Zooming In: SCREAM at 100 m Using Regional Refinement over the San Francisco Bay Area

Jishi Zhang[1], Peter Bogenschutz[1], Mark Taylor[2], and Philip Cameron-smith[1]

[1]Lawrence Livermore National Laboratory
[2]Sandia National Laboratories

Correspondence: Jishi Zhang (zhang73@llnl.gov) and Peter Bogenschutz (bogenschutz1@llnl.gov)

Abstract. Pushing global climate models to large-eddy simulation (LES) scales over complex terrain has remained a major challenge. This study presents the first known implementation of a global model—SCREAM (Simple Cloud-Resolving E3SM Atmosphere Model)—at 100 m horizontal resolution using a regionally refined mesh (RRM) over the San Francisco Bay Area. Two hindcast simulations were conducted to test performance under both strong synoptic forcing and weak, boundary-layer-driven conditions. We demonstrate that SCREAM can stably run at LES scales while realistically capturing topography, surface heterogeneity, and coastal processes. The 100 m SCREAM-RRM substantially improves near-surface wind speed, temperature, humidity, and pressure biases compared to the baseline 3.25 km simulation, and better reproduces fine-scale wind oscillations and boundary-layer structures. These advances leverage SCREAM's scale-aware SHOC turbulence parameterization, which transitions smoothly across scales without tuning. Performance tests show that while CPU-only simulations remain costly, GPU acceleration with SCREAMv1 on NERSC's Perlmutter system enables two-day hindcasts to complete in under two wall-clock days. Our results open the door to LES-scale studies of orographic flows, boundary-layer turbulence, and coastal clouds within a fully comprehensive global modeling framework.

1 Introduction

With the rise of high-performance computing, general circulation models (GCMs) are now able to operate at convection-permitting scales—often referred to as the *k*-scale. Meanwhile, many studies have employed limited-area cloud-resolving or mesoscale models at large-eddy simulation (LES) scales at O(100) m. These high-resolution simulations have substantially advanced our understanding of turbulence in complex terrain, a topic of importance for local microclimates, atmospheric exchanges over mountains, renewable energy, forest fires, urban meteorology, and more (De Wekker et al., 2018). Taking renewable energy as an example: wind energy relies on boundary-layer wind and turbulence for optimal turbine siting and resource assessment; hydropower depends on snowmelt, runoff, and transpiration; solar energy is influenced by terrain shading, cloud formation, and aerosol radiative effects. To accurately simulate such processes and guide energy infrastructure deployment, resolving lower-boundary heterogeneities is crucial. A robust way to improve model reliability is to increase resolution, thereby reducing the reliance on phenomenological parameterizations.

Because high-resolution observational networks remain sparse, LES simulations often serve as a benchmark for evaluating mesoscale models and GCMs. The convergence of LES solutions is attributed to Reynolds-number similarity: once the grid spacing is sufficiently fine to resolve the dominant large eddies (i.e., those not governed by viscous dissipation) and to capture the critical turbulent motions set by the flow's Reynolds number, the statistical nature of turbulence becomes stable, changing little with further refinement. Traditional LES configurations, however, typically focus on idealized flows. Recently, researchers have adapted LES turbulence parameterizations to include realistic topography and are more tightly coupled with an interactive land model, making them suitable for realistic flows over complex terrain (e.g., Chow et al., 2006; Zhong and Chow, 2012; Arthur et al., 2018).

Over the past two decades, global circulation models have experienced a dramatic increase in horizontal resolution—from around 100 km down to about 1 km, the so-called *k*-scale. Current *k*-scale global climate models (GCMs), which explicitly resolve deep convection and employ non-hydrostatic dynamics, have demonstrated substantial improvements over traditional O(100)-km models—advancements that cannot be achieved through refinement of physical parameterizations alone(e.g., Stevens et al., 2019; Caldwell et al., 2021). One might ask: is it feasible to push *k*-scale GCMs down to LES scales? The primary challenge is that boosting a global model from a few kilometers to O(100 m) resolution demands enormous computational resources—far more than that required by limited-area models. Nevertheless, regionally refined mesh (RRM) technology holds considerable promise in this context. RRM—also referred to as variable-resolution modeling in the CESM and MPAS-A communities—allows for a smooth transition from coarse to fine horizontal resolution over targeted regions, achieving much of the benefit of high-resolution simulations at a fraction of the computational cost of globally uniform high-resolution models. (e.g., Ringler et al., 2008; Harris and Lin, 2013; Zarzycki and Jablonowski, 2014; Huang and Ullrich, 2017; Tang et al., 2019, 2023; Rhoades et al., 2023).

The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), developed under the U.S. Department of Energy's Energy Exascale Earth System Model (E3SM) project, represents such a global 3.25 km convection-permitting model with RRM capabilities (Caldwell et al., 2021; Donahue et al., 2024). SCREAM-RRM has demonstrated good skill in simulating extreme weather events, such as derechos and atmospheric rivers (Liu et al., 2023; Bogenschutz et al., 2024), as well as in capturing the influence of complex topography and microclimates on regional climate across both historical and future periods in California (Zhang et al., 2024a).

Compared to the traditional approach of incrementally adding complexity to idealized LES models, increasing the resolution of a *k*-scale climate model to the LES scale offers unique advantages. Climate models inherently account for interactions between multiple components, including complex topography, interactive land processes, and turbulence-cloud interactions. Furthermore, SCREAM's hindcast configuration closely resembles operational forecasting, enabling realistic case studies by initializing from reanalysis and prescribing sea surface temperature (SST) and sea-ice cover from satellite observations. With a scale-aware turbulence scheme, it is conceptually feasible to transition from *k*-scale to LES resolution. Notably, the turbulence grey zone spanning O(1) km to O(100) m, features grid spacings comparable to the scale of major turbulent eddies (e.g., Langhans et al., 2012).

Physical parameterizations for microphysics and radiative transfer can, at least structurally, remain consistent across model resolutions down to micron scales—although accounting for terrain shading/reflection in radiation would be beneficial. Once Δx approaches 100 m, the effective resolution ($\approx$6–8Δx) can resolve deep convective elements. Indeed, Bogenschutz et al. (2023) demonstrated that SCREAM's turbulent transport in the idealized doubly periodic f-plane configuration (DP-SCREAM) maintains reasonable scale-awareness and scale-insensitivity down to 100 m, with simulation fidelity improving as resolution increases. These findings motivate further investigation into whether SCREAM can be extended to large-eddy simulation (LES) scales over complex terrain and how it performs in fully comprehensive configurations.

This paper documents our efforts to push SCREAM from 3.25 km to 100 m. We perform two hindcast simulations using SCREAM at a horizontal resolution of approximately 800 m over California, featuring an embedded high-resolution mesh refined to 100 m over the San Francisco Bay Area. While the focus is on the refined mesh over the Bay Area, to our knowledge, this represents the first attempt to test a GCM-based model at LES scales. While ICON (Icosahedral Nonhydrostatic) has also developed an LES version, it employs a conventional limited-area configuration (Dipankar et al., 2015). In contrast, a GCM like SCREAM must meticulously conserve energy, mass, and tracers while faithfully simulating a wide range of climate regimes. As a result, some of the challenges we identify may be specific to the GCM framework. At the same time, SCREAM-RRM avoids several limitations commonly associated with limited-area models, as discussed in this work. This paper details the modifications made to SCREAM's toolchain to support 100 m simulations, with the broader objective of enabling reproducible LES-scale SCREAM experiments across other regions of interest.

The remainder of this paper is organized as follows. Section 2 describes our methodology, including the design and setup of 100 m SCREAM-RRM for the San Francisco Bay Area, the hindcast case description, and the observational datasets used for model evaluation. Section 3 presents simulation results, including comparisons with the 3.25 km SCREAM-RRM and observations. Section 4 summarizes the improvements, discusses remaining issues and the broader implications of this work.

2 Methods

2.1 Introduction to SCREAM

The 100 m SCREAM-RRM in this study is developed based on SCREAM version 0 (Caldwell et al., 2021). The default horizontal resolution of SCREAM is 3.25 km and does not have parameterized deep convection. SCREAM uses a 128-layer hybrid vertical coordinate, with vertical thickness in the boundary layer increasing from about 30 m near the surface to approximately 200 m near the top of the boundary layer. The model top is located at around 40 km. SCREAM's dynamical core is built upon the High-Order Methods Modeling Environment (HOMME), employing nonhydrostatic fluid dynamics (Taylor et al., 2020). Among the available thermodynamic variable options in HOMME, this study adopts the default configuration using virtual potential temperature. HOMME supports multiple time-stepping approaches; by default, the Horizontally Explicit Vertically Implicit (HEVI) scheme (Satoh, 2002) is discretized using a Runge–Kutta IMplicit–EXplicit (IMEX) method (Steyer et al., 2019; Guba et al., 2020). This scheme combines the KGU53 third-order explicit Runge–Kutta method (Guerra and Ullrich, 2016) for most prognostic equations with a second-order implicit Runge–Kutta method for vertically propagating acoustic

waves. However, at extremely high resolution, the IMEX approach offers diminishing computational advantages and can encounter convergence challenges in the Newton iteration associated with its implicit component. Accordingly, we adopt the third-order KGU53 explicit scheme, which is commonly used in hydrostatic configurations but is also applicable to the non-hydrostatic mode, provided the time step remains below approximately 0.5 seconds. HOMME's dynamical core uses spectral elements with a 4×4 Gauss–Lobatto–Legendre (GLL) nodal grid within each element, referred to as GLL grids. The physical grid is subdivided into 2×2 grids, referred to as pg2 grids, uniformly distributed on each spectral element (Hannah et al., 2021).

SCREAM employs the Simplified Higher-Order Closure (SHOC) scheme as a unified parameterization for boundary-layer turbulence, cloud macrophysics, and shallow cumulus (Bogenschutz and Krueger, 2013). SHOC employs a double-Gaussian probability density function (PDF) to diagnose cloud fraction, cloud liquid water content, and buoyancy flux. The diagnosed buoyancy flux is then used to close the subgrid-scale (SGS) turbulent kinetic energy (TKE) equation. Higher-order moments for the PDF are diagnosed rather than predicted for computational efficiency. Vertical fluxes of turbulence are represented by downgradient diffusion (similar to Bretherton and Park (2009)). Eddy diffusivity is updated based on the predicted SGS TKE, consistent with Cheng et al. (2010), who suggested that when SGS TKE is accurately predicted, downgradient closures generally function well for k-scale models. Since SHOC is designed to operate across spatial scales ranging from a few hundred meters to several hundred kilometers, SHOC sets the maximum allowable value of (L) to the size of the horizontal grid mesh. In addition, L depends on SGS TKE, boundary-layer depth, eddy turnover timescale (boundary-layer depth divided by the convective velocity scale), and local stability, making it scale aware because SGS TKE typically increases with increasing Δx, thus increasing L. Phenomenologically, the maximum unresolved eddy size increases (or decreases) with the cloud depth (or buoyancy flux).

The cloud fraction and cloud liquid water diagnosed by SHOC are passed to the Predicted Particle Properties (P3) microphysics scheme (Morrison and Milbrandt, 2015) and the RTE+RRTMGP radiative transfer package (Pincus et al., 2019). The cloud droplet number concentration required by the microphysics scheme and gas optical properties for radiative transfer are provided by the Simple Prescribed Aerosols (SPA) module, which includes monthly aerosol climatology from 1° E3SM simulations. State variables and radiative fluxes at the lowest atmospheric level, updated by SHOC, microphysics, and radiation schemes are subsequently passed to the coupler, interacting with E3SM's land model (ELM) (Golaz et al., 2019) and the Los Alamos sea-ice model CICE4 (Hunke et al., 2008) with prescribed-ice and a data-ocean mode. Land surface fluxes, including momentum, sensible heat, and latent heat, rely on Monin–Obukhov similarity theory, iteratively solving for bare soil/canopy temperatures and surface fluxes via an energy-balance equation. Ocean surface fluxes are computed within the coupler, with multiple options provided. This study adopts the default bulk formula which iteratively solves stability, roughness, and the bulk transfer coefficients.

SCREAM version 1, implemented in C++ and referred to as EAMxx, has recently been documented and made publicly available (Donahue et al., 2024). Its global climatology closely resembles that of version 0 implemented in Fortran. When this study began, we only had CPU resources, and EAMxx was not yet fully operational. However, during manuscript preparation, we tested the 100 m EAMxx version of SCREAM-RRM on GPU nodes of the Perlmutter cluster at the National Energy

Research Scientific Computing Center (NERSC), and we present the associated performance benefits in Section 2.5. RRM configurations are largely compatible with both EAM and EAMxx, although EAMxx requires a different namelist structure and initial condition format.

2.2 San Francisco Bay Area 100 m SCREAM-RRM grid design

The level-1 refinement of the refined mesh used in this study covers the same domain as the 3.25 km California RRM employed by Zhang et al. (2024a) for future projections. However, in our configuration, the refined region has a horizontal resolution of 800 m over California, transitioning to 100 km outside the refined area. The 800 m refinement has been validated in Bogenschutz et al. (2024), capturing representative atmospheric river events. Building on the 800 m grid, a second layer of refinement at 100 m resolution is applied over the Bay Area, extending from Drakes Bay in the north to Santa Cruz in the south, westward to the Pacific Ocean off San Francisco, and eastward to Stockton (Fig. 1). This RRM grid achieves a maximum Dinv-based element distortion metric of 3.04, indicating high grid quality[1]). Notably, this grid yields a global refinement factor of 1000.

The 100 m region required careful design. Anticipating potential applications in wind energy assessment, we ensured that the city of Livermore was included within the refined domain, thereby capturing both the Altamont Pass and Shiloh wind farms. The selected 100 m region encompasses highly heterogeneous topography, ranging from Mount Diablo (~1000 m) and the Santa Cruz Mountains (100–1000 m) to the San Francisco Bay and the San Joaquin Valley, which lie near or below sea level (Fig. 1).

[1] See https://acme-climate.atlassian.net/wiki/spaces/DOC/pages/872579110/Running+E3SM+on+New+Atmosphere+Grids, last access: 22 April 2025

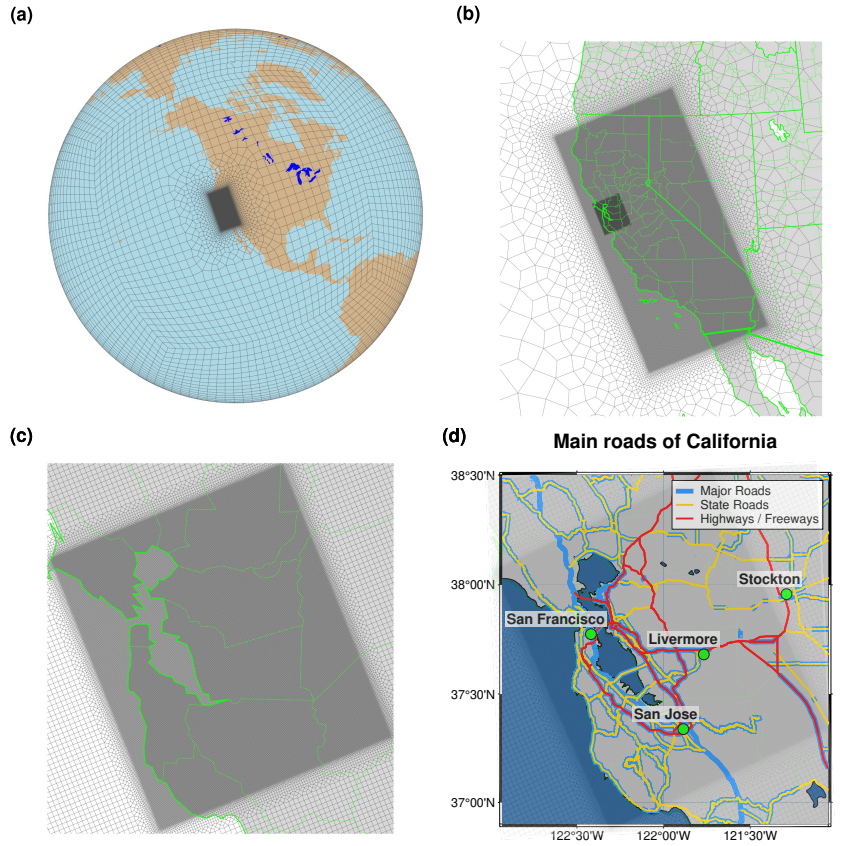

Figure 1. The regionally refined grid for (a) global view, (b) California view, (c) Bay Area view, and (d) Bay Area overlaid with roads in the 100 m SCREAM-RRM.

