U.S. DEPARTMENT OF ENERGY | Office of Science

DOE/SC-CM-24-004

FY 2024 Fourth Quarter Performance Metric: Evaluating the U.S. Water Cycle in Global Kilometer-Scale Simulations Produced on an Exascale Computer

L. Ruby Leung
Bryce Harrop
Zhe Feng
Lead Authors

Peter Caldwell
Chris Terai
Naser Mahfouz
Contributing Authors

October 2024

DISCLAIMER

This report was prepared as an account of work sponsored by the U.S. Government. Neither the United States nor any agency thereof, nor any of their employees, makes any warranty, express 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 U.S. Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the U.S. Government or any agency thereof.

Contents

1.0 Product Definition 1
2.0 Product Documentation 2
3.0 Results 3
3.1 Tropical Cyclones, Atmospheric Rivers, and Extratropical Cyclones 3
Tropical Cyclones 4
Atmospheric Rivers 5
Extratropical Cyclones 7
3.2 Mesoscale Convective Systems 8
4.0 Summary and Future Work 11
5.0 References 13

Figures

Figure 1. Storm frequency for the free running simulation (top row), nudged simulation (middle row) and ERA5 (bottom row) for TCs (left), ARs (middle), and ETCs (right). 4
Figure 2. Frequency distribution of TC 10-m wind speed (maximum value along each track measured in m/s) for free-running SCREAM (first column), nudged SCREAM (second column), ERA5 (third column), and IBTrACS (fourth column). 4
Figure 3. Precipitation in IMERG data (left), free-running SCREAM bias (middle), and nudged SCREAM bias (right) for the 13-month analysis period (top row), August-October 2019 (middle row), and November 2019-March 2020 (bottom row). 5
Figure 4. Distribution of occurrence counts for AR lifetime-averaged area (top row) and lifetime-averaged accumulated precipitation (bottom row) for free-running SCREAM (left column), nudged SCREAM (middle column), and ERA5 (right column). 6
Figure 5. AR area (left column), accumulated precipitation (middle column), and area-average precipitation (right column) for free-running SCREAM (top row), nudged SCREAM (middle row), and ERA5 (bottom row) 7
Figure 6. Frequency distribution of ETC 10-m wind speed (maximum value along each track measured in m/s) for free-running SCREAM (left column), nudged SCREAM (middle column), and ERA5 (right column). 7
Figure 7. MCS frequency (units: %) in March-April-May (MAM) of 2020 in (a) observation, (b) nudged simulation, and (c) free-running simulations and the difference between the (d) nudged and (e) free-running simulation compared to observation. 8
Figure 8. Same as Figure 7 but for June-July-August (JJA) of 2020. 8
Figure 9. Number of MCS in the Central U.S. for each month between September 2019 and August 2020 for observation (black), nudged simulation (orange), and free-running simulation (blue). 9
Figure 10. Similar to Figure 7 but for MAM MCS precipitation (units: mm day^{-1}). 9
Figure 11. Similar to Figure 10, but for JJA. 10

Figure 12. Spring (MAM) MCS characteristics depicted by violin plots of (a) MCS lifetime (units: hour), (b) cold cloud shield (CCS) area (units: km^2), 9c) minimum brightness temperature (T_b) (units: K), (d) movement speed (unit: m s^{-1}), (e) precipitation feature (PF) area (units: km^2), (f) PF rain rate (units: mm h^{-1}), (g) total rain volume (units: kg), and (h) heavy rain ratio (units: %).....11

Figure 13. Same as Figure 12 but for summer (JJA). 11

1.0 Product Definition

Global kilometer-scale models have emerged in the last decade (Stevens et al. 2019), holding great promise to improve modeling of water cycle processes and extreme events that have large societal impacts. By explicitly resolving deep convection, which is ubiquitous over the tropics and in the midlatitudes during the warm season, kilometer-scale models can produce more realistic probability distributions of rain rates, reducing the over-abundance of drizzle and the low-bias of intense rainfall in lower-resolution climate models (Satoh et al. 2019). Through changes in the latent heating and cloud-radiation effects, improving modeling of deep convection is also important for improving simulations of large-scale atmospheric circulations and, hence, the spatial distribution of clouds and precipitation (Zhou et al. 2022). Collectively, improvements in modeling the rain rate distributions and atmospheric circulations are expected to have notable impacts on how well the Earth's water cycle is simulated through more realistic representations of the energy balance at the top of the atmosphere and how precipitation is partitioned among evapotranspiration, runoff, and soil water storage at the land surface.

Recently, the Energy Exascale Earth System Model (E3SM) project (Leung et al. 2020) has developed a global kilometer-scale model called Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM) (Caldwell et al. 2021). The latest implementation, SCREAMv1 (Donahue et al. 2024), includes innovative algorithmic and software engineering advancements to run on Frontier, an exascale computer capable of performing more than one quintillion (10^{18}) calculations per second. This has resulted in groundbreaking performance, achieving computational throughput exceeding one simulated year per day for global simulations at 3.25-km grid spacing (Taylor et al. 2023). With this computational performance, it is now feasible to run longer simulations beyond the 40-day boreal winter and summer simulations reported in the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND), the first intercomparison project of global storm-resolving models (Stevens et al. 2019).

Extending the demonstration of the SCREAMv0 kilometer-scale modeling achieved through regional refinement summarized in the 2024 Second Quarter Performance Metric Report, this fourth quarter report focuses on the evaluation of the U.S. water cycle in SCREAMv1 global kilometer-scale simulations. While the SCREAMv0 regional refined simulations featured in the second quarter report are short simulations of 5-50 days produced on the National Energy Research Supercomputing Center (NERSC)'s Perlmutter using central processing units (CPUs), here SCREAMv1 is used to perform two 13-month-long, kilometer-scale global simulations on the graphics processing units (GPUs) of Frontier at the Oak Ridge Leadership Computing Facility (OLCF) and Perlmutter at NERSC. The two simulations were configured identically with prescribed sea surface temperature and sea-ice extent, except nudging is applied to one of the simulations to constrain its atmospheric circulations by a 50-km-resolution global reanalysis while the other is a free-running simulation after model initialization.

In this report, we track weather features in the 13-month-long SCREAMv1 simulations including atmospheric rivers (ARs), tropical cyclones (TCs), extratropical cyclones (ETCs) and mesoscale convective systems (MCSs) that are known to produce heavy precipitation and flooding in different regions of the U.S. The results for ARs, TCs, and ETCs are qualitatively compared with the performance of E3SMv1 at 10- km and 25-km resolutions (Harrop et al. 2023) and with similar results for E3SMv1

and other models reported in previous studies (e.g., Kim et al. 2022, Gore et al. 2023). Feng et al. (2023) tracked MCSs in the DYAMOND models, including SCREAMv0, and found that tropical MCSs are generally well simulated in global kilometer-scale simulations, despite noticeable differences among the models. Here the MCS results are also qualitatively compared with previous studies on evaluation of MCSs in E3SM and other models over the contiguous U.S. (CONUS).

