaciation to study ice-sheet stability and potential recovery; paleoclimate simulations of Northern Hemisphere ice-sheet inception at the end of the Last Interglacial period, and of Greenland ice-sheet evolution during the Holocene; studies of future Antarctic ice-sheet evolution and feedbacks with the atmosphere and ocean; simulations investigating the role of land processes and their role in climate variability and change; new last millennium simulations to investigate the transient interactions between the climate and an evolving dynamic ice sheet; and simulations related to improving our process-based understanding.

Our two overarching priorities are complimented by several additional goals that are pursued by multiple WGs. These efforts use either the existing CESM2 versions or target new functionalities for CESM3. They include:

- Producing scientifically supported, variable resolution / regionally-refined CESM configurations that can be used across various applications. These studies take advantage of the regional grid refinement capability of the atmosphere SE dycore. The application areas remain quite diverse. They include grid refinement over tropical regions to study tropical dynamics; over the southeast Pacific stratus region and bordering topography for process studies; over continental US (CONUS grid) to study atmospheric chemistry; over high mountain regions such as Himalayas to better represent complex terrain and heterogeneous processes; and over Greenland and Antarctica to similarly resolve steep topography, permitting realistic orographic precipitation and accurately representing narrow ablation zones. Additionally, a nested regional ocean downscaling capability in MOM6 is being pursued.

- Developing automated techniques for optimization of model parameters. These efforts will include Perturbed Parameter Ensemble simulations to quantify sensitivity of critical climate metrics to parameter choices in CAM6 physics; Perturbed Physics Tendency studies with stochastic physics; applying machine learning algorithms to eddy-permitting ocean simulations to derive mesoscale mixing emulators; and using machine learning to build stochastic parameterizations for sea-ice processes.

- Expanding CESM's capability to include easily-configured, coupled idealized modeling for use of the broader university community. A simplified modeling toolkit as part of the CESM simplified modeling hierarchy is being developed to allow CESM users to set up non-standard model configurations with greatly reduced effort. In addition to various simplified atmospheric model configurations that include an aquaplanet configuration and simplified physics and chemistry, analogous ocean model configurations with MOM6 is being developed. These simplified configurations can be used both as training tools for students and new model users and as a platform for process studies and parameterization development.

This proposal does not include requests for computational resources to perform global high-resolution (coupled) CESM2 simulations with the exception of a few limited and exploratory simulations. As part of a collaboration between the Texas A & M University (TAMU), Qingdao National Laboratory for Marine Science and Technology (QNLM), and NCAR under the International Laboratory for Earth System Prediction (iHESP) umbrella, an unprecedented set of high-resolution (0.25° in the atmosphere and land, and 0.1° in the ocean and sea-ice components) simulations have been (and are being) performed using CESM1.3, using external resources. These simulations include a 500-year PI control, an 1850-2100 transient, a 155-year 1950-control, and a 1950-2050 transient simulation with the latter two submitted to the CMIP6 High-Resolution MIP (HighResMIP) as CESM's contributions. The data sets from a subset of these simulations were released to the community in June 2020 with all to be released by the end of 2020. iHESP will continue to perform additional CESM1.3 high-resolution simulations over the next few years, including some case-based climate prediction simulations. With this background, during their Summer 2020 meetings, all CESM WGs discussed whether limited CESM resources should be invested in creating a similarly high-resolution CESM2 configuration. The feedback overwhelmingly indicated that CESM should not invest in this effort given that creating such a configuration will likely take a lot of human and computational resources for parameter tuning and scientific validation, and that iHESP simulations can be readily used to address many science questions. WGs suggested, however, that a functional high-resolution CESM2 can be made available for exploratory simulations to those interested. These suggestions and feedback from the WGs were further discussed at the 08 July 2020 Virtual CESM SSC Meeting. The SSC unanimously agreed with the WGs and recommended that CESM should focus on CESM3 development with its limited resources and consider creating a high-resolution CESM3 version in due time.

Working Group Research Objectives and Requests

Request for 217.0 M core-hours in Year 1 and 338.6 M core-hours in Year 2

In this section, we very briefly summarize the overall research goals and objectives specific to each WG. In addition, for each WG, we provide the requested computing allocation, split between development (D) and production (P) as well as between Years 1 and 2 – identified as D-Y1, D-Y2, P-Y1, and P-Y2, respectively. All the core-hour requests listed in Table 1 are in millions (M) of Cheyenne (equivalent) core-hours. Further details of each WG's request are available in the Supplementary Material document.

As indicated in the Model Performance section below, SEWG provided performance estimates for a wide range of CESM configurations to the WGs for their use in their requests. However, the choice for which estimate to use was left to the WGs to balance between simulation throughput and cost. The computational cost is also significantly impacted by the I/O requirements of a particular simulation. Therefore, while the majority of the requests use the same costs for a given model configuration, different estimates can also be found for the same model configuration across WG requests. In addition, the estimates for new model versions have relied on simple scaling arguments, considering changes, for example, in the number of vertical levels, tracers, or horizontal resolution. Finally, as we move towards CESM3, the new component models, such as MOM6, CICE6, or choice of an atmospheric dynamical core, and their physics and configurations have not been decided yet. As such, simulations involving new components and features use our best cost estimates based on various preliminary simulations.

	Year 1		Year 2		Totals		
	Dev	Prod	Dev	Prod	Year 1	Year 2	Y1 + Y2
Working Group							
AMWG	21.0	8.9	44.3	4.8	29.9	49.1	79.0
BGCWG	5.2	9.8	6.6	19.5	15.0	26.1	41.1
ChCWG	4.5	5.7	6.9	10.2	10.2	17.1	27.3
CVCWG		20.1		33.9	20.1	33.9	54.0
ESPWG	7.1	10.5	2.0	21.4	17.6	23.4	41.0
LIWG	6.0	9.0	9.0	16.5	15.0	25.5	40.5
LMWG	7.0	13.0	11.0	23.0	20.0	34.0	54.0
OMWG	10.9	9.1	15.3	18.6	20.0	33.9	53.9
PaleoWG	11.2	9.7	17.7	16.3	20.9	34.0	54.9
PCWG	3.0	12.0	13.9	11.6	15.0	25.5	40.5
SEWG	6.0		9.0		6.0	9.0	15.0
WAWG	9.7	17.6	12.5	14.6	27.3	27.1	54.4
Total WGs	**91.6**	**125.4**	**148.2**	**190.4**	**217.0**	**338.6**	**555.6**

Community Projects							
C1. World Avoided		9.2			9.2		9.2
C2. Single Forcings		3.2		10.9	3.2	10.9	14.1
C3. Wildfires		3.2		13.8	3.2	13.8	17.0
C4. Middle Atmosphere		3.2		3.8	3.2	3.8	7.0
C5. Ice Sheets		3.2		8.9	3.2	8.9	12.1
C6. Emerging Science		2.5		5.0	2.5	5.0	7.5
Total comm		**24.5**		**42.4**	**24.5**	**42.4**	**66.9**
Total WG+comm	**91.6**	**149.9**	**148.2**	**232.8**	**241.5**	**381.0**	**622.5**
Target					230.0	391.0	621.0

Table 1. Complete list of core-hour allocation requests for WG development and production projects and community projects. The entries are in millions (M) of Cheyenne core-hours. The requests for Year 1 and Year 2 as well as the total for both years are provided. Blank entries indicate 0.

