

An Evaluation of the Variable Resolution-CESM for Modeling California's Climate

Xingying Huang,¹ Alan M. Rhoades,¹ Paul A. Ullrich,¹ Colin M. Zarzycki,²

Abstract.

In this paper the recently developed variable-resolution option within the Community Earth System Model (VR-CESM) is assessed for long-term regional climate modeling of California at ~ 28 km and ~ 14 km horizontal resolutions. The mean climatology of near-surface temperature and precipitation is analyzed and contrasted with reanalysis, gridded observational datasets and a traditional regional climate model (RCM) – the Weather Research and Forecasting (WRF) model. Statistical metrics for model evaluation and tests for differential significance have been extensively applied. With only prescribed sea surface temperatures, VR-CESM tended to produce a warmer summer (by about 1 to 3 °C) and overestimated overall winter precipitation (about 25%-35%) compared to reference datasets. Increasing resolution from 28 km to 14 km did not produce a statistically significant improvement in the model results. By comparison, the analogous WRF climatology (constrained laterally and at the sea surface by ERA-Interim reanalysis) was ~ 1 to 3 °C colder than the reference datasets, underestimated precipitation by $\sim 20\%$ - 30% at 27 km resolution and overestimated precipitation by $\sim 65\%$ - 85% at 9 km. Overall, VR-CESM produced comparable statistical biases to WRF in key climatological quantities. This assessment highlights the value of variable-resolution global climate models (VRGCMs) in capturing fine-scale atmospheric processes, projecting future regional climate and addressing the computational expense of uniform-resolution global climate models.

1. Introduction

Global climate models (GCMs) have been widely used to simulate both past and future climate. Although these models have demonstrable success in representing large-scale features of the climate system, they are usually employed at relatively coarse resolutions ($\sim 1^\circ$), largely as a result of the substantial computational cost required at higher resolutions. Global climate reanalysis datasets, which assimilate climate observations using a global model, represent a best estimate of historical weather patterns. However, reanalysis datasets still cannot fulfill the needs of policymakers, stakeholders and researchers that require high-resolution regional climate data (<http://reanalyses.org/atmosphere/overview-current-reanalyses>). Regional features such as microclimates, land cover, and topography, are not well captured by either GCMs or reanalysis datasets [Leung *et al.*, 2003]. However, dynamical processes at unrepresented scales are significant drivers for local climate variability, especially over complex terrain [Soares *et al.*, 2012]. In order to capture fine-scale dynamical features, high horizontal resolution is needed for a more accurate representation of small-scale processes and interactions [Rauscher *et al.*, 2010]. With these enhancements, regional climate data is expected to be more useful for formulating climate adaptation and mitigation strategies locally.

In order to model regional climate at high spatial resolutions over a limited area, downscaling techniques have been developed, such as statistical and dynamical downscaling. Dynamical downscaling typically uses nested limited-area models (LAMs) or, more recently, variable-resolution enabled GCMs (VRGCMs) [Laprise, 2008]. In this context,

LAMs are typically referred to as regional climate models (RCMs) when used for climate study. Forced by the output either from GCMs or reanalysis datasets, RCMs have been widely used to capture physically consistent regional and local circulations at the needed spatial and temporal scales [Christensen *et al.*, 2007; Bukovsky and Karoly, 2009; Mearns *et al.*, 2012]. Recently, there has been a growing interest in the use of VRGCMs for modeling regional climate. Unlike RCMs, VRGCMs use a relatively coarse global model with enhanced resolution over a specific region [Staniforth and Mitchell, 1978; Fox-Rabinovitz *et al.*, 1997]. Strategies that have been employed for transitioning between coarse and fine-resolution regions within a VRGCM include grid stretching [Fox-Rabinovitz *et al.*, 1997; McGregor and Dix, 2008] and grid refinement [Ringler *et al.*, 2008; Skamarock *et al.*, 2012; Zarzycki *et al.*, 2014a]. VRGCMs have been demonstrated to be effective for regional climate studies and applications at a reduced computational cost compared to uniform GCMs [Fox-Rabinovitz *et al.*, 2006; Rauscher *et al.*, 2013; Zarzycki *et al.*, 2015].

Compared with RCMs, a key advantage of VRGCMs is that they use a single, unified modeling framework, rather than two separate models (GCM and RCM) with potentially disparate dynamics and physics parameterizations. RCMs may suffer from potential inconsistencies between the global and regional scales and lack two-way interactions at the nest boundary [Warner *et al.*, 1997; McDonald, 2003; Laprise *et al.*, 2008; Mesinger and Veljovic, 2013], which can be mitigated with the use of VRGCMs. VRGCMs also provide a cost-effective method of reaching high resolutions over a region of interest – the limited area simulations in this study at 28km and 14km resolution represent a reduction in required computation of approximately 10 and 25 times, respectively, compared to analogous globally uniform high-resolution simulations. For the purposes of this paper, we focus on the recently developed Community Earth System Model with variable-resolution option (VR-CESM) as our VRGCM of interest. This configuration is driven by the Community Atmosphere Model's (CAM's) Spectral Element (SE) dynamical core, which possesses attractive conservation and parallel

¹Department of Land, Air and Water Resources, University of California, Davis

²National Center for Atmospheric Research

scaling properties [Dennis et al., 2011; Taylor, 2011], as well as recently developed variable-resolution capabilities [Zarzycki et al., 2014a]. This model has been employed by Zarzycki et al. [2014b] to show that a high-resolution refinement patch in the Atlantic basin for simulating tropical cyclones represented significant improvements over the unrefined simulation. Zarzycki et al. [2015] also compared the large-scale climatology of VR-CESM 0.25° and uniform CESM at 1°, and found that adding a refined region over the globe did not noticeably affect the global circulation. Rhoades et al. [2015] has also assessed the use of VR-CESM for modeling Sierra Nevada mountain snowpack in the western United States.

However, for the purposes of long-term regional climate modeling, particularly in California, VR-CESM has yet to be rigorously evaluated. This paper aims to fill that gap by analyzing the performance of VR-CESM against gridded observational data, reanalysis data and in comparison to a traditional RCM. Our variable-resolution simulations are implemented with horizontal resolutions of 28km and 14km respectively, which are much more typical for dynamically downscaled studies. This paper focuses on California in the western United States as the study area. With its complex topography, coastal influence, and wide latitudinal extent, California is an excellent test bed for regional climate modeling. Further, an understanding of local climate variability is incredibly important for policymakers and stakeholders in California due to its vast agricultural industry, mixed demographics, and vulnerability to anthropogenically-induced climate change [Hayhoe et al., 2004; Cayan et al., 2008].

For comparison, the Weather Research and Forecasting (WRF, Skamarock et al. [2005]) model has been used for simulating California's climatology at 27km and 9km grid spacing. RCM simulations over California have also been conducted in previous studies and demonstrated the need for high spatial and temporal resolution to better address regional climate and extreme events, especially in the vicinity of complex topography where large climatological gradients are present [Leung et al., 2004; Kanamitsu and Kanamaru, 2007; Caldwell et al., 2009; Pan et al., 2011; Pierce et al., 2013]. In particular, Caldwell et al. [2009] presented results from WRF at 12km spatial resolution and showed that, although the RCM was effective at simulating the mean climate when compared with observations, some clear biases persisted (particularly an overestimation of precipitation).

This study focuses on the models' ability to represent current climate statistics, particularly those relevant to heat and precipitation extremes. We anticipate that this work will validate VR-CESM for modeling the mean regional climatology of California and will further motivate the adoption of variable-resolution modeling to study other local climatic processes. Our eventual goal is to utilize these models for assessing historical and future regional climate extremes.

This paper is organized as follows: Section 2 describes the model setup, datasets and methodology for evaluation and intercomparison. In section 3, simulation results are provided and discussed, with focuses on near-surface (2-meter) temperature and precipitation. Key results are summarized along with further discussion in section 4.

2. Models and Methodology

2.1. Simulation design

In this study, all global simulations use the Atmospheric Model Intercomparison Project (AMIP) experimental protocols [Gates, 1992]. These protocols are widely used and support climate model diagnosis, validation and intercomparison. AMIP experiments are constrained by realistic near-surface temperatures (SSTs) and sea ice from 1979 to near present without the added complexity of ocean-atmosphere

feedbacks in the climate system. In particular, observed SSTs and sea ice at 1° horizontal resolution are provided and updated following the procedure described by [Hurrell et al., 2008].

2.1.1. VR-CESM

CESM is a state-of-the-art Earth modeling framework managed by the National Center for Atmospheric Research (NCAR), consisting of coupled atmospheric, oceanic, land and sea ice models. For decades CESM has been used for modeling present and future global climate [Neale et al., 2010a; Hurrell et al., 2013]. The coupling infrastructure in CESM communicates the interfacial states and fluxes between each component model to ensure conservation. Since we follow AMIP protocols, only the atmosphere and land model are integrated dynamically. Here, CAM version 5 (CAM5) [Neale et al., 2010b] and the Community Land Model (CLM) version 4.0 [Oleson et al., 2010] are used. As mentioned earlier, the SE dynamical core is employed along with variable-resolution grid support. The FAMIPC5 (F_AMIP_CAM5) component set is chosen for these simulations.

For our study, the variable-resolution cubed-sphere grids are generated for use in CAM and CLM with the open-source software package SquadGen [Ulrich, 2014; Guba et al., 2014]. The grids used in this study are depicted in Figure 1. The maximum horizontal resolution on these grids is 0.25° (~28km) and 0.125° (~14km) respectively, with a quasi-uniform 1° mesh over the remainder of the globe. Grids are constructed using a paving technique with a 2:1 spatial resolution ratio, so two transition layers are required from 1° to 0.25°, and one additional transition from 0.25° to 0.125°. In our study, and previous studies (e.g. Zarzycki et al. [2015]), general circulation patterns (e.g., wind, pressure and precipitation) do not exhibit apparent artifacts in the variable-resolution transition region, and the design of the SE dynamical core ensures that dry air and tracer mass are conserved globally [Taylor and Fournier, 2010]. Simulations are performed over the time period from 1979-01-01 to 2005-12-31 (UTC) and year 1979 is discarded as a spin-up period. This 26-year time period is chosen to provide an adequate sampling of inter-annual variability, to limit computational cost, and to coincide with the satellite era where adequate high-quality gridded and reanalysis datasets are available.

Variable-resolution topography files were produced by sampling the National Geophysical Data Center (NGDC) 2-min (~4 km) Gridded Global Relief Dataset (ETOPO2v2), followed by the application of a differential smoothing technique as described in Zarzycki et al. [2015]. Using this technique, the c parameter from their Eq. (1) was adjusted to reduce noise in the vertical pressure velocity field. The grid-scale topography is depicted in Figure 2, including the topography of uniform CESM at 1°. As we can see, higher resolution provides clear improvement in the representation of regional topography, which is necessary for the correct treatment of fine-scale dynamic processes strongly influenced by complex terrain. Topography at very coarse resolution (~1°) cannot represent local details like the shape of valleys or mountain peaks, resulting in the loss of regional climate patterns.

Land surface datasets, including plant functional types, at 0.5° were adopted. Greenhouse gas (GHG) concentrations and aerosol forcings are prescribed based on historical observations. CAM and CLM tuning parameters are not modified from their default configurations.

2.1.2. WRF

WRF has been widely used over the past decade for modeling regional climate [Lo et al., 2008; Leung and Qian, 2009; Soares et al., 2012; Sun et al., 2015]. In our study, the fully compressible non-hydrostatic WRF model (version

3.5.1) with the Advanced Research WRF (ARW) dynamical core is used. WRF is a limited area model that supports nested domains with a typical refinement ratio of 3:1. The simulation domains of WRF are depicted in Figure 3. Two WRF simulations, representing finest grid resolutions of 27 km and 9 km, are conducted. For the WRF 27km simulation, one domain is used. For the WRF 9km simulation, two domains are used, with the outer domain at 27 km (same as the WRF 27km) and an inner nested domain at 9 km horizontal grid resolution. For both simulations, grids are centered on California and have 120×110 and 151×172 grid points, respectively. At all lateral boundaries, 10 grid points are used for relaxation to the coarse solution. In order to reduce the drift between forcing data and modeling output, grid nudging [Stauffer and Seaman, 1990] is applied to the outer domain every 6 hours at all levels except the planetary boundary layers (PBL), as suggested by Lo *et al.* [2008]. This setup uses 41 vertical levels with model top pressure at 50 hPa.

Additionally, the following physics parameterizations are employed: WSM (WRF Single-Moment) 6-class graupel microphysics scheme [Hong and Lim, 2006], Kain-Fritsch cumulus scheme [Kain, 2004], CAM shortwave and longwave radiation schemes [Collins *et al.*, 2004]. These settings are chosen by assessing the results from several common parameterization combinations over one-year simulation, with comparing to gridded observations. For the boundary layer, the Yonsei University scheme (YSU) [Hong *et al.*, 2006] is used, and the Noah Land Surface Model [Chen and Dudhia, 2001] is applied. Both are chosen as they are common for climate applications that balance long-term reliability and computational cost. Although many other options and combinations of parameterizations are available for configuring WRF (and others have tackled a complete assessment of these options for particular problems), our choices are made simply to represent a typical WRF configuration.

ECMWF Reanalysis (ERA-Interim) data at both the surface and multiple pressure-levels provides initial and lateral conditions for the domains. The lateral conditions and SSTs are updated every 6 hours. ERA-Interim reanalysis (~ 80 km) has been widely used and validated for its reliability as forcing data [Dee *et al.*, 2011]. WRF simulations are conducted over the same time period as VR-CESM (i.e., 1979-01-01 through 2005-12-31 UTC). Again, the year 1979 is used as a spin-up period and is discarded for purposes of analysis. Notably, the ~ 9 km resolution employed in the innermost domain is finer than most previous studies for long-term climate.

The topography employed for the 27 km and 9 km simulations is interpolated from USGS (United States Geological Survey) elevation data with 10-min (~ 20 km) and 2-min (~ 4 km) resolution, respectively. The post-processed grid-scale topography is contrasted in Figure 2. Elevation differences between VR-CESM and WRF are irregular and relatively small, except over the Central Valley where VR-CESM has consistently higher values than WRF. This indicates a different methodology for preparation of the topography dataset and may also be partly due to the use of the USGS elevation instead of NGDC elevation datasets.

2.2. Methodology

Near-surface temperature and precipitation have been analyzed over California to assess the performance of VR-CESM in representing the mean climatology. Specifically, our evaluation focuses on daily maximum, minimum and average near-surface temperatures (T_{max} , T_{min} and T_{avg}) and daily precipitation (Pr). These variables are key in a baseline climate assessment due to their close relationship with water resources, agriculture and health. In this context, the

biggest impact of weather on California is through heat and precipitation extremes. Since heat extremes dominate during the summer season, we focus on June, July and August (JJA) for assessment of temperature. On the other hand, since the vast majority of precipitation in California occurs in the winter season, December-January-February (DJF) is emphasized.