Notably, the 100 m Bay Area RRM is the first known attempt to refine a global mesh down to 100 m resolution. Our primary goal is to verify the feasibility of this configuration by demonstrating its stability and its ability to produce physically credible solutions. Although our 100 m coverage cannot be extensive, our high resolution mesh covers an area of 150 × 150 km^2—comparable to or larger than typical DP-SCREAM or large LES domains. The 100 m RRM contains 1,333,296 physical columns—more than twice the number in the 800 m California RRM (587,904 columns; Bogenschutz et al., 2024) and approximately 20 times that in the 3.25 km RRM (67,872 columns; Zhang et al., 2024a)—yet it accounts for only 5.3% of the total columns in the global 3.25 km SCREAM configuration.

The land model is executed on the same grid as the atmosphere, while the ocean component uses the MPAS oRRS18to6v3 mesh—a publicly available mesh developed by the E3SM community—which provides 18 km resolution in the tropics, gradually refining to 6 km near the poles. We generated the atmosphere/land RRM grid with SQuadGen version 1.2.2 (SQuadGen, Ullrich and Roesler (2024); https://github.com/ClimateGlobalChange/squadgen, last access: 25 January 2025), with the following command:

```
./SQuadGen --output CAne32x128_Altamont100m_v2.g \
           --refine_rect "-119.8,31,-118.8,43,7;-122.3,36.9,-121.7,38.5,10" \
           --refine_level 10 --resolution 32 --smooth_type SPRING \
           --lat_ref 38 --lon_ref -116 --orient_ref 20
```

The grid configuration files, including the domain files and mapping files for atmosphere/land/ocean, topography, land surface data, and dry deposition file, were generated using the E3SM standalone toolchains. Topography generation is discussed in detail in the next section.

2.3 Topography generation

As part of "Running E3SM on New Atmosphere Grids" toolchain (https://acme-climate.atlassian.net/wiki/spaces/DOC/pages/872579110/Running+E3SM+on+New+Atmosphere+Grids, last access: 25 January 2025), topography generation has been modified from the NCAR topography workflow (Lauritzen et al., 2015). In the default E3SM topography workflow, new grids are generated based on the ne3000 (3.25 km) cubed-sphere topography, which is derived from the USGS GTOPO30 digital elevation model (DEM) at 30-arc-second (~1 km) resolution. This is because most newly generated grids have coarser resolutions than SCREAM's default 3.25 km resolution. Therefore, the first step in our workflow was to obtain higher-resolution DEM data and interpolate it onto the cubed-sphere grid.

For regional modeling (such as WRF), DEM data finer than 100 m resolution is often readily available. For global grids, however, the highest publicly available DEM we found is USGS GMTED2010 at 7.5-arc-second (250 m) resolution, which lacks coverage poleward of 56° latitude (and Greenland). We obtained a publicly available 250 m GMTED2010 dataset (https://topotools.cr.usgs.gov/gmted_viewer/gmted2010_global_grids.php, last access: 25 January 2025), using a GIS tool to merge all tiles into a global lat-lon NetCDF file. GMTED2010 also provides 15-arc-second (500 m) global DEM data (Danielson and Gesch, 2011). We obtained a global 500 m GMTED2010 netcdf file by incorporating the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD44W land–water mask from J. van Geffen (personal communication, courtesy of Maarten Sneep, Koninklijk Nederlands Meteorologisch Instituut). To create a global 250 m DEM grid, we bilinearly regridded the 500 m GMTED2010 dataset and extracted the areas poleward of 56° latitude to merge with the 250 m GMTED2010 dataset. However, due to memory limitations, the 250 m DEM could not be converted to the cubed-sphere using the bin_to_cube tool in the topography workflow (e.g., at least 507 GB of memory was required). Thus, our highest feasible DEM resolution is 500 m.

Using the 500 m DEM, we successfully generated an 800 m base cubed-sphere topography by fixing a bug in E3SM's bin_to_cube tool and switching the NetCDF output format to NF_64BIT_DATA. We also conducted additional preprocessing steps (e.g., reclassifying the land–water mask), although these are not required in the current E3SM workflow. The 800 m base cubed-sphere topography was subsequently mapped onto the target 100 m RRM GLL (dynamical core) and pg2 (physics) grids using the cube_to_target tool, which currently offers the most efficient and practical approach. During this step, we encountered a bug where the generated RRM exhibited anomalously large negative topography values at specific, non-adjacent grid points.

As a temporary solution, we replaced these outliers by taking the arithmetic mean of their neighbors. Further investigation revealed negative weights in the overlapping regions between the base cubed-sphere grid and the target grid, particularly when the target grid was smaller than the base cubed-sphere grid. We identified a corner-case condition in the cube_to_target tool's line-integration logic that mischaracterized very short segments (where dx, dy ≈ 0). This condition was originally intended to handle lines nearly parallel to latitude (i.e., when dy < fuzzy_width, which defaults to 1e-8). Disabling this condition eliminates the occurrence of negative weights. Since exact integration is not essential in this context, we recommend commenting out this condition in future use.

The remainder of the topography workflow followed the E3SM Version 3 topography best practices (https://acme-climate.atlassian.net/wiki/spaces/DOC/pages/2712338924/V3+Topography+GLL+PG2+grids, last access: 25 January 2025), with the exception that we omitted the second cube_to_target call to accelerate file generation. This step is only required for the sub-grid scale gravity wave drag and turbulent mountain stress parameterizations, as it computes elevation variances across scales above and below a critical wavenumber. Since SCREAM version 0—the model used in this study—does not include either of these parameterizations, this step was not necessary. According to linear theory, the separation scale between gravity waves and drag generated by unstratified turbulent flow depends on obstacle horizontal scale (Lauritzen et al., 2015). As resolution increases to 100 m, the small-scale drag induced by turbulent flow is expected to be explicitly resolved rather than parameterized.

One of the most crucial steps in the workflow was the smoothing of the RRM topography. Using the recommended six smoothing iterations led to an immediate crash at the model's first timestep, even after significantly reducing the dynamical core timestep to one-tenth of the Courant–Friedrichs–Lewy (CFL) limit and to one-tenth of the value used in the 100 m DP-SCREAM simulations (Bogenschutz et al., 2023). These early tests were conducted on a bomb cyclone hindcast event that made landfall in California, characterized by strong synoptic winds. In contrast, a sensitivity test using a finer 100 m mesh—which did not include the steep topographic gradients present in the RRM configuration examined in this study—ran successfully with the six smoothing iterations. Thus, we suspect that steep topography is the culprit for numerical instability. Eventually, we doubled the smoothing factor to 1—matching the V2 topo guidelines—and successfully ran the bomb cyclone case at a 0.025 s dynamical core timestep. Further doubling the smoothing factor to 24 allowed the simulation to run with a timestep of 0.05 s. Notably, in a sensitivity experiment we conducted, switching the dynamical core to hydrostatic mode allowed the model to run with a topography smoothing factor of 6 and a time step of 0.25 s.

With regards to the non-hydrostatic dynamical core, fixing the negative weighting problem mentioned earlier slightly delayed instability onset, enabling simulations to reach model step 8 before crashing when using a 0.05 s timestep. Similar challenges with numerical stability due to increased resolution over complex topography have been documented in the LES community (e.g., Connolly et al., 2021). Steeper topographic gradients at finer resolutions necessitate shorter time steps than what guidance from the CFL criterion might suggest. Here, we needed a tenfold reduction in timestep, implying additional adjustments may be required for SCREAM's non-hydrostatic dynamical core at 100 m (e.g., a scale-aware sponge layer). Alternatively, regridding a coarser DEM onto a finer grid might introduce blocky mountains or steep gradients at the boundaries around DEM elements that exacerbate instabilities.

Figure 2 compares the topography of the 3.25 km California RRM and the 100 m Bay Area RRM.

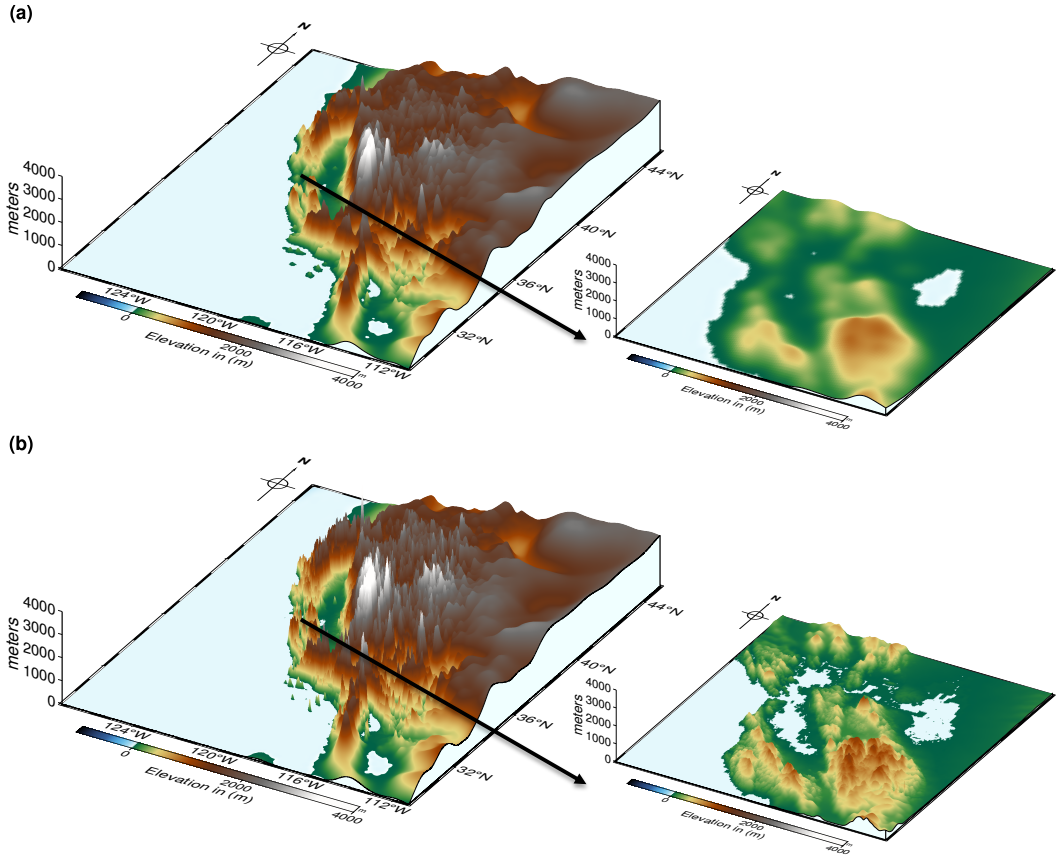

Figure 2. The topography of the (a) 3.25 km California RRM and (b) 100 m Bay Area RRM. Each is shown with both a full California overview and a bird's-eye view of the Bay Area.

2.4 Dynamical core configuration

The final namelist for the dynamical core configuration is as follows:

- Topography smoothing: 12 iterations
- Turn on topography improvement options: pgrad_correction=1, hv_ref_profiles=6
- Dycore timestep: se_tstep = 0.025 s, which is 1/10th of what we would expect from the CFL limit and 1/10th of what was used in in Bogenschutz et al. (2023) for 100 m DP-SCREAM simulations
- Time-stepping method: tstep_type = 5 (fully 3rd order accurate KGU53 explicit method)
- Hyperviscosity coefficient: nu = 3.4e-7 1/s (10 times the default value)
- Top-of-model sponge layer coefficient: nu_top = 333.34

- Tracer advection timestep: dt_tracer_factor = 8
- Lagrangian vertical level remap timestep: dt_remap_factor = 2
- Hyperviscosity subcycles: hypervis_subcycle = hypervis_subcycle_q = hypervis_subcycle_tom = 1
- Hyperviscosity scaling: hypervis_scaling = 3.0

Among these, topography smoothing and the dynamical core timestep are most critical for model stability.

2.5 Throughput

Because of the extremely small dynamical core time step (0.025 seconds) needed to ensure numerical stability, and limitations on the total number of nodes per job submission, each simulation achieves a throughput of approximately one simulated hour per wall-clock hour. On Livermore Computing's Ruby cluster, each submission utilizes 180 Intel Xeon CLX-8276L (2.2 GHz) CPUs, each with 28 cores and 192 GiB of memory, resulting in a throughput of approximately 0.16 simulated days per day (SDPD). Accounting for both runtime and queue delays, a single hindcast typically requires about one month to complete.

While preparing this manuscript, performance scaling tests were conducted for SCREAM version 1 (C++ implementation) on GPU nodes of NERSC's Perlmutter system. Each node contains four NVIDIA A100 GPUs (Ampere architecture, 40 GiB each) and four HPE Slingshot 11 NICs. We performed one-hour SCREAM-RRM simulations at 100 m resolution, scaling from 32 to 1024 nodes. The performance results are shown in Fig. 3. These findings underscore the critical importance of GPU acceleration for enabling ultra-high-resolution simulations. While such runs are nearly infeasible on CPU-only systems—requiring up to a month to complete—they become substantially more tractable on GPU platforms, with total runtimes reduced to less than two wall-clock days. We note that we used SCREAMv0 in this work because the full RRM capabilities for SCREAMv1 were not yet scientifically validated at the time this study was performed.

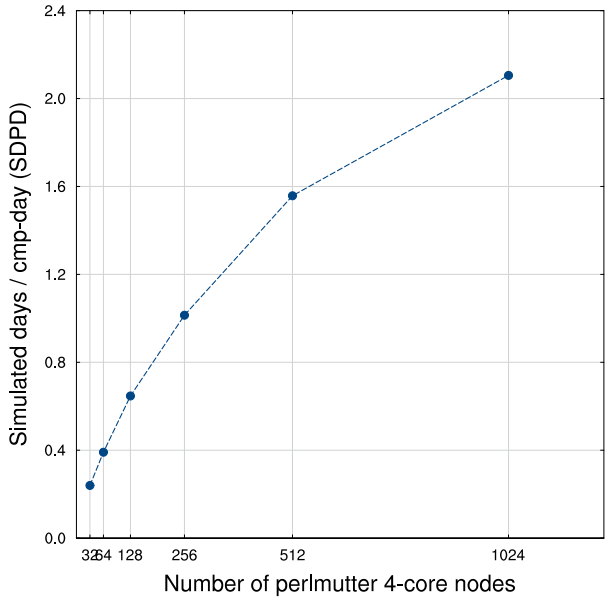

Figure 3. Performance scaling of the 100 m Bay Area SCREAMv1-RRM using the GPU nodes of the NERSC Perlmutter machine.

3 Experiment design

California exhibits a diverse range of microclimates, shaped by distinct seasonal weather patterns. During winter, the region experiences hydrological events primarily driven by atmospheric rivers and midlatitude cyclones, characterized by strong synoptically forced winds. In contrast, summer conditions are dominated by marine stratocumulus clouds, influenced by large-scale subsidence associated with persistent high-pressure ridges, and are typically governed by weak synoptic forcing and locally driven circulations. In this study, we selected two representative 2-day cases corresponding to these typical weather conditions to evaluate the performance of the 100 m SCREAM-RRM under strong and weak synoptic flows.

3.1 Case description

3.1.1 Storm2008

The first case is initialized at 00Z on January 4, 2008, and runs for 48 hours, capturing “Storm 3" from the January 2008 North American storm complex. This storm underwent explosive intensification, producing wind gusts exceeding 44 $\mathrm{m\,s^{-1}}$, total precipitation amounts up to 250 mm, and snowfall accumulations reaching 1800 mm in parts of California. We refer to this case as Storm2008. Meteorological observations from National Oceanic and Atmospheric Administration (NOAA) (https://tidesandcurrents.noaa.gov/met.html, last access: 25 January 2025) indicate average gusts near 20 m/s and mean wind speeds around 15 m/s at Redwood City station (station 9414523), with a minimum sea-level pressure below 1000 hPa. Storm 3 was noteworthy for reaching 956 hPa sea-level pressure, making it the strongest storm ever recorded on the West Coast at that

time. Figure 4 illustrates the atmospheric river impacting California at 2008-01-04 18Z, as simulated by the 3.25 km (left) and 100 m (right) SCREAM-RRMs, along with the associated precipitation patterns.