Atmosphere Model Working Group (AMWG)

D-Y1: **21.0 M;** D-Y2: **44.3 M;** P-Y1: **8.9 M;** P-Y2: **4.8 M; Total: 79.0 M**

AMWG utilizes CSL resources primarily for the development of the CESM Community Atmosphere Model (CAM) and associated capabilities. This encompasses the advancement of both the representation of the unresolved physical processes in parameterization schemes and the dynamical core processes, including tracer transports. It also covers sensitivity experiments aimed at understanding many interactions among the represented physical and dynamical processes across climate regimes and multiple timescales.

In this allocation cycle, AMWG development activity will focus on the following broad areas: 1) deliver a unified atmospheric model that incorporates a well-resolved stratosphere as well as improved resolution of boundary layer turbulent processes; 2) examine new physics and their impacts on simulated climate; 3) produce scientifically-supported, regionally-refined model configurations for domains of interest in process studies, including the southeast Pacific stratus region and bordering topography; 4) continue with evaluation of new candidate atmospheric dynamical cores – MPAS and FV3, noting that these are in addition to the SE dycore already in use; and 5) explore higher horizontal resolution in the context of a new vertical grid and potential physic modifications.

Production activities will continue to explore physical processes in CAM6. These efforts will include Perturbed Parameter Ensemble simulations to quantify sensitivity of critical climate metrics to parameter choices in CAM6 physics and additional climate sensitivity experiments to further our investigation of higher climate sensitivity in CESM2 simulations.

Biogeochemistry Working Group (BGCWG)

D-Y1: **5.2 M;** D-Y2: **6.6 M;** P-Y1: **9.8 M;** P-Y2: **19.5 M; Total: 41.1M**

The goal of BGCWG is to produce a state-of-the-art Earth system model for the research community that includes terrestrial and marine ecosystem biogeochemistry. This model is used to explore ecosystem and biogeochemical dynamics and feedbacks in the Earth system under past, present, and future climates. Land and ocean ecosystems influence climate through a variety of biogeophysical and biogeochemical pathways. Interactions between climate and ecosystem processes, especially in response to human modification of ecosystems and atmospheric CO_2 growth, produce a rich array of climate forcings and feedbacks that amplify or diminish climate change.

Better understanding of ecosystem and biogeochemical dynamics and feedbacks with respect to a changing climate requires an expansion of current CESM land and ocean model capabilities. For this purpose, BGCWG development activities are focused on: continued development of the Newton-Krylov fast spin up technique; continued development of biogeochemical parameterizations; porting of MARBL to MOM6; coupling across components and understanding interactions; and automated techniques for the optimization of model parameters. Production runs address fully-coupled carbon cycle experiments and single component experiments with well-established models. Computing resources are also requested to address the following overarching production goals: examination of ocean ecosystems with resolved mesoscale eddies; additional carbon cycle sensitivity experiments; and evaluation of new biogeochemistry developments in a production context.

Chemistry Climate Working Group (ChCWG)

D-Y1: **4.5 M;** D-Y2: **6.9 M;** P-Y1: **5.7 M;** P-Y2: **10.2 M; Total: 27.3 M**

The goal of ChCWG is to continue development of the representation of chemistry and aerosols in CESM and to further our understanding of the interactions between gas-phase chemistry, aerosols, and climate, using multiple horizontal and vertical model resolutions. The scientific motivation for these developments is the need to understand present-day and future air quality for multiple scales, and to understand the role of climate change on tropospheric composition. The development and production simulations requested here will lead to improving the representation and chemical forecasts of tropospheric composition and air quality, and will allow us to participate in multi-model intercomparison activities.

The development request contains simulations to assist in the ongoing improvement in the representation of tropospheric chemistry and aerosols in CESM (CAM-chem and Whole Atmosphere Community Climate Model, WACCM). New chemistry mechanisms, both more complex and simpler, will be tested. Testing of two photolysis schemes, which take into account the attenuation by aerosols, will be performed. New aerosol schemes will be tested, including updated dust schemes, marine organic aerosols, formation of secondary organic aerosols, and the CARMA sectional aerosol scheme in the current code base. The combination of these new developments of the chemical mechanism, photolysis, and aerosols will be fully tested and evaluated. This development is tied closely to the work beginning in the development of Multi-Scale Infrastructure for Chemistry and Aerosols (MUSICA), and will include testing of regionally refined grids over various parts of the globe, as well as globally uniform high-resolution simulations. The development of chemical data assimilation continues to be a priority and some resources are allocated for that work. On the production side, simulations of CAM-chem-SE(CSLAM) will be performed to study trends in atmospheric composition seen in observations. In addition, simulations will be performed in support of model intercomparison activities such as the Community Climate Model Initiative (CCMI) and World Meteorological Organization (WMO) ozone assessments. Daily chemical forecasts will also be produced to support community air quality research.

Climate Variability and Change Working Group (CVCWG)

D-Y1: **0.0 M**; D-Y2: **0.0 M**, P-Y1: **20.1 M;** P-Y2: **33.9 M; Total: 54.0 M**

The goals of CVCWG are to understand and quantify contributions of natural and anthropogenically-forced patterns of climate variability and change in the 20th and 21st centuries and beyond by means of simulations with CESM and its component models. With these model simulations, researchers will be able to investigate mechanisms of climate variability and change, as well as detect and attribute past climate changes, and project and predict future changes. The CVCWG simulations are motivated by broad community interest and are widely used by the national and international research communities. The highest priority for the CVCWG simulations is given to runs that directly benefit the CESM community. CVCWG is also a central element of the Department of Energy (DOE) - NCAR Cooperative Agreement and provides an interface with national (e.g., U.S. National Assessment) and international (e.g., Intergovernmental Panel on Climate Change, IPCC) climate-change assessment activities. CVCWG does not lead model development but instead performs production runs and analyzes model simulations.

The proposed simulations for this request can be roughly divided into three categories: 1. simulations that will be of broad use to the research community and complement the existing and forthcoming datasets that are being made available through CMIP6 and the CESM2 Large Ensemble (CESM2-LENS); 2. targeted simulations to address specific topics in climate variability and change; and 3. exploratory simulations to gain understanding of how the representation of climate variability may be influenced by simulating the system at higher resolution than has previously been used on the timescales relevant for climate variability.