In order to adequately account for natural variability of the mean climate, the simulation period must be chosen appropriately [Solomon, 2007]. However, the number of simulated years required for adequate climate statistics depends greatly on the regional climate variability and spatial scale. Past studies have used average weather conditions over a 30-year period to ensure sufficient statistics and to avoid imprinting from annual variability [Dinse, 2009]. To check that our 26-year simulation period is sufficient, we have examined the interannual variability of mean temperature and precipitation in all simulations and observations over 5, 10, 20 and 25 seasons or years. We observe that for climatological mean temperature and precipitation, the relevant statistics are effectively converged for a 20-year sample, suggesting that our simulation period is sufficient to adequately capture the variability of these quantities.

The results in section 4 are obtained from simulated and observed data over the period 1980 to 2005. All datasets have been linearly de-trended at each grid point so as to facilitate averaging of all simulation years. It is found that, for annual and JJA near-surface temperature (T_{max} , T_{min} and T_{avg}), a statistically significant trend is present under the two-tailed t-statistic with a significance level of 0.05. The average magnitude of the linearly fit trend is about 1.3 K over 26 years for T_{max} , T_{min} or T_{avg} . No statistically significant trend has been detected for precipitation.

California consists of a diverse variety of climate regions as a consequence of its rugged topography and large latitudinal extent. The distinct character of these regions is poorly captured in typical coarse global climate simulations [Abatzoglou *et al.*, 2009; Caldwell *et al.*, 2009]. In order to assess the performance of VR-CESM within each region, the state has been divided into five climate divisions, including the Central Valley (CV), Mountain Region (MR), North Coast (NC), South Coast (SC), and Desert Region (DR). The spatial extent of these divisions is depicted in Figure 3. These five divisions are determined loosely based on the results of Abatzoglou *et al.* [2009] and the climate divisions used by the California Energy Commission. To restrict the analysis in each division, simulations and datasets have been masked to restrict climate variables to each division.

Standard statistical measures have been used to quantify the model performance in comparison with the reference datasets. These include the root-mean-square deviation (RMSD), mean signed difference (MSD), mean relative absolute difference (MRD), and sample standard deviation (s). Further, spatial correlation is assessed by computing Pearson product-moment coefficient of linear correlation between climatological means from models and reference datasets. Mathematically, these quantities are written as

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N (v_i - \hat{v}_i)^2} \quad (1)$$

$$MSD = \frac{1}{N} \sum_{i=1}^N (v_i - \hat{v}_i) \quad (2)$$

$$s = \sqrt{\frac{1}{M-1} \sum_{j=1}^M (v_j - \bar{v})^2} \quad (3)$$

$$MRD = \left(\sum_{i=1}^N |v_i - \hat{v}_i| \right) / \left(\sum_{i=1}^N \hat{v}_i \right). \quad (4)$$

where v_i and \hat{v}_i are values from the simulation output and reference dataset, respectively; i is the grid-point index and N is the total number of grid points over specific regions; j is the simulation year index, M is the total number of simulated years and \bar{v} is the mean value over all years. Grid-point differences are calculated by remapping the reference datasets to the model's output grid using bilinear interpolation. Remapping using patch-based interpolation has also been tested and nearly identical results have been observed. When necessary, the statistical quantities are further averaged over each division.

Throughout the remainder of this paper, student's t-test has been used to test whether two sets of annual-, seasonal- or monthly-average data are the same. F-test is applied to test whether the sample variances are equal. These tests are used only when the sample population can be described adequately by a normal distribution, where normality is assessed under the Anderson-Darling test. When the sample populations do not approximately follow a normal distribution, Mann-Whitney-Wilcoxon (MWW) test and Levene's test are employed in lieu of the t-test and F-test, respectively. All statistical tests are evaluated at the $p = 0.05$ significance level.

Complementary results to this study are provided in the online supplement, including the original grid-refined mesh files, the sensitivity of climatological statistics to choice of time period, the observed time trend, and other seasons not addressed in this paper and corresponding statistics metric tables. Results are also provided with comparison of VR-CESM to the output from a globally uniform CESM run at 0.25° spatial resolution with the finite volume (FV) dynamical core [Wehner et al., 2014].

2.3. Gridded and Reanalysis Datasets

Reanalysis and gridded observational datasets of the highest available quality are employed (see Table 1). Differences between gridded observations can be due to the choice of meteorological stations, interpolation techniques, elevation models and processing algorithms. Consequently, the use of multiple reference datasets is necessary to understand the underlying uncertainty in the observational data. Detailed descriptions of these datasets are as follows.

NARR: The North American Regional Reanalysis (NARR) is the NCEP (National Centers for Environmental Prediction) high-resolution reanalysis product that provides dynamically downscaled data over North America at ~ 32 km resolution and 3-hourly intervals from 1979 through present [Mesinger et al., 2006]. We note that some inaccuracies have also been identified in NARR, particularly in precipitation fields [Bukovsky and Karoly, 2007].

NCEP CPC: This dataset provides gauge-based analysis of daily precipitation from the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC). It is a suite of unified precipitation products obtained by combining all information available at CPC via the optimal interpolation objective analysis technique. The gauge analysis covers the Conterminous United States with a fine-resolution at 0.25° from 1948-01-01 to 2006-12-31.

PRISM: The Parameter-elevation Regressions on Independent Slopes Model (PRISM) [Daly et al., 2008] supports a 4 km gridded dataset obtained by taking point measurements and applying a weighted regression scheme that accounts for many factors affecting the local climatology. The datasets include total precipitation and minimum/maximum, (derived) mean temperatures and dew-points. Monthly climatological variables are available for 1895 through 2014 from the PRISM Climate Group (Oregon State University, <http://prism.oregonstate.edu>, created 4 Feb 2004). Notably, PRISM is the United States

Department of Agriculture's official climatological dataset. PRISM is used as our primary reference dataset for model performance evaluation.

UW: The UW daily gridded meteorological data is obtained from the Surface Water Modeling group at the University of Washington [Maurer et al., 2002; Hamlet and Lettenmaier, 2005]. UW incorporates topographic corrections by forcing the long-term average precipitation to match that of the PRISM dataset. The temperature dataset is produced in a similar fashion as precipitation, but uses a simple 6.1 K/km lapse rate for topographic effect. The dataset is provided at 0.125° horizontal resolution covering the period 1949 to 2010.

Daymet: Daymet is an extremely high resolution (1 km) gridded dataset with daily outputs of total precipitation, humidity, and minimum/maximum temperature covering 1980 through 2013 [Thornton et al., 1997, 2014]. The dataset is produced using an algorithmic technique that ingests point station measurements in conjunction with a truncated Gaussian weighting filter. Some adjustments are made to account for topography. Daymet is available through the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC).

To assess differences in these data products, we have calculated the MSD values among PRISM, UW and Daymet for seasonally averaged JJA T_{max} , T_{min} and DJF Pr over the five divisions and tabulated these results in Table 2. Student's t-test is employed to determine significances of differences. For T_{max} and T_{min} , gridded observational datasets are different from each other over some divisions. The most pronounced divergences occur in the NC region, with MSD values reaching up to $\sim 4^\circ\text{C}$, although differences are also apparent for MR T_{min} . Clearly, UW and Daymet have a colder climatology than PRISM. NARR, as a reanalysis dataset, is different from the others over most divisions, with overall larger T_{min} and smaller T_{max} . For precipitation, essentially no significant differences are present, especially among PRISM, UW and Daymet. NARR and CPC (not shown) seem to have slightly lower precipitation values than others.

3. Results

3.1. Temperature

The mean JJA T_{max} , T_{min} and T_{avg} climatology over the simulation period, together with PRISM and NARR reference data, is plotted in Figure 4. UW and Daymet have not been plotted here since they are visually indistinguishable to PRISM everywhere except for NC, where UW and Daymet exhibit lower temperatures (see Table 2). Statistical measures over California are tabulated in Table 3. In general, all simulations have captured the spatial climate patterns exhibited by PRISM, with high spatial correlations (>0.95), especially for T_{max} and T_{avg} . Nonetheless, several clear biases (relative to PRISM) are present in these simulations, as discussed below.

- **T_{max} :** When compared with the reference datasets, VR-CESM showed a warm bias of about 2 to 3°C in T_{max} over much of the inland domain (CV and MR) and a 2 to 3°C cool bias along the coast, although the coastal bias is reduced by $\sim 0.5^\circ\text{C}$ at 14 km resolution. This is in contrast with WRF, which produced an overall colder climate everywhere except the CV. This bias is especially pronounced for the WRF 9 km simulation, which was approximately 3°C cooler than PRISM. T_{max} within the CV has been overestimated by all the simulations. This likely represents a systematic issue with high-resolution models with respect to California. Possible reasons for this overestimation are discussed at the end of this section.

- T_{min} : VR-CESM showed a strong warm bias in T_{min} (~ 3 to 4°C), with a particularly large overestimation over Nevada ($> 5^\circ\text{C}$). WRF also exhibited a warm bias, but of a much smaller magnitude (~ 2 to 3°C). However, the pattern of T_{min} presented in Figure 4 in both WRF simulations suggests a cooler interior to the CV and warmer perimeter, which is not supported by observations.

- T_{avg} : The warm bias of T_{min} and T_{max} by VR-CESM resulted in a similar overestimation of T_{avg} . For WRF, underestimation of T_{max} and overestimation of T_{min} led to an overall closer match to T_{avg} over most of the domain, but is indicative of a suppressed diurnal cycle.

Compared with the reference datasets over California, VR-CESM 0.125° produced the lowest RMSD values for T_{max} , whereas WRF had smallest RMSD for T_{min} . However, in both cases the RMSD was around 2°C . Notably, T_{min} from VR-CESM matched much more closely with NARR, although this is likely indicative of a related warm bias in NARR. In fact, closer examination of the differences among VR-CESM, WRF and NARR marine near-surface temperature patterns indicated that CESM and NARR have T_{min} values that are approximately 2°C larger than WRF. Since coastal near-surface temperature is strongly influenced by ocean SSTs, this difference is likely a key driver of the warm bias in CESM. The Delta breeze effect, which is associated with a sea breeze circulation that brings relatively cool and humid marine air into the interior CV from the San Francisco Bay area, was apparent in all runs. It is especially encouraging that VR-CESM generally performed as well as WRF, in comparison with reference datasets, even though VR-CESM was not constrained or nudged at the lateral boundaries of the high-resolution domain.

The spatial standard deviation of JJA T_{max} , T_{min} and T_{avg} from models and PRISM is presented in Figure 5. In PRISM, the CV had smaller variability than surrounding regions, although the difference is small ($\sim 0.2^\circ\text{C}$). Further, areas with rougher topography did exhibit somewhat higher variability than smoother locations. Interestingly, the higher resolution (0.125°) VR-CESM simulation also matched the spatial pattern and magnitude of standard deviation observed in PRISM, especially for T_{min} and T_{avg} . However, in WRF and VR-CESM 0.25°, the variability is largely consistent across different divisions, and the values are around 0.5 to 1.5°C for all of the datasets, except for the high Sierras in the WRF 9km simulation which showed enhanced variability ($\sim 2^\circ\text{C}$). Compared with reference datasets, the RMSD values of VR-CESM and WRF 27km are ~ 0.1 - 0.2°C , and ~ 0.2 - 0.3°C for WRF 9km.

The seasonal cycle of monthly mean T_{avg} in each division is shown in Figure 6 for simulations and for reference data from PRISM and NARR along with the associated 95% confidence interval. PRISM and NARR match closely almost everywhere except in the summer season of NC, SC and CV, indicative of underlying observational uncertainty. This difference is likely due to the discrepancy in assimilating the coastal cooling effect. In general, model results match closely with reference data with no larger than a 2°C absolute difference, with the largest errors occurring in the summer and winter seasons. Compared with PRISM, VR-CESM overpredicts summer T_{avg} in all divisions except NC and SC, and underpredicts winter T_{avg} in all divisions. This corresponds to a larger annual temperature range. WRF has better performance in preserving the monthly trend when compared with CESM, with about 1°C underestimation over all seasons. There is no clear improvement in the seasonal cycle across resolutions.

Variability in monthly average T_{avg} is expressed by the interannual standard deviation of monthly T_{avg} over the 26-year period and is plotted in Figure 7 for all divisions. Generally, standard deviation is between 1 to 2°C . Among all models, WRF 27km is closest to PRISM (when assessed

using RMSD over the 12 month period). WRF 9km is also relatively close to PRISM, but exhibits an unusual $\sim 1^\circ\text{C}$ increase in variability in January and February (statistically significant at the 0.05 level). VR-CESM exhibits a weaker correlation with PRISM in all divisions with enhanced variability in DJF and weakened variability in April and May at both resolutions, and in the fall season in the 0.125° simulation. Nonetheless, the standard deviation from all datasets is statistically equivalent to the reference (with $p > 0.05$), except for the WRF 9km results in January and February.

Due to the impact of summer heat waves, we now focus on T_{max} over JJA. In Figure 8, the frequency distribution of T_{max} using all JJA daily values at each gridpoint over 26 years is depicted for models and reference data from UW and Daymet. PRISM is not included since it only deviates from UW and Daymet in the coastal divisions (NC and SC). In these divisions PRISM is similar in character to UW but shifted several degrees towards warmer temperatures. Properties of the frequency distribution, including average, variability, skewness and Kurtosis are tabulated in Table 4. As exemplified by the similarity in the moments of the distribution, VR-CESM clearly captures the general distribution of T_{max} . Outside of the CV, skewness and kurtosis measures match closely between VR-CESM and the UW dataset. In the NC and SC, Daymet overestimates the frequency of very cold days leading to deviation in the moments from UW. Consistent with the observations in Figure 4, outside of the CV, WRF tends to be cooler in general and VR-CESM tends to be warmer. In NC and SC, all models more accurately capture the frequency of high T_{max} days than low T_{max} days. Enhanced frequency of cool T_{max} values appears to be the primary driver in overestimation of sample variance in these divisions. For both VR-CESM and WRF there is no apparent improvement in statistics at higher resolutions.

In the CV, models show a clear warm bias and underestimated skewness, associated with a long forward tail and temperatures approaching near 50°C . As discussed earlier, all models overestimate T_{max} over CV. In order to further assess the accuracy of the gridded observations, we examine the T_{max} data directly from recorded weather station measurements over the CV (obtained from Global Historical Climate Network, provided by the NOAA/NCDC, <http://www.ncdc.noaa.gov/>). The results validate that T_{max} values above 45°C are rare (although station observations suggest these days may be slightly more frequent than suggested by UW and Daymet). The warm bias associated with the aforementioned extreme hot days in both VR-CESM and WRF is likely correlated with overly dry summertime soil moisture, as discussed in *Caldwell et al.* [2009]. This could be caused by the lack of accurate land surface treatment in climate models – for example, *Bonfils and Lobell* [2007] found that irrigation in CV has significantly decreased summertime maximum temperatures, especially in heavily-irrigated areas. Other studies have also found the cooling effects of irrigation over CV, such as *Kueppers et al.* [2007].