Comparison of the two SCREAMv1 global kilometer-scale simulations with observations shows that the model can produce realistic statistics and characteristics for all four types of storms. More specifically, the observed spatial distributions of storm frequency over CONUS and the surrounding oceans are generally well reproduced. By constraining the atmospheric circulations towards those depicted by a global reanalysis, storm frequency is generally better simulated in the nudged simulation compared to the free-running simulation, underscoring the importance of correctly modeling the large-scale circulations for skillful simulation of storm frequency. While this is true for ARs, ETCs, and MCSs, the nudged simulation produced TC frequency comparable to the reanalysis, but both are biased low compared to observations, while the free-running simulation better captured the higher TC frequency and intensity in observations. The latter suggests that by better resolving TC processes and their influence on the atmospheric environments, kilometer-scale simulations are more skillful for TC modeling than a global reanalysis with data assimilation, but at a coarser resolution. As for storm characteristics such as storm area and precipitation, both simulations are generally skillful and comparable, although some deficiencies have also been identified for future improvements. Through qualitative comparison with climate simulations at 25-50-km grid spacing reported in the literature, the SCREAMv1 global kilometer-scale simulations clearly demonstrate superior skill in modeling the four types of storms (ARs, ETCs, TCs, and MCSs) that are the main causes of extreme precipitation and flooding.

2.0 Product Documentation

This report documents our effort to evaluate the U.S. water cycle simulated by SCREAMv1 configured at 3.25-km grid spacing globally, with 128 vertical levels and a model top at 40 km. Simulations were performed with a model top-level timestep of 100s and a dynamic time step of 8.33s. Descriptions of the model processes, model coupling, and computational strategies are provided by Donahue et al. (2024). Two 13-month-long simulations have been performed on the Frontier and Perlmutter GPUs. The simulations cover August 1, 2019 to September 1, 2020, with the simulation period chosen for its neutral phase in the El Niño-Southern Oscillation (ENSO), a dominant interannual mode of climate variability that has important impacts on the U.S. water cycle. These simulations were initialized using the atmospheric conditions from the European Center for Medium-Range Forecast Reanalysis version 5 (ERA5) at 25-km grid spacing (Hersbach et al. 2020), with sea surface temperature (SST) and sea-ice extent prescribed as the lower boundary conditions from the UK Met Office Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) at 0.1^{o}x0.1^{o} and daily resolution. Other than constrained by the observed SST and sea-ice extent, the control simulation is a free-running simulation with interactive land and sea-ice thermodynamics after the model initialization.

Noting that model biases in simulating water cycle processes such as storms and precipitation may result from model biases in simulating the large-scale atmospheric circulations as well as model biases in simulating local processes such as turbulence, clouds, and land-atmosphere interactions (Feng et al. 2022, Harrop et al. 2023), the control simulation is compared with another simulation in which model biases in the large-scale atmospheric circulations are suppressed. The latter is achieved by configuring a simulation

that is identical to the control simulation except that the winds are nudged towards those of the Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) (Gelaro et al. 2017) at 50-km resolution throughout the 13-month simulation. These simulations with and without nudging, referred to as "nudged" and "free-running", respectively, are evaluated against observations and compared to understand the role of large-scale circulation biases in modeling storm frequency and characteristics.

Storms including ARs, TCs, ETCs, and MCSs are tracked in both simulations. The TempestExtremes feature tracking tool (Ullrich et al. 2021) is applied to track ARs, TCs, and ETCs using model outputs (sea level pressure, geopotential height at 200 and 500 hPa, column integrated moisture flux, 10-m wind speed, and surface geopotential) on a coarsened grid (ne120) corresponding to roughly 25-km grid spacing. For MCSs, the PyFLEXTRKR feature tracking tool (Feng et al. 2023) is applied to model outputs (hourly brightness temperature and precipitation) interpolated to a 10-km latitude-longitude grid. Various statistics are calculated for comparison with those of the ARs, TCs, and ETCs tracked in ERA5 and described in Harrop et al. (2023). TC statistics are further compared against those of the International Best Track Archive for Climate Stewardship (IBTrACS) observations (Knapp et al. 2010). For MCSs, statistics are compared against those derived from the 10-km, hourly combined geostationary satellite infrared brightness temperature (Tb) and GPM IMERG precipitation data set (Feng et al. 2021a). The simulations are evaluated using observations for the same overlapping period. Since each simulation covers only one year, we caution about the statistical robustness of the model-observation differences and recommend that further evaluation using decade-long simulations be performed soon.

3.0 Results

3.1 Tropical Cyclones, Atmospheric Rivers, and Extratropical Cyclones

In this section we present an examination of the TCs, ARs, and ETCs and compare their representation in the free-running and nudged SCREAMv1 simulations. We begin by looking at the frequency of occurrence of each of these storm types in our region of interest during the simulation period. Here, frequency of occurrence is measured as the total number of 6-hourly time points that contain a storm type at a given location. Figure 1 shows these storm occurrence frequencies for each storm type (as different columns) and for the free-running (top row) and nudged (middle row) simulations. For comparison, ERA5 results for the same period are shown in the bottom row of Figure 1. There are several similarities between the two simulations and ERA5, but several key differences also stand out. The first is the significant increase in TC frequency in the free-running simulation relative to the nudged simulation and ERA5, particularly over the northeastern Pacific. The second is a reduced Pacific AR frequency in the free-running simulation, while an increase in Atlantic ARs is simultaneously observed. Finally, ETCs occur more frequently in ERA5 than either SCREAM simulation. In the following sub-sections, we will examine each storm type in greater detail to better understand these differences.

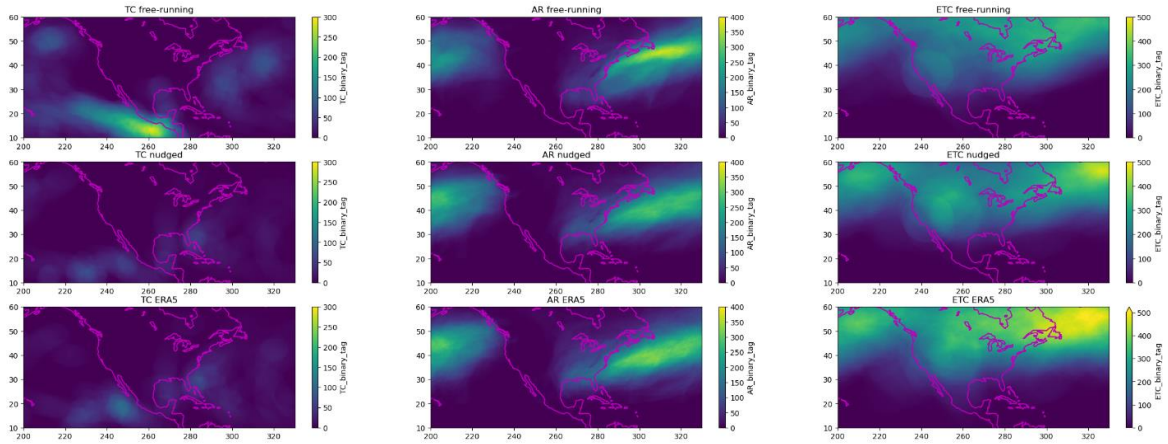

Figure 1. Storm frequency for the free running simulation (top row), nudged simulation (middle row) and ERA5 (bottom row) for TCs (left), ARs (middle), and ETCs (right). The occurrence counts are the number of 6-hourly time points that contain the relevant storm type at a given location.

