

¹ **High-resolution regional climate model evaluation using variable-resolution**

² **CESM over California**

³ Xingying Huang, * Alan M. Rhoades and Paul A. Ullrich

⁴ *Department of Land, Air and Water Resources, University of California, Davis*

⁵ Colin M. Zarzycki

⁶ *National Center for Atmospheric Research*

⁷ *Corresponding author address: Xingying Huang, Department of Land, Air and Water Resources,

⁸ University of California Davis, Davis, CA 95616.

⁹ E-mail: xyhuang@ucdavis.edu

ABSTRACT

10 Understanding the effect of climate change at regional scales remains a
11 topic of intensive research. Due to computational constraints, the fine hor-
12 izontal resolutions required to reach regional scales have been largely out of
13 reach for current global climate models. However, high resolution is needed
14 to represent fine-scale processes and topographic forcing, which are signif-
15 icant drivers of local climate variability. Although regional climate models
16 (RCMs) have been widely used at these scales, variable-resolution global cli-
17 mate models (VRGCMs) have arisen as an alternative for studying regional
18 weather and climate. In this paper, the recently developed variable-resolution
19 option within the Community Earth System Model (CESM) is assessed for
20 long-term regional climate modeling. The mean climatology of temperature
21 and precipitation, across California's diverse climate zones, is analyzed and
22 contrasted with the Weather Research and Forcasting (WRF) model (as a tra-
23 ditional RCM), regional reanalysis, gridded observational datasets and a uni-
24 form high-resolution CESM with the finite volume (FV) dynamical core. The
25 results show that variable-resolution CESM is competitive in representing re-
26 gional climatology on both annual and seasonal time scales. This assessment
27 adds value to the use of VRGCMs for projecting climate change over the
28 coming century and improve our understanding of both past and future re-
29 gional climate related to fine-scale processes. This assessment is also relevant
30 for addressing the scale limitation of current RCMs or VRGCMs when next-
31 generation model resolution increases to $\sim 10\text{km}$ and beyond.

³² **1. Introduction**

³³ Global climate models (GCMs) have been widely used to simulate both past and future climate.

³⁴ Although GCMs have demonstrated the capability to successfully represent large-scale features

³⁵ of the climate system, they are usually employed at coarse resolutions ($\sim 1^\circ$), largely due to com-

³⁶ putational limitations. Global climate reanalysis datasets, which assimilate climate observations

³⁷ using a global model, represent a best estimate of historical weather patterns, but still cannot

³⁸ fulfill the need of high resolution in the aspect of regional climate (<http://reanalyses.org/atmosphere/overview-current-reanalyses>). Consequently, regional climate is not well cap-

³⁹ tured by either GCMs or global reanalysis datasets. However, dynamical processes at unrep-

⁴⁰ resented scales are significant drivers for regional and local climate variability, especially over

⁴¹ complex terrain (Soares et al. 2012). In order to capture these fine-scale dynamical features, high

⁴² horizontal resolution is needed to allow for a more accurate representation of fine-scale forcings,

⁴³ processes and interactions (Leung et al. 2003a; Rauscher et al. 2010). With these enhancements,

⁴⁴ regional climate data is expected to be more usable for policy makers and local stakeholders in

⁴⁵ formulating climate adaptation and mitigation strategies.

⁴⁶

⁴⁷ In order to model regional climate at high spatial and temporal resolution over a limited area,

⁴⁸ downscaling techniques have been developed. Downscaling techniques can largely be catego-

⁴⁹ rized into statistical downscaling and dynamical downscaling. Dynamical downscaling is popular

⁵⁰ and commonly employed using nested limited-area models (LAMs) or using variable-resolution

⁵¹ enabled GCMs (VRGCMs) to model regional scales (Laprise 2008). In this context, LAMs are

⁵² typically referred to as regional climate models (RCMs) when used to study climate. Forced by

⁵³ output of GCMs or reanalysis data, RCMs have been widely used, particularly to capture physi-