3.2. Precipitation

California's Mediterranean climate is associated with heavy precipitation in winter months and drier conditions in summertime. Agricultural and urban water use in California thus depends on accumulation of wintertime precipitation, which accounts for approximately half of total annual average precipitation as we calculated.

The long-term average climatology of DJF and annual daily Pr over 26 years from simulations and reference datasets (including PRISM and NARR) is depicted in Figure 9. Other reference datasets are almost the same as PRISM. Statistical quantities over California are given in Table 5. We can see that precipitation is heavily influenced

by orography, leading to most accumulation occurring along the NC and MR. As with temperature, the model results match the spatial patterns of the PRISM, with high correlation coefficients (>0.94).

For DJF Pr, especially along the western edge of the Sierra Nevada and into the CV, VR-CESM overestimates total precipitation ($\sim 25\%$ - 35%) relative to PRISM (see MRD in Table 5), particularly for the coarser resolution (28 km) simulation. This difference is statistically significant over the western edge of the Sierra Nevada compared to PRISM at the 95% level for VR-CESM 0.25°. VR-CESM 0.125° performs better and produces far more realistic (and less scale sensitive) precipitation over the Sierra Nevada with improved treatment of orographic effects. On the other hand, precipitation is slightly underestimated relative to PRISM along the NC (with a statistically significant difference), particularly near the Oregon border. There are also notable differences between WRF 27km and WRF 9km that suggest a strong resolution dependence in the underlying microphysics. For DJF Pr, WRF 27km underestimates precipitation along the NC (by about 20%-30%), but fairly accurately captures precipitation in the CV; whereas WRF 9km greatly overestimates precipitation (by about 65%-85%) along the NC and MR (see MRD in Table 5). Using Table 5 as a guide, VR-CESM 0.125° performs better than VR-CESM 0.25° and WRF 27km with RMSD values around 1.2 mm/day over DJF. Since we expect most of this improvement is due to a better representation of topography at 14km, this result suggests that the default physical parameterization suite in CESM is fairly resolution insensitive. WRF 9km is significantly different from PRISM over the MR and part of NC, and the potential reasons are discussed at the end of this section.

Interannual variability of precipitation was calculated for the models and PRISM using the standard deviation of annual and DJF precipitation and depicted in Figure 10. In general, precipitation variability exhibits a similar pattern to the precipitation intensity. The spatial pattern of variability agrees well between models and PRISM, with the closest match provided by VR-CESM 0.125° and WRF 27km. Standard deviation is $\sim 50\%$ higher for WRF 9km, consistent with overestimated precipitation intensity. VR-CESM 0.25° also tends to overestimate variability in the southern Sierra Nevada, likely due to over enhanced orographic uplift from the relatively coarse topography (relative to 0.125°). Comparing with all the gridded observations, RMSD values are ~ 0.7 - 0.9 mm/day for VR-CESM, ~ 0.5 - 0.7 mm/day for WRF 27km, and ~ 1.7 - 2.0 mm/day for WRF 9km.

The annual cycle of precipitation averaged over each month and region for the models and reference datasets (taking PRISM and NARR as representative of all datasets) is presented in Figure 11. The 95% confidence intervals of UW and PRISM are also depicted; differences between models and reference datasets are statistically significant when simulation results appear outside of the highlighted region. In general, the overall monthly climatology is consistent between models and reference datasets, with highest precipitation values occurring over winter and lowest values over summer. Nonetheless, the largest deviations occur during the winter season. WRF 27km is drier than PRISM and UW with relative differences ranging from $\sim 10\%$ - 40% , whereas WRF 9km is far wetter with relative differences reaching up to 40%-80% over these five divisions. VR-CESM tracks well with observed precipitation with $\sim 10\%$ - 20% relative difference everywhere except in the CV, where precipitation is overestimated in the rainy seasons by about 70%-80%. From the MWW test, VR-CESM and WRF 27km are not significantly different from reference datasets in most divisions, except over the CV in late winter to spring for VR-CESM 0.25°, and the NC winter and spring, and DR's winter for WRF 27km. The magnitude of precipitation

in WRF 9km is significantly different from the reference datasets over most divisions, except DR and SC's winter and spring. Nonetheless, the strong seasonal dependence on precipitation is apparent with extremely dry conditions during summer months. A slight increase in summertime precipitation is apparent in the DR, indicating the North American monsoon. We also observe that the peak month for precipitation tends to occur earlier in VR-CESM, particularly at 0.125°, compared with the reference. VR-CESM also exhibits some unexpected jaggedness (particularly December for VR-CESM 0.25° and February for VR-CESM 0.125°), likely due to an issue with capturing the seasonality of moisture transport over the Pacific. This issue being driven by variability outside of the high resolution domain seems corroborated by the observation that WRF correlates strongly with the reference datasets (even though the reported magnitude is incorrect).

The monthly cycle of sample standard deviation is depicted in Figure 12. The variability in observations has a similar monthly trend as precipitation rate, with overall values from 0 to 4 mm/day. Generally, higher interannual variability occurs over locations with higher mean precipitation (see Figure 11), also observed by previous studies (for example, *Duffy et al.* [2006]). Compared with observations, VR-CESM and WRF27km exhibited no more than 1 mm/day larger variability in the rainy season except over the CV. WRF 9km again showed enhanced variability (~ 1.5 mm/day more) during the wintertime over most divisions. The main cause of the interannual variability of precipitation over California is the El Niño-Southern Oscillation (ENSO), which varies the amount of moisture flux transported to this region. To assess deviations from observations, the Levene test was used in place of the F-test due to non-normality of precipitation. Under this test WRF 9km was significantly different from both reference datasets over summer and winter periods in MR, and over July to September in DR and SC. VR-CESM 0.25° was significantly different within the CV over the winter season.

The frequency distribution of DJF Pr has been constructed from rainy days ($Pr \geq 0.1$ mm/day) for the simulations and reference datasets and depicted in Figure 13. Since the frequency of precipitation is very similar across all reference datasets, only UW and CPC are included. Generally, VR-CESM matches closely with observations everywhere except in the CV. In the CV, WRF 27km appears to better capture high-intensity precipitation events, but performs poorly on low-intensity events ($Pr < 20$ mm/day). The underestimation of rainfall frequency in WRF 27km appears consistent across divisions. WRF 9km produces a significantly better treatment of low-intensity events, but greatly overestimates the frequency of high-intensity events ($Pr > 20$ mm/day). For strong precipitation events, VR-CESM matches closely to observations everywhere except the CV.

The overestimation of precipitation for WRF at high resolution has also been found in previous studies. Although not as pronounced as WRF 9km here, *Caldwell et al.* [2009] demonstrated that WRF at 12km largely overestimated the precipitation over California's mountainous regions (however, this paper did employ a different set of parameterizations and had a different spatial extent of mountain region). Further discussion can be found in former studies that employ different microphysics schemes (and so produce a wide range of precipitation magnitudes) [*Jankov et al.*, 2005; *Chin et al.*, 2010; *Caldwell et al.* [2009] also argued that the bias comes from a variety of sources, rather than simply different choices of sub-grid scale parameterizations. The exact cause of this overprediction has yet to be identified in the literature and a comprehensive analysis of the cause of these errors is beyond the scope of this paper.

A concise summary of model performance for JJA near-surface temperature and DJF Pr over California is provided by the Taylor diagram (Figure 14). This diagram includes the spatially centered correlation between the simulated and observed fields, the RMS variability of simulations normalized by that in the observations, and mean differences from a reference data [Taylor, 2001]. In general, all models correlate well with PRISM (used as the reference dataset in this diagram). As previously mentioned, WRF is better at modeling T_{min} than VR-CESM; however, VR-CESM has smaller deviation for T_{max} . WRF 9km performs badly for precipitation with nearly a two times larger normalized standard deviation than other simulations. Overall, VR-CESM provides a competitive representation of the regional climatology over California and has errors which are comparable to WRF. Comparing resolutions, there is only minor improvement in VR-CESM 0.125° relative to 0.25° in simulating T_{max} and Pr.

4. Discussions and summary

The need for high-resolution model data to address regional climate change and extreme events has motivated the development of new modeling tools. Our study investigated the use of a variable-resolution GCM (i.e., VR-CESM) as an alternative approach for two-way dynamically downscaled climate modeling. The performance of VR-CESM was evaluated for modeling California's unique regional climate. This relatively new technique has been evaluated against gridded observational datasets, reanalysis data and the WRF model (forced with ERA-Interim data at lateral boundaries).

Based on 26 years of high-resolution historical climate simulations (1980-2005), we analyzed the mean climatology of California across its climate divisions in terms of both near-surface temperature and precipitation. Generally, when compared with gridded observational datasets, both VR-CESM and WRF adequately represented regional climatological patterns with high spatial correlations (>0.94). Uncertainty between reference datasets exists, and is statistically significant over some climate divisions, making it necessary to utilize more than one high-quality observational products in the model evaluation. Overall, we found that VR-CESM showed comparable performance to WRF for regional climate modeling at spatial resolutions of 10-30 km.

Simulated temperature was assessed in terms of mean climatology of T_{min} , T_{max} and T_{avg} and interannual monthly-averaged variability of T_{avg} . Deviations between the models and the reference datasets are apparent, but their character is different between VR-CESM and WRF. During the summer period, VR-CESM produced a 2 to 3 °C warmer climate than observations, especially in the CV. On the other hand, WRF exhibited a colder (~ 2 °C) T_{max} over most divisions (except the CV), but was only a little warmer in T_{min} . Overall, VR-CESM was more accurate in reproducing mean climatology of T_{max} , whereas WRF was better at modeling T_{min} and T_{avg} . WRF modeled the annual cycle of T_{avg} better than VR-CESM with about a 1 °C overall underestimation. VR-CESM overestimated T_{avg} by 2 °C over the summer season and underestimated T_{avg} by 2 °C over the winter season, indicating a larger annual temperature range over most divisions. Higher resolution (0.125°) VR-CESM captures the spatial pattern of annual variability for near-surface temperature pattern shown in PRISM. Both WRF and VR-CESM well represent variability in monthly average T_{avg} over each climate division, except for the WRF 9km in January and February.

Temperatures were also further investigated in terms of the climatology of JJA T_{max} , due to its relevance to summertime heat waves. Both models successfully simulated the spatial character of JJA T_{max} , although both also had

an apparent warm bias over the CV. The failure to correctly capture CV T_{max} is likely caused in part by the lack of irrigation cooling over this division in both models. Future work will address this issue by applying irrigation model to VR-CESM so as to figure out the role irrigation plays in regulating T_{max} and its frequency distribution. Notably, the high-resolution (0.125°) VR-CESM simulation accurately captured the spatial pattern of interannual variability for near-surface temperature when compared with PRISM. Both WRF and VR-CESM represented variability in monthly average T_{avg} well over each climate division, except for the WRF 9km simulation in January and February where variability was greatly overestimated.

Precipitation was assessed in terms of mean climatology, interannual monthly-averaged variability and frequency of precipitation intensity. In general, VR-CESM matched closely with PRISM everywhere except for an overestimation of DJF Pr (about 25%-35%) along the western flank of the Sierra Nevada and into the CV. Increasing the spatial resolution to 0.125° produced some reduction in this overestimation (about 10%) likely due to improved treatment of orographic effects. WRF 27km underestimated DJF precipitation (by about 20%-30%) along the NC and MR (where almost all the precipitation appears), whereas WRF 9km showed a large overestimation (about 65%-85%). The standard deviation of precipitation ranged from 0 to 6 mm/day, with generally higher interannual variability over locations of higher mean precipitation. When assessing the frequency of strong precipitation events, VR-CESM matched closely to the UW dataset everywhere except the CV.

Higher resolution (0.125°) VR-CESM did produce better results when assessing JJA T_{max} and precipitation (along with their variability), compared with the coarser resolution run. However, the improvements are not statistically significant over most of the study area. The largest improvement at higher resolution was in the spatial character of precipitation, driven primarily with a better representation of the underlying topography. Notably, this result highlights the relative insensitivity to resolution in VR-CESM's physical parameterizations. This may be an advantageous result for multi-scale modelers interested in climate applications. Correctly simulating precipitation is vital to properly representing snowpack, which is of critical importance to water availability in the western United States [Bales et al., 2006; Wise, 2012; Rhoades et al., 2015]. Decreased scale sensitivity implies the result will be more independent of the choice of grid resolution.

For WRF, when resolution is increased to 9km, the model produces vastly overestimated precipitation, as previous studies have also found when using RCMs for fine-scale regional simulations. Although the convective parameterization was not disabled (as is suggested for some models below 10km resolution), the effect of this change would likely not modify the results since almost all of the precipitation comes from resolved (large-scale) condensation (not shown). In this sense, precipitation modeling bias of WRF is more strongly related with resolved-scale processes and the choice of microphysics scheme plays a major role, motivating the need for more work on scale-aware parameterizations [O'Brien et al., 2013].

In summary, VR-CESM demonstrated competitive utility for studying high-resolution regional climatology when compared to a regional climate model (WRF). Deviations exhibited within these models are not indicative of deep underlying problems with the model formulation, but one should nonetheless be aware of these biases when using these models for climate studies. This study suggests that VRGCMs are, in general, useful tools for assessing climate change over the coming century. As the need for assessments of regional climate change increases, alternative modeling strategies, including VRGCMs will be needed to improve our understanding of the effects of fine-scale processes representation in regional climate regulation. Future work will focus on the capability of the variable resolution system to correctly capture the features of discrete, extreme heat and precipitation events.

Acknowledgments.

The authors would like to thank Dr. Travis O'Brien for the many useful conversations on climate model assessment. We also want to thank Xue Meng Chen for her assistance in performing the WRF simulations. We acknowledge the work done to create the following datasets used in this study including PRISM, UW, Daymet, NARR and CPC (provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>). The simulation data used is available by request at xyhuang@ucdavis.edu. This project is supported in part by the University of California, Davis and by the Department of Energy "Multiscale Methods for Accurate, Efficient, and Scale-Aware Models of the Earth System" program.