Tropical Cyclones

To better understand the differences in TCs across the two simulations and reanalysis, we look at a distribution of the maximum 10-m wind speed along each TC track. For tracking TCs we use a minimum threshold of 10 m/s for winds, so we include tropical depressions, tropical storms, and tropical cyclones together. Figure 2 shows the distributions of TCs in each simulation (left two columns), ERA5 (third column), and the IBTrACS observations (fourth column) for comparison. The blue bars in each panel correspond to the August 2019-August 2020 period of the SCREAM simulations for a fair comparison. The orange, empty histogram overlaid in each panel provides the climatological IBTrACS data for 1979-2023.

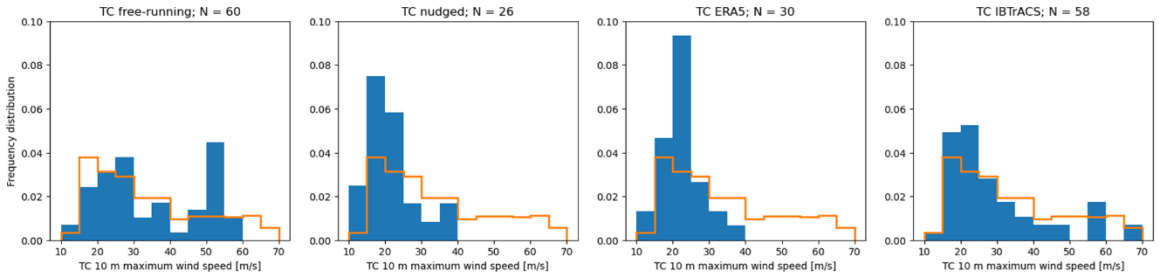

Figure 2. Frequency distribution of TC 10-m wind speed (maximum value along each track measured in m/s) for free-running SCREAM (first column), nudged SCREAM (second column), ERA5 (third column), and IBTrACS (fourth column). The blue bars in each panel correspond to the August 2019-August 2020 period. The orange, empty histogram in each panel provides the distribution for the IBTrACS data for 1979-2023. The TC count is provided in each panel's title.

The first takeaway from Figure 2 is that the larger number of TCs in the free-running SCREAM shown in Figure 1 (nearly double the count of the nudged SCREAM and ERA5) agrees better with observations, showing that this is not a shortcoming of the model, but an improvement. ERA5 struggles to generate sufficient TC counts, and nudging to its data inhibits SCREAM from generating TCs as well. Additionally, ERA5 struggles to produce strong storms. During the 13-month period examined here,

ERA5 produces no TCs with 10-m wind speeds reaching 40 m/s (approximately Category 2 wind speeds). Nudging the SCREAM model creates this same limitation in representing strong TCs in SCREAM. When SCREAM is free running, we see more storms in the 40-60-m/s max wind speed range. Compared to IBTrACS, the free-running SCREAM clearly performs the best, which agrees with recent findings from Baker et al. (2024), who showed that simulated TCs become much more realistic at a resolution of approximately 5 km.

The free-running SCREAM simulation does, however, appear to produce too many strong TCs relative to what is observed. This over-abundance of strong TCs may be related to an environment that is too favorable for precipitation. Figure 3 shows the SCREAM simulations compared against NASA's Integrated Multi-satellite Retrievals for GPM (IMERG) product. The top row uses the full 13-month record of the simulations (with a matching period used for IMERG), and the middle row shows the August through October period of 2019 when TCs are more active. Precipitation is biased high in the free-running SCREAM simulation, and the bias is stronger than normal during the period when TCs are active. Nudging is shown to substantially reduce this bias in precipitation, suggesting that the large-scale circulation in the free-running SCREAM may be too favorable for convection to develop. Future work is needed, but we hypothesize that this environment that is too favorable for precipitation contributes to the over-abundance of strong TCs. Future development efforts for SCREAM should seek to reduce the precipitation and strong TC biases, while taking care not to degrade the TC count, which is currently well represented by the model.

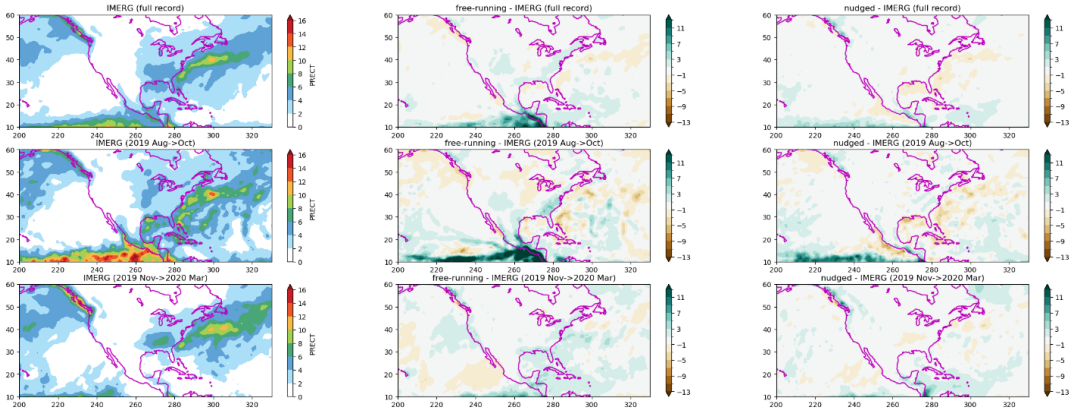

Figure 3. Precipitation in IMERG data (left), free-running SCREAM bias (middle), and nudged SCREAM bias (right) for the 13-month analysis period (top row), August-October 2019 (middle row), and November 2019-March 2020 (bottom row). All panels display precipitation with units of mm/day.

Atmospheric Rivers

ARs are tracked as shapes where we can quantify their area and accumulated precipitation. Figure 4 shows the area of ARs for the free-running SCREAM, nudged SCREAM, and ERA5 (top row, left to right), as well as the accumulated precipitation for the same simulations and reanalysis (bottom row). Unlike TCs where free-running SCREAM produced more storms than ERA5, free-running SCREAM produces fewer ARs than ERA5. As already noted from Figure 1, the reduction in AR count is primarily realized over the Pacific Ocean. Despite the total AR count being lower in the free-running SCREAM, the model still does a reasonable job of capturing the distribution of AR areas and accumulated precipitation values compared to ERA5. When nudged, SCREAM also does a good job of capturing the distributions

of AR area and accumulated precipitation. This performance is not terribly surprising given that the spatial scales of ARs are well represented with both SCREAM's and ERA5's horizontal resolution. The bottom row of Figure 4 shows the precipitation biases in the two SCREAM simulations relative to IMERG for November 2019 through March 2020. It is worth pointing out that despite having a low bias in AR count, the free-running SCREAM has a smaller positive bias in orographic precipitation than the nudged SCREAM experiment. A single year of simulation makes it difficult to know whether the changes in this bias result from a model bias or simply interannual variability. Future analyses are needed to better understand the relationship between these large-scale weather patterns and their associated precipitation.