Earth System Prediction Working Group (ESPWG)

D-Y1: **7.1 M**; D-Y2: **2.0 M**, P-Y1: **10.5 M;** P-Y2: **21.4 M; Total: 41.0 M**

Our new Working Group on Earth System Prediction focuses on understanding the processes responsible for predictability on scales from subseasonal to decadal and filling a very much needed niche of providing the community a framework for performing and analyzing initialized predictions of the Earth system as well as serving as a community nexus for related research.

Experiments proposed by ESPWG focus on facilitating community efforts to understand the sources of predictability on timescales from subseasonal to decadal, and exploring the many uncertainties associated with Earth system prediction system design, such as land, ocean, atmosphere initialization, drift, bias, etc. Key science questions that the group will address are a) how and at what time scales do individual Earth system states (land, atmosphere, ocean, and sea-ice) affect predictability, and b) how do various aspects of prediction system design (such as ensemble size, initialization technique, and model structural characteristics, including resolution and physical parameterization) impact prediction skill. In the next two years, the group's developmental efforts will focus on case study sensitivity tests to efficiently assess the impact on subseasonal to decadal skill of different choices for: initialization, model physics, ensemble size, and atmospheric vertical resolution. These efforts will include tests of initialization using data-constrained atmosphere and ocean states produced by ensemble data assimilation (DA) and work to optimize DA initialization in the presence of model biases. The proposed production simulations will focus on understanding predictability on seasonal to multi-year timescales using CESM2, as well as a detailed assessment of sources of predictability and the role of stochastic physics on the subseasonal timescale.

Land Ice Working Group (LIWG)

D-Y1: **6.0 M;** D-Y2: **9.0 M;** P-Y1: **9.0 M;** P-Y2: **16.5 M; Total: 40.5 M**

The main objectives of LIWG during this allocation cycle are to develop new physics parameterizations in the Community Ice Sheet Model (CISM), to extend ice-sheet coupling to marine-based ice sheets in CESM2, and to carry out coupled ice-sheet and climate simulations with interactive Greenland, Antarctic, and paleo ice sheets. A major goal is to understand and simulate land-ice evolution on time scales of decades to millennia, and thus to provide scientifically sound estimates of future sea-level rise and associated uncertainties.

CISM development priorities include subglacial hydrology, improved treatments of ice fracture and calving, and new parameterizations and reduced-order ocean models to compute sub-ice-shelf melt rates. These developments allow the representation of mountain glaciers, particularly in High Mountain Asia and more complete coupling of

CISM with POP (initially) and MOM6 (in a later stage). Part of the development allocation will be used to test and carry out a variety of fully-coupled CESM simulations, with the aim of providing more realistic initial conditions and ice sheet–ocean interactions for simulations of future Greenland and Antarctic ice-sheet evolution. Also, CISM applications to paleo ice-sheets will be tested, including the large Northern Hemisphere ice sheets that advance and retreat during glacial cycles. All these efforts support coupled modeling of marine ice sheets, a major goal for CESM3. On the production side, we will run CESM2 with multiple CISM domains (e.g., both Greenland and Antarctica), with two-way coupling between the ocean and ice-sheet models. High-priority coupled simulations include: multi-century simulations of future Greenland deglaciation, to study ice sheet stability and potential recovery; paleoclimate simulations of Northern Hemisphere ice-sheet inception at the end of the Last Interglacial period, and of Greenland ice-sheet evolution during the Holocene; studies of future Antarctic ice-sheet evolution and feedbacks with the atmosphere and ocean; and simulations of glacier surface mass balance in High Mountain Asia, using a new variable-resolution grid.

Land Model Working Group (LMWG)

D-Y1: **7.0 M;** D-Y2: **11.0 M;** P-Y1: **13.0 M;** P-Y2: **23.0M; Total: 54.0M**

The goals of LMWG are to advance the state-of-the-art in modeling Earth's land surface, its ecosystems, watersheds, and socioeconomic drivers of global environmental change, and to provide a comprehensive understanding of the interactions among physical, chemical, biological, and socioeconomic processes by which people and ecosystems affect, adapt to, and mitigate global environmental change. Land biogeophysical and biogeochemical processes are intimately linked and therefore it is not possible to separate land biogeophysics development from land biogeochemistry development. As in our previous requests, land biogeochemistry model development has been included in the LMWG request.

LMWG continues to pursue an ambitious program of model development. In particular, there are several large development projects that are underway, including a multi-layer canopy scheme, a representative hillslope hydrology model, and a CLM FATES configuration. These projects will continue into this allocation request along with other development projects on water management and agriculture model development. In addition, LMWG in collaboration with land modeling scientists across NCAR continue to work towards unifying land modeling activities across NCAR within CTSM (of which CLM5 is the current climate configuration of the broader CTSM). Parameter estimation/calibration and sensitivity assessments are increasingly important LMWG activities. Within this allocation request is support for an unprecedented parameter exploration of the full CLM. Simulations investigating the role of land processes and their role in climate variability and change in support of LMWG research are also requested.

Ocean Model Working Group (OMWG)

D-Y1: **10.9 M;** D-Y2: **15.3 M;** P-Y1: **9.1 M;** P-Y2: **18.6 M; Total: 53.9 M**

The primary goals of OMWG are to advance the capability and fidelity of the CESM ocean component in support of specific science objectives of the broader CESM community and to conduct curiosity driven research using CESM to advance our understanding of ocean processes, the role of the ocean in the Earth system, and its interactions with other Earth system components.

OMWG continues in the transition of the dynamical core of the ocean component, moving from the Parallel Ocean Program version 2 (POP2) to MOM6. During the current allocation cycle, substantial progress was made toward developing a "workhorse" version of the ocean component based on MOM6. This configuration has a nominal resolution of 2/3°, somewhat higher resolution than the nominal 1° resolution POP2 but still firmly in the mesoscale eddy-parameterized regime. Because of differences in the numerical methods and interactions between resolved and parameterized processes in MOM6 as compared to POP2 much of the experience developed in tuning POP2 at similar resolutions did not transfer directly to MOM6. A functional version of the workhorse CESM(MOM6) model has been released in CESM2.2, but much work remains to be completed in developing, implementing, and tuning parameterizations for the MOM6 framework as we progress towards CESM3. Nearly all of the development core-hours requested in this proposal are to be directed at continuing this effort. The development falls into two categories. The first category is completing porting and adaption of parameterizations developed or used in POP2 to MOM6 and developing an understanding of the differences in sensitivity of the simulations in the two models. The second category are new parameterizations that are becoming available through new science and/or that are facilitated by the additional accuracy and flexibility of the MOM6 Arbitrary-Lagrangian-Eulerian (ALE) vertical coordinate and underlying numerical algorithms. At the present time two alternative vertical coordinate designs: a geopotential coordinate designated z_*, and a hybrid isopycnal-geopotential coordinate are being assessed. The evaluation of the relative advantages and disadvantages of these two alternatives will continue into the proposed effort. In addition to the development of the workhorse MOM6, development of an eddy-resolving configuration of MOM6 is underway and will continue in the proposed effort, initially on the same 0.1° grid as used in the eddying POP2 model.