⁵⁴ cally consistent regional and local circulations at the needed spatial and temporal scales (Leung

55 et al. 2003a; Christensen et al. 2007; Bukovsky and Karoly 2009; Mearns et al. 2012). However,
56 since RCMs only allow one-way coupling with the global domain, there has been growing interest
57 in VRGCMs for modeling regional climate. This approach uses a relatively coarse global model
58 with enhanced resolution over a specific region (Staniforth and Mitchell 1978; Fox-Rabinovitz
59 et al. 1997). Different strategies have been employed for transitioning between coarse and fine
60 resolution regions such as stretched-grid models or grid refinement (Fox-Rabinovitz et al. 1997;
61 Ringler et al. 2008; Skamarock et al. 2012). VRGCMs have been demonstrated to be effective
62 for regional climate studies and applications at a reduced computational cost compared to uni-
63 form GCMs (Fox-Rabinovitz et al. 2001, 2006; Rauscher et al. 2013; Zarzycki et al. 2015). Fox-
64 Rabinovitz et al. (2000) found that the stretched-grid version of a GCM not only captured large-
65 scale meteorological patterns but also mesoscale features particularly associated with orographic
66 forcing.

67 Compared with RCMs, a key advantage of VRGCMs is that they use a single, unified modeling
68 framework, rather than a separate GCM and RCM. Thus, VRGCMs avoid potential inconsistency
69 between the global and regional domains, and naturally support two-way interaction between these
70 domains without the need for nudging (Warner et al. 1997; McDonald 2003; Laprise et al. 2008;
71 Mesinger and Veljovic 2013). However, in order to obtain deeper insight into the performance
72 of these two modeling approaches, it is necessary to compare them directly. For the purposes of
73 this paper, we will focus on the recently developed variable-resolution Community Earth System
74 Model (varres-CESM) as our VRGCM of interest. Although CESM has been used heavily for
75 uniform resolution modeling, variable-resolution in the Community Atmosphere Models (CAM)
76 Spectral Element (SE) dynamical core has only been recently developed (Dennis et al. 2011; Tay-
77 lor 2011). Zarzycki et al. (2014) used varres-CESM to show that a high-resolution refinement
78 patch in the Atlantic basin when simulating topical cyclones represented significant improvements

79 over the unrefined simulation. Zarzycki et al. (2015) also compared the large-scale features of
80 varres-CESM 0.25° and uniform CESM at 1° , and found that adding a refined region over the
81 globe did not noticeably affect the global circulation.

82 However, varres-CESM has yet to be rigorously investigated for long-term regional climate
83 simulation (Taylor and Fournier 2010; Zarzycki et al. 2014). This paper is the first to investigate
84 whether a VRGCM can show similar or improved ability to model regional climate relative to a
85 traditional RCM **This sentence paints the paper as more competitive than it's supposed to be.** The
86 goal of this paper is to evaluate the performance of varres-CESM against gridded observational
87 data, reanalysis data and in comparison to a traditional RCM. In addition, comparisons are drawn
88 with a costly uniform high-resolution CESM simulation (Wehner et al. 2014a). Our variable-
89 resolution simulations will focus on relatively high resolutions for climate assessment, namely
90 28km and 14km grid spacing, which are much more typical for dynamically downscaled studies.
91 For comparison with the more widely used RCM method, the Weather Research and Forecast-
92 ing (WRF) model will be applied at 27km and 9km grid spacing (Skamarock et al. 2005). The
93 study focuses on models' ability to represent current climate statistics, particularly those relative
94 to climate extremes. We anticipate that this assessment will add value in modeling mean regional
95 climatology and improve our understanding about the effects of multi-scale processes in regional
96 climate regulation. Our eventual goal is to utilize these models for assessing future climate and
97 regional climate extremes.