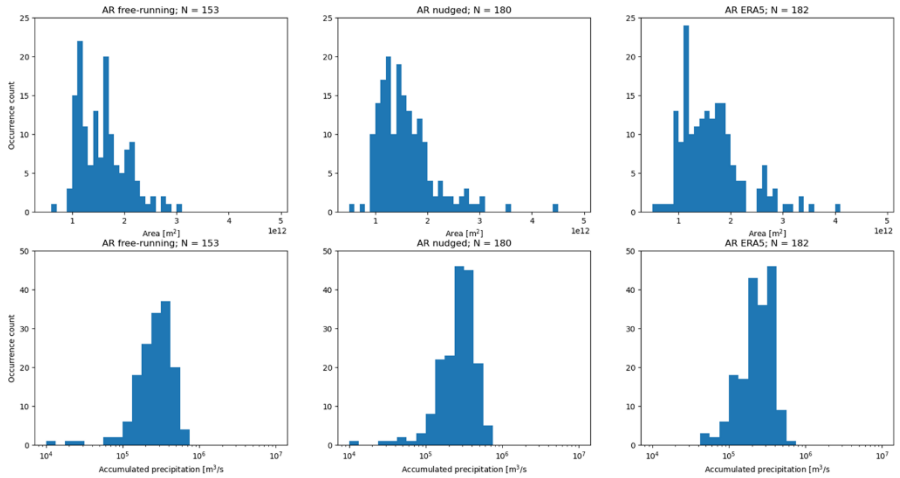

Figure 4. Distribution of occurrence counts for AR lifetime-averaged area (top row) and lifetime-averaged accumulated precipitation (bottom row) for free-running SCREAM (left column), nudged SCREAM (middle column), and ERA5 (right column). "Accumulated precipitation" is the total precipitation summed spatially across all pixels tagged as part of the AR object. AR counts are provided in each panel's title. Area has units of m^2 and accumulated precipitation has units of m^3/s.

To better understand how ARs are represented, we examine their evolution across the simulations and ERA5. Figure 5 shows the life cycle of each AR, presented with normalized time units so that a time of 0 is the AR start and time 100 is the AR end. The growth and decay of AR area is roughly symmetric when averaged across all ARs in ERA5, and this feature is well captured by both SCREAM experiments. Precipitation tends to be strongest early in the AR's lifetime in ERA5 (roughly 10-15% of the way through the storm's life) and decays gradually from there. Again, this precipitation feature is captured by both SCREAM experiments.

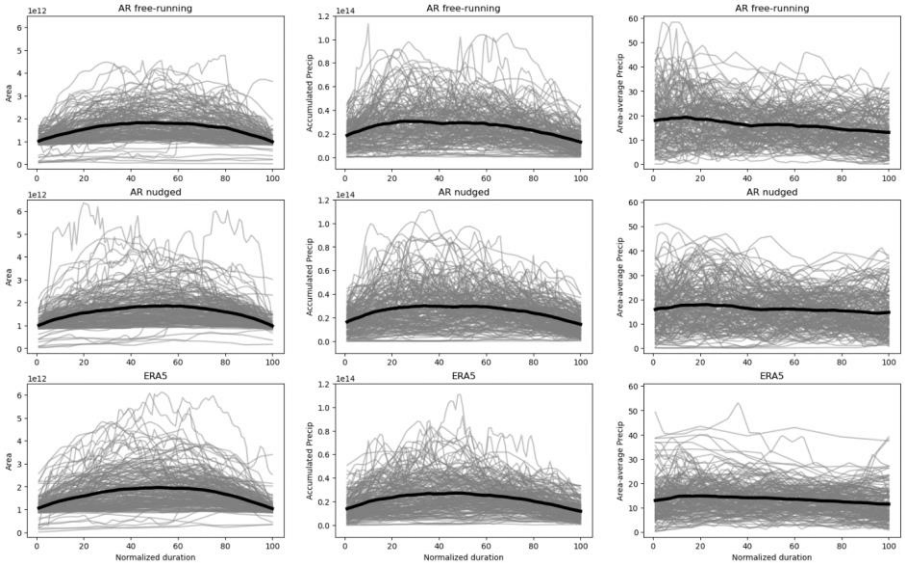

Figure 5. AR area (left column), accumulated precipitation (middle column), and area-average precipitation (right column) for free-running SCREAM (top row), nudged SCREAM (middle row), and ERA5 (bottom row). Each gray line corresponds to an individual AR, while the black line is the average of all ARs. Area has units of m^2, accumulated precipitation has units of m^2mm/day, and area-average precipitation has units of mm/day.

Extratropical Cyclones

Like ARs, ETCs exist on a spatial scale that is expected to be well represented by both SCREAM and ERA5. Figure 6 shows the 10-m maximum along track wind speeds as done for TCs. Again, like ARs, there are fewer ETCs in the free-running SCREAM experiment than in ERA5. It is surprising that there are also fewer ETCs in the nudged SCREAM experiment. It is unclear why this discrepancy exists between the nudged simulation and ERA5, and future work is needed to uncover its cause. The other feature of note when comparing SCREAM to ERA5 is the shift in the wind speed distribution toward stronger values in the free-running experiment. Chen et al. (2024) showed that ERA5 tends to underestimate strong near-surface winds relative to surface observations, so it is reasonable to expect that the shift in wind speeds is an improvement in their representation. Future work using observations to compare with the simulated winds in SCREAM can help to assess the magnitude of any biases associated with simulated ETCs.

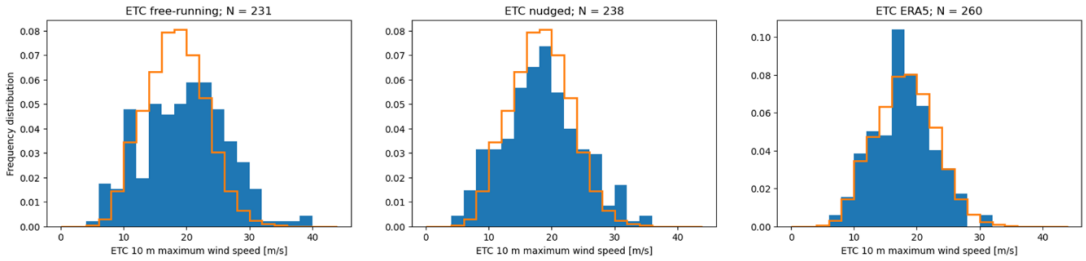

Figure 6. Frequency distribution of ETC 10-m wind speed (maximum value along each track measured in m/s) for free-running SCREAM (left column), nudged SCREAM (middle column), and ERA5 (right column). The blue bars in each panel correspond to the August 2019-August 2020 period. The orange, empty histogram in each panel provides the distribution for ERA5 data for 1979-2020. The ETC count is provided in each panel's title.