References

- Abatzoglou, J. T., K. T. Redmond, and L. M. Edwards (2009), Classification of regional climate variability in the state of California, *Journal of Applied Meteorology and Climatology*, 48(8), 1527–1541.
- Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, R. Rice, and J. Dozier (2006), Mountain hydrology of the western united states, *Water Resources Research*, 42(8).
- Bonfils, C., and D. Lobell (2007), Empirical evidence for a recent slowdown in irrigation-induced cooling, *Proceedings of the National Academy of Sciences*, 104(34), 13,582–13,587.
- Bukovsky, M. S., and D. J. Karoly (2007), A brief evaluation of precipitation from the North American Regional Reanalysis, *Journal of Hydrometeorology*, 8(4), 837–846.
- Bukovsky, M. S., and D. J. Karoly (2009), Precipitation simulations using WRF as a nested regional climate model, *Journal of Applied Meteorology and Climatology*, 48(10), 2152–2159.
- Caldwell, P. (2010), California wintertime precipitation bias in regional and global climate models, *Journal of Applied Meteorology and Climatology*, 49(10), 2147–2158.
- Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala (2009), Evaluation of a WRF dynamical downscaling simulation over California, *Climatic Change*, 95(3-4), 499–521.
- Cayan, D. R., A. L. Luers, G. Franco, M. Hanemann, B. Croes, and E. Vine (2008), Overview of the California climate change scenarios project, *Climatic Change*, 87(1), 1–6.
- Chen, F., and J. Dudhia (2001), Coupling an advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity, *Monthly Weather Review*, 129(4), 569–585.
- Chin, H.-N. S., P. M. Caldwell, and D. C. Bader (2010), Preliminary study of California wintertime model wet bias, *Monthly Weather Review*, 138(9), 3556–3571.
- Christensen, J. H., B. Hewitson, A. Busuioc, A. Chen, X. Gao, R. Held, R. Jones, R. K. Kolli, W. Kwon, R. Laprise, et al. (2007), Regional climate projections, *Climate Change, 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, University Press, Cambridge, Chapter 11, pp. 847–940.
- Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, J. T. Kiehl, B. Briegleb, C. Bitz, S. Lin, et al. (2004), Description of the NCAR community atmosphere model (CAM 3.0), *Tech. rep.*, Tech. Rep. NCAR/TN-464+STR.
- Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P. Pasteris (2008), Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States, *International Journal of Climatology*, 28(15), 2031–2064.
- Dee, D., S. Uppala, A. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. Balmaseda, G. Balsamo, P. Bauer, et al. (2011), The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597.
- Dennis, J., J. Edwards, K. J. Evans, O. Guba, P. H. Lauritzen, A. A. Mirin, A. St-Cyr, M. A. Taylor, and P. H. Worley (2011), CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model, *International Journal of High Performance Computing Applications*, p. 1094342011428142.
- Dinse, K. (2009), Climate variability and climate change: what is the difference, *Book Climate Variability and Climate Change: What is the Difference*.
- Duffy, P., R. Arritt, J. Coquard, W. Gutowski, J. Han, J. Iorio, J. Kim, L.-R. Leung, J. Roads, and E. Zeledon (2006), Simulations of present and future climates in the western United States with four nested regional climate models, *Journal of Climate*, 19(6), 873–895.
- Fox-Rabinovitz, M., J. Côté, B. Dugas, M. Déqué, and J. L. McGregor (2006), Variable resolution general circulation models: Stretched-grid model intercomparison project (SGMIP), *Journal of Geophysical Research: Atmospheres (1984–2012)*, 111(D16).
- Fox-Rabinovitz, M. S., G. L. Stenchikov, M. J. Suarez, and L. L. Takacs (1997), A finite-difference GCM dynamical core with a variable-resolution stretched grid, *Monthly Weather Review*, 125(11), 2943–2968.
- Gates, W. L. (1992), AMIP: The Atmospheric Model Intercomparison Project, *Bulletin of the American Meteorological Society*, 73, 1962–1970.
- Guba, O., M. A. Taylor, P. A. Ullrich, J. R. Overvelt, and M. N. Levy (2014), The spectral element method on variable resolution grids: evaluating grid sensitivity and resolution-aware numerical viscosity, *Geoscientific Model Development*, 7, 2803–2816, doi:10.5194/gmd-7-2803-2014.
- Hamlet, A. F., and D. P. Lettenmaier (2005), Production of Temporally Consistent Gridded Precipitation and Temperature Fields for the Continental United States*, *Journal of Hydrometeorology*, 6(3), 330–336.
- Hayhoe, K., D. Cayan, C. B. Field, P. C. Frumhoff, E. P. Maurer, N. L. Miller, S. C. Moser, S. H. Schneider, K. N. Cahill, E. E. Cleland, et al. (2004), Emissions pathways, climate change, and impacts on California, *Proceedings of the National Academy of Sciences of the United States of America*, 101(34), 12,422–12,427.
- Hong, S.-Y., and J.-O. J. Lim (2006), The WRF single-moment 6-class microphysics scheme (WSM6), *Asia-Pacific Journal of Atmospheric Sciences*, 42(2), 129–151.
- Hong, S.-Y., Y. Noh, and J. Dudhia (2006), A new vertical diffusion package with an explicit treatment of entrainment processes, *Monthly Weather Review*, 134(9), 2318–2341.
- Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski (2008), A new sea surface temperature and sea ice boundary dataset for the Community Atmosphere Model, *Journal of Climate*, 21(19), 5145–5153.
- Hurrell, J. W., M. M. Holland, P. R. Gent, S. Ghan, J. E. Kay, P. Kushner, J.-F. Lamarque, W. G. Large, D. Lawrence, K. Lindsay, et al. (2013), The Community Earth System Model: A framework for collaborative research, *Bulletin of the American Meteorological Society*, 94(9), 1339–1360.
- Jankov, I., W. A. Gallus Jr, M. Segal, B. Shaw, and S. E. Koch (2005), The impact of different WRF model physical parameterizations and their interactions on warm season MCS rainfall, *Weather and Forecasting*, 20(6), 1048–1060.
- Kain, J. S. (2004), The Kain-Fritsch convective parameterization: An update, *Journal of Applied Meteorology*, 43(1), 170–181.
- Kanamitsu, M., and H. Kanamaru (2007), Fifty-seven-year California Reanalysis Downscaling at 10 km (CaRD10). Part I: System detail and validation with observations, *Journal of Climate*, 20(22), 5553–5571.
- Kueppers, L. M., M. A. Snyder, and L. C. Sloan (2007), Irrigation cooling effect: Regional climate forcing by land-use change, *Geophysical Research Letters*, 34(3).
- Laprise, R. (2008), Regional climate modelling, *Journal of Computational Physics*, 227(7), 3641–3666.
- Laprise, R., R. De Elia, D. Caya, S. Biner, P. Lucas-Picher, E. Diaconescu, M. Leduc, A. Alexandru, and L. Separovic (2008), Challenging some tenets of regional climate modelling, *Meteorology and Atmospheric Physics*, 100(1-4), 3–22.
- Leung, L. R., and Y. Qian (2009), Atmospheric rivers induced heavy precipitation and flooding in the western US simulated by the WRF regional climate model, *Geophysical Research Letters*, 36(3).
- Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby (2003), Regional climate research: needs and opportunities, *Bulletin of the American Meteorological Society*, 84(1), 89–95.

- Leung, L. R., Y. Qian, X. Bian, W. M. Washington, J. Han, and J. O. Roads (2004), Mid-century ensemble regional climate change scenarios for the western United States, *Climatic Change*, 62(1-3), 75–113.
- Lo, J. C.-F., Z.-L. Yang, and R. A. Pielke (2008), Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, *Journal of Geophysical Research: Atmospheres (1984–2012)*, 113(D9).
- Maurer, E., A. Wood, J. Adam, D. Lettenmaier, and B. Nijssen (2002), A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States*, *Journal of Climate*, 15(22), 3237–3251.
- McDonald, A. (2003), Transparent boundary conditions for the shallow-water equations: testing in a nested environment, *Monthly Weather Review*, 131(4), 698–705.
- McGregor, J. L., and M. R. Dix (2008), An updated description of the conforml-cubic atmospheric model, in *High resolution numerical modelling of the atmosphere and ocean*, pp. 51–75, Springer.
- Mearns, L. O., R. Arritt, S. Biner, M. S. Bukovsky, S. McGinnis, S. Sain, D. Caya, J. Correia Jr, D. Flory, W. Gutowski, et al. (2012), The North American regional climate change assessment program: overview of phase I results, *Bulletin of the American Meteorological Society*, 93(9), 1337–1362.
- Mesinger, F., and K. Veljovic (2013), Limited area NWP and regional climate modeling: a test of the relaxation vs Eta lateral boundary conditions, *Meteorology and Atmospheric Physics*, 119(1-2), 1–16.
- Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P. C. Shafran, W. Ebisuzaki, D. Jovic, J. Woollen, E. Rogers, E. H. Berbery, et al. (2006), North American regional reanalysis, *Bulletin of the American Meteorological Society*, 87, 343–360.
- Neale, R. B., C.-C. Chen, A. Gettelman, P. H. Lauritzen, S. Park, D. L. Williamson, A. J. Conley, R. Garcia, D. Kinnison, J.-F. Lamarque, et al. (2010a), Description of the NCAR community atmosphere model (CAM 5.0), *NCAR Tech. Note NCAR/TN-486+STR*.
- Neale, R. B., C.-C. Chen, A. Gettelman, P. H. Lauritzen, S. Park, D. L. Williamson, A. J. Conley, R. Garcia, D. Kinnison, J.-F. Lamarque, D. Marsh, M. Mills, A. K. Smith, S. Tilmes, F. Vitt, P. Cameron-Smith, W. D. Collins, M. J. Iacono, R. C. Easter, X. Liu, S. J. Ghan, P. J. Rasch, and M. A. Taylor (2010b), Description of the NCAR Community Atmosphere Model (CAM 5.0), *NCAR Technical Note NCAR/TN-486+STR*, National Center for Atmospheric Research, Boulder, Colorado.
- O'Brien, T. A., F. Li, W. D. Collins, S. A. Rauscher, T. D. Ringler, M. Taylor, S. M. Hagos, and L. R. Leung (2013), Observed scaling in clouds and precipitation and scale incognizance in regional to global atmospheric models, *Journal of Climate*, 26(23), 9313–9333.
- Oleson, K., D. Lawrence, G. Bonan, M. Flanner, E. Kluzek, P. Lawrence, S. Levis, S. Swenson, P. Thornton, A. Dai, M. Decker, R. Dickinson, J. Feddema, C. Heald, F. Hoffman, J. Lamarque, N. Mahowald, G. Niu, T. Qian, J. R. Anderson, S. Running, K. Sakaguchi, A. Slater, R. Stockli, A. Wang, Z. Yang, X. Zeng, and X. Zeng (2010), Technical description of version 4.0 of the Community Land Model (CLM), *NCAR Technical Note NCAR/TN-478+STR*, National Center for Atmospheric Research, Boulder, Colorado, doi:10.5065/D6FB50WZ.
- Pan, L.-L., S.-H. Chen, D. Cayan, M.-Y. Lin, Q. Hart, M.-H. Zhang, Y. Liu, and J. Wang (2011), Influences of climate change on California and Nevada regions revealed by a high-resolution dynamical downscaling study, *Climate Dynamics*, 37(9-10), 2005–2020.
- Pierce, D. W., T. Das, D. R. Cayan, E. P. Maurer, N. L. Miller, Y. Bao, M. Kanamitsu, K. Yoshimura, M. A. Snyder, L. C. Sloan, et al. (2013), Probabilistic estimates of future changes in California temperature and precipitation using statistical and dynamical downscaling, *Climate Dynamics*, 40(3-4), 839–856.
- Rauscher, S. A., E. Coppola, C. Piani, and F. Giorgi (2010), Resolution effects on regional climate model simulations of seasonal precipitation over Europe, *Climate Dynamics*, 35(4), 685–711.
- Rauscher, S. A., T. D. Ringler, W. C. Skamarock, and A. A. Mirin (2013), Exploring a global multiresolution modeling approach using aquaplanet simulations, *Journal of Climate*, 26(8), 2432–2452.
- Rhoades, A. M., X. Huang, P. A. Ullrich, and C. M. Zarzycki (2015), Characterizing Sierra Nevada Snowpack Using Variable-Resolution CESM, *J. Appl. Meteor. Clim.*, p. Under minor revision.
- Ringler, T., L. Ju, and M. Gunzburger (2008), A multiresolution method for climate system modeling: application of spherical centroidal Voronoi tessellations, *Ocean Dynamics*, 58(5-6), 475–498.
- Skamarock, W., J. Klemp, J. Dudhia, D. Gill, and D. Barker (2005), Coauthors, 2008: A Description of the Advanced Research WRF Version 3. NCAR Technical Note, *Tech. rep.*, NCAR/TN-475+STR.
- Skamarock, W. C., J. B. Klemp, M. G. Duda, L. D. Fowler, S.-H. Park, and T. D. Ringler (2012), A multiscale nonhydrostatic atmospheric model using centroidal Voronoi tessellations and C-grid staggering, *Monthly Weather Review*, 140(9), 3090–3105.
- Soares, P. M., R. M. Cardoso, P. M. Miranda, J. de Medeiros, M. Belo-Pereira, and F. Espirito-Santo (2012), WRF high resolution dynamical downscaling of ERA-Interim for Portugal, *Climate dynamics*, 39(9-10), 2497–2522.
- Solomon, S. (2007), *Climate change 2007-the physical science basis: Working group I contribution to the fourth assessment report of the IPCC*, vol. 4, Cambridge University Press.
- Staniforth, A. N., and H. L. Mitchell (1978), A variable-resolution finite-element technique for regional forecasting with the primitive equations, *Monthly Weather Review*, 106(4), 439–447.
- Stauffer, D. R., and N. L. Seaman (1990), Use of four-dimensional data assimilation in a limited-area mesoscale model. Part I: Experiments with synoptic-scale data, *Monthly Weather Review*, 118(6), 1250–1277.
- Sun, F., D. Walton, and A. Hall (2015), A Hybrid Dynamical-Statistical Downscaling Technique, Part II: End-of-Century Warming Projections Predict a New Climate State in the Los Angeles Region, *Journal of Climate*, (2015).
- Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres*, 106(D7), 7183–7192, doi:10.1029/2000JD900719.
- Taylor, M. A. (2011), Conservation of mass and energy for the moist atmospheric primitive equations on unstructured grids, in *Numerical Techniques for Global Atmospheric Models*, pp. 357–380, Springer.
- Taylor, M. A., and A. Fournier (2010), A compatible and conservative spectral element method on unstructured grids, *Journal of Computational Physics*, 229(17), 5879–5895.
- Thornton, P. E., S. W. Running, and M. A. White (1997), Generating surfaces of daily meteorological variables over large regions of complex terrain, *Journal of Hydrology*, 190(3), 214–251.
- Thornton, P. E., M. M. Thornton, B. W. Mayer, N. Wilhelm, Y. Wei, R. Devarakonda, and R. B. Cook (2014), Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2, *Tech. rep.*, Oak Ridge National Laboratory (ORNL).
- Ullrich, P. A. (2014), SquadGen: Spherical Quadrilateral Grid Generator, <http://climate.ucdavis.edu/squadgen.php>, accessed: 2015-02-11.
- Warner, T. T., R. A. Peterson, and R. E. Treadon (1997), A tutorial on lateral boundary conditions as a basic and potentially serious limitation to regional numerical weather prediction, *Bulletin of the American Meteorological Society*, 78(11), 2599–2617.
- Wehner, M. F., K. Reed, F. Li, Prabhat, J. Bacmeister, C.-T. Chen, C. Paciorek, P. Gleckler, K. Sperber, W. D. Collins, A. Gettelman, and C. Jablonowski (2014), The effect of horizontal resolution on simulation quality in the Community Atmosphere Model, CAM5.1, *J. Model. Earth. Sys.*, doi: 10.1002/2013MS000276.
- Wise, E. K. (2012), Hydroclimatology of the us intermountain west, *Progress in Physical Geography*, 36(4), 458–479.
- Zarzycki, C. M., M. N. Levy, C. Jablonowski, J. R. Overfelt, M. A. Taylor, and P. A. Ullrich (2014a), Aquaplanet experiments using CAM's variable-resolution dynamical core, *Journal of Climate*, 27(14), 5481–5503.
- Zarzycki, C. M., C. Jablonowski, and M. A. Taylor (2014b), Using Variable-Resolution Meshes to Model Tropical Cyclones in the Community Atmosphere Model, *Monthly Weather Review*, 142(3), 1221–1239.