98 This paper focuses on California (CA) as the study area. With its complex topography, coastal
99 influence, and wide latitudinal range, California is an excellent test bed for regional climate model-
100 ing. Further, an understanding of local climate variability is incredibly important for policymakers
101 and stakeholders in California due to its vast agricultural industry, mixed demographics, and vul-
102 nerability to anthropogenically-induced climate change (Hayhoe et al. 2004; Cayan et al. 2008).

103 RCM simulations over California have also been conducted in previous studies and demonstrated
104 the need for high resolution to better study regional climate and extreme events, especially over
105 complex topography with large climatological gradients (Leung et al. 2004; Kanamitsu and Kana-
106 maru 2007; Caldwell et al. 2009; Pan et al. 2011; Pierce et al. 2013). In particular, Caldwell et al.
107 (2009) presented results from WRF at 12km spatial resolution and showed that although the RCM
108 was effective at simulating the mean climate when compared with observations, some clear biases
109 persisted (such as overestimation of precipitation).

110 This paper is organized as follows. Section 2 describes the model setup, verification data and
111 evaluation methods. In section 3, simulation results are provided and discussed, with a focus on
112 2 meter temperature (Ts) and precipitation (Pr). Key results are summarized along with further
113 discussion in section 4.

114 **2. Models and Methodology**

115 *a. Simulation design*

116 In this study, all global simulations use the AMIP (Atmospheric Model Intercomparison Project)
117 protocols, which is a standard experimental regulation for global atmospheric GCMs (AGCMs)
118 (Gates 1992). AMIP simulation supports climate model diagnosis, validation and intercompar-
119 ison. AMIP experiment is constrained by realistic sea-surface temperatures (SSTs) and sea ice
120 from 1979 to near present without the added complexity of ocean-atmosphere feedbacks in the
121 climate system (<http://www-pcmdi.llnl.gov/projects/amip/NEWS/overview.php>). Ob-
122 served monthly mean SSTs and sea ice at 1 degree are provided and updated following a procedure
123 described by (Hurrell et al. 2008).

124 1) VARRES-CESM

125 CESM is a state-of-the-art Earth modeling framework managed by the National Center for At-
126 mospheric Research (NCAR), consisting of coupled atmospheric, oceanic, land and sea ice mod-
127 els. CESM has been frequently used for modeling present and future global climate (Neale et al.
128 2010a; Hurrell et al. 2013).

129 The coupling infrastructure in CESM allows the interfacial states and fluxes between the various
130 component models are communicated and the fluxed quantities are conserved. Since we follow
131 AMIP protocols in this study, active communication occurs only between atmospheric and land
132 model. Here, CAM version 5 (CAM5) (Neale et al. 2010b) and the Community Land Model
133 (CLM) version 4 (Oleson et al. 2010) are used. As mentioned earlier, SE was used as the dynam-
134 ical core in CAM along with variable-resolution grid support. The FAMIPC5 (F_AMIP_CAM5)
135 compset was chosen for these simulations.

136 For our study, the variable-resolution cubed-sphere grids are generated for use in CAM and CLM
137 with the open-source software package SQuadGen (Ullrich 2014). The grids used are depicted in
138 Figure 1. The maximum horizontal resolution on these grids are 0.25 degree ($\sim 28\text{km}$) and 0.125
139 degree ($\sim 14\text{km}$) respectively, with a quasi-uniform 1 degree mesh over the remainder of the globe.
140 Grids are constructed using a paving technique with a 2:1 aspect ratio, so two transition layers are
141 required from 1 degree to 0.25 degree, and one additional transition from 0.25 degree to 0.125
142 degree. The global circulation patterns (e.g. wind, pressure and precipitation) are not affected by
143 the variable-resolution grid mesh, and showed conserved results over the transition boundary as
144 described in (Zarzycki et al. 2015). Simulations cover the time period from 1979-01-01 to 2005-
145 12-31 (UTC), although the year 1979 was discarded as the spin-up period. This time period was

146 chosen to provide an adequate sampling of annual variability, to limit computational cost, and for
147 a period where adequate reanalysis data is available for comparison.