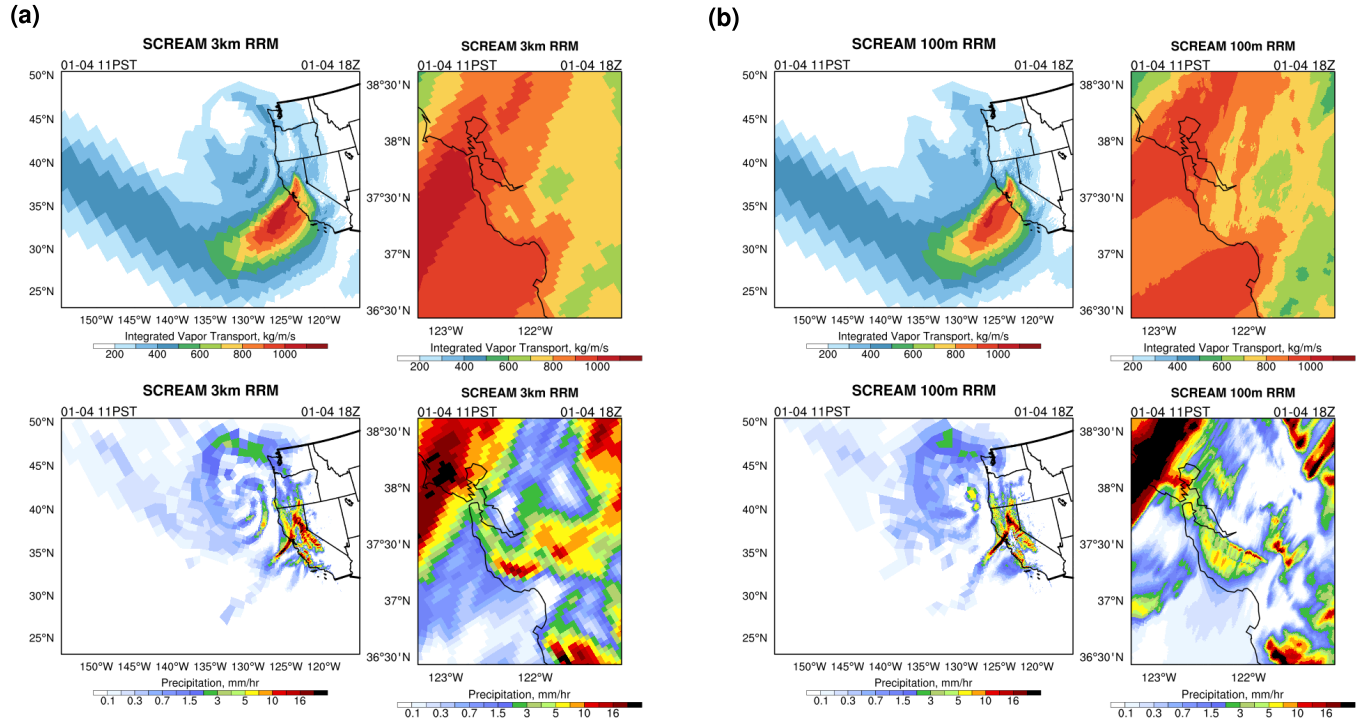

Figure 4. (a) Integrated vapor transport (top) and precipitation (bottom) in the East Pacific (left) and a zoomed-in view of the Bay Area (right) in the 3.25 km California SCREAM-RRM at 2008-01-04 18Z. (b) Same as (a), but for the 100 m Bay Area SCREAM-RRM.

3.1.2 Stratocumulus2023

The second case, which we refer to as Stratocumulus2023, covers a weak synoptic wind condition dominated by marine stratocumulus. This case was initialized at 00Z on 10 July 2023 and simulated for 48 hours. It is representative of typical summer conditions, featuring marine stratocumulus clouds that form inland during the night and gradually retreat toward the coast by midday. The Central Valley experiences arid, Mediterranean-like summer conditions, which generally inhibit marine stratocumulus from penetrating far inland. However, these clouds often reach the interior valleys of the San Francisco Bay Area (e.g., Livermore). Locals often note the contrast: overcast, dewy, and cool conditions in the early morning, which recede by noon to give way to hot, clear skies. We selected this case based on the historical temperature records in Livermore and GOES-W 0.6 μm visible imagery for North America (https://satcorps.larc.nasa.gov/, last access: 25 January 2025) from 2020 to 2023. Concurrent NOAA observations in Redwood City recorded average wind speeds of 5–10 $\mathrm{m\,s^{-1}}$, with gusts of comparable magnitude. The winds exhibited a pronounced diurnal cycle. Under weak synoptic forcing, topography and boundary-layer turbulence strongly modulate the flow. Figure 5 shows liquid water path and cloud fraction of the coastal stratocumulus near

San Francisco at 2023-07-11 18Z simulated by both versions of SCREAM-RRM. Since there is no data for this day from North America GOES-W, we compare to the GOES-East/West Merged CONUS LWP-Best Estimate.

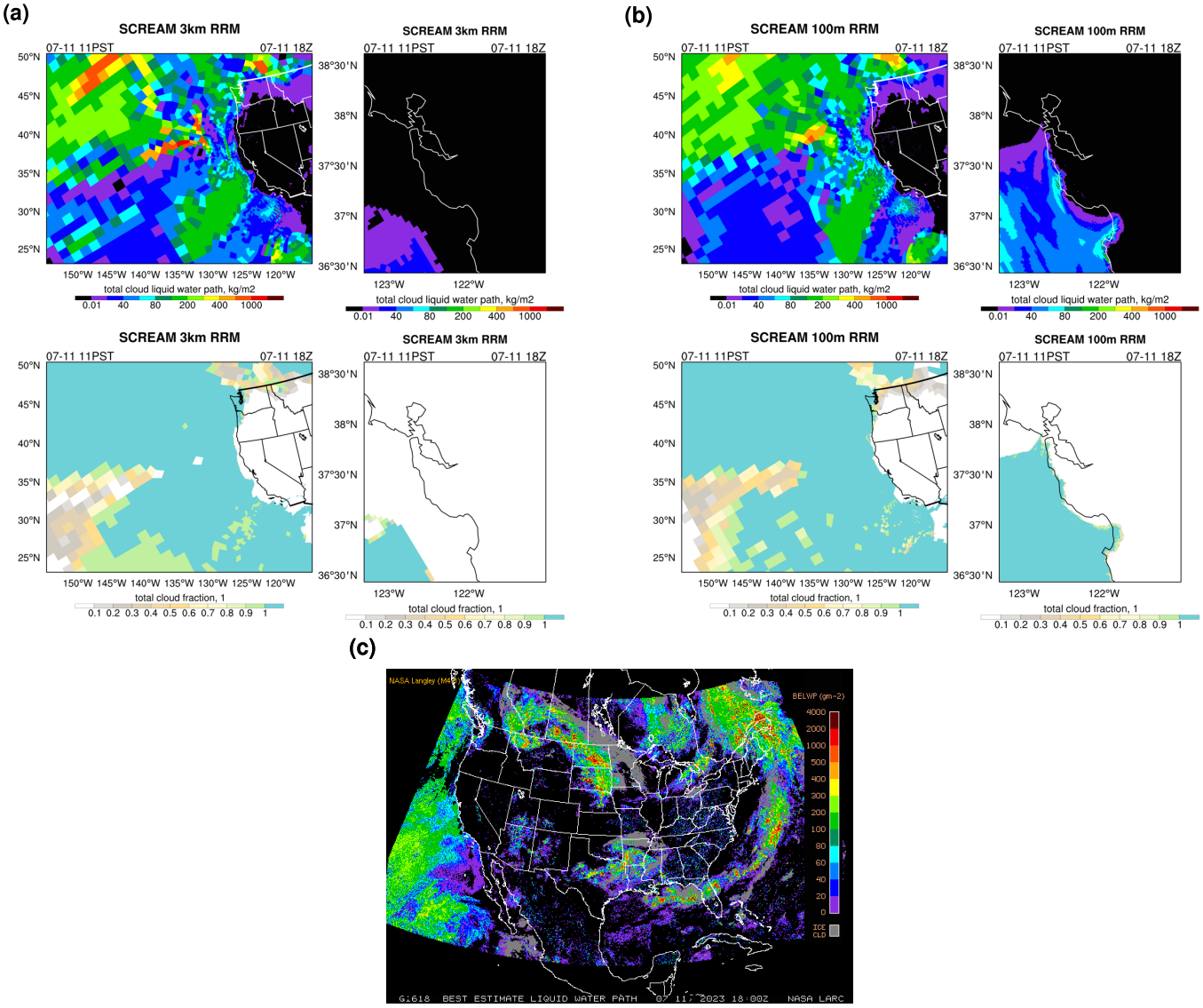

Figure 5. (a) and (b) Similar to Fig. 4, but for total cloud liquid water path (top) and total cloud fraction (bottom). (c) GOES-East/West Merged CONUS LWP-Best Estimate for North America.

3.2 Initialization

285 The atmosphere initial conditions were derived from the Fifth Generation European Centre for Medium-Range Weather Forecasts Reanalysis (ERA5) (Hersbach et al., 2020) High Resolution (HRES) global reanalysis data, which has a native horizontal

resolution of 31 km (0.28125°). The variables are stored with T639 triangular truncation (spectral resolution) or a reduced Gaussian grid with a resolution of N320. Vertically, it uses a 137-layer hybrid pressure/sigma coordinate that affords ~20 m resolution near the surface and ~130 m at 850 hPa, closely matching SCREAM's default 128-layer vertical coordinate in the lower and middle troposphere. We chose this native-grid ERA5 dataset over the more commonly used pressure-level lat-lon dataset to better align with SCREAM's vertical grid, especially given the critical role that fine vertical resolution plays in stratocumulus simulations (e.g., Lee et al., 2022; Bogenschutz et al., 2023).

Since we were using ERA5 HRES native-grid data, the CAPT package (https://github.com/PCMDI/CAPT, last access: 25 January 2025) was used to generate the atmosphere initial conditions, including horizontal interpolation onto SCREAM's GLL grid, adjustment of surface pressure, and vertical interpolation onto the 128-layer hybrid coordinate. Alternatively, the HICCUP package (https://github.com/E3SM-Project/HICCUP, last access: 25 January 2025), which is frequently used in the E3SM/SCREAM community (e.g., Zhang et al., 2024a; Bogenschutz et al., 2024; Zhang et al., 2024b), would be convenient if using ERA5 pressure-level data. CAPT requires Earth System Modeling Framework (ESMF) software (https://earthsystemmodeling.org/regrid/, last access: 25 January 2025) for horizontal interpolation, and the target GLL description file used for ESMF can be obtained from the E3SM's homme tool. HICCUP performs horizontal interpolation by calling netCDF Operator (NCO) (Zender, 2008), which has several built-in horizontal mapping software including NCO, TempestRemap (Ullrich and Taylor, 2015), ESMF, and Climate Data Operator (CDO) (https://code.mpimet.mpg.de/projects/cdo, last access: 25 January 2025).

CAPT's vertical interpolation follows European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) Documentation CY23R4 (https://www.ecmwf.int/sites/default/files/elibrary/2003/77032-ifs-documentation-cy23r4-part-vi-technical-and-computational-procedures_1.pdf, last access: 25 January 2025), applying linear/quadratic/linear+quadratic interpolation for temperature/moist variables/wind, respectively. HICCUP calls NCO, using the linear to log(pressure) vertical interpolation by default. Both CAPT and HICCUP apply surface adjustment based on the Trenberth et al. (1993) procedure, which calculates the model surface pressure based on the difference in surface geopotential height (PHIS) between ERA5 and the model topography and assume dry hydrostatic lapse rate. After subtly modifying HICCUP's surface adjustment code—specifically in the choice of PHIS used in the calculation—we verified that the resulting initial conditions successfully passed a one-hour test under the previously described dynamical core configuration. For comparison, the simulation initialized with native-grid data using CAPT showed less pronounced surface adjustments and weaker horizontal grid imprinting than the one initialized with pressure-level data using HICCUP. These differences reflect the inherent distinctions between using pressure-level versus native-grid ERA5 data; however, they do not affect the model's numerical stability.

A physically consistent land initial condition for a hindcast typically requires a multi-year spin-up driven by reanalysis atmospheric forcing. Following the workflow used in betacast (Zarzycki et al., 2014), we conducted a 5-year land-only (i.e., I-compset) simulation using ERA5 atmospheric forcing to generate the land initial condition. This I-compset simulation did not use a cold start; instead, it was initialized from an interpolated land restart file derived from a well-spun 1° E3SMv1 historical simulation (1850–2015), using the state at 2015-01-01 00 UTC. The interpolation process required two existing land restart files on the 3.25 km and 800 m CARRM grids, obtained from prior one-hour test simulations initialized with a cold-start land

state. The ERA5 forcing data, provided at 0.25° resolution and 6-hour intervals, include precipitation, downward shortwave and longwave radiation, 2-meter temperature and humidity, 10-meter wind, and surface pressure. Following the Betacast workflow, these data were converted into the data-atmosphere stream file format used by E3SM I-compsets.

3.3 SST and sea ice extent

SST and sea ice extent were obtained from NOAA Optimal Interpolation (OI) data (Reynolds et al., 2007). The OISST data were preprocessed by filling missing land values using a relaxed Poisson equation, while preserving their native 0.25° resolution. The data were then reformatted into the data-ocean stream file format (https://esmci.github.io/cime/versions/ufs_release_v1.1/html/data_models/data-ocean.html, last accessed: 25 January 2025). During runtime, SCREAM interpolates the SST and sea ice fields to the model's ocean grid.

3.4 Observation data

Within the 100 m refinement region, we use the following in-situ observations for validation:

- World Meteorological Organization (WMO) Station Data: 6 h meteorological data from Meteomanz (http://www.meteomanz.com/, last access: 25 January 2025), including near-surface temperature, wind, and relative humidity.
- Integrated Surface Dataset (ISD) from the National Centers for Environmental Information (NCEI) website (https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database, last access: 25 January 2025), which offers 30 min, 1 h, 6 h, or longer intervals. Variables include near-surface temperature, wind, and humidity.
- NOAA Tides and Currents Meteorological Observations from NOAA Center for Operational Oceanographic Products and Services stations (https://tidesandcurrents.noaa.gov/, last access: 25 January 2025), providing 6 min surface pressure, near-surface wind and temperature.
- Integrated Global Radiosonde Archive (IGRA) from NCEI (https://www.ncei.noaa.gov/products/weather-balloon/integrated-global-radiosonde-archive, last access: 25 January 2025), at irregular temporal resolution, providing vertical profiles of pressure, temperature, humidity, and wind. This is the only publicly available source of vertical profiles for our analysis.

To our knowledge, no publicly available gridded, observation-driven meteorological products for California offer spatial resolutions finer than 3.25 km. For instance, both the Parameter-elevation Regressions on Independent Slopes Model (PRISM) and the High-Resolution Rapid Refresh (HRRR)—which may be considered a form of reanalysis—have resolutions of 4 km. These datasets are inappropriate for evaluating numerical models at 100 m (cf. Lundquist et al., 2019) and are thus excluded from this study.

For the Stratocumulus2023 case, due to the lack of coastal soundings for cloud hydrometeors, we compare SCREAM-RRM to GOES-East/West LWP-Best Estimate imagery provided by the Satellite Cloud and Radiation Property Retrieval System (SatCORPS) https://satcorps.larc.nasa.gov/, last access: 25 January 2025), as shown in Fig. 5.