New capabilities and flexibility of the MOM6 framework have opened opportunities for configurations of the ocean component model that address science priorities in the CESM community that had not previously been practical with POP2. In particular, two completely new modeling capabilities have been implemented in CESM with MOM6. First, a nested regional ocean downscaling capability has been implemented with MOM6 as both the parent coarse resolution and the child fine resolution model. The initial testbed for this capability is a domain encompassing the eastern tropical Pacific. This development effort will continue, tracking parameterization developments in the high-resolution global configurations and applications to additional regions. The second new capability is a fully coupled "aquaplanet" configuration with MOM6 as the dynamical ocean component as part of the CESM simplified modeling hierarchy. In this configuration the CESM(MOM6) resolution is 2° and the parameterization choices have been made with simplicity and computational efficiency as a priority. During the proposal period, work will focus on developing additional template configurations and spun-up baseline solutions that will serve the CESM community as starting points for fundamental research in climate dynamics.

While we build toward the future CESM3 model with MOM6, research that leverages high value experiments and ensembles with the CESM(POP2) model will continue. These include contributions to international model intercomparison efforts where POP2 provides our baseline contribution as well as filling out a suite of experiments on the interaction of wind-waves and sea-ice conducted in collaboration with PCWG. Notably, much of this work will be conducted with the eddying version of the model in both forced and coupled configurations. A renewed effort in ocean and coupled data assimilation will continue with POP2 in the near term due to the availability of covariance statistics from the CESM2-LENS. Finally, we include a modest request to support experiments with a Large Eddy Simulation (LES) turbulence model. This effort will provide "truth data" for development of a new ocean surface boundary layer scheme under development by OMWG.

Paleoclimate Working Group (PaleoWG)

D-Y1: **11.2 M;** D-Y2: **17.7 M;** P-Y1: **9.7 M;** P-Y2: **16.3 M; Total: 54.9 M**

PaleoWG is a consortium of scientists engaged in modeling to understand Earth's past climates and provide a long-term perspective on climate system feedbacks and processes that underlie the transient nature of climate change. PaleoWG conducts simulations for specific past climate states, designed to explore the relationships between climate forcings, such as variations in atmospheric greenhouse gases, the presence of large continental-scale ice sheets, solar variability and volcanic activity, and feedbacks and processes that control equilibrium climate sensitivity and climate responses on a range of temporal and spatial scales. Assessing model simulations for these out-of-sample climate states against paleoclimate reconstructions based on geological and geochemical records is an important element of the PaleoWG activities.

The overall development goal is to provide the community with an expanded set of capabilities in CESM suitable for application to a wide range of paleoclimate research problems. To support the broader paleoclimate community, new configurations and capabilities of CESM2 will be implemented and tested through paleoclimate applications. A high priority is to continue to maintain and expand the water isotope tracing capability of CESM, which is essential for directly using the paleoclimate geochemical archive to inform paleoclimate simulations, as well as providing additional constraints on hydrological processes in modern climate. PaleoWG will test paleoclimate applications using MOM6. Coupled WACCM6 configurations will be tested for application to both high CO2 and glacial paleoclimate states. The capability of two-way coupling between climate and CISM2 within CESM2 will be investigated for running multi-millennial simulations at a lower resolution with a focus on the transient ice sheet-climate interactions. High resolution atmosphere-land-only paleoclimate simulations will be run with the isotope-enabled CESM1.2 for improved comparisons to proxy data. A perturbed parameter ensemble will also be performed to explore the parameter space of CAM6 with a focus on the ability of CESM2 to simulate past cold and warm extreme climates.

The overall production goals of PaleoWG focus on providing CESM paleoclimate simulations in support of community research into the fundamental questions of paleoclimate science. Production simulations proposed here will complement simulations completed using our current allocation with CESM2. These include two additional last millennium simulations with CESM2(WACCM6ma) and two simulations using CESM2(CISM2) to investigate the transient interactions between the climate and an evolving dynamic ice sheet.

Polar Climate Working Group (PCWG)

D-Y1: **3.0 M;** D-Y2: **13.9 M;** P-Y1: **12.0 M;** P-Y2: **11.6 M; Total: 40.5 M**

PCWG is a consortium of scientists who are interested in understanding and modeling Arctic and Antarctic climate and its relationship to global climate. To enable polar science within PCWG and the CESM project as a whole, we request computing resources for both polar-specific CESM parameterization development and for polar-specific CESM scientific research. We anticipate that these results will provide new understanding of polar climate processes.

Our overall development objective is to ensure that CESM has state-of-the-art abilities to simulate polar climate. Towards this goal, improvements to numerous aspects of the sea ice model within CESM will be developed and tested. A large part of proposed development runs will be dedicated to updating the version of the sea ice component from the CICE Consortium model CICE5 to CICE6. This new version of the CICE model will include a number of new physics developments as well as a column model component for the vertical thermodynamics, known as Icepack. In addition to assessing the impacts of new physics in CICE6, several new parameterizations and capabilities will be incorporated and tested as well. These include details of coupling with MOM6; stochastic parameterizations; and exploratory high-resolution experiments.

The overarching PCWG production goal is to enable important and topical polar science research using CESM. This includes experiments of value to a large number of researchers that are related to polar prediction, integrating models and observations to enhance process understanding, understanding the response of sea ice to a variety of forcings, and understanding coupled system interactions and feedbacks at both poles. The proposed experiments make use of CESM2 configurations, including CESM2 tuned ice albedo experiments that were performed by PCWG under the current allocation. This will allow for the diagnoses of important climate processes relative to the large number of simulations available for the CESM2-LENS and also enhanced understanding of new interactions within CESM2.

Software Engineering Working Group (SEWG)

D-Y1: **6.0 M;** D-Y2: **9.0M;** P-Y1: **0.0 M;** P-Y2: **0.0 M; Total: 15.0 M**

The role of SEWG is to coordinate the computational development of the CESM model components, oversee the evolving design of CESM as new model components, new model grids, and new model physics are added to the system, and at the same time engineer the model system to obtain optimal throughput and efficiency. This continues to be particularly challenging as the number of model configurations, model complexity, and model resolutions are rapidly increasing. Numerous tests are carried out for each new CESM revision on all production platforms to ensure required functionality (such as exact restart capability), correct results (such as bit-for-bit reproducibility where it is expected), tracking of memory and performance metrics (to determine if these have changed relative to the previous revision), and other key production requirements (such as optimizing performance of new revisions, especially where new component science has been introduced). This testing also ensures the robustness of model infrastructure development, such as improvements to the model driver, coupler, tools, and scripts. Computing time is requested to carry out these important functions throughout various CESM versions that will be generated, both in the CESM2 series and on the path towards CESM3.