148 Variable-resolution topography files were produced by sampling the National Geophysical Data
149 Center (NGDC) 2-min (\sim 4 km) Gridded Global Relief Dataset (ETOPO2v2) topography dataset,
150 followed by application of a differential smoothing technique as described in Zarzycki et al.
151 (2015). Using this technique, the c parameter from Eq. (1) was adjusted to reduce noise in the
152 vertical pressure velocity field. The grid-scale topography is depicted in Figure 3. The higher
153 resolution simulations provide a much finer representation of regional topography. This is impor-
154 tant for understanding local climate since topography is an important driver for fine-scale dynamic
155 processes, especially over complex terrain.

156 Land surface datasets, and plant functional types, were created using the 0.5 degree reference
157 datasets. Greenhouse gas (GHG) concentrations and aerosol forcing are prescribed based on his-
158 torical observations. SSTs and ice coverage are supplied by the 1 degree Hadley Centre Sea Ice
159 and Sea Surface Temperature dataset (HadISST) (Hurrell et al. 2008). CAM and CLM tuning
160 parameters are not modified from their default configuration.

161 <http://www.cesm.ucar.edu/experiments/cesm1.0/#rcp>

162 2) UNIFORM CESM

163 Output from a globally uniform CESM run at 0.25° spatial resolution is utilized for comparison.
164 It helps us to see if variable-resolution CESM, which is at much lower computation cost than
165 uniform one, can show comparable performance in modeling mean climatology (Bacmeister et al.
166 2014). This globally uniform simulation uses the CAM5-FV (finite volume) dynamical core and is
167 described in additional detail in Wehner et al. (2014b). I'm also not sure if uniform CESM should
168 even be included. Maybe it should be removed. Huang: no info about CLM

¹⁶⁹ 3) WRF

¹⁷⁰ WRF has been widely used over the past decade for modeling regional climate (Lo et al. 2008;
¹⁷¹ Leung and Qian 2009; Soares et al. 2012), and so represents an adequate platform for assessing
¹⁷² climatology. In our study, the fully compressible non-hydrostatic WRF model in version 3.5.1
¹⁷³ with the Advanced Research WRF (ARW) dynamical solver is used. WRF is a limited area model
¹⁷⁴ that supports nested domains with a typical refinement ratio of 3:1. The simulation domains of
¹⁷⁵ WRF are depicted in Figure 2. Two WRF simulations, representing finest grid resolutions of 27km
¹⁷⁶ and 9km, are conducted. For the WRF 27km simulation, one domain is used. For the WRF 9km
¹⁷⁷ simulation, two nested domains are used with the outer domain at 27km (same as the WRF 27km)
¹⁷⁸ and inner domain at 9km horizontal grid resolution. Two-way nesting is enabled by overwriting
¹⁷⁹ coarse grid data by averaged fine grid data (Skamarock and Klemp 2008). For both regional
¹⁸⁰ simulations, grids are centered on CA and have 120×110 and 151×172 grid points, respectively.
¹⁸¹ Around the boundaries, 10 grid points are used for lateral relaxation. In order to reduce the drift
¹⁸² between forcing data and RCM, grid nudging (Stauffer and Seaman 1990) was applied to the outer
¹⁸³ domain every 6 hours at all levels except the planetary boundary layer (PBL), as suggested by Lo
¹⁸⁴ et al. (2008). This setup uses 41 vertical levels with model top pressure at 50 hPa.

¹⁸⁵ Additionally, we used the following physics parameterizations: WSM (WRF Single-Moment)
¹⁸⁶ 6-class graupel microphysics scheme (Hong and Lim 2006), Kain-Fritsch cumulus scheme (Kain
¹⁸⁷ 2004), CAM shortwave and longwave radiation schemes (Collins et al. 2004). These settings
¹⁸⁸ are supported by the one-year test running result with different options of cumulus scheme and
¹⁸⁹ radiation schemes. For the boundary layer, the Yonsei University scheme (YSU) (Hong et al.
¹⁹⁰ 2006) and the Noah Land Surface Model (Chen and Dudhia 2001) were used. Both were chosen

191 as they are common for climate applications that balance long-term reliability and computational
192 cost.