Most temperature/humidity sensors are mounted at a height of 2 m, while anemometers are deployed near 10 m. Hence, we evaluate 2 m temperature, 2 m relative humidity, 10 m wind speed, and surface pressure. Each site is matched to the nearest model gridpoint using sklearn.neighbors.BallTree (https://scikit-learn.org/1.6/, last access: 25 January 2025) within the 100 m bounding box (W: 237.07, E: 238.95, S: 36.9, N: 38.5; see Fig. 1). Note that due to the oblique orientation of the 100 m refinement area, the grid resolution at the box boundaries can locally approach 800 m, e.g., Stockton in Meteomanz/ISD. Figure 6 shows the station distributions in each dataset. Although Meteomanz and ISD datasets include precipitation, their sparse spatial coverage and coarse temporal resolution limit their usefulness for meaningful precipitation comparisons. For example, no precipitation rates exceeding 5mm/hr were recorded during the Storm2008 case. The Tides and Currents data were downloaded using noaa_coops package (https://github.com/GClunies/noaa_coops, last access: 25 January 2025), while other datasets were retrieved directly from their websites.

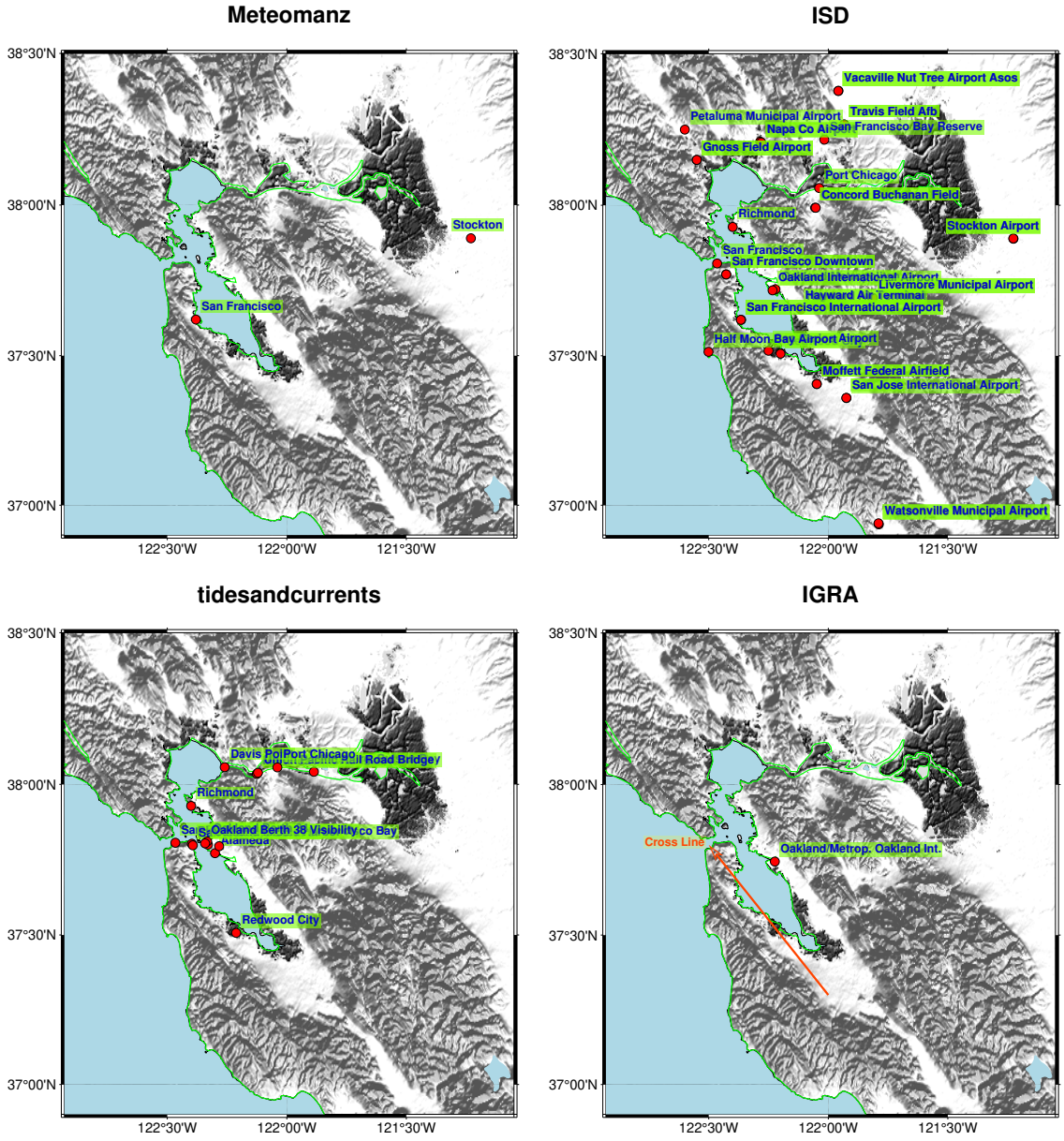

Figure 6. The spatial distribution of Meteomanz, ISD, Tides and Currents, and IGRA stations within the Bay Area, in that order.

The first three datasets (Meteomanz, ISD, and Tides and Currents) were used to calculate skill scores, including the Pearson pattern correlation coefficient, root mean squared error (RMSE), and bias. Meteomanz are available every 6 h, and Tides and Currents provide data every 6 min; the model's 5 min instantaneous outputs are aggregated to align with these intervals. ISD lacks consistent temporal resolution and is reindexed to 30 min by filling in the missing value in the metric calculations. For comparisons with IGRA, model time slices were aligned with the elapsed time of the radiosonde launches. The IGRA

soundings were vertically interpolated to SCREAM's 128-layer pressure coordinate, defined by reference midpoint pressures ranging from 998.5 hPa near the surface to 2.6 hPa at the model top. Soundings plotted were visualized using the MetPy package (https://github.com/Unidata/MetPy, last access: 25 January 2025).

4 Results

We begin the results section with a visual overview of the SCREAM 100 m Bay Area RRM (referred to as "BA-100m") simulated vertical motion and cross sections. These high-resolution snapshots highlight the 100 m RRM's ability to resolve fine-scale processes such as topography-induced lifting, boundary-layer structure, and wave activity. A particularly notable feature is the presence of high-frequency wind speed oscillations, which closely resemble those seen in observations.

We then conduct a more detailed evaluation of the performance of BA-100m and the SCREAM 3.25 km CA RRM (referred to as "CA-3km") under two contrasting weather regimes: the synoptically forced Storm2008 case and the locally driven Stratocumulus2023 case. Specifically, we calculated the metrics of these simulations to in-situ observations of near-surface temperature, wind speed, relative humidity, and surface pressure. The spatial distribution of these variables is generally smooth across a few hundred meters, so direct comparison with station data is feasible for the 100 m and 3.25 km RRMs. In addition to the overall skill scores, we particularly emphasize SCREAM's ability to capture the temporal evolution of meteorological factors at each site, especially by utilizing the 6 min high-frequency observation data from Tides and Currents. Furthermore, we compare the sounding data at Oakland International Airport station, focusing on the differences in temperature and humidity in the boundary layer.

4.1 An initial look at the 100 m SCREAM-RRM runs

Videos A1 and A2 show the evolution of vertical velocity and precipitation at upper (332 hPa), lower-middle (808 hPa), and boundary layer (951 hPa) levels in the BA-100m and CA-3km simulations. Each frame represents a 10-minute interval, capturing the temporal progression of atmospheric features at high resolution.

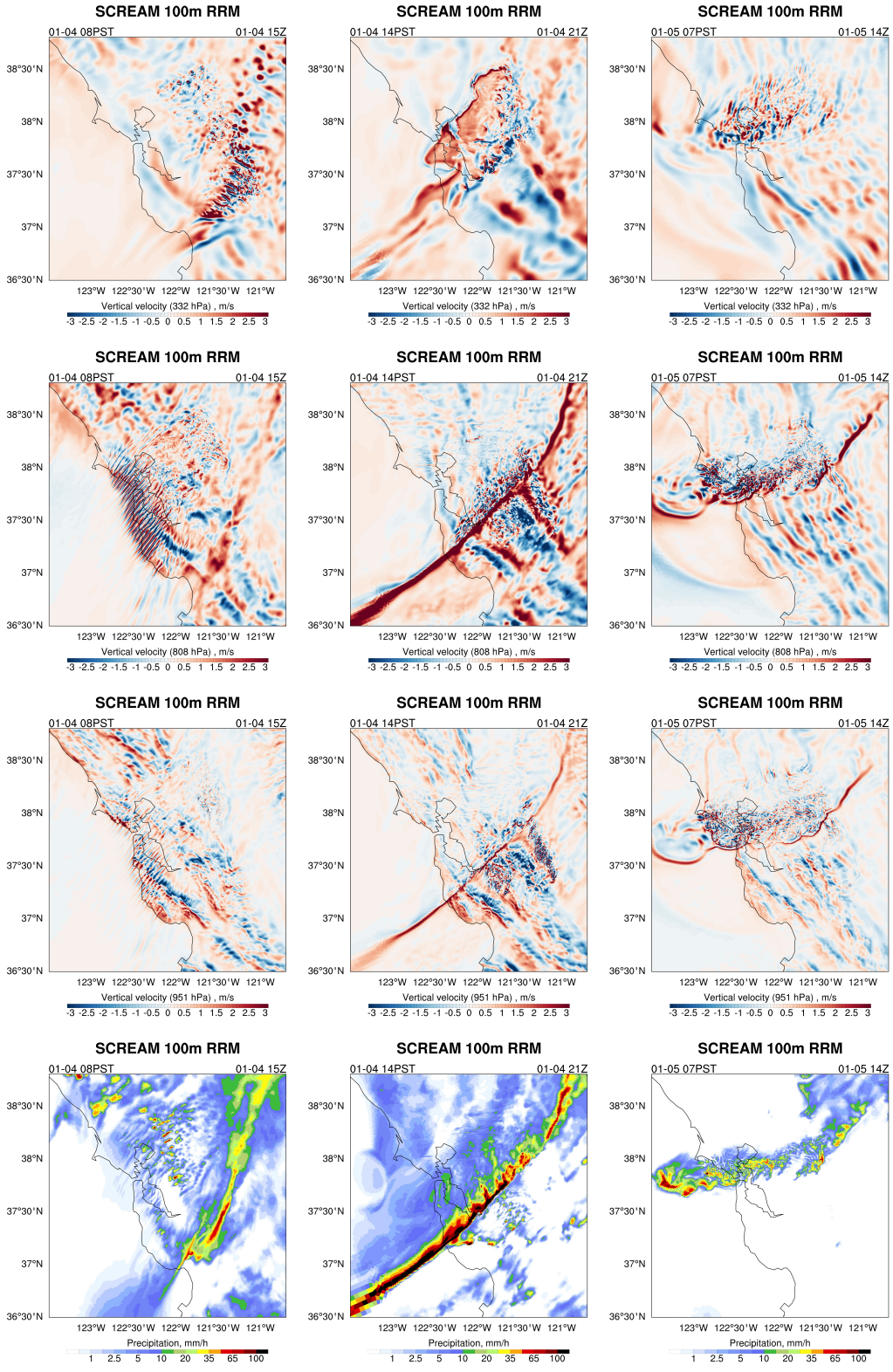

Figure 7. Vertical velocity at 332 hPa (top row), 808 hPa (second row), and 951 hPa (third row), along with the total precipitation rate at the surface (bottom row), at 2008-01-04 08PST (left), 2008-01-04 14PST (middle), and 2008-01-05 07PST (right), in the 100 m Bay Area SCREAM-RRM simulation.

BA-100m shows remarkable detail (See Video A1 in Video Supplement, Fig. 7). One question is how much of this rich detail is realistic? Accurate observations of vertical velocity, gravity waves, and precipitation are notoriously difficult, making it hard to directly assess the objective full picture of the "true value". However, these images serve as a reminder of several fundamental characteristics of fluid dynamics. In the Storm2008 case, several distinct features emerge:

- A persistent background mountain-scale standing wave pattern, with a wavelength that increases with altitude.
- Strong updrafts ahead of the cold front, as high-speed winds interact with coastal topography, producing significant turbulence and wave instabilities (both convective and shear-driven) downstream.
- A mountain-gap wind generated by the Venturi effect, with an elongated updraft extending from the mountain pass.
- A vortex originating over the ocean that amplifies, deforms, and undergoes nonlinear fragmentation after colliding with the Marin Hills—intensifying turbulence. Prior to landfall, the vortex induces substantial divergence and subsidence due to evaporative cooling from precipitation, resulting in a well-defined cold pool and gust front.

While the meticulous depiction of these features and their phenomenological plausibility is encouraging, it does not necessarily validate the intensity of the perturbations. Due to the lack of direct observations, quantifying the realism of these perturbations in each local region remains difficult. However, available routine observations provide some insights (see next section). For instance, while the simulated cold pool characteristics align with the simulated temperature decrease and pressure increase over the last 12 hours, no such signal is evident in the observations, raising questions about the validity of this specific process. The encouraging aspect is that no obviously suspicious numerical artifacts were detected.

Stratocumulus2023 provides a more intuitive example, demonstrating the seamless transition of the RRM grid across different scales without spurious fluctuations at the boundaries (See Video A2 in Video Supplement, Fig. 8). The vertical motion of the boundary layer exhibits a pronounced diurnal variation. After sunrise, as the surface heats up, the boundary layer deepens, where the turbulence intensifies substantially compared to nighttime. The horizontal half-wavelength of fluctuations in the 100 m refinement region is about 1 km, corresponding roughly to SCREAM's effective resolution of about 6 Δx (600 m) (Caldwell 2021). The transition from the 100 m refinement region to the 800 m refinement region is rapid, with the wavelength increasing to about 10 km. This transition appears natural and does not introduce numerical artifacts. Similarly, the transition from the ocean to the coastal ranges due to topography leads to a wavelength shift from greater than 10 km to 1 km, also appearing physically plausible. During the first night of the simulation, stratocumulus is developing near the San Francisco coast (Video S2).

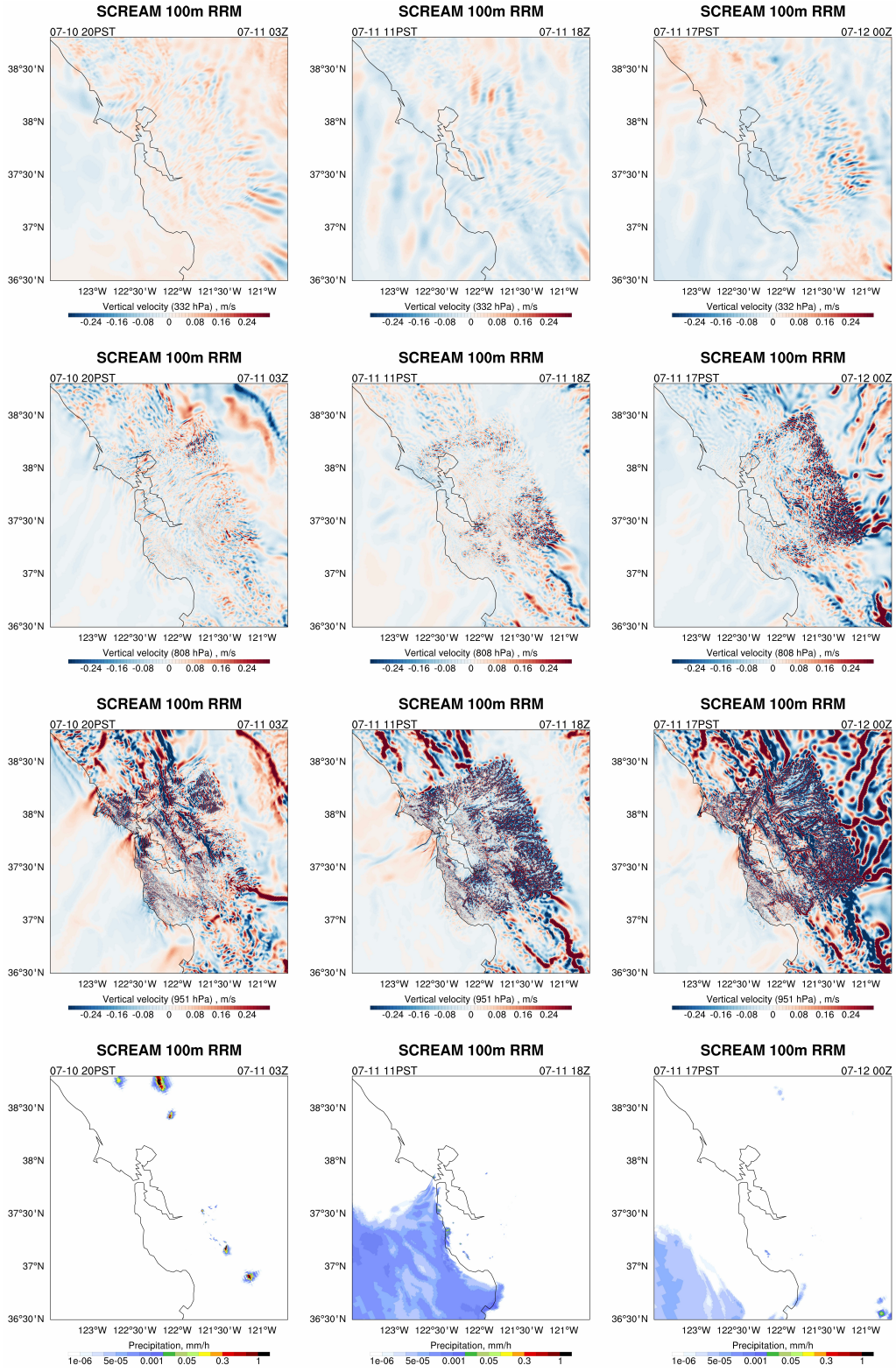

Figure 8. Same as Fig. 7 but for the Stratocumulus2023 event at 2023-07-10 20PST (left), 2023-07-11 11PST (middle), and 2023-07-11 17PST (right).