3.2 Mesoscale Convective Systems

Previous studies have identified significant negative biases in MCS frequency and precipitation in simulations at grid spacings of 25-50 km (e.g., Feng et al. 2021b, Leung et al. 2022). At 3.25-km grid spacing, SCREAM reproduced the observed spatial pattern of MCS frequency across the CONUS quite well, highlighting more frequent occurrence of MCSs east of the Rocky Mountains (Figures 7 and 8). However, the free-running simulation produced MCSs too frequently in the Northern Great Plains and southeastern U.S. during spring (Figure 7) and not enough in the Central Great Plains and the Midwest as well as too frequently in the northeastern U.S. during summer (Figure 8). These biases are significantly reduced in the nudged simulation in both seasons, suggesting a major role of large-scale circulation biases in limiting skillful modeling of MCS frequency in the free-running simulation.

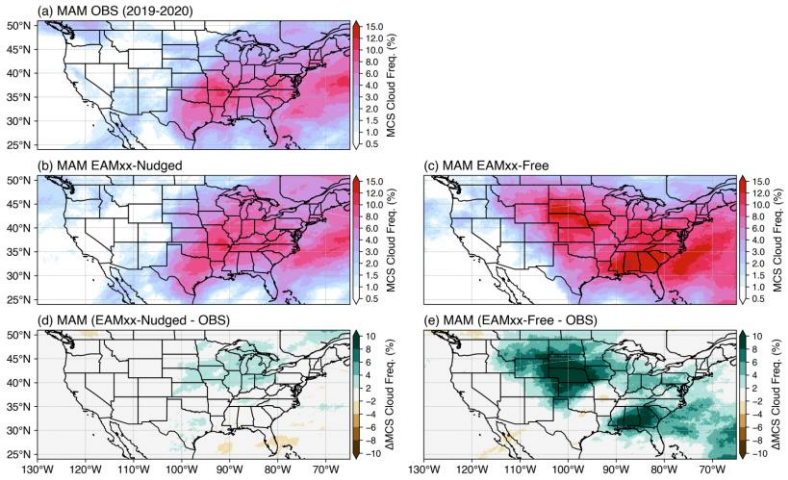

Figure 7. MCS frequency (units: %) in March-April-May (MAM) of 2020 in (a) observation, (b) nudged simulation, and (c) free-running simulations and the difference between the (d) nudged and (e) free-running simulation compared to observation. MCS frequency, also labeled as MCS cloud frequency, is defined by how often an MCS cloud shield mask is identified by the MCS tracking algorithm at each hour and each grid point.

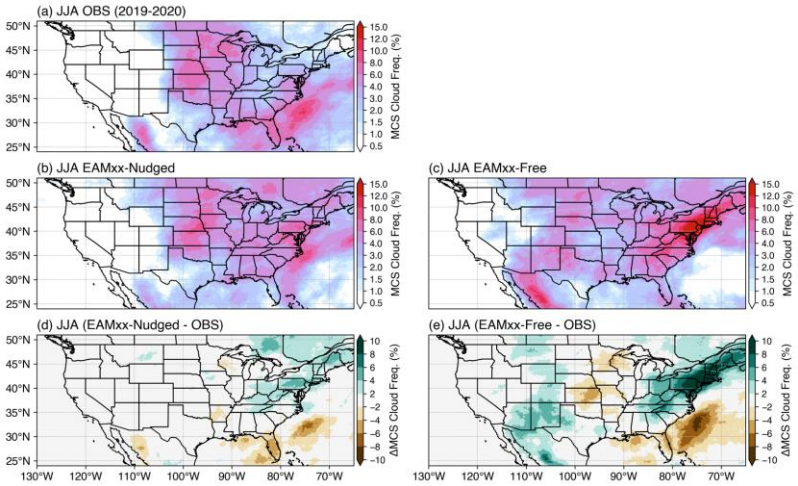

Figure 8. Same as Figure 7 but for June-July-August (JJA) of 2020.

Figure 9 summarizes the analysis of MCS frequency by comparing the observed and simulated monthly total MCS number over the Central U.S. The nudged simulation reproduced the monthly variations of MCS number found in observations, with a peak occurrence of 40 (42) MCSs in July and a total of 176 (178) MCSs between March and October in the simulation (observation). The free-running simulation also realistically captured the seasonal variability except for a significantly larger and smaller number of MCSs in May and July, respectively, compared to observation. The simulation produced a total of 196 MCSs, which is 10% higher than the observed MCS number.

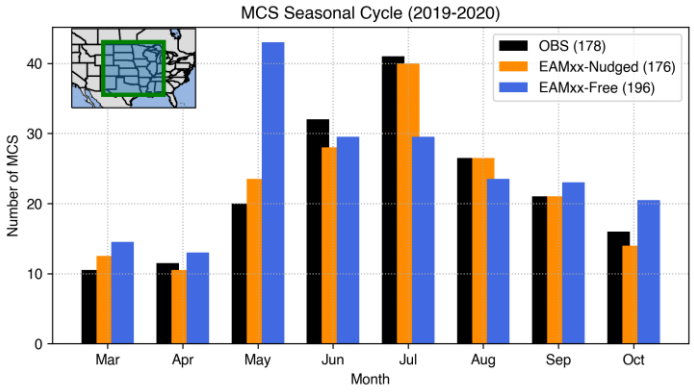

Figure 9. Number of MCS in the Central U.S. for each month between September 2019 and August 2020 for observation (black), nudged simulation (orange), and free-running simulation (blue).

Consistent with the biases in MCS number, the free-running simulation features an over-abundance of MCS precipitation in the Northern Great Plains and southeastern U.S. during spring (Figure 10), while MCS precipitation is too low in the Central Great Plains and the Midwest but too high in the northeastern U.S. during summer (Figure 11). Nudging results in a much better simulation of the spatial distribution of MCS precipitation. However, both simulations feature large negative biases in MCS precipitation in the northern Gulf of Mexico, consistent with negative biases in total precipitation there (not shown).

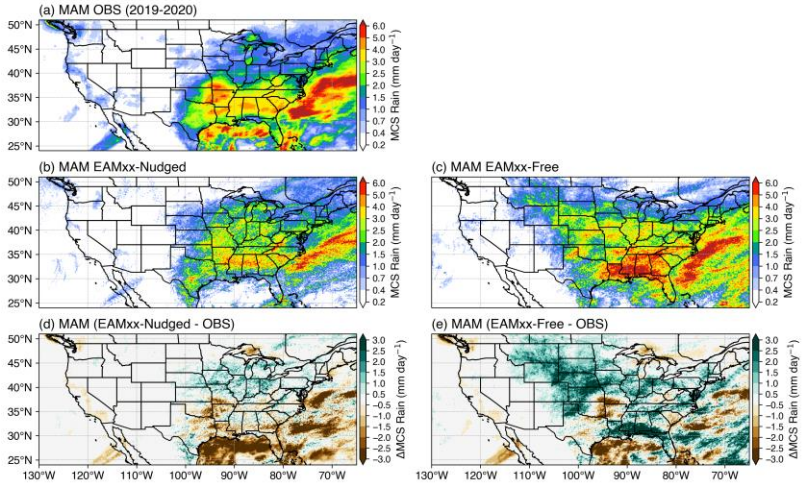

Figure 10. Similar to Figure 7 but for MAM MCS precipitation (units: mm day^{-1}).