Zarzycki, C. M., C. Jablonowski, D. R. Thatcher, and M. A. Taylor (2015), Effects of localized grid refinement on the general circulation and climatology in the Community Atmosphere Model, *Journal of Climate*, 28(7), 2777–2803.

Corresponding author: Xingying Huang, Department of Land, Air and Water Resources, University of California Davis, Davis, CA 95616, USA. (xy-huang@ucdavis.edu)

Table 1. Reanalysis and gridded observational datasets used in this study.

Data source	Variables used	Spatial resolution	Temporal resolution
NARR	Pr, T_s	32 km	daily, 3-hourly
NCEP CPC	Pr	~28 km (0.25°)	daily
UW	Pr, T_{min} , T_{max}	~14 km (0.125°)	daily
PRISM	Pr, T_{min} , T_{max} , T_{avg}	4 km	monthly
Daymet	Pr, T_{min} , T_{max}	1 km	daily

Table 2. MSD (left column minus top row) of JJA temperature (°C) and DJF precipitation (mm/day) between all reference datasets. Statistically significant differences are emphasized (95% confidence level).

JJA T_{min}	PRISM					UW					Daymet					
	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	
UW	-2.0	-0.5	-0.3	-0.4	-2.3											
Daymet	-1.8	0.2	-0.2	-0.3	-4.6	0.2	0.8	0.1	0.1	-2.3						
NARR	2.5	2.5	3.1	-0.2	2.0	4.5	3.1	3.4	0.2	4.2	4.3	2.3	3.2	0.1	6.5	
JJA T_{max}	PRISM					UW					Daymet					
	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	
UW	0.3	-0.1	-0.2	-0.7	-2.4											
Daymet	-0.3	0.4	-0.2	0.2	-3.7	-0.6	0.5	0.1	0.9	-1.3						
NARR	-1.3	1.2	-0.1	-0.6	-3.7	-1.7	1.2	0.1	0.1	-1.3	-1.1	0.7	0.1	-0.7	0.0	
DJF Pr	PRISM					UW					Daymet					
	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	MR	CV	DR	SC	NC	
UW	-0.2	0.1	0.0	-0.2	-0.2											
Daymet	-0.1	-0.4	-0.0	0.2	0.5	0.1	-0.5	-0.0	0.3	0.8						
NARR	-0.5	-0.4	-0.1	-0.1	-1.3	-0.3	-0.5	-0.1	0.1	-1.1	-0.4	-0.0	-0.1	-0.3	-1.8	
CPC	-0.8	0.0	0.0	-0.1	-1.1	-0.5	-0.1	0.0	0.1	-0.9	-0.6	0.4	0.0	-0.3	-1.7	

Table 3. RMSD (°C), MSD (°C) and Spatial Correlation (Corr) for seasonally-averaged daily JJA temperatures over California.

RMSD	UW		PRISM			Daymet	
	T_{max}	T_{min}	T_{max}	T_{min}	T_{avg}	T_{max}	T_{min}
VR-CESM 0.25°	2.32	3.75	2.92	3.12	2.60	2.81	3.93
VR-CESM 0.125°	1.90	3.63	2.45	2.94	2.18	2.48	3.70
WRF 27km	2.31	2.74	2.93	2.25	2.17	2.51	2.99
WRF 9km	3.32	2.94	3.49	1.84	1.77	3.20	2.94
Uniform CESM 1°	3.06	4.59	3.62	3.43	3.16	3.58	5.07
MSD	UW		PRISM			Daymet	
	T_{max}	T_{min}	T_{max}	T_{min}	T_{avg}	T_{max}	T_{min}
VR-CESM 0.25°	0.98	2.91	0.61	1.73	0.82	1.18	2.88
VR-CESM 0.125°	0.65	2.85	0.20	1.66	0.58	0.82	2.74
WRF 27km	-0.58	0.82	-0.95	-0.36	-0.77	-0.39	0.79
WRF 9km	-2.28	1.86	-2.72	0.67	-1.14	-2.10	1.76
Uniform CESM 1°	0.82	3.03	0.60	1.76	1.08	1.24	3.38
Corr	UW		PRISM			Daymet	
	T_{max}	T_{min}	T_{max}	T_{min}	T_{avg}	T_{max}	T_{min}
VR-CESM 0.25°	0.99	0.98	0.99	0.98	0.99	0.99	0.97
VR-CESM 0.125°	0.99	0.98	0.99	0.98	0.99	0.99	0.98
WRF 27km	0.99	0.98	0.99	0.98	0.99	0.99	0.97
WRF 9km	0.99	0.98	0.99	0.99	0.99	0.99	0.98
Uniform CESM 1°	0.99	0.96	0.99	0.97	0.99	0.99	0.95

Table 4. The first four moments of the JJA T_{max} frequency in each climate division. Column titles refer to the Average (Avg), Variance (Var), Skewness (Skew) and Kurtosis (Kurt).

	CV				MR			
	Avg	Var	Skew	Kurt	Avg	Var	Skew	Kurt
UW	32.6	24.8	-0.8	0.9	26.7	33.2	-0.4	0.3
Daymet	32.7	23.5	-0.9	1.5	25.9	39.3	-0.5	0.5
VR-CESM 0.25°	34.1	26.2	-0.4	0.2	28.1	27.6	-0.4	0.3
VR-CESM 0.125°	34.3	28.5	-0.5	0.4	27.2	30.0	-0.4	0.3
WRF 27km	33.9	34.8	-0.5	0.2	24.9	34.8	-0.3	0.0
WRF 9km	32.4	33.1	-0.7	0.6	22.4	38.5	-0.5	0.6

	NC				SC				DR			
	Avg	Var	Skew	Kurt	Avg	Var	Skew	Kurt	Avg	Var	Skew	Kurt
UW	25.9	30.4	0.1	-0.5	25.9	30.4	0.1	-0.5	37.0	22.9	-0.6	0.7
Daymet	26.5	30.1	-0.3	0.4	26.5	30.1	-0.3	0.4	37.0	24.3	-0.6	0.6
VR-CESM 0.25°	26.4	37.4	0.1	-0.7	26.4	37.4	0.1	-0.7	37.6	19.0	-0.5	0.8
VR-CESM 0.125°	26.3	37.4	0.1	-0.6	26.3	37.4	0.1	-0.6	37.3	21.3	-0.5	0.4
WRF 27km	26.0	36.7	-0.1	-0.5	26.0	36.7	-0.1	-0.5	36.5	22.6	-0.6	0.5
WRF 9km	24.9	32.6	0.0	-0.6	24.9	32.6	0.0	-0.6	34.4	24.4	-0.5	0.4

Notes: If skew > 0 [skew < 0], the distribution trails off to the right [left]. If kurtosis > 0 [< 0], a sharper [flatter] peak compared to a normal distribution (leptokurtic and platykurtic, respectively) is expected.

Table 5. RMSD (mm/day), MSD (mm/d), MRD, Spatial Correlation (Corr) for averaged precipitation over California

Annual	CPC				UW			
	RMSD	MSD	MRD	Corr	RMSD	MSD	MRD	Corr
VR-CESM 0.25°	0.61	0.39	0.30	0.98	0.62	0.29	0.29	0.96
VR-CESM 0.125°	0.47	0.21	0.24	0.98	0.53	0.12	0.24	0.97
WRF 27km	0.42	-0.21	0.21	0.97	0.58	-0.31	0.24	0.97
WRF 9km	2.23	1.49	0.97	0.95	2.05	1.39	0.85	0.96
Uniform CESM 1°	1.97	-1.57	0.99	0.94	2.31	-1.70	0.99	0.91

	PRISM				Daymet			
	RMSD	MSD	MRD	Corr	RMSD	MSD	MRD	Corr
VR-CESM 0.25°	0.72	0.20	0.31	0.95	0.57	0.19	0.25	0.97
VR-CESM 0.125°	0.62	0.05	0.26	0.96	0.50	0.03	0.22	0.97
WRF 27km	0.77	-0.40	0.27	0.96	0.65	-0.41	0.27	0.97
WRF 9km	1.89	1.32	0.78	0.97	2.01	1.31	0.76	0.96
Uniform CESM 1°	2.53	-1.83	0.99	0.90	2.31	-1.80	0.99	0.93

DJF	CPC				UW			
	RMSD	MSD	MRD	Corr	RMSD	MSD	MRD	Corr
VR-CESM 0.25°	1.49	0.99	0.36	0.97	1.45	0.67	0.33	0.95
VR-CESM 0.125°	1.19	0.64	0.29	0.97	1.23	0.35	0.27	0.96
WRF 27km	0.89	-0.38	0.21	0.97	1.29	-0.69	0.26	0.96
WRF 9km	4.26	2.61	0.86	0.95	3.84	2.32	0.70	0.95
Uniform CESM 1°	3.97	-3.12	0.99	0.93	4.80	-3.50	0.99	0.90

	PRISM				Daymet			
	RMSD	MSD	MRD	Corr	RMSD	MSD	MRD	Corr
VR-CESM 0.25°	1.65	0.58	0.35	0.94	1.35	0.51	0.28	0.96
VR-CESM 0.125°	1.40	0.29	0.29	0.95	1.17	0.21	0.25	0.96
WRF 27km	1.55	-0.79	0.28	0.96	1.35	-0.85	0.28	0.96
WRF 9km	3.57	2.26	0.66	0.96	3.80	2.18	0.65	0.95
Uniform CESM 1°	5.07	-3.65	0.99	0.90	4.69	-3.65	0.99	0.93

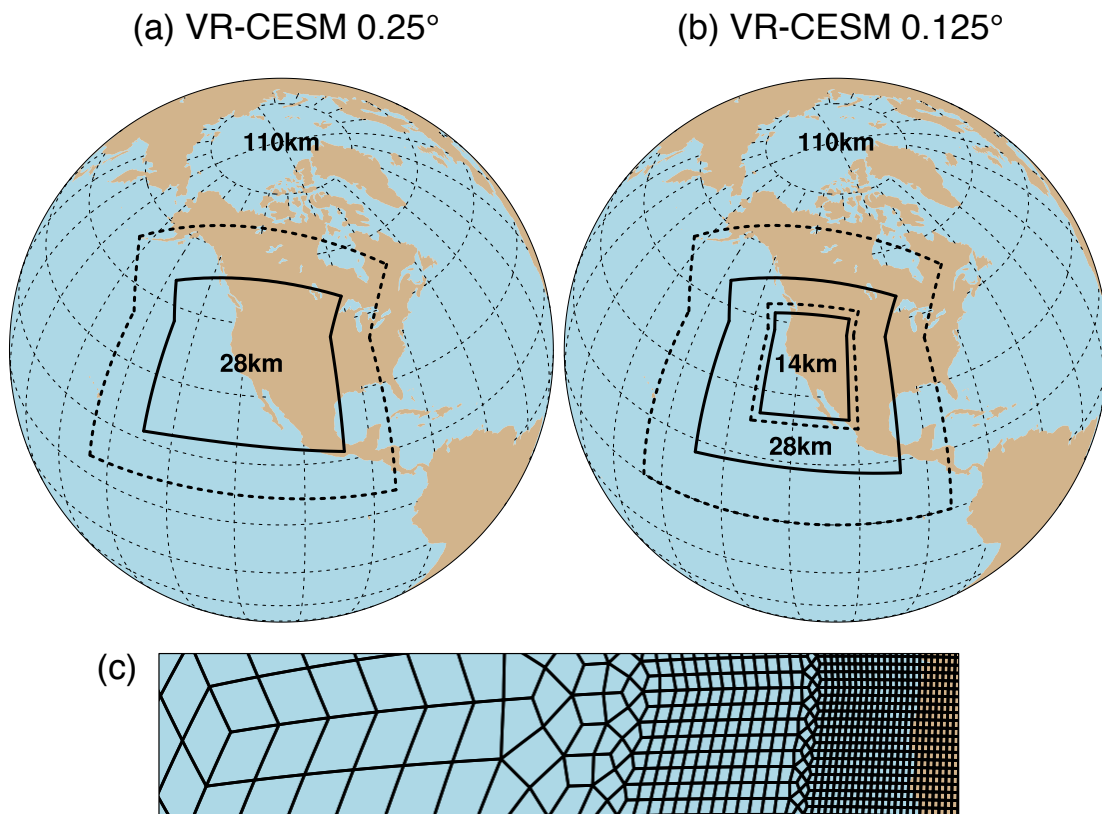


Figure 1. The approximate grid spacing in the (a) VR-CESM 0.25° and (b) VR-CESM 0.125° meshes used in this study. (c) A depiction of the transition from the global 1° resolution mesh through two layers of refinement to 0.25° and again to 0.125° .