Cross sections provide additional insight into the small-scale vertical structures captured by the 100 m simulation. In Storm2008, strong updraft zones emerge across the coastal ranges and extend deep into the troposphere (Fig. 9a). While cloud formation generally aligns with areas of upward motion, the strongest clouds do not always coincide with the most intense updrafts—likely reflecting the influence of horizontal advection of condensate and drag effects from falling precipitation. Additionally, a pair of oppositely rotating horizontal vortices is evident—one at mid-levels downstream of the frontal zone and another farther downstream in the lower troposphere—suggesting the presence of both convective and shear-driven instabilities (e.g., Sun et al., 2015). In contrast, Stratocumulus2023 is characterized by weaker topographic forcing and more thermally driven turbulence (Fig. 9b). The short horizontal wavelengths of boundary-layer fluctuations remain intact to mid-levels, with limited vertical propagation due to weak ambient wind. These results demonstrate that SCREAM-BA-100m physically represents complex vertical transport and wave dynamics across different flow regimes.

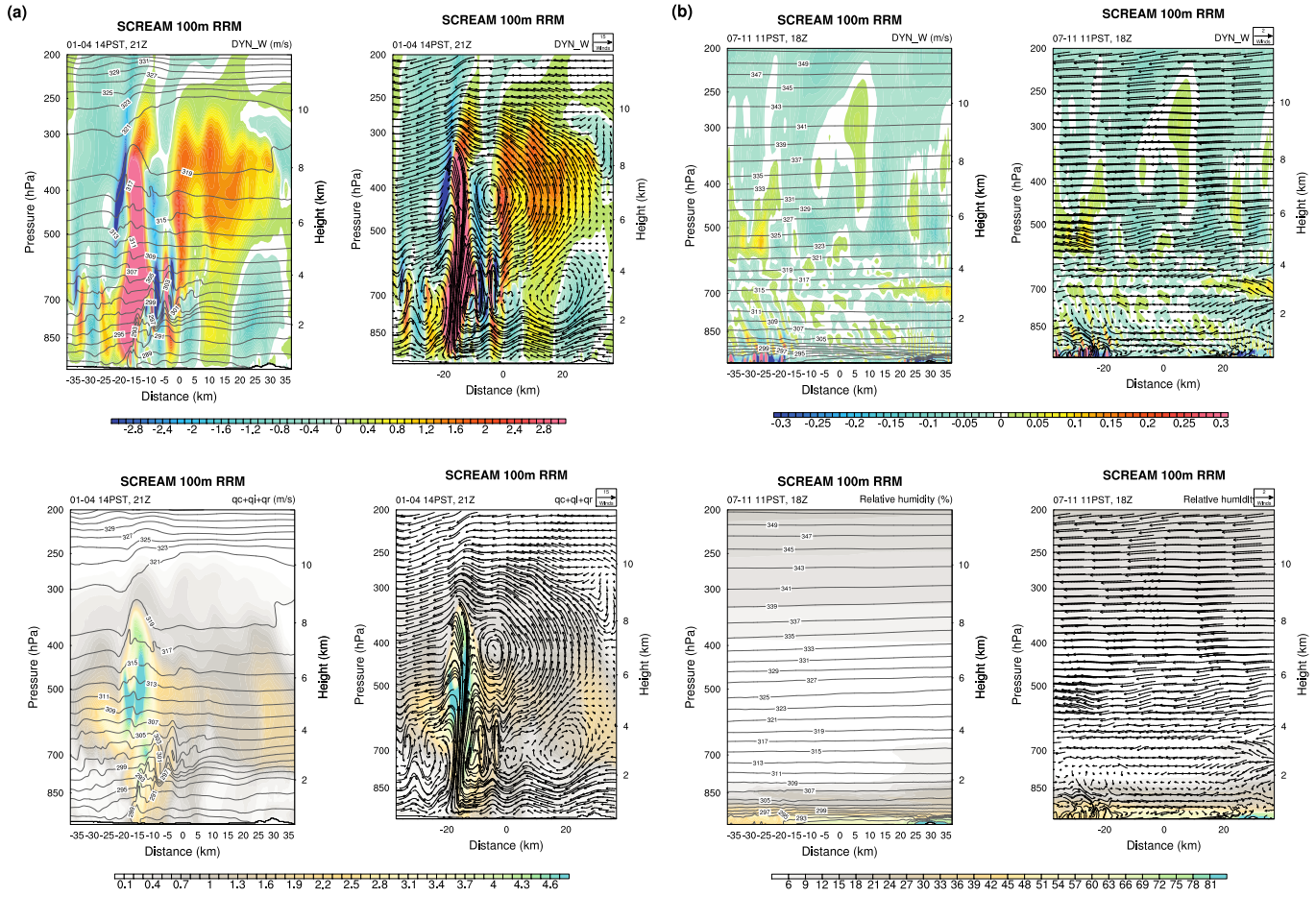

Figure 9. Cross sections aligned parallel to the Santa Cruz Mountains extend from the southeast (238E, 37.3N) to the northwest (237.5E, 37.8N), spanning 74 km along the x-axis. See Fig. 6 for the cross section line (shown in orangered) on the IGRA station map. The top row shows vertical velocity (shading), overlaid with potential temperature (contours, left) or wind vectors (right) at 2008-01-04 14PST. Horizontal winds are adjusted to be parallel to the cross-section, and vertical velocity is amplified by a factor of 10. The bottom row is the same as the top row, except that the shading shows total cloud hydrometeors (liquid + ice + rain).

4.2 In-situ evaluation

4.2.1 Storm2008

Figure 10 presents the overall skill scores—correlation coefficients, root mean square errors (RMSE), and biases—for near-surface temperature, wind speed, relative humidity, and surface pressure from the SCREAM-RRM simulations of the Storm2008 event. These metrics are evaluated against observations from Meteomanz, ISD, and Tides and Currents stations, all of which are located within the 100 m rectangular domain shown in Fig. 6. Among these sources, Meteomanz provides data from only two stations, whereas Tides and Currents offers the most extensive coverage. The RMSE and bias for all variables show sub-

stantial improvement in the BA-100m simulation compared to CA-3km. In terms of correlation, temperature exhibits a slight decrease in BA-100m, while relative humidity shows an increase. Wind speed correlations decrease relative to both ISD and Tides and Currents observations, and surface pressure shows a modest reduction in correlation when evaluated against Tides and Currents data.

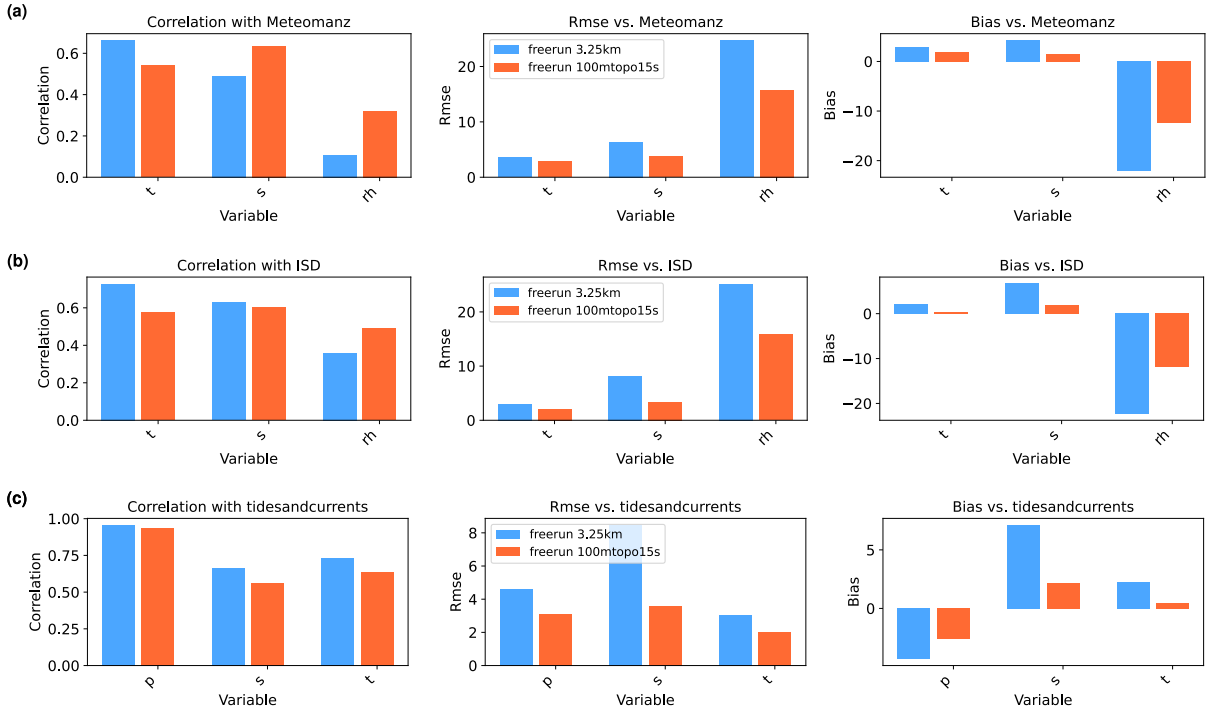

Figure 10. The skill scores for near-surface temperature (t), wind speed (s), relative humidity (rh), and surface pressure (p), compared to (a) Meteomanz, (b) ISD, and (c) Tides and Currents observations, are presented as overall metrics: Pearson pattern correlation coefficient (left), root mean squared error (middle), and bias (right). The blue and orangered bars indicate simulation results from the 3.25 km California SCREAM-RRM and the 100 m Bay Area SCREAM-RRM, respectively.

A more intuitive understanding can be gained from the detailed temporal evolution of each station (Fig. 11, Fig. 12). Observations show a notable drop in air pressure, wind speed, and temperature after the cold front passage. Temperature observations at multiple stations exhibit a roughly four-hour period of rising followed by falling values. The previous evolutions shown in Video A1 in the Video Supplement suggest this behavior may be associated with weak pre-frontal convergence in the region. Compared to the observations, CA-3km exhibits a warm bias, a high wind speed bias, a dry bias, and a lower air pressure bias. BA-100m compares better with the observations for all variables, reducing all biases. Note that Stockton station is located just outside the 100 m domain, so a comparatively modest improvement there is expected.

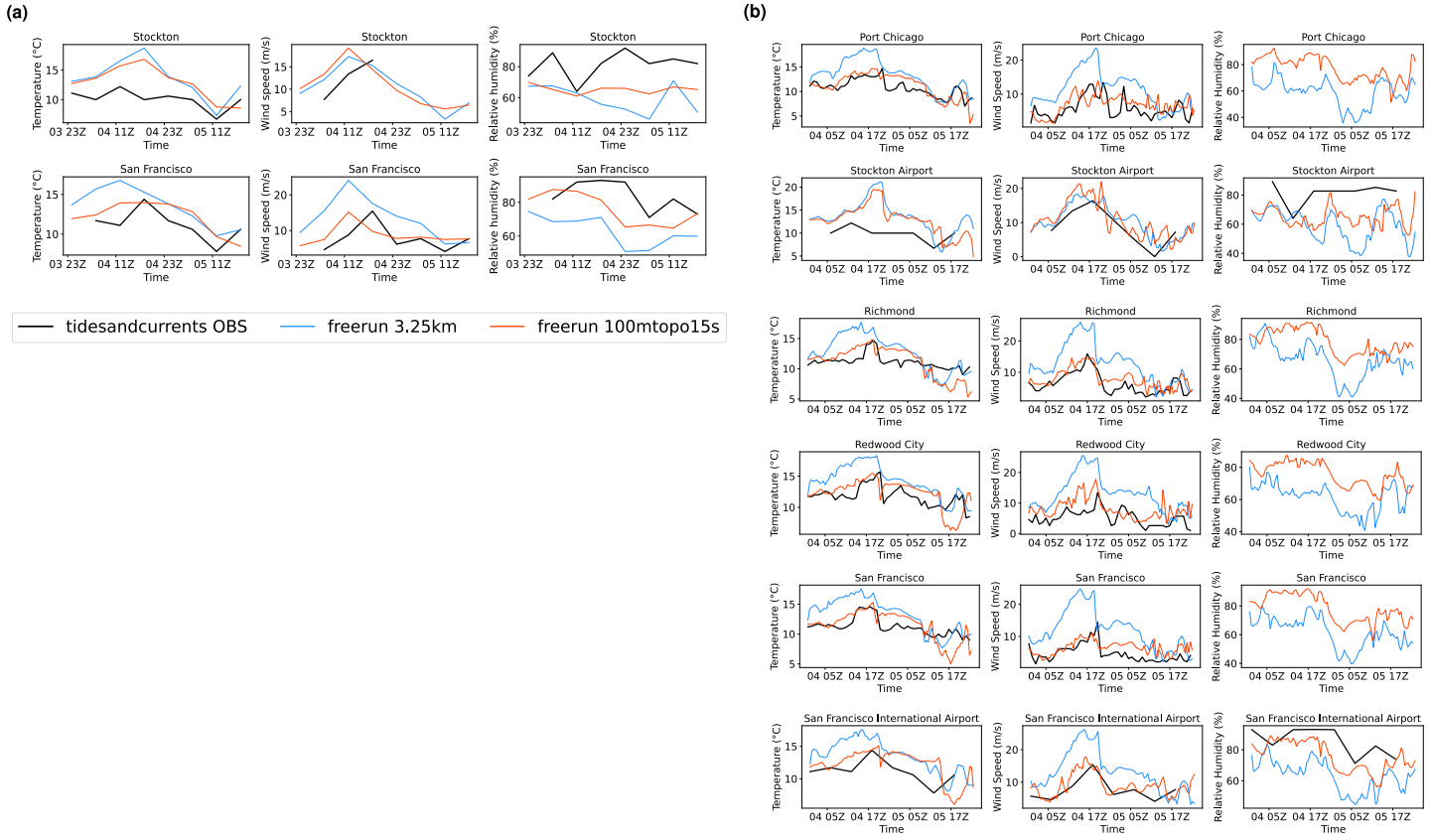

Figure 11. (a) Time series at each station for the Storm2008 event, with black, blue, and orangered lines representing Meteomanz observations, the 3.25 km California SCREAM-RRM simulation, and the 100 m Bay Area SCREAM-RRM simulation, respectively. (b) Same as (a), but for ISD observations.