Whole Atmosphere Working Group (WAWG)

D-Y1: **9.7 M;** D-Y2: **12.5 M;** P-Y1: **17.6 M;** P-Y2: **14.6 M; Total: 54.4 M**

WAWG research plan promotes the development of a unified sun-to-earth modeling framework with WACCM and WACCM-X, a version of WACCM extended throughout the thermosphere, to 500 km.

The development proposal includes exploring higher vertical and horizontal resolutions, both in the context of WACCM and WACCM-X. Resources are also requested to evaluate and tune SE versions of WACCM that include regional refinement. On the production side, simulations will contribute to various studies of QBO and other high-vertical resolution simulations for basic research. Production will also include WACCM simulations of the Last Millennium with interactive volcanic aerosols derived from emissions, and studies of space climate with WACCM-X.

Community Projects

Request for 24.5 M core-hours in Year 1 and 42.4 M core-hours in Year 2

Due to their popularity and success, *Community Projects* have become a traditional part of our CESM CSL allocation requests. These projects represent large simulations that are of substantial interest to multiple WGs, and thus, to the broader CESM community, including international partners and collaborators. They complement our community efforts, pushing community-wide scientific objectives further. In this request, we identified 6 projects, ranging from 7.0 to 14.1 M core-hours (see Table 1). The process for selection consisted of a call for proposals; a review of feasibility and cross-WG interest by the WG co-chairs and the CESM Chief Scientist; and finally a review by the CESM SSC. The sixth project labeled as *Emerging Science* is a modest request that can be allocated to a few relevant emerging science topics that need to be addressed in a timely manner during the course of this allocation cycle.

In the following, summaries of these projects are provided. Further details of each community project request are available in the Supplementary Material document.

C1. World-Avoided Mini-MIP (9.2 M core-hours):

Requested by: To support CESM's participation in the World-Avoided Mini-MIP organized by the U.S. Climate Modeling Summit with the U.S. modeling centers participating.

The purpose of these simulations is to identify the chemistry (including air quality and health) and climate impacts of the 1970 Clean Air Act, including its amendments. This is part of a coordinated effort between all the U.S. climate modeling centers, organized by the U.S. Climate Modeling Summit sponsored by the USGCRP's Interagency Group on Integrative Modeling. The simulations will require use of the CESM2(WACCM6) transient version at the nominal 1° horizontal resolution. These World-Avoided sensitivity simulations will be run for the 1970-2014 period. To identify robust signals, an ensemble size of 10 is recommended. Accompanying control simulations are being carried out separately at no cost to this request.

C2. Additional Single Forcing Experiments to Complement CESM2 Large Ensemble (14.1 M core-hours)

Requested by: CVCWG and community members

The forthcoming CESM2 100-member Large Ensemble (CESM2-LENS) will be a very widely used resource. These simulations are being performed in South Korea in collaboration with the Institute for Basic Science Center for Climate Physics in Busan. The scientific value of this new CESM2-LENS, which uses all forcings from SSP3-7.0, can be greatly increased if additional simulations can be used to further our understanding of its behavior. An important complement is *single forcing* simulations in which individual natural or anthropogenic forcings are imposed in isolation to identify the roles of individual forcings and their non-linear interactions in producing the overall climate changes seen in CESM2-LENS. Here, the relevant forcings to consider are: greenhouse gas forcing (GHG); industrially emitted aerosols (AER); biomass burning aerosols (BMB); Ozone (O3); volcanic eruptions (VOL); and solar (SOL). Based on experience from the CESM1 single forcing experiments, at least 15 members are necessary to tease out the forced response. We propose to run the following four single forcing ensembles for the 1850-2050 period using the CESM2(CAM6) configuration at the 1° resolution as in CESM2-LENS: (A) GHG-only; (B) AER-only; (C) BMB-only; and (D) O3 + VOL + SOL only. These four sets would require 60 201-year simulations. Fortunately, 40 of these simulations are being performed using our NCAR Strategic Capability allocation. We, therefore, request resources here only for the remaining 20 simulations to complete our 60-member *single forcing* set.

C3. Wildfires (17.0 M core-hours)

Requested by: ChCWG, LIWG, LMWG, and PCWG

This project will investigate how wildfires change in a warming climate and the impacts of fire variability and change on the climate system. This is a timely topic given events like the recent Australian and Siberian wildfires and the apparent sensitivity of the high latitude climate to the details of the biomass burning emissions in CESM2 historical and scenario simulations. The project will use two types of simulations: i) with prescribed variations in biomass burning emissions to test the sensitivity to those emissions, and ii) with biomass burning emissions produced prognostically by the fire model to investigate the simulation of changing fires and their impacts. The use of these different types of simulations will allow us to assess the direct influence of changes in fire emissions and to investigate feedbacks associated with changing fires. Our proposed integrations will allow us to address a number of science questions that include: How and why are fires likely to change in a warming climate and what are the climate impacts? How does the choice of fire emissions in pre-industrial, historical, and future climate affect the simulation of climate variability and change and the quantification of internal variability? and How does variability and change in fire emissions affect air quality? All the proposed simulations will be performed using either CESM2(CAM6) or CESM2(WACCM6) configurations at the nominal 1° horizontal resolution.

C4. Development and Evaluation of CESM2(WACCM6-MA) (7.0 M core-hours)

Requested by: WAWG, CVCWG, and Geoengineering Large Ensemble (GLENS) collaborators

CESM2(WACCM6) was developed and run for CMIP6 at 1° horizontal resolution with the FV dycore and with full chemical representation of the troposphere, stratosphere, mesosphere, and lower thermosphere (TSMLT1). In addition, a lower-cost version was developed at reduced, i.e., 2° horizontal resolution with more limited Middle Atmosphere (MA) chemistry, reducing model cost by a factor of ~8. This version also ran the CMIP6 DECK and historical simulations, and proved adequate for many climate applications, e.g., Last Millennium paleoclimate. In this request, we propose to develop and evaluate an intermediate-cost version, using MA chemistry at 1° resolution (FV1x1). This MA version, WACCM6-MA, will retain features of WACCM6-TSMLT1 such as an interactive QBO and the regional predictability, while reducing cost by a factor of ~2.3 with respect to the full WACCM6. It will be comparable in features and cost to the version of CESM1(WACCM) that has been used extensively for stratospheric aerosol geoengineering research, including GLENS. We request computing time to perform 4 simulations required to evaluate the model's climate and establish baseline control simulations for additional studies. We will run one historical simulation (1850-2014), followed by two future scenarios: SSP2-4.5 and SSP5-8.5 (2015-2100). In addition, we will continue the ongoing preindustrial control simulation for 250 years past the point from which the historical simulation branched, in order to assess any long-term climate drifts.