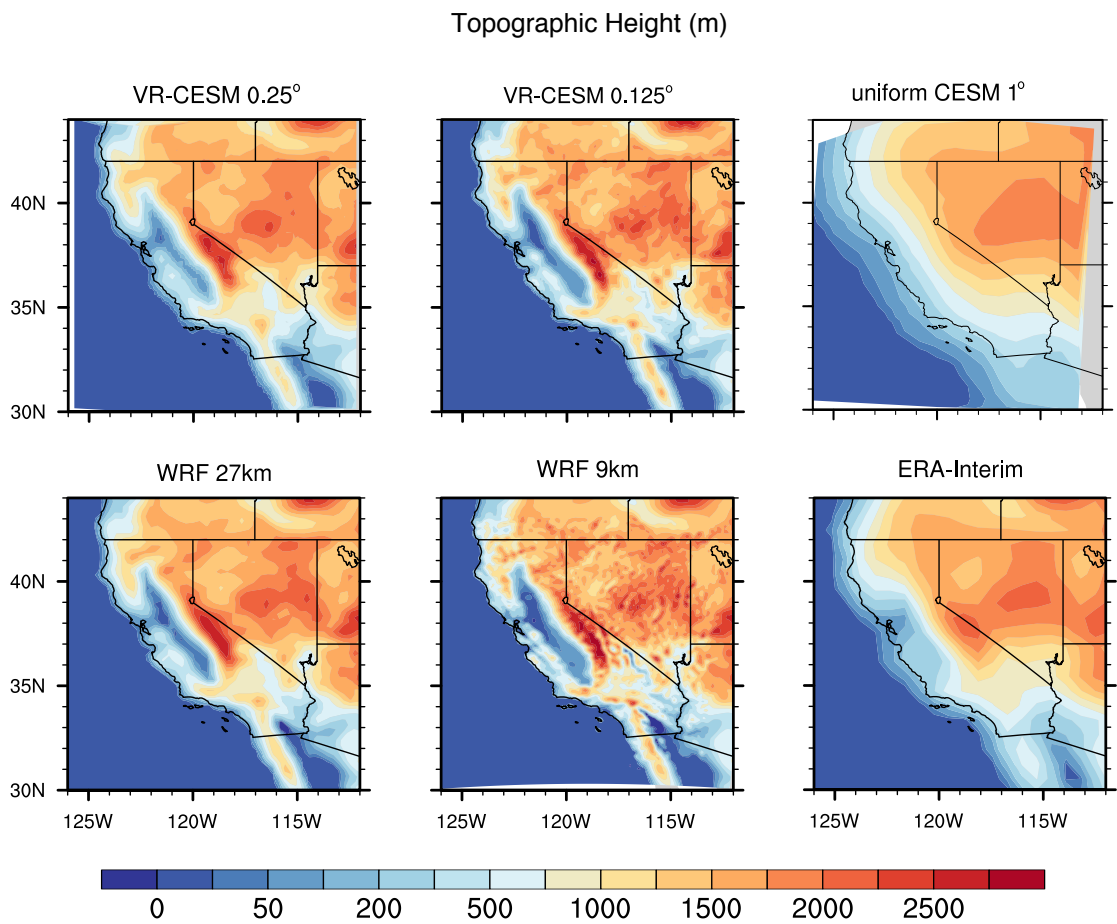


Figure 2. Topographic heights (from top left to bottom right) for VR-CESM 0.25°, VR-CESM 0.125°, uniform CEM 1°, WRF 27km, WRF 9km and ERA-Interim ($\Delta x \sim 80$ km).

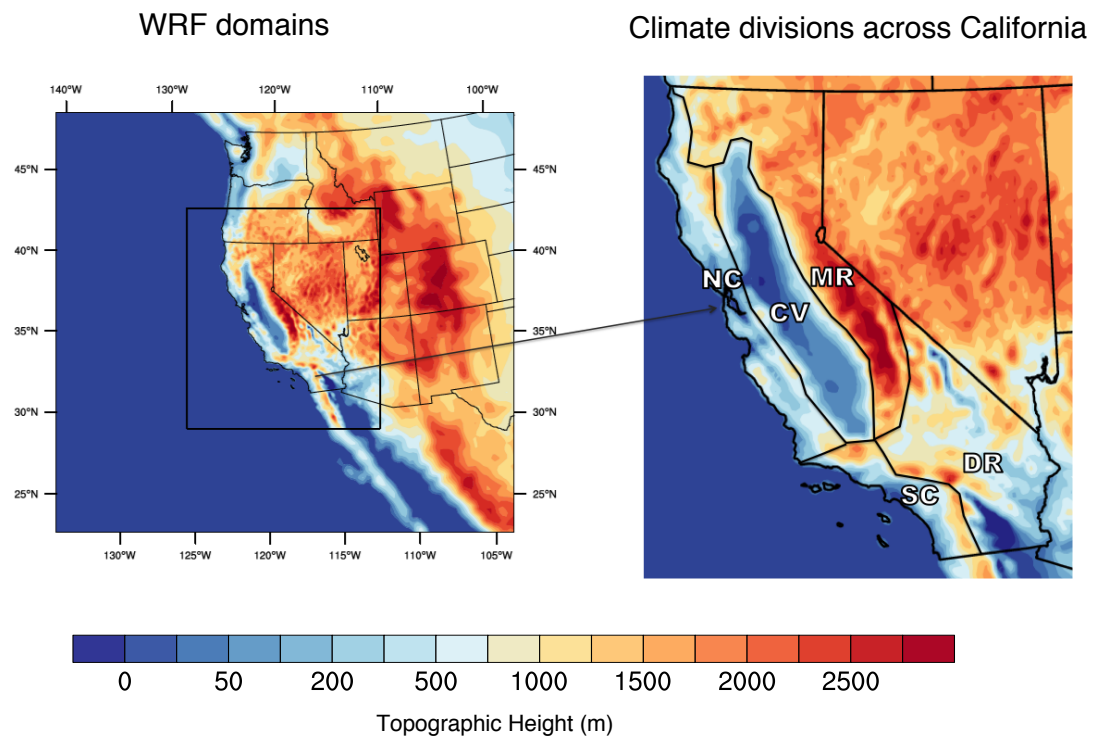


Figure 3. Left: WRF 27km (entire plot region) and WRF 9km (solid black box) simulation domains; Right: five climate divisions for California. Both plots are overlaid with WRF model topography.

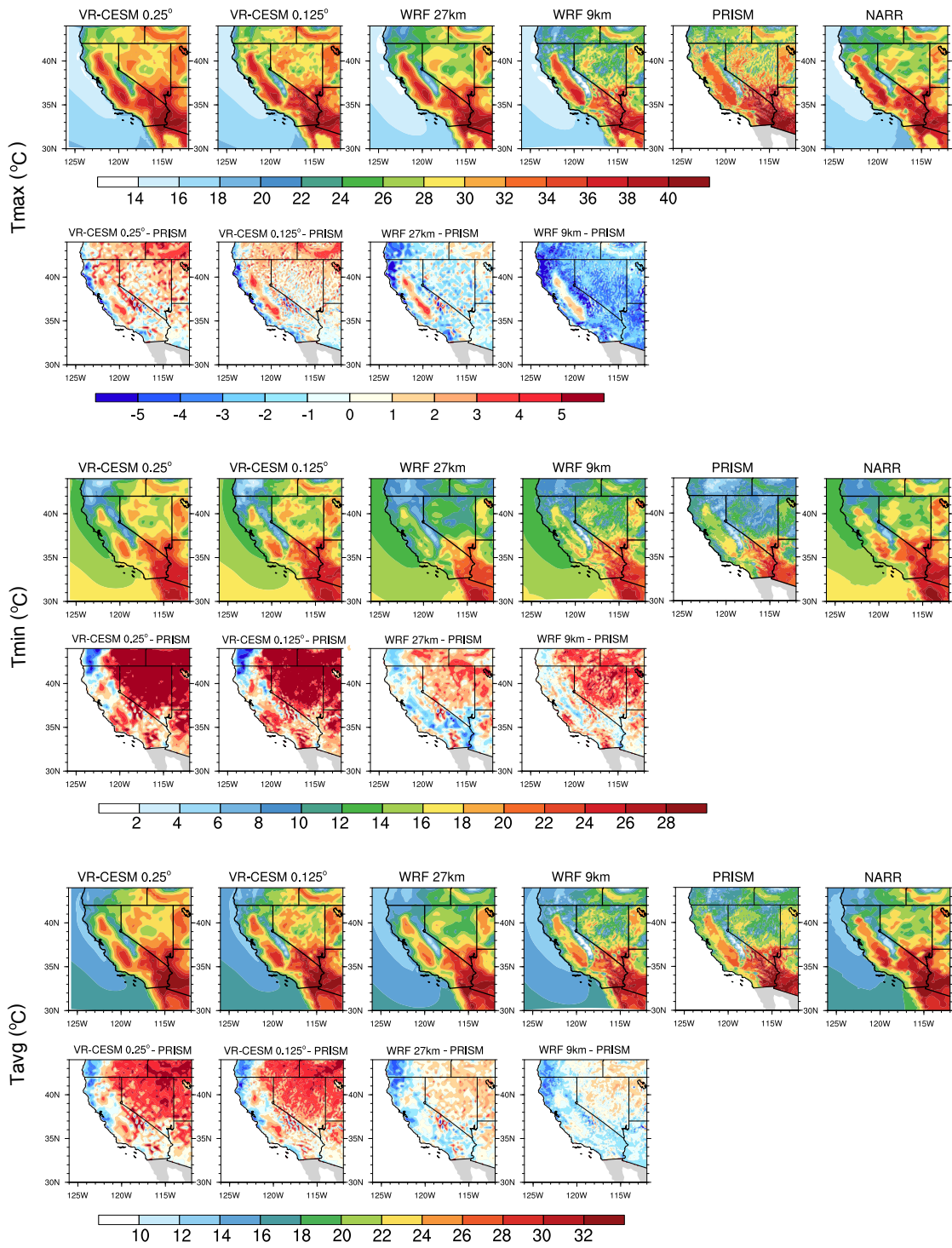


Figure 4. JJA averaged daily T_{max} , T_{min} and T_{avg} from models and reference datasets, and differences (sharing the same legend) between model results and PRISM.

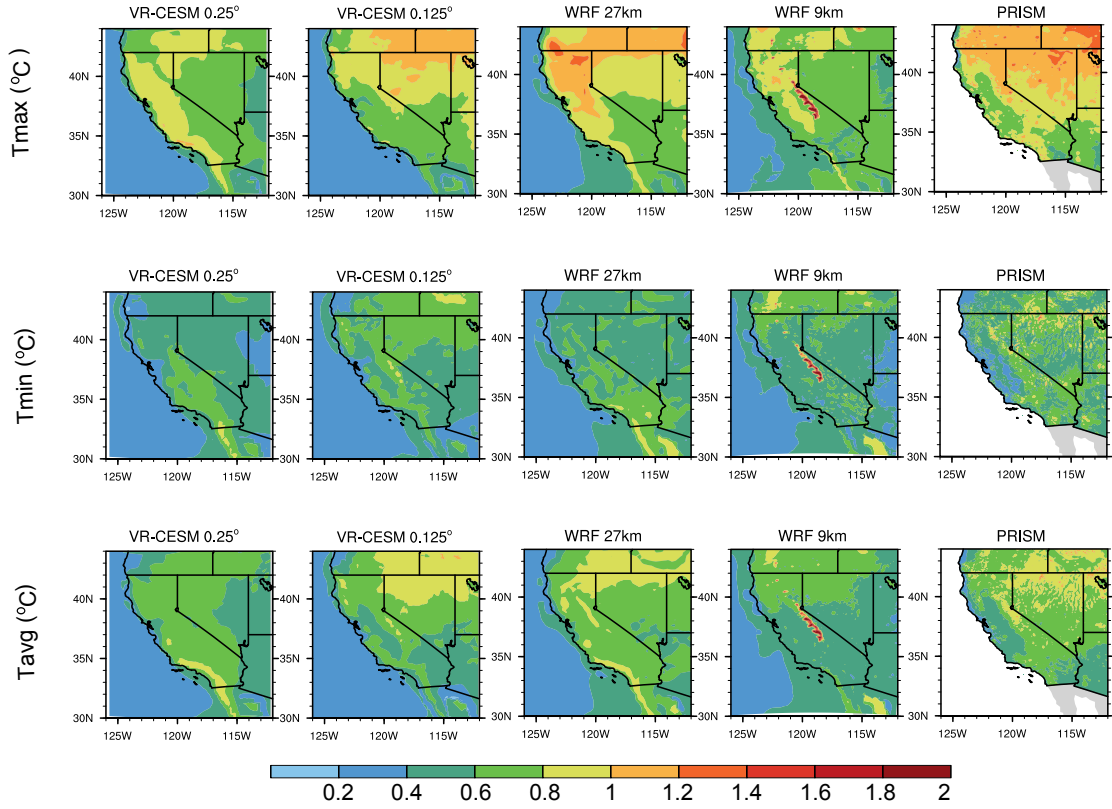


Figure 5. Sample standard deviation of JJA average daily T_{max} , T_{min} and T_{avg} from model results and PRISM.

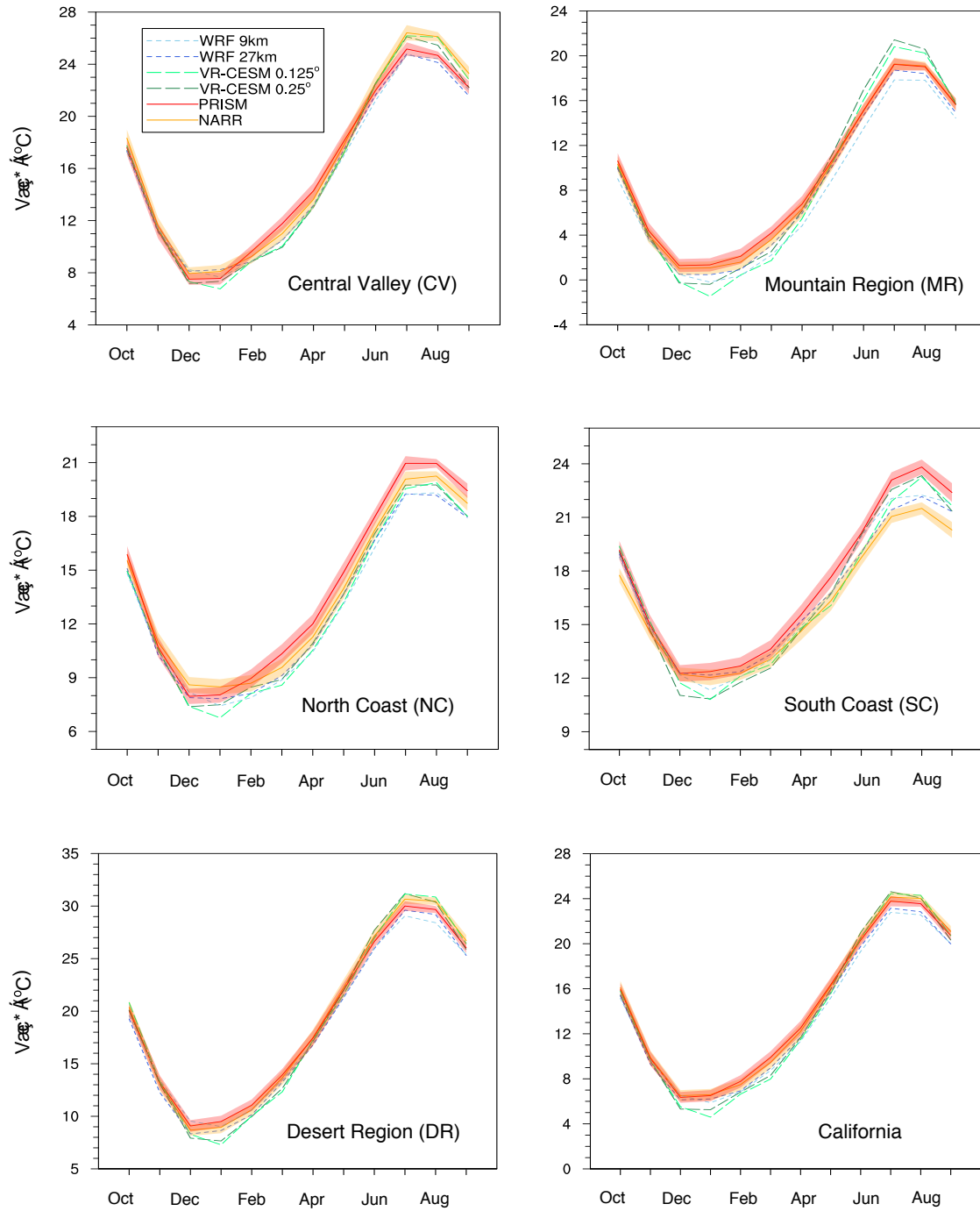


Figure 6. Seasonal cycle of monthly-average T_{avg} for each climate division. The shading corresponds to the 95% confidence interval of PRISM and NARR.