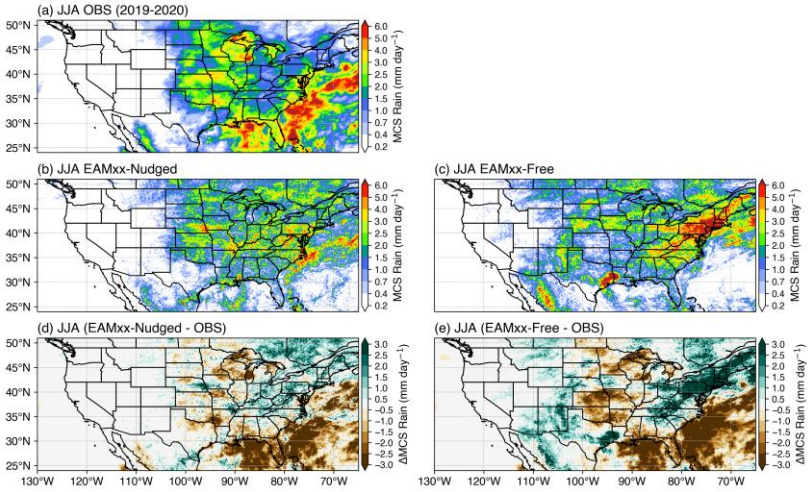

Figure 11. Similar to Figure 10, but for JJA.

Lastly, we evaluate how well SCREAM simulates MCS characteristics including the MCS lifetime, cold cloud shield (CCS) area (size of the stratiform anvil), minimum brightness temperature (an indication of the depth of the convective core), movement speed, precipitation feature (PF) area and rain rate, total rain volume, and heavy rain ratio. Figures 12 and 13 summarize the MCS characteristics using kernel density estimates (KDEs) constructed based on the MCSs tracked in each data set in spring and summer. During spring (Figure 12), both simulations generally reproduce the observed KDEs of the MCS lifetime, the CCS area, brightness temperature, and total rain volume. However, the MCSs in both simulations produce much higher PF mean rain rate and heavy rain ratio but smaller PF area compared to the observed MCSs. Biases in the PF area and rain rate offset each other to produce more realistic total rain volume. Similarity between the biases in the two simulations suggests a major contribution from model physics that influence the MCS precipitation characteristics. On the contrary, the nudged simulation shows a larger bias in MCS movement speed compared to the free-running simulation, suggesting modulations of the MCS movement speed by nudging the large-scale circulations. Similar biases of small PF area and high PF rain rate and heavy rain ratio are also found for MCSs in both simulations during summer (Figure 13). As in spring, the nudged simulation also continues to produce higher MCS movement speed than observation and the free-running simulation. The degradation of MCS movement speed in the nudged simulations suggest complex interactions between MCS internal dynamics and large-scale circulations that should be further investigated.

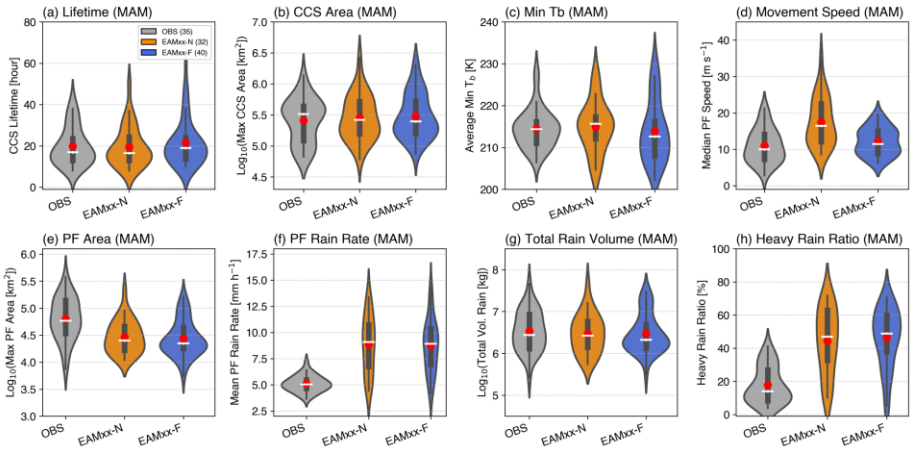

Figure 12. Spring (MAM) MCS characteristics depicted by violin plots of (a) MCS lifetime (units: hour), (b) cold cloud shield (CCS) area (units: km^2), 9c) minimum brightness temperature (T_b) (units: K), (d) movement speed (unit: m s^{-1}), (e) precipitation feature (PF) area (units: km^2), (f) PF rain rate (units: mm h^{-1}), (g) total rain volume (units: kg), and (h) heavy rain ratio (units: %). KDEs for observation, nudged simulation, and free-running simulation are shown in grey, orange, and blue, respectively. Also shown within each KDE are the median (white horizontal line), mean (red dot), interquartile range (vertical black bars), and 1.5x interquartile range (vertical black lines).

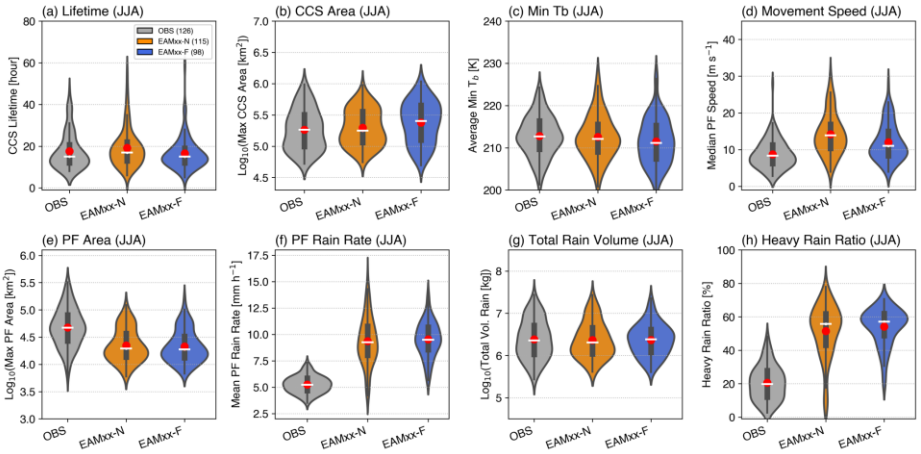

Figure 13. Same as Figure 12 but for summer (JJA).