C5. Simulating the Climate of Greenland and Antarctic Ice Sheets Using VR-CESM (12.1 M core-hours)

Requested by: AMWG, CVCWG, LIWG, and PCWG

We request resources to simulate and understand processes controlling the evolution of both the Greenland Ice Sheet (GrIS) and Antarctic Ice Sheets (AIS) using a Variable Resolution (VR-CESM) configuration of the coupled model. In a collaboration between AMWG, CVCWG, LIWG, and PCWG, the Arctic grid is currently being coupled to the 1° POP2 and the 4-km CISM (see Accomplishments Report and the cover images for some preliminary results). This configuration is unique for its ability to resolve complex interactions at high-resolution between sea-ice, ocean, ice sheets and the atmosphere, and is being spun up to provide a dynamic representation of GrIS during the preindustrial period. Here, we seek to continue this successful cross-WG collaboration through branching off the preindustrial simulation to run the fully-coupled Arctic grid configuration over the historical period (1850-2020), and into a future RCP-SSP scenario (2020-2200) for a total of 350 years. In addition to providing a cutting-edge projection of the contribution of GrIS to future global sea level, it also provides an opportunity to understand resolution dependent meteorological phenomena such as atmospheric rivers, polar lows, and katabatic winds, and how they fare against observations. Additional resources are also requested for developing and using an Antarctic grid. This application of VR-CESM to the Southern Hemisphere high latitudes will be the first of its kind. The focus of the simulations will be to provide a high-resolution data set of contemporary Antarctic snowfall and surface mass balance, which will be compared in detail with default CESM2 output and available in-situ and remote sensing observations. Furthermore, we will study the impact of Southern Ocean conditions on Antarctic atmosphere, surface melt, and snowfall. For the Antarctic grid, resources for 4 40-year AMIP-style VR-CESM sensitivity simulations are requested.

C6. Emerging Science (7.5 M core-hours)

Requested by: The CESM Chief Scientist

During the course of our allocation cycles, several science topics naturally emerge that can be addressed using various CESM configurations. The topics in this category are usually time critical. They include, for example, climate impacts of reductions in emissions due to COVID-19 and deciphering impacts of details of how biomass burning emissions are constructed in CMIP6. Clearly, such topics are of broad community interest, with both national and international assessment implications. Therefore, it behooves CESM to respond to such emerging science topics in a timely fashion where a modest amount of computer time can be allocated for these purposes. Furthermore, if additional computational resources would be necessary for coupled CESM3 simulations as discussed in our Summary Plan above, some resources would be provided from this allocation.

Data Management and Archival Needs

As part of this CESM CSL request, each WG and Community Project generated estimates of their long-term (order 5 years) storage needs on the Campaign Storage (CS) associated with each proposed development and production experiment set (listed in the Supplementary Material in detail). These estimates are summarized in Table 2 in Terabytes (TB). *It is important to stress that these estimates are only for our anticipated output volume to be stored on CS for long-term archival and that the actual data volume produced by the proposed simulations will be higher*. *However, many of the data sets, particularly pertaining to the development simulations, will not have any archival value.* As in computational time requests, storage estimates for a given configuration may differ depending on the particular output requirements of a simulation. As in the previous CSL proposals, we follow the CESM Data Management and Data Distribution Plan (available at http://www.cesm.ucar.edu/management/docs/data.mgt.plan.2011.pdf) that has development and production data stored and distributed via different strategies, with each tailored to suit the different user needs. Note that the present estimates are for lossless compression and take into account improvements in data compression from the use of the netCDF-4 standard detailed in our last request. As shown in Table 2, we anticipate archiving about 6.25 PB of data from the simulations performed at NWSC.

	Year 1		Year 2		Totals		
	Dev	Prod	Dev	Prod	Year 1	Year 2	Y1 + Y2
Working Group							
AMWG	268.0	24.0	387.0	36.0	292.0	423.0	715.0
BGCWG	0.0	157.0	30.5	396.0	157.0	426.5	583.5
ChCWG	0.0	45.0	0.0	70.0	45.0	70.0	115.0
CVCWG	0.0	293.3	0.0	445.4	293.3	445.4	738.7
ESPWG	1.5	282.0	1.0	155.5	283.5	156.5	440.0
LIWG	0.0	89.6	0.0	233.8	89.6	233.8	323.4
LMWG	0.0	19.7	0.0	102.0	19.7	102.0	121.7
OMWG	279.6	100.9	339.3	328.1	380.5	667.4	1047.9
PaleoWG	96.0	109.0	124.0	245.0	205.0	369.0	574.0
PCWG	23.3	190.2	32.8	192.2	213.5	225.0	438.5
SEWG	0.0	0.0	0.0	0.0	0.0	0.0	0.0
WAWG	33.3	233.0	23.0	81.0	266.3	104.0	370.3
Total WG	**701.7**	**1543.7**	**937.6**	**2285.0**	**2245.4**	**3222.6**	**5468.0**

Community Projects							
C1. World Avoided	0.0	198.0	0.0	0.0	198.0	0.0	198.0
C2. Single Forcings	0.0	50.0	0.0	169.0	50.0	169.0	219.0
C3. Wildfires	0.0	42.0	0.0	179.2	42.0	179.2	221.2
C4. Middle Atmosphere	0.0	43.0	0.0	52.0	43.0	52.0	95.0
C5. Ice Sheets	0.0	7.0	0.0	18.0	7.0	18.0	25.0
C6. Emerging Science	0.0	10.0	0.0	10.0	10.0	10.0	20.0
Total comm	**0.0**	**350.0**	**0.0**	**428.2**	**350.0**	**428.2**	**778.2**
Total WG+comm	**701.7**	**1893.7**	**937.6**	**2713.2**	**2595.4**	**3650.8**	**6246.2**
Non-NWSC simulations		750.0		750.0	750.0	750.0	1500.0
Total w/ non-NWSC	701.7	2623.7	937.6	3463.2	3345.4	4400.8	7746.2

Table 2. Complete list of storage estimates for WG development and production projects and community projects for archival on CS. The entries are in TB. The estimates for Year 1 and Year 2 as well as the total for both years are provided.

a. Data archiving

Out of many PBs of total data volume expected to be generated from all proposed experiments, we anticipate archiving about 6.25 PB on CS. Our archival strategy will be as follows:

Development: Output data will be primarily stored on the requested glade partition (see below). As shown in Table 2, the majority of the WG development simulations will require no long-term storage. However, we anticipate storing ~1.64 PB from development simulations, representing ~26% of our total storage request. Our experience with the CESM2 development cycle taught us that multi-year development work requires keeping a modest set of reference solutions. We similarly expect archiving some of our CESM3 development simulations, forming the basis of our development storage request. Also based on experience, we expect that 1.64 PB will account for approximately 20% of the development output, meaning that 80% of the generated output has no archival value. In addition, the data will likely be removed 36 months after creation, unless retention is requested from the relevant WG co-chairs. *One-off* development experiments will be removed more quickly at the PIs and / or WG co-chairs' discretion.