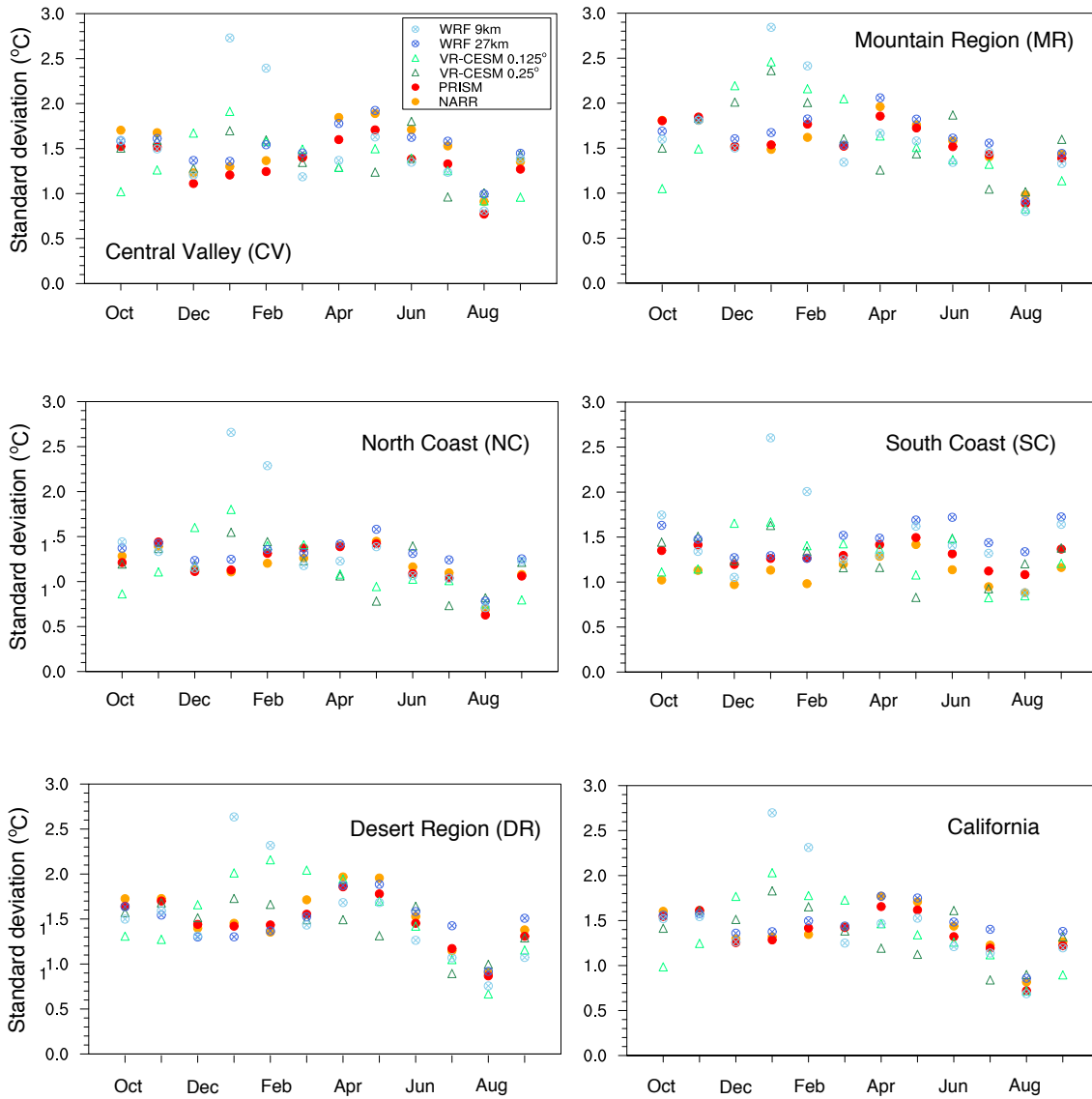


Figure 7. Standard deviation values of monthly-average T_{avg} for each climate division.

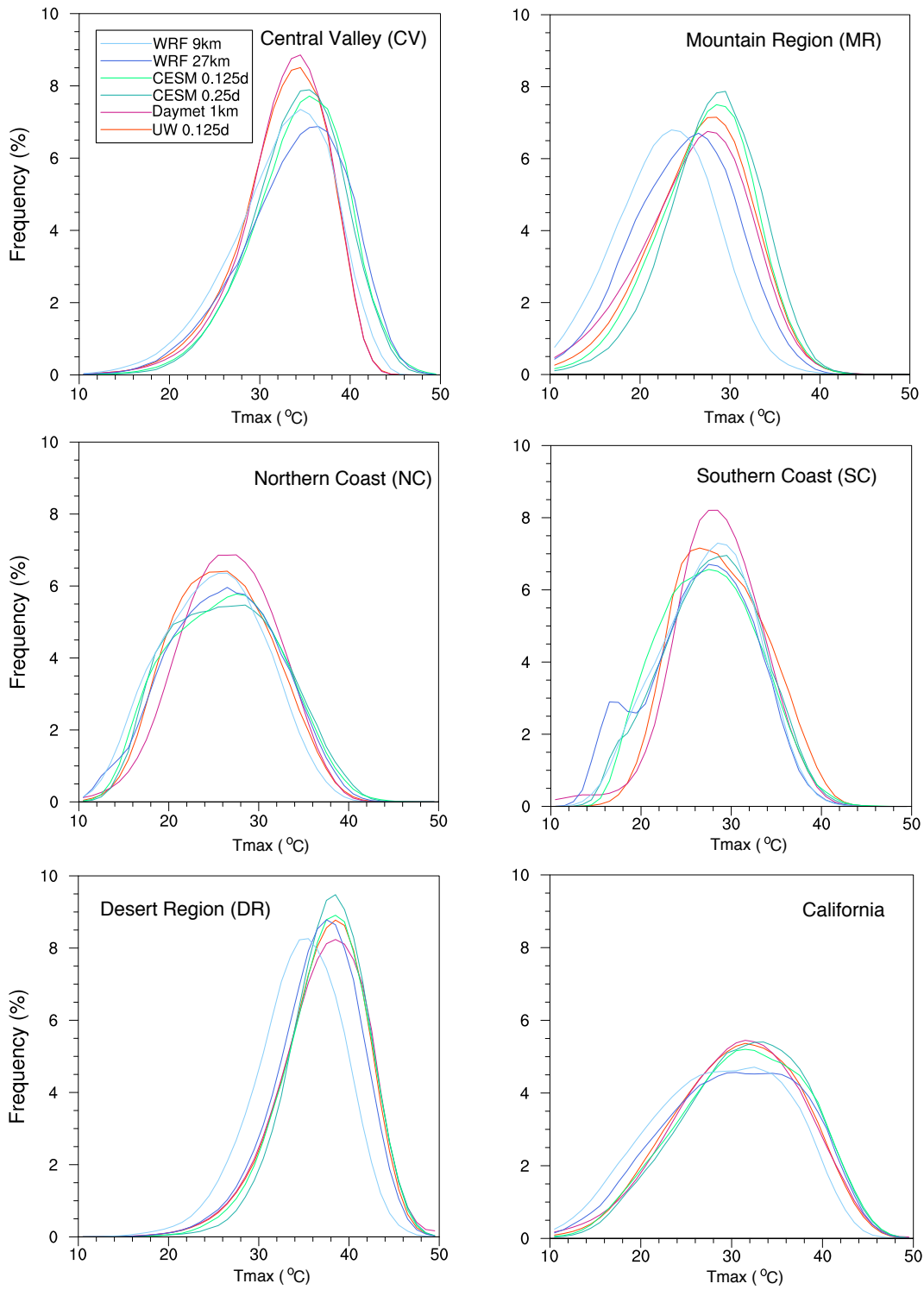


Figure 8. Frequency distribution of JJA daily T_{max} over the simulation period 1980-2005.

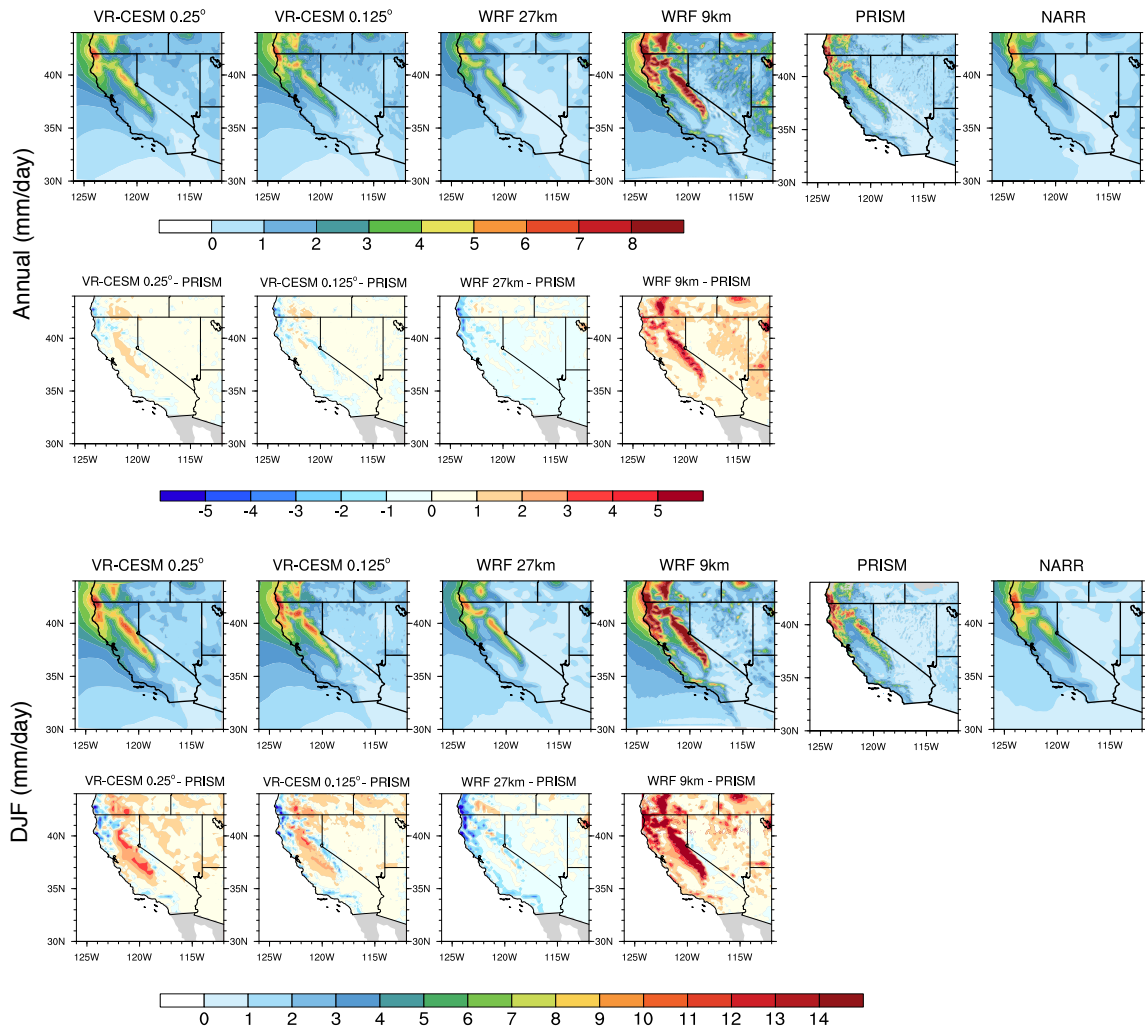


Figure 9. Annual and DJF precipitation from model results and reference datasets, and differences (sharing the same legend) between model results and PRISM.

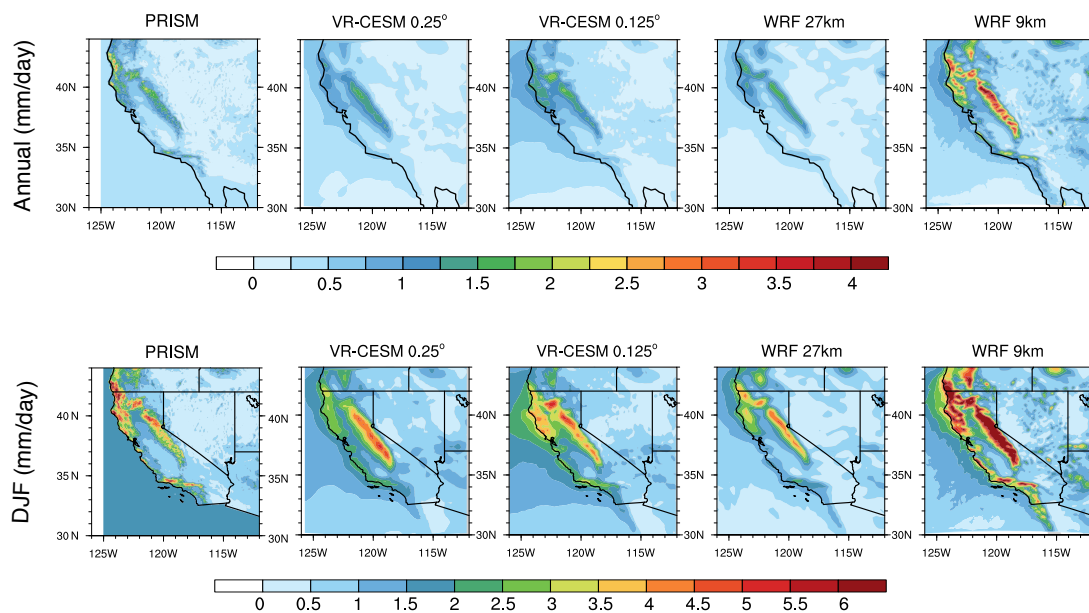


Figure 10. Sample standard deviation of Annual and DJF precipitation from models and PRISM.