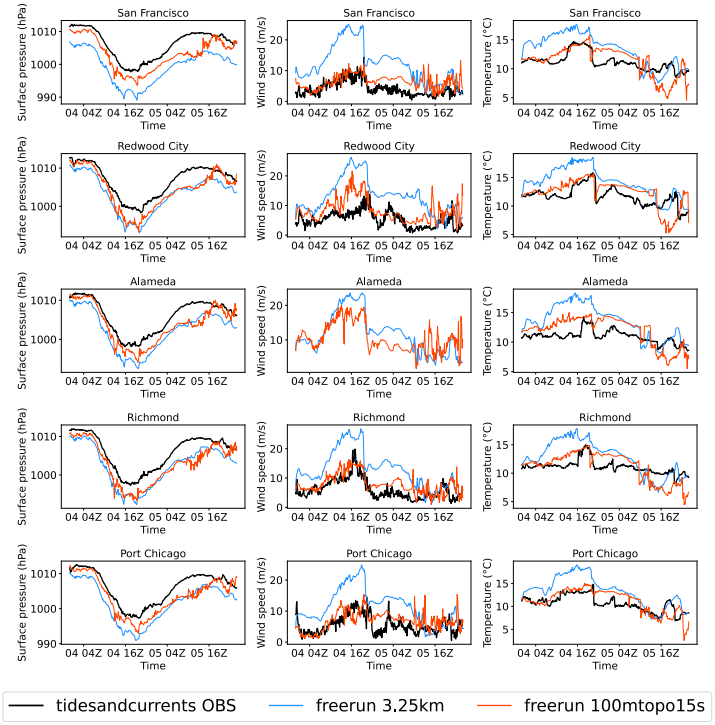

Figure 12. Same as Fig. 11 but for tidesandcurrents observations.

Despite the overall improvements, BA-100m still exhibits biases in simulated surface pressure, relative humidity, and temperature on the second day. Notably, it fails to capture the observed four-hour temperature rise preceding the cold front. In the latter half of the day, the simulated temperature phase diverges markedly from observations, likely reflecting the reduced predictability of this bomb cyclone under free-running simulation conditions.

The improvement in wind speed simulation with BA-100m is particularly noteworthy. Aside from a spurious peak occurring after the 36th simulation hour—absent in observations—the enhanced performance persists throughout most of the period. Furthermore, the model captures high-frequency wind speed oscillations that align well with Tides and Currents observations, suggesting a realistic representation of small-scale gravity wave activity. However, BA-100m overestimates wind speeds at Redwood City during the first day's daytime hours and underestimates the post-frontal wind speed reduction, particularly at the Port Chicago station.

A comparison with sounding observations from Oakland International Airport at the model initialization time reveals a notable bias introduced by the ERA5 initial conditions (Fig. 13). While upper-level comparisons are less reliable—due to increasing temporal mismatch as the pilot balloon ascends and potential trajectory shifts under strong synoptic winds—substantial deviations are evident below 700 hPa. Specifically, the dew point is up to 3 K lower below 900 hPa and up to 3 K higher above that level, while temperatures in the boundary layer are consistently 1–2 K higher than observed.

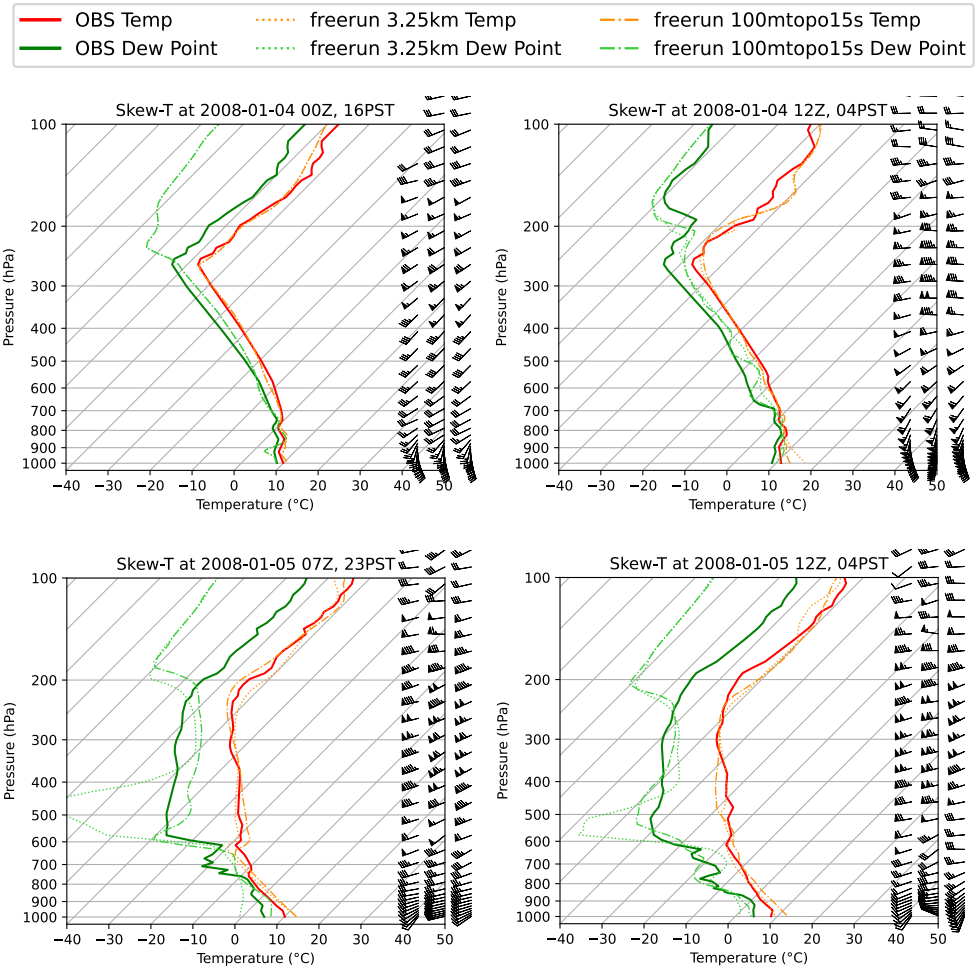

Figure 13. Soundings at Oakland International Airport for the Storm2008 event. Red and green lines represent air temperature and dew point from IGRA observations (thick solid), the 3.25 km California RRM (semi-transparent dotted), and the 100 m Bay Area RRM (semi-transparent dotted-dashed), respectively.

The dew point and temperature biases in the initial conditions are consistent with those identified in the station-based evaluation, suggesting that some of the persistent warm and dry biases in SCREAM originate from the ERA5 initial conditions. In the CA-3km simulation, a clear trade-off between dry and warm biases emerges: when the dry bias is more pronounced, the warm bias tends to be reduced (e.g., 01-05 07Z and 12Z), and vice versa (e.g., 01-04 12Z). In contrast, the BA-100m simulation exhibits a substantial reduction in the dry dew point bias below 900 hPa—particularly near the surface and most notably during the second night. As the simulation progresses, differences between CA-3km and BA-100m become more pronounced at all vertical levels, with BA-100m generally aligning more closely with observations, except for a warm temperature bias at 12Z on the second day.

In addition, CA-3km shows a pronounced wind speed overestimation near 800 hPa. As the simulation progresses, wind differences between BA-100m and CA-3km increase, with BA-100m more closely matching observations throughout the lower to mid-troposphere (Fig. 13). By 12Z on the second day, CA-3km exhibits a near-surface wind direction error of almost 90°, while BA-100m remains largely aligned with the observations.

In the Storm2008 case, the BA-100m simulation significantly reduces the high wind speed bias, as well as the warm and dry biases observed in the 3.25 km simulation. These improvements may stem from enhanced resolution of turbulent kinetic energy (TKE) and more effective turbulent mixing, along with potential feedbacks from precipitation processes—since BA-100m can resolve much shorter wavelengths and more accurately align precipitation with local orographic features. The improved wind speed simulation may also contribute to reductions in temperature and humidity biases, as weaker near-surface winds allow for the localized accumulation of cooler, moister air masses associated with evaporative cooling from precipitation. The remaining warm and dry biases are likely influenced, at least in part, by inaccuracies in the initial conditions—a plausible outcome given the 31 km native resolution of the ERA5 reanalysis. Furthermore, San Francisco Bay and San Pablo Bay are treated as ocean grid points in the model, with surface temperatures prescribed by 1° SST data. This coarse representation fails to capture the actual temperature distribution near adjacent stations and represents an additional source of bias beyond the initial conditions.

4.2.2 Stratocumulus2023

Figure 14 presents the overall skill scores for near-surface temperature, wind speed, relative humidity, and surface pressure for the Stratocumulus2023 event. Under the locally driven flow conditions of Stratocumulus2023, the CA-3km simulation exhibits substantially lower RMSE and bias in wind speed and surface pressure compared to the Storm2008 case. The BA-100m simulation demonstrates an overall reduction in RMSE and bias for all variables and all datasets. Correlation coefficients also improve for most variables in BA-100m, with the exception of wind speed, which exhibits a reduced correlation with Meteomanz and Tides and Currents.

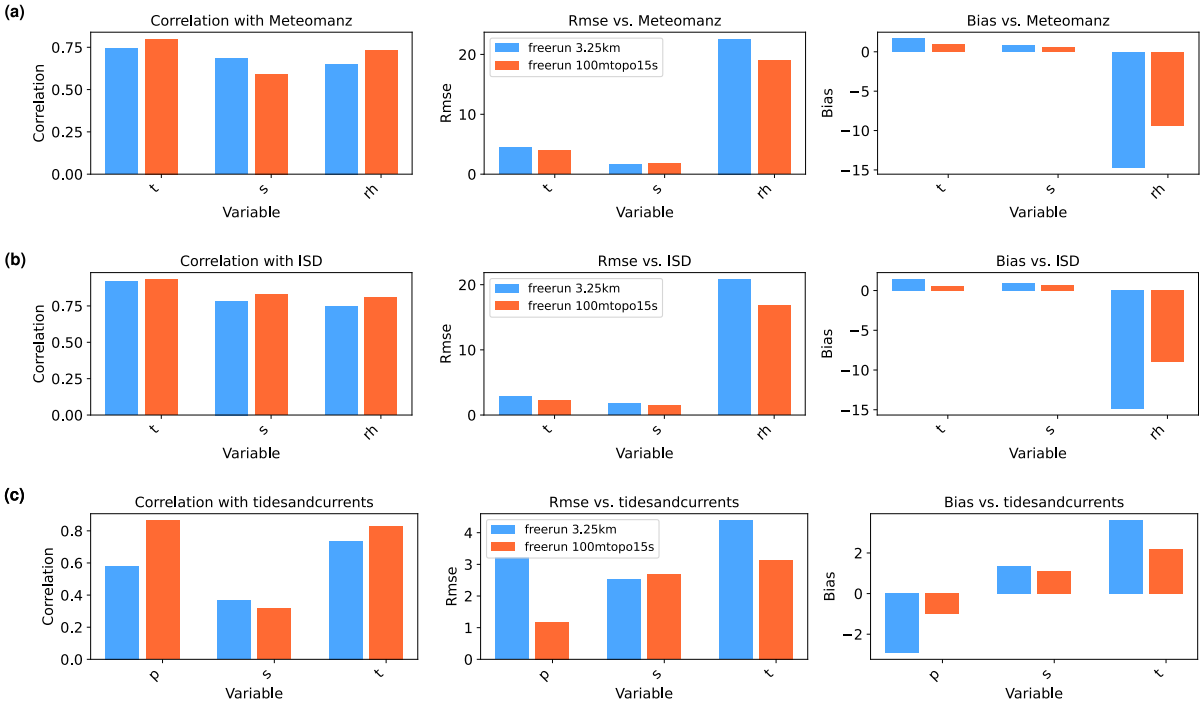

Figure 14. Same as Fig. 10 but for the Stratocumulus2023 event.

A detailed station-by-station comparison with Tides and Currents observations (Fig. 16) reveals that while wind speed performance in BA-100m degrades at a few stations—such as San Francisco Pier 1 and Port Chicago—most stations exhibit notable improvements relative to CA-3km. The frequency of wind speed oscillations in BA-100m more closely matches observations, although the simulated oscillations are less pronounced and less continuous than those observed. Surface pressure bias is also substantially reduced in BA-100m, typically decreasing from approximately 3 hPa to below 1 hPa.

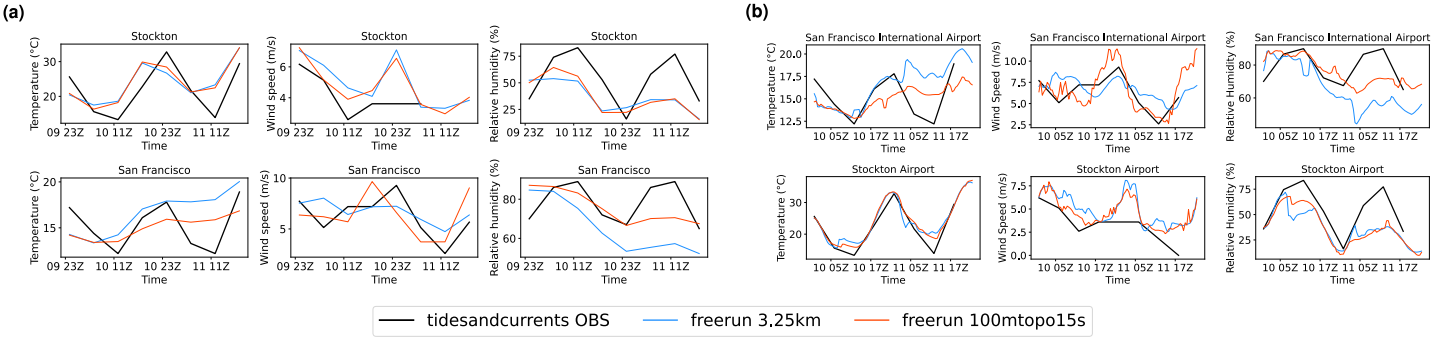

Figure 15. Same as Fig. 11 but for the Stratocumulus2023 event.

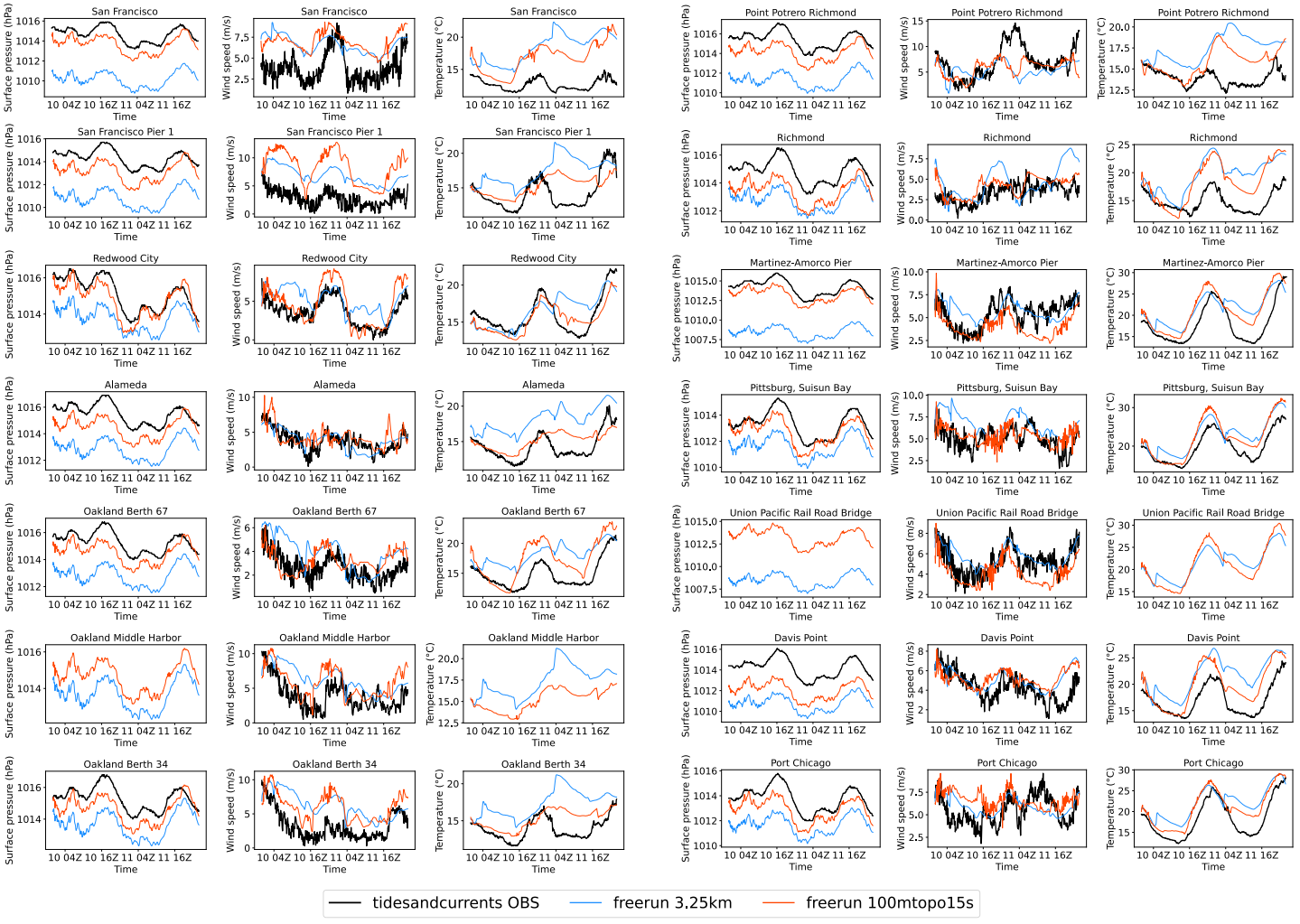

Figure 16. Same as Fig. 12 but for the Stratocumulus2023 event.