Production and Community Projects: The majority of our long-term storage request with ~4.6 PB is for our production and community projects. Initially, output data will be stored on glade scratch space (model *history* data), and the majority will be converted to *timeseries* format, using lossless compression. These timeseries data will then be archived to CS. After a period of five years, we will evaluate if certain data sets should be kept for a longer period of time. These data will then be gradually cut back to 50% of their initial volume over a period of three additional years, based on usage and anticipated demand. This data level will be maintained for three more years. Afterward, each WG will determine what data are to be removed and at what rate, as the archived data is gradually reduced to an acceptable level, as determined by data archiving costs at the time.

Archive Management: A majority of the experiments will make full use of the existing CESM Experiment Database (see https://csegweb.cgd.ucar.edu/expdb2.0/cgi-bin/expList.cgi). This database contains details about the run configuration and establishes provenance. The database application runs an automated monthly email reminder script triggered off dates stored in the database fields; as such, it will be used to remind all affiliated users with the experiment, including scientific leads and software engineers, to prune their data from CS according to the CESM Data Management and Data Distribution Plan. The CISL SAM (Systems Accounting Manager) website will also be used to assist in managing CESM data.

b. Data distribution

Development: In general, output data will be made available only to the WG members that are directly involved with the experiments. For WG members that do not have access to CSL resources, these data will be made available via the Climate Data Gateway (CDG), formerly known as the Earth System Grid (ESG), or the NCAR Data Sharing Service, as appropriate.

Production and Community Projects: Output data will be made available according to the guidelines established by the CESM Data Management and Data Distribution Plan, which was formulated by the CESM SSC, NCAR, and NSF. Initially, access is restricted to the WG members directly involved with the experiments. After a period of no more than 12 months following creation, these data will be made available to the community via the CDG. Data sets from any remaining / proposed CMIP6-related simulations will be made available to the NSF-funded CISL CMIP Analysis Platform (https://www2.cisl.ucar.edu/resources/cmip-analysis-platform) and the ESGF in accordance with the CMIP protocol and requirements.

c. Data analysis and visualization request

The simulations produced under development and production CSL resources will require considerable analysis and visualization. For these needs, we request access to the Casper data analysis and visualization (DAV) cluster. This will require standard interactive access to these clusters for the WG members who have CSL access and for additional participants who are helping in the analysis of these simulations. Currently this includes about 150-200 participating scientists but is subject to change with changing WG members and involvement. We request an allocation of 300,000 hours.

d. GLADE project file space (total request: 2.4 PB)

In order to minimize the usage of CS for storing development results, we request 2.4 PB of CESM GLADE project space, a doubling from our current 1.2 PB allocation. This estimate is based on our previous experience and our anticipated use. This will also enable efficient access to highly utilized CESM simulation outputs and forcing data used in coupled integrations. It will also allow for the post-processing of community project integrations. The requested increase is vital to expedite analysis of simulations, especially those with many ensemble members. This space is collectively managed by the CESM WGs.

e. Campaign Storage space (total request: 6.25 PB + 1.5 PB)

We request 6.25 PB of CS space as summarized in Table 2 and also as discussed above in section (a) for our simulations performed at NWSC. However, we note that we are also performing CESM simulations on non-NWSC supercomputers. In particular, as introduced in C2 above, the new 100-member CESM2-LENS simulations are being performed in South Korea in collaboration with the Institute for Basic Science Center for Climate Physics in Busan. We estimate a data volume of order 4-4.5 PB (with lossless compression) from these simulations. We can accommodate about 1.5 PB on our *paid* Campaign Storage space that was bought about 2 years ago with funding from the NCAR directorate. Because these simulations will be of significant interest to the broader CESM community – following the extremely successful CESM1-LENS effort, we are exploring the feasibility of bringing additional data sets to the US so that we can serve them to the NSF-supported CESM user community. For this purpose, we would like to request an additional 1.5 PB of storage on CS to partially accommodate these data sets.

f. Lossy data compression

As indicated above, our data storage estimates use lossless compression which results in an approximately 50% reduction in the data volume for the fully-coupled system. To reduce our data footprint even further, we are pursuing lossy compression techniques in collaboration with Allison Baker (CISL), Dorit Hammerling (Colorado School of Mines, CSM), and Alex Pinard (CSM). Previous work by Baker was very promising, but focused only on the atmospheric model. The new effort will consider all model components, starting with the ocean model, and investigate feasibility of various levels (aggressiveness) of lossy compression. Relevant WGs will participate in the evaluation and acceptability of lossy compressed data sets. If a level of lossy compression is deemed scientifically acceptable, it will be adopted for use in future CESM data storage, perhaps during the lifetime of this request.

Model Performance

Many of the production simulations during the 2-year period of performance of this request will be done with versions of the model that have very similar performance characteristics to the CESM2.1 version used for CMIP6 simulations. Substantial work went into optimizing the performance and throughput of that model version. Primarily by optimizing the layout of each component and also by improving vectorization and reducing communication costs, a very respectable rate of approximately 30 model years per wall-clock day on 4320 cores on Cheyenne for the 1° workhorse configuration (Fig. 1) was achieved. The details of this model performance effort were already provided in our previous request, and as such they are not repeated here for brevity. The SEWG has provided performance estimates for a wide range of configurations to the WGs for their use in their requests. These estimates are available at https://csegweb.cgd.ucar.edu/timing/cgi-bin/timings.cgi.

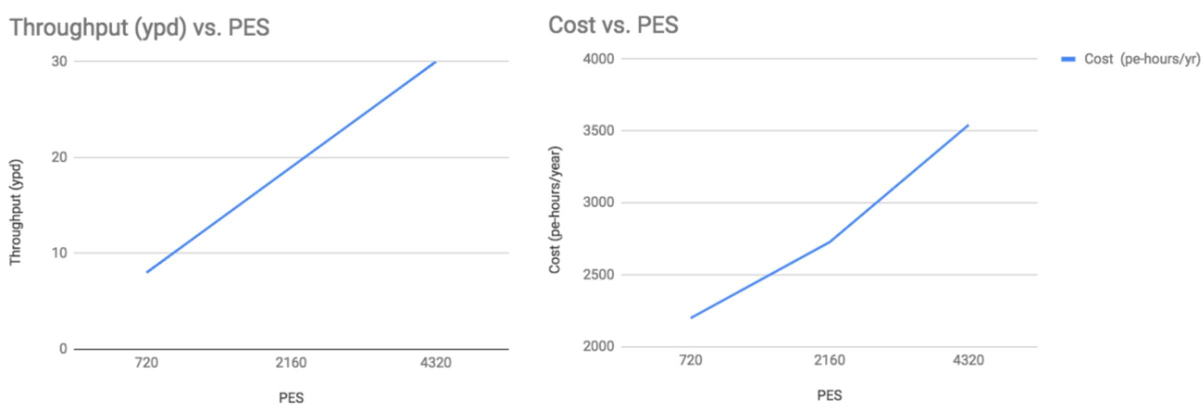


Figure 1. CESM2 nominal 1° version throughput (simulation years per wall-clock day) and cost (core-hours per simulation year) as a function of processor count.