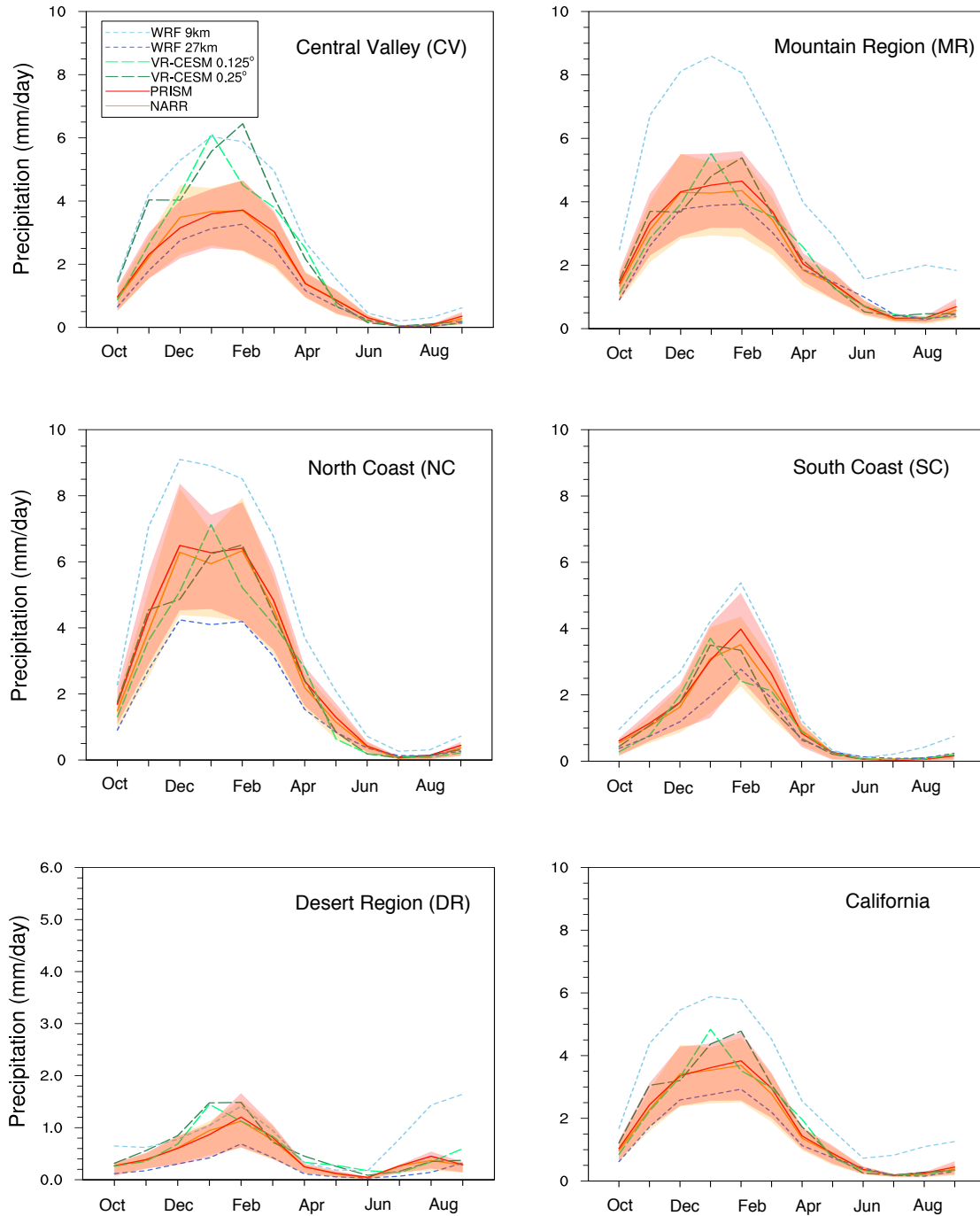


Figure 11. As Figure 6, but for monthly-average total precipitation. The shading refers to the 95% confidence interval of PRISM and UW.

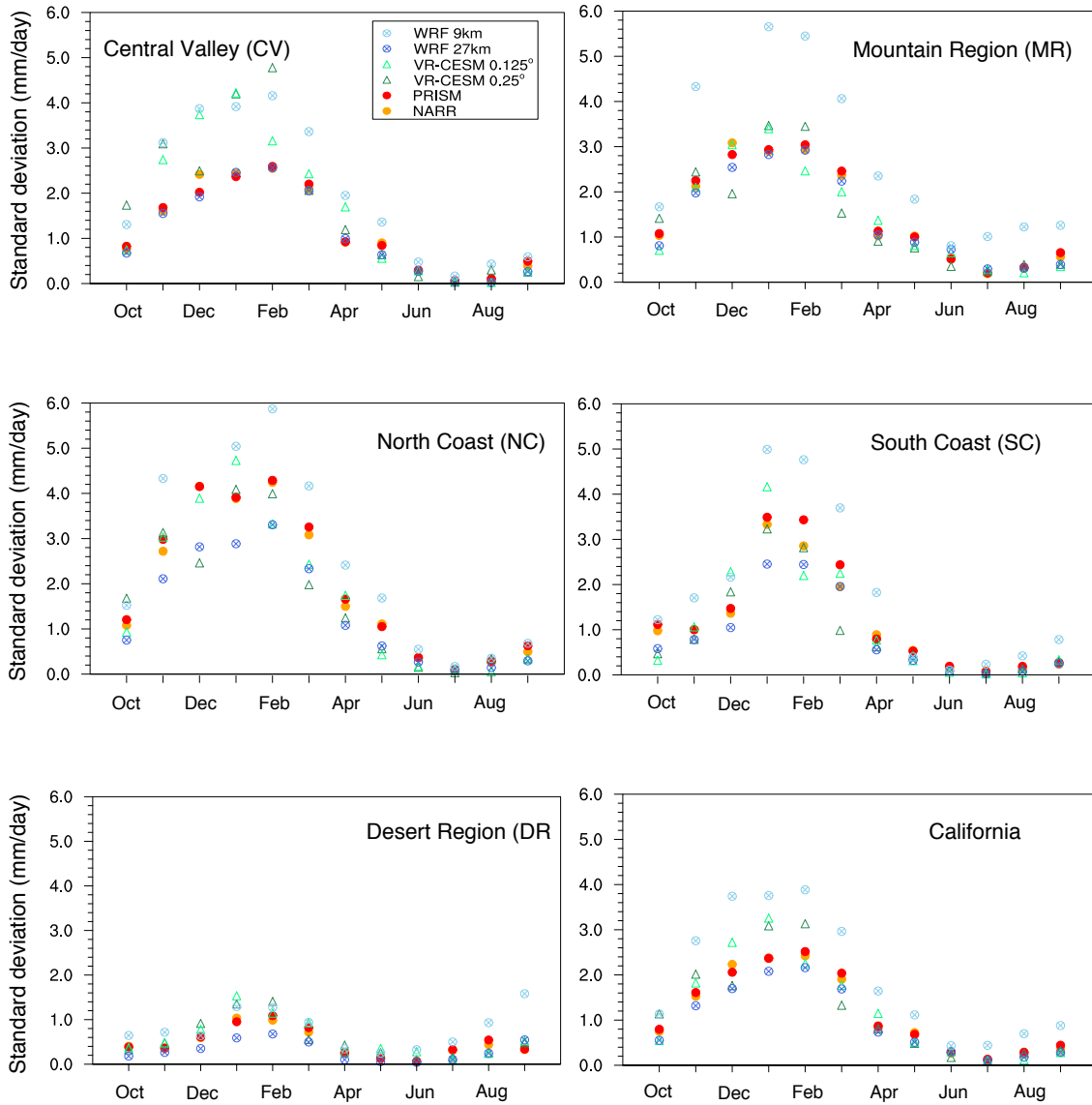


Figure 12. As Figure 7, but for monthly-average total precipitation.

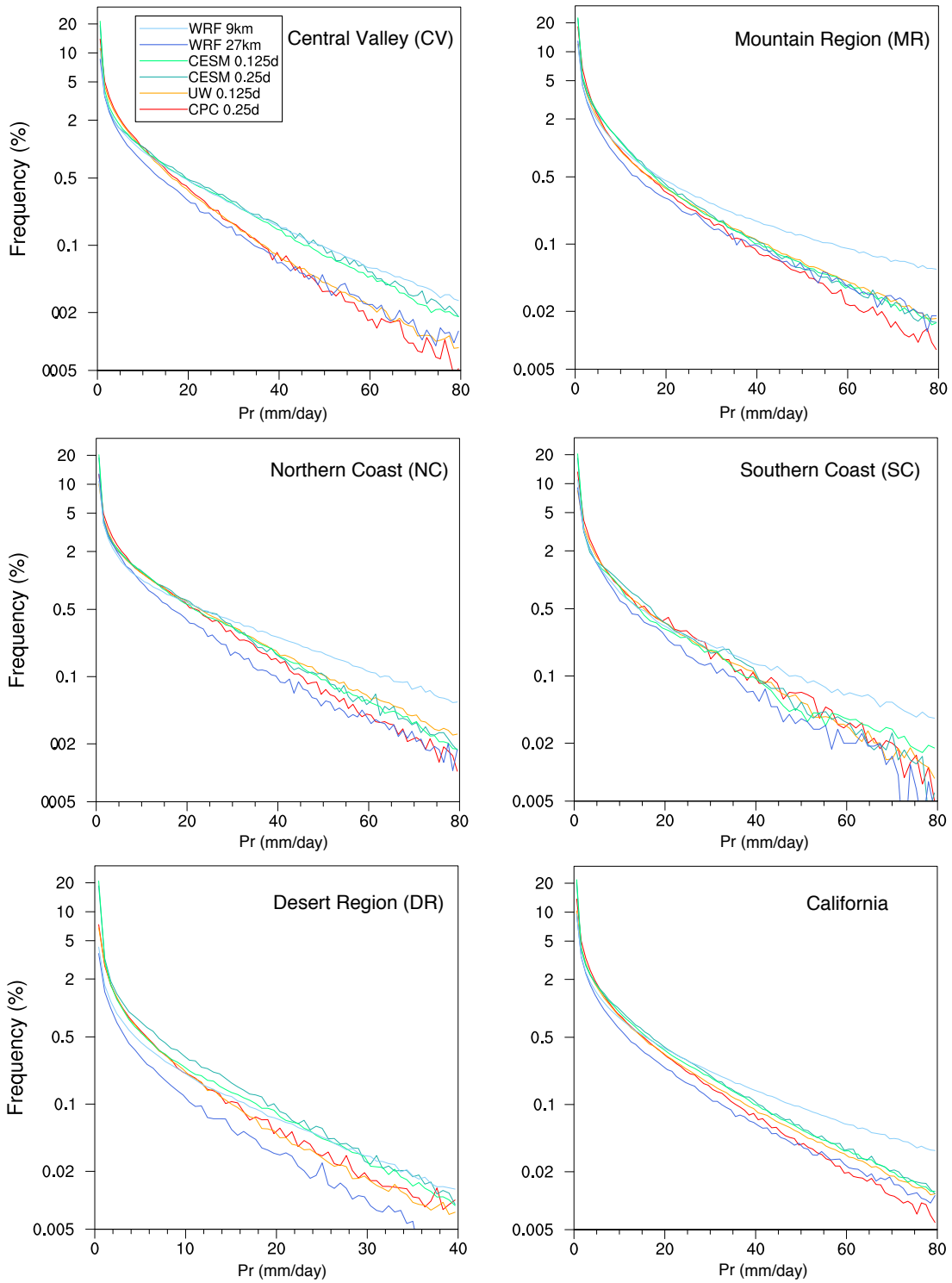


Figure 13. As Figure 8, but for DJF Pr (note that the vertical scale is logarithmic).

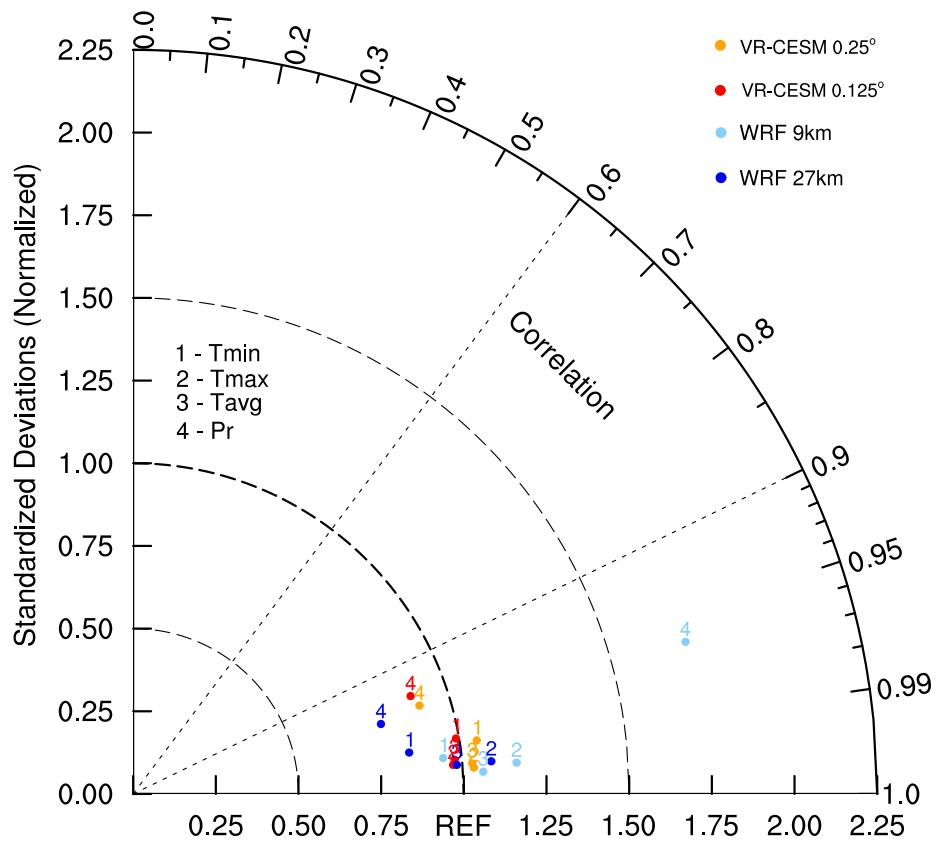


Figure 14. Taylor diagram of the annual climatology for the entire California region using the PRISM dataset as reference.