The CA-3km simulation exhibits an unrealistic temperature jump during the 4th hour of the simulation, characterized by an initial cooling followed by rapid warming—a feature consistently reflected across all Tides and Currents stations. This behavior may be associated with mesoscale wave responses triggered by the coarse representation of topography. The artificial temperature jump is followed by a sustained warm bias in CA-3km, reaching 2–4 K during the first day. On the second day, this bias becomes more pronounced, with some sites exhibiting temperature overestimations of up to 8 K. In contrast, the BA-100m simulation maintains a temperature bias below 1 K during the first 12 hours of nighttime. However, during the daytime, BA-100m shows a more rapid and exaggerated temperature increase compared to observations, leading to an overestimation of the daytime maximum temperature. Notably, BA-100m also develops a substantial warm bias on the second day, though it is approximately half the magnitude of that in CA-3km.

This behavior is likely linked to the simulation of coastal stratocumulus west of San Francisco, as illustrated in Fig.5. The CA-3km simulation fails to reproduce coastal stratocumulus on the second day, whereas BA-100m captures the feature, albeit with a smaller spatial extent than indicated by satellite observations and with a faster inland retreat. This improved representation is consistent with enhanced relative humidity in San Francisco, as observed in the Meteomanz and ISD data (Fig.15).

Unlike Storm2008, the initial conditions of Stratocumulus2023 show much better agreement with IGRA sounding observations(Fig. 17). This improved alignment is partly due to the calm weather and slowly evolving synoptic pattern, which limits the horizontal drift of the balloon, even at higher altitudes. Nevertheless, some discrepancies remain. Distinct peaks in dew point and temperature are evident in the IGRA soundings at specific pressure levels (e.g., 940 hPa, 750 hPa, and 500 hPa), but these features are absent in the model initial conditions, which appear considerably smoother. This mismatch is likely attributable to the limited vertical resolution in both the IFS model used to generate ERA5 and in SCREAM itself.

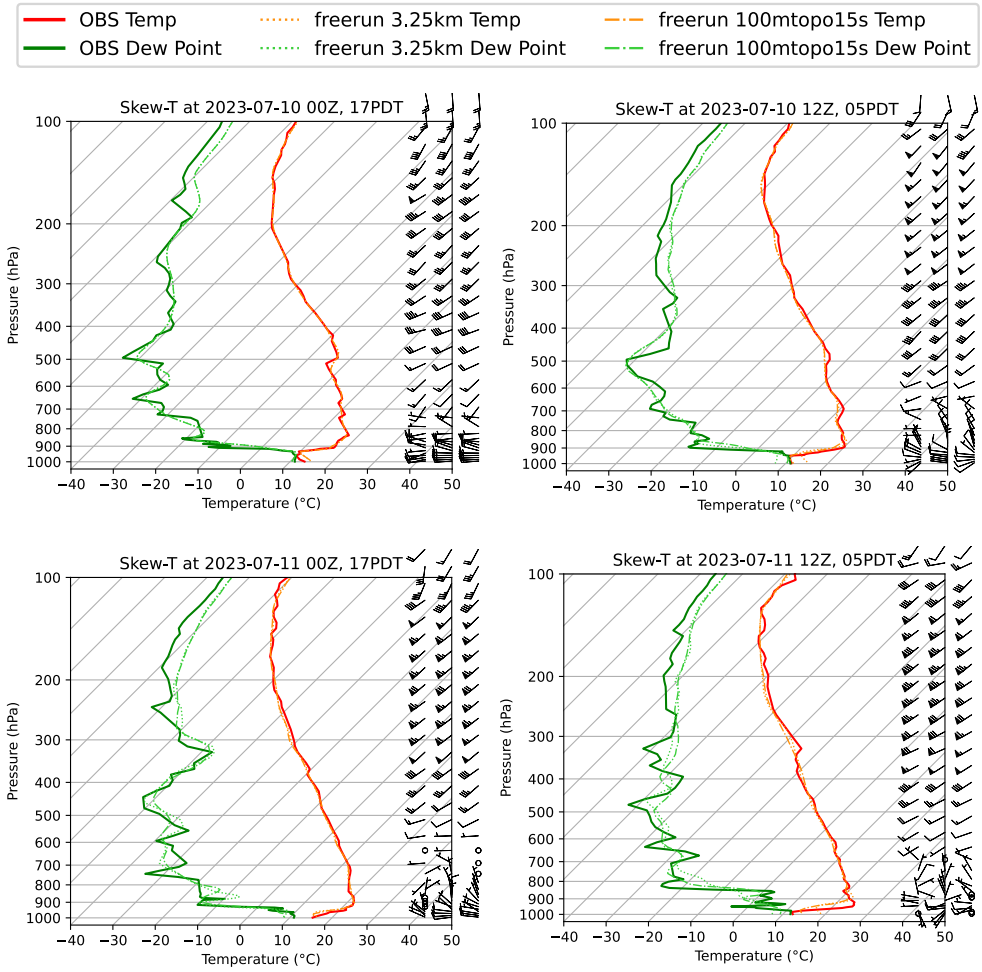

Figure 17. Same as Fig. 13 but for the Stratocumulus2023 event.

At the Oakland International Airport station, the model initial condition shows a dew point bias within 1 K below 950 hPa near the surface, while the temperature is 1–2 K higher than observed. This is consistent with the warm bias at initialization time reported in the Tides and Currents data from Oakland Berth 67 / Berth 34. The difference between CA-3km and BA-100m grows modestly over time but remains substantially smaller than the disparity observed in the Storm2008 case. Below 500 hPa, BA-100m aligns more closely with observations, particularly during the second night (07-11 12Z). The dew point decreases more sharply from the surface to approximately 850–800 hPa, although the gradient is still less pronounced than in the observed profile.

From the top of the boundary layer to near the surface, BA-100m shows markedly improved agreement with observations during the nighttime, with minimal temperature and dew point biases at 07-10 12Z. By 07-11 12Z, a modest bias develops, with the dew point 2 K lower and the temperature 2 K higher than observed. In contrast, the CA-3km simulation exhibits a dew point bias 2–3 K larger than that in BA-100m, along with a warm bias of approximately 4 K.

We speculate that the improved performance in this locally driven flow regime primarily results from enhanced representation of turbulent mixing. The observed daytime warming and wind speed increase on the first day suggest that continental heating may be overestimated. It remains unclear to what extent this bias arises from the radiation scheme, which does not currently incorporate topographic shading or surface reflections. Additionally, the sea surface temperature along the San Francisco coastline may be insufficiently resolved, contributing further to the discrepancies.

Building on the resolution sensitivity study of DP-SCREAM in Bogenschutz et al. (2023), SCREAM exhibits characteristics of a scale aware model. As horizontal resolution increases, the partitioning between subgrid-scale (SGS) and resolved turbulence diminishes. This scale awareness, inherent to the SHOC parameterization (Bogenschutz and Krueger, 2013), enables SCREAM to operate effectively at 100 m resolution without the need for parameter tuning. Moreover, simulations at 100 m resolution largely avoid the challenges associated with the turbulence gray zone, where grid spacing becomes comparable to the dominant turbulence length scale. In this transitional regime, SGS turbulent transport should ideally be modeled in three dimensions, requiring a full turbulent stress tensor scheme (e.g., Wyngaard, 2004; Chow et al., 2019; Honnert et al., 2020). However, SHOC currently represents only one-dimensional vertical turbulent transport.

In summary, our findings reinforce the scale awareness previously demonstrated in DP-SCREAM and highlight SCREAM's strong performance at large-eddy simulation (LES) resolution, particularly when coupled with realistic topography and surface heterogeneity.

5 Conclusions

This study presents the first known implementation of a global model—specifically SCREAM—at 100 m horizontal resolution within a regionally refined mesh (RRM) configuration, applied over the San Francisco Bay Area. Using two hindcast cases—one a dynamically active bomb cyclone and the other a weakly forced marine stratocumulus regime—we demonstrate that SCREAM can stably operate at large-eddy simulation (LES) scales while interacting with realistic topography, land sur-

face heterogeneity, and coastlines. This advancement bridges the gap between idealized LES setups and full-complexity global models, offering new opportunities for process-level investigation at resolutions approaching turbulence-resolving scales.

The 100 m SCREAM-RRM simulation (BA-100m) substantially reduces key near-surface biases relative to the 3.25 km SCREAM-RRM (CA-3km), particularly for surface wind speed, temperature, relative humidity, and pressure. These improvements are especially evident in the Storm2008 case, where the high wind speed bias is mitigated and high-frequency oscillations in wind speed—consistent with observations—are captured. In the Stratocumulus2023 case, boundary-layer evolution and coastal cloud retreat are better represented, highlighting SCREAM's potential in both synoptically and locally driven flow regimes.

SCREAM's performance at 100 m resolution is enabled by the scale-aware SHOC turbulence parameterization, which transitions smoothly from kilometer to LES-scale without tuning. SHOC's dependence on subgrid turbulent kinetic energy and cloud-layer depth allows the model to appropriately reduce subgrid contributions as resolution increases, effectively resolving larger eddies. Furthermore, simulations at 100 m resolution largely bypass the turbulence gray zone, although future work may benefit from expanding SCREAM's current 1-D vertical subgrid turbulence treatment to a full 3-D stress tensor formulation.

In addition to scientific advances, this study highlights important performance considerations. On CPU-only systems such as Livermore Computing's Ruby cluster, simulations at 100 m resolution advance at roughly one simulated hour per wall-clock hour, requiring nearly a month to complete a two-day hindcast. However, performance scaling tests conducted on the GPU-accelerated SCREAMv1 (EAMxx) running on NERSC's Perlmutter system demonstrate a dramatic improvement, reducing the runtime to under two wall-clock days. This shift represents a crucial step toward making ultra-high-resolution simulations viable for broader research and applications.

Overall, this work demonstrates that SCREAM can be successfully extended to 100 m resolution within a global model framework using a regionally refined grid. This capability opens the door to simulating boundary-layer turbulence, orographic flows, coastal cloud systems, and convective processes at previously inaccessible scales in Earth system models. As observational constraints and diagnostic tools improve, future work should explore applications in wind energy siting, hydrological extremes, and complex terrain validation. Further development to improve topography preprocessing, initial condition fidelity, and numerical stability will enhance the robustness and reproducibility of LES-scale SCREAM simulations across diverse regions.

Code and data availability. The SCREAM 100 m San Francisco Bay Area Regionally Refined Model 0.0 version code, in addition to the model output, can be found at Zhang (2025) (https://doi.org/10.5281/zenodo.15288872, last access: 26 April 2025). The SCREAM 100 m San Francisco Bay Area RRM (Regionally Refined Model) source code is also available on GitHub at: 1) https://github.com/jsbamboo/E3SM/tree/jzhang/RRM_SCREAMv0_BA100m (for SCREAMv0, last access: 26 April 2025) and a maint branch (BA100m-SCREAMv0-v0.0; https://github.com/jsbamboo/E3SM/releases/tag/BA100m-SCREAMv0-v0.0, last access: 26 April 2025), and 2) https://github.com/jsbamboo/E3SM/tree/jzhang/RRM_SCREAMv1_dev (for SCREAMv1/EAMxx, last access: 26 April 2025) and a maint branch (BA100m-EAMxx-v0.0; https://github.com/jsbamboo/E3SM/releases/tag/BA100m-EAMxx-v0.0, last access: 26 April 2025).

Video supplement. Video A1. From left to right: vertical velocity at 332 hPa, 808 hPa, and 951 hPa, along with the surface total precipitation rate, showing 10-minute evolution during the Storm2008 event simulated by the San Francisco Bay Area 100 m SCREAM-RRM. Each frame was produced using ncvis (https://github.com/SEATStandards/ncvis, last access: 22 April 2025) with a quadtree-based sampling strategy, optimized for unstructured cubed-sphere output. Video available at: https://portal.nersc.gov/archive/home/z/zhang73/www/videos_scream-BA100mRRM-v0.0/A1_CA100m_vs_CA3km_Storm2008_2.mp4.

Video A2. Same as Video A1 but for the Stratocumulus2023 event. Video available at: https://portal.nersc.gov/archive/home/z/zhang73/www/videos_scream-BA100mRRM-v0.0/A2_CA100m_vs_CA3km_Stratocumulus2023_2.mp4.

Please note that the videos will be posted on the AV portal of the German National Library of Science and Technology (TIB) if the article is accepted.

Author contributions. PB, JZ, and PCS conceived the study and drafted the computing-resource allocation proposal that enabled this work. JZ and PB designed the San Francisco Bay Area RRM grid and configured the simulation files and model code. JZ conducted the simulations with support from PB and MT. MT performed additional idealized dynamical core experiments. JZ and PB developed the analysis plan. PCS and PB secured the project funding. JZ and PB defined the scope of the paper and prepared the first draft. All co-authors contributed to the manuscript.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Acknowledgements. The authors thank Jos van Geffen for sharing the 500 m GMTED2010 data, which had been post-processed into a global latitude-longitude format and served as a crucial input for generating the 800 m topography on the cubed sphere. We thank Katie Lundquist for sharing a 250 m GMTED2010 tile dataset, and Kristin Chang for guidance on using GIS tools to combine another public 250 m tiles. Although the 250 m GMTED2010 version was ultimately not used, it helped us understand the original data format and may help future simulations. We acknowledge Hsi-Yen Ma for an early demonstration of CAPT's usage and for maintaining the CAPT package, Walter Hannah for maintaining the HICCUP package, and Colin Zarzycki for maintaining the betacast package. These reanalysis-based atmospheric and land-surface initial condition generation tools have significantly reduced the workload needed for climate models to settle into their initial conditions. We are grateful to Peter Hjort Lauritzen for helping debug the cube_to_target topography generation tool and for identifying the root cause of the issue. Although our simulations were already complete by then, this bug fix will benefit future simulations. We thank Chris Golaz and Robert Arthur for insightful discussions regarding the sounding dataset and the gray zone of turbulence. We appreciate Paul Ullrich for maintaining the ncvis software for native-grid visualization, which has made our rapid inspection of high-resolution E3SM outputs far more efficient and improved our day-to-day sanity.

This research is supported by the Energy Exascale Earth System Model (E3SM) project (https://e3sm.org/), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research.

This work is also supported by LLNL LDRD project [22-SI-008], “Climate Resilience for National Security” in the early stage. LLNL Institutional Computing Grand Challenge program provides the computing support for this work. Work at LLNL was performed under the auspices of the U.S. DOE by LLNL under contract (grant no. DE-AC52-07NA27344; IM release: LLNL-JRNL-2005295).

Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC (NTESS), a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration (DOE/NNSA) under contract DE-NA0003525. This written work is authored by an employee of NTESS. The employee, not NTESS, owns the right, title and interest in and to the written work and is responsible for its contents. Any subjective views or opinions that might be expressed in the written work do not necessarily represent the views of the U.S. Government. The publisher acknowledges that the U.S. Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this written work or allow others to do so, for U.S. Government purposes. The DOE will provide public access to results of federally sponsored research in accordance with the DOE Public Access Plan. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

Arthur, R. S., Lundquist, K. A., Mirocha, J. D., and Chow, F. K.: Topographic effects on radiation in the WRF Model with the immersed boundary method: Implementation, validation, and application to complex terrain, Monthly Weather Review, 146, 3277–3292, 2018.