4.0 Summary and Future Work

This report evaluates four different types of storms in two 13-month-long SCREAM simulations performed on the GPUs of the Frontier exascale computer and Perlmutter. To put the above results in context, it is useful to qualitatively compare the km-scale model skill in simulating AR, ETC, TC, and MCS with previous studies documenting some aspects of the performance of E3SM or similar models at 25-100-km resolutions. For ARs, the 2023 Q1 metric report evaluated ARs in the E3SM-HR and E3SM-LR simulations at 25-km and 100-km resolutions, respectively (Caldwell et al. 2019). The analysis noted an over-abundance of ARs in the Eastern Pacific relative to ERA5 while AR precipitation intensity along the U.S. west coast is significantly underestimated in both simulations except in southern California. Here, AR frequency in the free-running SCREAM simulation is much lower than observed in

the Eastern Pacific, while nudging can largely reduce the bias, suggesting some challenges for the free-running SCREAM simulation to reproduce the observed large-scale circulation that supports the development of ARs. In contrast with E3SM-HR and E3SM-LR, both SCREAM simulations much more realistically capture the AR precipitation intensity over land, highlighting the positive impact of km-scale models in simulating heavy precipitation associated with orographic effects in the U.S. west coast.

For TCs, Balaguru et al. (2020) evaluated TCs in E3SM-HR and E3SM-LR and noted the significant impacts of increasing resolution from 100 km to 25 km in improving modeling of TCs. Despite the obvious improvements, however, E3SM-HR still underestimates TC frequency and intensity in the Eastern Pacific. In contrast, the free-running SCREAM simulated a comparable number of TCs observed during the simulation period, and the simulated TC intensity also matches well with the observations. As noted above, nudging the simulated winds towards MERRA-2 suppresses the formation of TCs and their intensity, suggesting that interactions between TCs and their environments play an important role in TC development and intensification.

For ETCs, Gore et al. (2023) determined the large-scale meteorological patterns of ETC in the U.S. from observations and evaluated how well E3SM-HR and E3SM-LR, among other models in the Coupled Model Intercomparison Project Phase 6 (CMIP6), captured these patterns and the associated ETC frequency and intensity. Overall, E3SM-HR and E3SM-LR well simulated ETC frequency and intensity, with < 5% bias in ETC frequency and < 7% bias in ETC intensity. Here, for a single year, SCREAM shows slightly degraded performance compared to E3SM-HR and E3SM-LR in terms of ETC frequency (11% and 8% underestimations in the free-running and nudged simulations, respectively). The free-running simulation has relatively more intense ETCs compared to ERA5, and this bias is reduced in the nudged simulation. As relatively larger-scale features, the ability of models to capture ETCs appears to be sensitive to how well atmospheric circulations are simulated.

For MCSs, several studies adopted the same methodology for evaluating MCSs in climate simulations as described in this report. These studies consistently identified significant underestimation of MCS numbers in climate simulations at resolutions of 25-50 km in the central U.S. during the warm season. For example, simulations by the Model for Prediction Across Scales (MPAS) at both 25-km and 50-km resolution show a negative bias in MCS number by ~75% and MCS precipitation by ~52% (Feng et al. 2021b). In the 2020 Q2 metric report, MCSs simulated by E3SM at 25-km resolution were compared against observations over the central U.S. The simulation underestimated MCS number by ~35% and MCS precipitation by ~60% in spring and ~90% in summer. SCREAM is performing notably better, with the free-running simulation overestimating MCS number by ~10% during the warm season, although larger biases are found in May (~100% overestimation) and July (~25% underestimation). Nudging largely reduced the biases in simulating both MCS number and precipitation, highlighting the importance for improving the large-scale circulation that provides favorable environments for the development of MCS in the U.S.

Overall, analysis of four storm types that are major drivers of extreme precipitation and floods in the U.S. in observations and in SCREAM simulations at 3.25-km resolution provides evidence of skillful simulations of storm frequency and intensity. Qualitative comparison of model skill between the SCREAM simulations and the high-resolution simulations at 25-50-km resolution reported in previous studies show clear advantages of km-scale modeling for simulating storm intensity, while storm frequency is generally comparable or improved. Comparison of the free-running and nudged SCREAM simulations underscores the need for improving modeling of the large-scale circulation, which has large

impacts on simulating storm frequency, but does not have an obvious dependence on model resolution. As atmospheric circulation is highly variable across a wide range of timescales, the preliminary findings reported here should be further investigated using longer simulations, such as the decadal simulations being performed by the E3SM project using SCREAMv1. Lastly, atmospheric circulations that influence storms in the U.S. are sensitive to tropical heating through its impact on atmospheric overturning circulations and teleconnections. As tropical heating is strongly influenced by physical processes in the atmosphere, ultimately improving and better calibrating physics parameterizations such as cloud microphysics and atmospheric turbulence that are not resolved even at km scale is critical for fully realizing the value of km-scale models for climate predictions and projections.

5.0 References

Baker, AJ, B Vannière, and PL Vidale. 2024. "On the Realism of Tropical Cyclone Intensification in Global Storm-Resolving Climate Models." *Geophysical Research Letters* 51(17): e2024GL109841, https://doi.org/10.1029/2024gl109841

Balaguru, K, LR Leung, LP Van Roekel, J-C Golaz, PA Ullrich, PM Caldwell, SM Hagos, BE Harrop, and A Mametjanov. 2020. "Characterizing tropical cyclones in the Energy Exascale Earth System Model version 1." *Journal of Advances in Modeling Earth Systems* 12(8): e2019MS002024, https://doi.org/10.1029/2019MS002024

Caldwell, PM, CR Terai, B Hillman, ND Keen, P Bogenschutz, W Lin, H Beydoun, M Taylor, L Bertagna, AM Bradley, TC Clevenger, AS Donahue, C Eldred, J Foucar, J-C Golaz, O Guba, R Jacob, J Johnson, J Krishna, W Liu, K Pressel, AG Salinger, B Singh, A Steyer, P Ullrich, D Wu, X Yuan, J Shpund, H-Y Ma, and CS Zender. 2021. "Convection-permitting simulations with the E3SM global atmosphere model." *Journal of Advances in Modeling Earth Systems* 13(11): e2021MS002544, https://doi.org/10.1029/2021MS002544

Chen, TC, F Collet, and A Di Luca. 2024. "Evaluation of ERA5 precipitation and 10-m wind speed associated with extratropical cyclones using station data over North America." *International Journal of Climatology* 44(3): 729–747, https://doi.org/10.1002/joc.8339

Donahue, AS, PM Caldwell, L Bertagna, H Beydoun, PA Bogenschutz, AM Bradley, TC Clevenger, J Foucar, C Golaz, O Guba, W Hannah, BR Hillman, JN Johnson, N Keen, W Lin, B Singh, S Sreepathi, MA Taylor, J Tian, CR Terai, PA Ullrich, X Yuan, and Y Zhang. 2024. "To exascale and beyond—The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales." *Journal of Advances in Modeling Earth Systems* 16(7): e2024MS004314, https://doi.org/10.1029/2024MS004314