Some new development runs, as well as some production runs later in the two-year allocation period, will use new versions of components or entirely new components that have very different performance characteristics from CESM2.1 as we move towards CESM3. There are still large uncertainties and many unknowns about the exact configurations of these components and therefore about their performance characteristics. Therefore, we are not attempting to provide detailed model performance estimates for these future model versions. Instead, the computational requests in this proposal account for a general increase in model cost based on rough estimates of what is expected, considering, e.g., changes in resolution, atmospheric dynamical core, and the new ocean model component. Once these new configurations have been determined, SEWG will work to optimize their performance as they have done in the past for all CESM versions. Finally, as the new NWSC-3 Supercomputer becomes available, SEWG will dedicate resources to optimize CESM, if needed.

Summary

Earth system models are the most powerful tools for meeting the intellectual challenge of understanding the climate and the Earth system: They are the only scientific tool capable of integrating numerous physical, chemical, and biological processes that determine past, present, and future climates. Furthermore, they are critical for testing hypotheses, confirming understanding, and for making predictions of use to society and policy makers. Among a multitude of such Earth system models, CESM is rather unique because its development and applications are determined through strong partnerships with scientists from universities, national laboratories, and other research organizations. As such, CESM enables the investigations of new scientific problems through collaborations with a community continuously increasing in size, empowering many new partnerships.

With the completion of the CESM2 CMIP6 simulations and release of both the data sets and model code, this CSL request has two, forward-looking overarching priorities. They are: i) creation of the next generation CESM version, CESM3; and performing and providing simulations with CESM2 to enable and support community-driven science to answer fundamental Earth system related questions. On the path towards CESM3, all component models have plans for major model development efforts that include selection of a new atmospheric dynamical core, increasing atmospheric model height and its vertical resolution, completing the transition to MOM6, as well as numerous advancements in the land, sea-ice, and land-ice model components. During the lifetime of this request, we expect that a vast majority of the proposed developments will be completed and preliminary CESM3 versions will be available in Spring-Summer 2022. The two overarching priorities are complemented by several additional goals that are pursued by multiple WGs. These include: production of scientifically supported, variable resolution / regionally refined CESM configurations that can be used across various applications; development of automated techniques for optimization of model parameters; and expanding CESM's capabilities to include easily-configured, coupled idealized modeling for use of the broader university community.

To achieve these priorities and goals, the present proposal is subdivided into two sub-requests, totaling 622.5 M Cheyenne (equivalent) core-hours. The first is for WGs' development and production projects for 239.8 M and 315.8 M core-hours for a total of 555.6 M core-hours. The second request for 66.9 M core-hours is for *Community Projects*, containing simulations that are of interest to multiple WGs, and thus, to the broader CESM community as well as to the USGCRP's Interagency Group on Integrative Modeling. The first 18 months of our 2-year request is for Cheyenne with the remainder on the new NWSC-3 Supercomputer. As such, we are aware of the need to transition to the new Supercomputer in the middle of our second year allocation. Although we do not anticipate major disruptions in our simulations during this transition period because NWSC-3 will likely be very similar to Cheyenne in many aspects, SEWG will dedicate resources (in collaboration with CISL) to ensure a smooth transition.

We request 6.25 PB on CS for long-term archival of our data sets after lossless compression. About 26% of this request is to support our model development efforts. Our experience with the CESM2 development cycle taught us that multi-year development work requires storage of a modest set of reference solutions. We similarly expect archiving some of our CESM3 development simulations. For our storage, as before, we will follow CESM Data Management and Distribution Plan. To minimize / optimize storage on CS, we request 2.4 PB on GLADE as project space. This request is again based on our previous experience and anticipates the analysis needs of simulations, especially with many ensemble members.

The CSL computational resources remain indispensable to carry out our ambitious scientific and model development agenda as we move towards CESM3. The goals and priorities outlined in this proposal emanate directly from the community of scientists who participate in the CESM project through the WGs and the SSC. They were developed, refined, and prioritized after a several month process with the goal of producing a coherent and coordinated plan for the use of the CSL resources over the upcoming period of performance. All of the proposed experiments will fill important development and science needs and contribute to the overall project priorities. However, the tables of individual WG development and production simulations presented in the Supplementary Material include a prioritization of proposed simulations in the event that a reduced allocation is awarded or NWSC-3 does not come online as planned.

References

Danabasoglu, G., and Coauthors, 2020: Community Earth System Model version 2 (CESM2). *J. Adv. Model. Earth Sys.*, **12**, e2019MS001916, doi: 10.1029/2019MS001916.

Eyring, V., and Coauthors, 2016: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, **9**, 1937-1958, doi: 10.5194/gmd-9-1937-2016.

Fasullo, J. T., 2020: Evaluating simulated climate patterns from the CMIP archives using satellite and reanalysis datasets. *Geosci. Model Dev.* (submitted).

Gent, P. R., and Coauthors, 2011: The Community Climate System Model version 4. *J. Climate*, **24**, 4973-4991, doi: 10.1175/2011JCLI4083.1.

Hurrell, J., and Coauthors, 2013: The Community Earth System Model: A framework for collaborative research. *Bull. Amer. Meteor. Soc.*, **94,** 1339-1360, doi: 10.1175/BAMS-D-12-00121.1.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble project: A community resource for studying climate change in the presence of internal climate variability. *Bull. Amer. Meteor. Soc.*, **96**, 1333-1349, doi: 10.1175/BAMS-D-13-00255.1.

Knutti, R., D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and how we got there. *Geophys. Res. Lett.*, **40**, 1194–1199, doi: 10.1002/grl.50256.

Otto-Bliesner, B. L., and Coauthors, 2016: Climate variability and change since 850 CE: An ensemble approach with the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **97**, 735-754, doi: 10.1175/BAMS-D-14-00233.1.

Yeager, S. G., and Coauthors, 2018: Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model. *Bull. Amer. Meteor. Soc.*, **99**, 1867-1886, doi: 10.1175/BAMS-D-17-0098.1.