Bogenschutz, P., Zhang, J., Tang, Q., and Cameron-Smith, P.: Atmospheric River Induced Precipitation in California as Simulated by the Regionally Refined Simplified Convective Resolving E3SM Atmosphere Model, Geoscientific Model Development, 2024.

Bogenschutz, P. A. and Krueger, S. K.: A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models, Journal of Advances in Modeling Earth Systems, 5, 195–211, https://doi.org/10.1002/jame.20018, 2013.

Bogenschutz, P. A., Eldred, C., and Caldwell, P. M.: Horizontal Resolution Sensitivity of the Simple Convection-Permitting E3SM Atmosphere Model in a Doubly-Periodic Configuration, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022ms003466, 2023.

Bretherton, C. S. and Park, S.: A New Moist Turbulence Parameterization in the Community Atmosphere Model, Journal of Climate, 22, 3422–3448, https://doi.org/10.1175/2008jcli2556.1, 2009.

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., Beydoun, H., Taylor, M., Bertagna, L., Bradley, A. M., Clevenger, T. C., Donahue, A. S., Eldred, C., Foucar, J., Golaz, J. C., Guba, O., Jacob, R., Johnson, J., Krishna, J., Liu, W., Pressel, K., Salinger, A. G., Singh, B., Steyer, A., Ullrich, P., Wu, D., Yuan, X., Shpund, J., Ma, H. Y., and Zender, C. S.: Convection-Permitting Simulations With the E3SM Global Atmosphere Model, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/10.1029/2021ms002544, 2021.

Cheng, A. N., Xu, K. M., and Stevens, B.: Effects of Resolution on the Simulation of Boundary-layer Clouds and the Partition of Kinetic Energy to Subgrid Scales, Journal of Advances in Modeling Earth Systems, 2, https://doi.org/Artn 3 10.3894/James.2010.2.3, 2010.

Chow, F., Schär, C., Ban, N., Lundquist, K., Schlemmer, L., and Shi, X.: Crossing Multiple Gray Zones in the Transition from Mesoscale to Microscale Simulation over Complex Terrain, Atmosphere, 10, https://doi.org/10.3390/atmos10050274, 2019.

Chow, F. K., Weigel, A. P., Street, R. L., Rotach, M. W., and Xue, M.: High-resolution large-eddy simulations of flow in a steep Alpine valley. Part I: Methodology, verification, and sensitivity experiments, Journal of Applied Meteorology and Climatology, 45, 63–86, https://doi.org/Doi 10.1175/Jam2322.1, 2006.

Connolly, A., Chow, F. K., and Hoch, S. W.: Nested Large-Eddy Simulations of the Displacement of a Cold-Air Pool by Lee Vortices, Boundary-Layer Meteorology, 178, 91–118, https://doi.org/10.1007/s10546-020-00561-6, 2021.

Danielson, J. J. and Gesch, D. B.: Global multi-resolution terrain elevation data 2010 (GMTED2010), Report 2331-1258, US Geological Survey, 2011.

De Wekker, S. F. J., Kossmann, M., Knievel, J. C., Giovannini, L., Gutmann, E. D., and Zardi, D.: Meteorological Applications Benefiting from an Improved Understanding of Atmospheric Exchange Processes over Mountains, Atmosphere, 9, https://doi.org/ARTN 371 10.3390/atmos9100371, 2018.

Dipankar, A., Stevens, B., Heinze, R., Moseley, C., Zängl, G., Giorgetta, M., and Brdar, S.: Large eddy simulation using the general circulation model ICON, Journal of Advances in Modeling Earth Systems, 7, 963–986, https://doi.org/10.1002/2015ms000431, 2015.

Donahue, A. S., Caldwell, P. M., Bertagna, L., Beydoun, H., Bogenschutz, P. A., Bradley, A., Clevenger, T. C., Foucar, J. G., Golaz, J.-C., and Guba, O.: To exascale and beyond–The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM), a performance portable global atmosphere model for cloud-resolving scales, Authorea Preprints, 2024.

Golaz, J. C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., Abeshu, G., Anantharaj, V., Asay-Davis, X. S., Bader, D. C., Baldwin, S. A., Bisht, G., Bogenschutz, P. A., Branstetter, M., Brunke, M. A., Brus, S. R., Burrows, S. M., Cameron-Smith, P. J., Donahue, A. S., Deakin, M., Easter, R. C., Evans, K. J., Feng, Y., Flanner, M., Foucar, J. G., Fyke, J. G., Griffin, B. M., Hannay, C., Harrop, B. E., Hoffman, M. J., Hunke, E. C., Jacob, R. L., Jacobsen, D. W., Jeffery, N., Jones, P. W., Keen, N. D., Klein, S. A., Larson, V. E., Leung, L. R., Li, H. Y., Lin, W. Y., Lipscomb, W. H., Ma, P. L., Mahajan, S., Maltrud, M. E., Mametjanov, A., McClean, J. L., McCoy, R. B., Neale, R. B., Price, S. F., Qian, Y., Rasch, P. J., Eyre, J. E. J. R., Riley, W. J., Ringler, T. D., Roberts, A. F., Roesler, E. L., Salinger, A. G., Shaheen, Z., Shi, X. Y., Singh, B., Tang, J. Y., Taylor, M. A., Thornton, P. E., Turner, A. K., Veneziani, M., Wan, H., Wang, H. L., Wang, S. L., Williams, D. N., Wolfram, P. J., Worley, P. H., Xie, S. C., Yang, Y., Yoon, J. H., Zelinka, M. D., Zender, C. S., Zeng, X. B., Zhang, C. Z., Zhang, K., Zhang, Y., Zheng, X., Zhou, T., and Zhu, Q.: The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution, Journal of Advances in Modeling Earth Systems, 11, 2089–2129, https://doi.org/10.1029/2018ms001603, 2019.

Guba, O., Taylor, M. A., Bradley, A. M., Bosler, P. A., and Steyer, A.: A framework to evaluate IMEX schemes for atmospheric models, Geoscientific Model Development, 13, 6467–6480, https://doi.org/10.5194/gmd-13-6467-2020, 2020.

Guerra, J. E. and Ullrich, P. A.: A high-order staggered finite-element vertical discretization for non-hydrostatic atmospheric models, Geoscientific Model Development, 9, 2007–2029, https://doi.org/10.5194/gmd-9-2007-2016, 2016.

Hannah, W. M., Bradley, A. M., Guba, O., Tang, Q., Golaz, J. C., and Wolfe, J.: Separating Physics and Dynamics Grids for Improved Computational Efficiency in Spectral Element Earth System Models, Journal of Advances in Modeling Earth Systems, 13, https://doi.org/ARTN e2020MS002419 10.1029/2020MS002419, 2021.

Harris, L. M. and Lin, S.-J.: A Two-Way Nested Global-Regional Dynamical Core on the Cubed-Sphere Grid, Monthly Weather Review, 141, 283–306, https://doi.org/https://doi.org/10.1175/MWR-D-11-00201.1, 2013.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., and Schepers, D.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Honnert, R., Efstathiou, G. A., Beare, R. J., Ito, J., Lock, A., Neggers, R., Plant, R. S., Shin, H. H., Tomassini, L., and Zhou, B. W.: The Atmospheric Boundary Layer and the "Gray Zone" of Turbulence: A Critical Review, Journal of Geophysical Research-Atmospheres, 125, https://doi.org/ARTN e2019JD030317 10.1029/2019JD030317, 2020.

Huang, X. and Ullrich, P. A.: The Changing Character of Twenty-First-Century Precipitation over the Western United States in the Variable-Resolution CESM, Journal of Climate, 30, 7555–7575, https://doi.org/https://doi.org/10.1175/JCLI-D-16-0673.1, 2017.

Hunke, E., Lipscomb, W., Turner, A., Jeffery, N., and Elliott, S.: CICE: The Los Alamos sea ice model, documentation and software, Report, version 4.0, Tech. Rep. LA-CC-06-012, Los Alamos National Laboratory, 2008.

Langhans, W., Schmidli, J., and Schär, C.: Mesoscale Impacts of Explicit Numerical Diffusion in a Convection-Permitting Model, Monthly Weather Review, 140, 226–244, https://doi.org/10.1175/2011mwr3650.1, 2012.

Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., and Taylor, M. A.: NCAR_Topo (v1.0): NCAR global model topography generation software for unstructured grids, Geoscientific Model Development, 8, 3975–3986, https://doi.org/10.5194/gmd-8-3975-2015, 2015.

Lee, H., Bogenschutz, P., and Yamaguchi, T.: Resolving Away Stratocumulus Biases in Modern Global Climate Models, Geophysical Research Letters, 49, https://doi.org/10.1029/2022gl099422, 2022.

Liu, W., Ullrich, P. A., Li, J., Zarzycki, C., Caldwell, P. M., Leung, L. R., and Qian, Y.: The June 2012 North American Derecho: A Testbed for Evaluating Regional and Global Climate Modeling Systems at Cloud-Resolving Scales, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022ms003358, 2023.

Lundquist, J., Hughes, M., Gutmann, E., and Kapnick, S.: Our Skill in Modeling Mountain Rain and Snow is Bypassing the Skill of Our Observational Networks, Bulletin of the American Meteorological Society, 100, 2473–2490, https://doi.org/10.1175/bams-d-19-0001.1, 2019.

Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, Journal of the Atmospheric Sciences, 72, 287–311, https://doi.org/10.1175/jas-d-14-0065.1, 2015.

Pincus, R., Mlawer, E. J., and Delamere, J. S.: Balancing Accuracy, Efficiency, and Flexibility in Radiation Calculations for Dynamical Models, Journal of Advances in Modeling Earth Systems, 11, 3074–3089, https://doi.org/https://doi.org/10.1029/2019MS001621, 2019.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G.: Daily high-resolution-blended analyses for sea surface temperature, Journal of climate, 20, 5473–5496, 2007.

Rhoades, A. M., Zarzycki, C. M., Inda-Diaz, H. A., Ombadi, M., Pasquier, U., Srivastava, A., Hatchett, B. J., Dennis, E., Heggli, A., McCrary, R., McGinnis, S., Rahimi-Esfarjani, S., Slinskey, E., Ullrich, P. A., Wehner, M., and Jones, A. D.: Recreating the California New Year's Flood Event of 1997 in a Regionally Refined Earth System Model, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2023ms003793, 2023.

Ringler, T., Ju, L., and Gunzburger, M.: A multiresolution method for climate system modeling: application of spherical centroidal Voronoi tessellations, Ocean Dynamics, 58, 475–498, https://doi.org/10.1007/s10236-008-0157-2, 2008.

Satoh, M.: Conservative scheme for the compressible nonhydrostatic models with the horizontally explicit and vertically implicit time integration scheme, Monthly Weather Review, 130, 1227–1245, https://doi.org/Doi 10.1175/1520-0493(2002)130<1227:Csftcn>2.0.Co;2, 2002.

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

Steyer, A., Vogl, C. J., Taylor, M., and Guba, O.: Efficient IMEX Runge-Kutta methods for nonhydrostatic dynamics, arXiv preprint arXiv:1906.07219, 2019.

Sun, J. L., Nappo, C. J., Mahrt, L., Belusic, D., Grisogono, B., Stauffer, D. R., Pulido, M., Staquet, C., Jiang, Q. F., Pouquet, A., Yagüe, C., Galperin, B., Smith, R. B., Finnigan, J. J., Mayor, S. D., Svensson, G., Grachev, A. A., and Neff, W. D.: Review of wave-turbulence interactions in the stable atmospheric boundary layer, Reviews of Geophysics, 53, 956–993, https://doi.org/10.1002/2015rg000487, 2015.

Tang, Q., Klein, S. A., Xie, S. C., Lin, W. Y., Golaz, J. C., Roesler, E. L., Taylor, M. A., Rasch, P. J., Bader, D. C., Berg, L. K., Caldwell, P., Giangrande, S. E., Neale, R. B., Qian, Y., Riihimaki, L. D., Zender, C. S., Zhang, Y. Y., and Zheng, X.: Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling, Geoscientific Model Development, 12, 2679–2706, https://doi.org/10.5194/gmd-12-2679-2019, 2019.

Tang, Q., Golaz, J.-C., Van Roekel, L. P., Taylor, M. A., Lin, W., Hillman, B. R., Ullrich, P. A., Bradley, A. M., Guba, O., Wolfe, J. D., Zhou, T., Zhang, K., Zheng, X., Zhang, Y., Zhang, M., Wu, M., Wang, H., Tao, C., Singh, B., Rhoades, A. M., Qin, Y., Li, H.-Y., Feng, Y., Zhang, Y., Zhang, C., Zender, C. S., Xie, S., Roesler, E. L., Roberts, A. F., Mametjanov, A., Maltrud, M. E., Keen, N. D., Jacob, R. L., Jablonowski, C., Hughes, O. K., Forsyth, R. M., Di Vittorio, A. V., Caldwell, P. M., Bisht, G., McCoy, R. B., Leung, L. R., and Bader, D. C.: The fully coupled regionally refined model of E3SM version 2: overview of the atmosphere, land, and river results, Geoscientific Model Development, 16, 3953–3995, https://doi.org/10.5194/gmd-16-3953-2023, 2023.

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., and Eldrid, C.: An Energy Consistent Discretization of the Nonhydrostatic Equations in Primitive Variables, Journal of Advances in Modeling Earth Systems, 12, https://doi.org/ARTN e2019MS001783 10.1029/2019MS001783, 2020.

Trenberth, K. E., Berry, J. C., and Buja, L. E.: Vertical interpolation and truncation of model-coordinate data, National Center for Atmospheric Research, Climate and Global Dynamics Division, https://doi.org/https://doi.org/10.5065/D6HX19NH, 1993.

Ullrich, P. and Roesler, E.: ClimateGlobalChange/squadgen: v1.2.2 (v1.2.2), https://doi.org/https://doi.org/10.5281/zenodo.13241731, 2024.

Ullrich, P. A. and Taylor, M. A.: Arbitrary-order conservative and consistent remapping and a theory of linear maps: Part I, Monthly Weather Review, 143, 2419–2440, 2015.

Wyngaard, J. C.: Toward numerical modeling in the "terra incognita", Journal of the Atmospheric Sciences, 61, 1816–1826, https://doi.org/Doi 10.1175/1520-0469(2004)061<1816:Tnmitt>2.0.Co;2, 2004.

Zarzycki, C. M. and Jablonowski, C.: A multidecadal simulation of Atlantic tropical cyclones using a variable-resolution global atmospheric general circulation model, Journal of Advances in Modeling Earth Systems, 6, 805–828, https://doi.org/https://doi.org/10.1002/2014MS000352, 2014.

Zarzycki, C. M., Jablonowski, C., and Taylor, M. A.: Using variable-resolution meshes to model tropical cyclones in the Community Atmosphere Model, Monthly Weather Review, 142, 1221–1239, 2014.

Zender, C. S.: Analysis of self-describing gridded geoscience data with netCDF Operators (NCO), Environmental Modelling & Software, 23, 1338–1342, 2008.

Zhang, J.: Code and Model Data for SCREAM 100 m San Francisco Bay Area Regionally Refined Model 0.0 version (0.0) [Data set]., https://doi.org/10.5281/zenodo.15288872, 2025.

Zhang, J., Bogenschutz, P., Tang, Q., Cameron-smith, P., and Zhang, C.: Leveraging regional mesh refinement to simulate future climate projections for California using the Simplified Convection-Permitting E3SM Atmosphere Model Version 0, Geoscientific Model Development, 17, 3687–3731, 2024a.

Zhang, J., Caldwell, P. M., Bogenschutz, P. A., Ullrich, P. A., Bader, D. C., Duan, S., and Beydoun, H.: Through the lens of a kilometer-scale climate model: 2023 Jing-Jin-Ji flood under climate change, Authorea Preprints, 2024b.

Zhong, S. and Chow, F. K.: Meso-and fine-scale modeling over complex terrain: Parameterizations and applications, pp. 591–653, Springer, 2012.