Feng, Z, LR Leung, N Liu, J Wang, RA Houze, J Li, JC Hardin, D Chen, and J Guo. 2021a. "A Global High-Resolution Mesoscale Convective System Database Using Satellite-Derived Cloud Tops, Surface Precipitation, and Tracking." *Journal of Geophysical Research – Atmospheres* 126(8): e2020JD034202, https://doi.org/10.1029/2020JD034202

Feng, Z, F Song, K Sakaguchi, and LR Leung. 2021b. "Evaluation of mesoscale convective systems in climate simulations: Methodological development and results from MPAS-CAM over the United States." *Journal of Climate* 34(7): 2611–2633, https://doi.org/10.1175/JCLI-D-20-0136.1

Feng, Z, J Hardin, HC Barnes, J Li, LR Leung, A Varble, and Z Zhang. 2023. "PyFLEXTRKR: a flexible feature tracking Python software for convective cloud analysis." *Geoscientific Model Development* 16(10): 2753–2776, https://doi.org/10.5194/gmd-16-2753-2023

Feng, Z, LR Leung, J Hardin, CR Terai, F Song, and P Caldwell. 2023. "Mesoscale convective systems in DYAMOND global convection-permitting simulations. " *Geophysical Research Letters* 50(4): e2022GL102603, https://doi.org/10.1029/2022GL102603

Gelaro, R, W McCarty, MJ Suarez, R Todling, A Molod, L Takacs, CA Randles, A Darmenov, MG Bosilovich, R Reichle, K Wargan, L Coy, R Cullather, C Draper, S Akella, V Buchard, A Conaty, AM da Silva, W Gu, G-K Kim, R Koster, R Lucchesi, D Merkova, JE Nielsen, G Partyka, S Pawson, W Putnam, M Rienecker, SD Schubert, M Sienkiewicz, and B Zhao. 2017. "The modern-era retrospective analysis for research and applications, Version 2 (MERRA-2)." *Journal of Climate* 30(14): 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1

Gore, MJ, CM Zarzycki, and MM Gervais. 2023." Connecting large-scale meteorological patterns to extratropical cyclones in CMIP6 climate models using self-organizing maps." *Earth's Future* 11(8): e2022EF003211, https://doi.org/10.1029/2022EF003211

Harrop, BE, K Balaguru, J-C Golaz, LR Leung, S Mahajan, AM Rhoades, PA Ullrich, C Zhang, X Zheng, T Zhou, PM Caldwell, ND Keen, and A Mametjanov. 2023. "Evaluating the water cycle over CONUS at the watershed scale for the Energy Exascale Earth System Model version 1 (E3SMv1) across resolutions." *Journal of Advances in Modeling Earth Systems* 15(11): e2022MS003490, https://doi.org/10.1029/2022MS003490

Hersbach, H, B Bell, P Berrisford, S Hirahara, A Horanyi, J Muñoz-Sabater, J Nicolas, C Peubey, R Radu, D Schepers, A Simmons, C Soci, S Abdalla, X Abellan, G Balsamo, P Bechtold, G Biavati, J Bidlot, M Bonavita, G De Chiara, P Dahlgren, D Dee, M Diamantakis, R Dragani, J Flemming, R Forbes, M Fuentes, A Geer, L Haimberger, S Healy, RJ Hogan, E Hólm, M Janisková, S Keeley, P Laloyaux, P Lopez, C Lupu, G Radnoti, P de Rosnay, I Rozum, F Vamborg, S Villaume, and J-N Thépaut. 2020. "The ERA5 global reanalysis." *Quarterly Journal of the Royal Meteorological Society* 146(730): 1999–2049, https://doi.org/10.1002/qj.3803

Kim, S, LR Leung, B Guan, and JCH Chiang. 2022. "Atmospheric river representation in the Energy Exascale Earth System Model (E3SM) version 1.0." *Geoscientific Model Development* 15(14): 5461–5480, https://doi.org/10.5194/gmd-15-5461-2022

Knapp, KR, MC Kruk, DH Levinson, HJ Diamond, and CJ Neumann. 2010. "The international best track archive for climate stewardship (IBTrACS) unifying tropical cyclone data." *Bulletin of the American Meteorological Society* 91(3): 363–376, https://doi.org/10.1175/2009BAMS2755.1

Leung, LR, DC Bader, MA Taylor, and RB McCoy. 2020. "An introduction to the E3SM special collection: Goals, science drivers, development, and analysis." *Journal of Advances in Modeling Earth Systems* 12(11): e2019MS001821, https://doi.org/10.1029/2019MS001821

Leung, LR, WR Boos, JL Catto, C DeMott, G Martin, JD Neelin, TA O'Brien, S Xie, Z Feng, N Klingaman, Y-H Kuo, RW Lee, C Martinez-Villalobos, S Vishnu, M Priestley, C Tao, and Y Zhou. 2022. "Exploratory precipitation metrics: Spatiotemporal characteristics, process-oriented, and phenomena-based." *Journal of Climate* 35(12): 3659–3686, https://doi.org/10.1175/JCLI-D-21-0590.1

Satoh, M, B Stevens, F Judt, M Khairoutdinov, S-J Lin, WM Putman, and P Düben 2019. "Global cloud resolving models." *Current Climate Change Reports* 5: 172–184, https://doi.org/10.1007/s40641-019-00131-0

Stevens, B, M Satoh, L Auger, J Biercamp, CS Bretherton, X Chen, P Düben, F Judt, M Khairoutdinov, D Klocke, C Kodama, L Kornblueh, S-J Lin, P Neumann, WM Putnam, N Röber, R Shibuya, B Vanniere, PL Vidale, N Wedi, and L Zhou. 2019. "DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains." *Progress in Earth and Planetary Science* 6(1): 1–17, https://doi.org/10.1186/s40645-019-0304-z

Taylor, MA, PM Caldwell, L Bertagna, C Clevenger, AS Donahue, JG Foucar, O Guba, BR Hillman, N Keen, J Krishna, MR Norman, S Sreepathi, CR Terai, JB White III, D Wu, AG Salinger, RB McCoy, LR Leung, and DC Bader. 2023. "The Simple Cloud-Resolving E3SM Atmosphere Model running on the Frontier Exascale system." *SC'23: Proceedings of the International Conference for High-Performance Computing, Networking, Storage and Analysis*, November 2023, Article No. 7, 1–11, https://doi.org/10.1145/3581784.3627044

Ullrich, PA, CM Zarzycki, EE McClenny, MC Pinheiro, AM Stansfield, and KA Reed. 2021. "TempestExtremes v2. 1: A community framework for feature detection, tracking, and analysis in large datasets." *Geoscientific Model Development* 14(8): 5023–5048, https://doi.org/10.5194/gmd-14-5023-2021

Zhou, W, LR Leung, and J Lu. 2022. "Linking Large-Scale Double-ITCZ Bias to Local-Scale Drizzling Bias in Climate Models." *Journal of Climate* 35: 7965–7979, https://doi.org/10.1175/JCLI-D-22-0336.1

U.S. DEPARTMENT OF
ENERGY

Office of Science