193 ECMWF Reanalysis (ERA-Interim) data at the surface and on pressure-levels provides initial
194 and lateral conditions for the domains. The lateral conditions and SSTs were updated every 6
195 hours. ERA-Interim reanalysis (~ 80 km) has been widely used and validated for its reliability
196 as forcing data (Dee et al. 2011). WRF simulations are conducted over the same time period as
197 varres-CESM. Again, the year 1979 is used as a spin-up period and is discarded for purposes of
198 analysis. Notably, the ~ 10 km resolution is actually finer than most previous studies for long-term
199 climate.

200 The topography for 27km and 9km simulations are interpolated from USGS (U.S. Geologi-
201 cal Survey) elevation data with 10-min (~ 20 km) and 2-min (~ 4 km) resolution, respectively.

202 The post-processed grid-scale topography is contrasted in Figure 3. Comparing ERA-Interim to
203 varres-CESM, it is obvious that high resolution ($\sim 0.25^\circ$) is needed to represent regionally diverse
204 topographical features. Also, topography exhibited more details when resolution went to 0.125°
205 from 0.25° . Elevation differences between varres-CESM and WRF are irregular and relatively
206 small, except over the Central Valley where varres-CESM has consistent higher values than WRF.
207 This indicates a different methodology for preparation of the topography dataset and may also
208 partly due to the different data sources.

209 *b. Methodology*

210 Near surface (2 meter) temperature and precipitation have been analyzed over California to as-
211 sess the performance of varres-CESM in representing mean climatology. Specifically, our evalua-
212 tion focuses on daily maximum, minimum and average 2m temperatures (Tmax, Tmin and Tavg)
213 and daily precipitation (Pr). These variables are key for a baseline climate assessment, as a con-

214 sequence of their close relationship with water resources, agriculture and health. In this context,
215 the biggest impact of weather on California is through heat and precipitation extremes. Since heat
216 extremes dominate during the summer season, we focus on June, July and August (JJA) for as-
217 sessment of temperature. On the other hand, since the vast majority of precipitation in CA occurs
218 in the winter season (together with the accumulation of snowpack (Rhoades et al. 2015)), pre-
219 cipitation over December-January-February (DJF) is emphasized. Future work will focus on the
220 capability of the variable resolution system to correctly capture the frequency and intensity of heat
221 and precipitation extremes.

222 In order to adequately account for natural variability in the mean climate, the simulation period
223 must be chosen appropriately (Solomon 2007). However, the number of simulated years required
224 for adequate climate statistics depends greatly on the regional climate variability and spatial scale.
225 Past studies have used average weather conditions over a 30-year period to ensure sufficient statis-
226 tics and avoid imprinting from annual variability (Dinse 2009). To ensure that our 26-year sim-
227 ulation period is sufficient, **we have studied the variability of mean temperature and precipitation**
228 **in both simulations and observations over 5, 10, 20 and 25 seasons or years. The results showed**
229 **that variability becomes generally stable when time period reaches 20 or 25 years, stating that our**
230 **simulation period is long enough to adequately capture the variability of recent mean climatology.**

231 The results in the following section are obtained from simulated and observed data over the
232 period 1980-2005. All datasets have been de-trended **at each grid point** so that simulation years can
233 be averaged. In each case it is found that for temperature a statistically significant trend is present
234 under the two-tailed t-statistic with a significance level of 0.05. **The average magnitude of the**
235 **trend is about 1.3 K over 26 years.** No statistically significant trend was detected for precipitation.
236 California's rugged topography and large latitudinal extent had led to a diverse variety of climate
237 regions that are poorly captured in typical coarse global climate simulations. In order to assess the

238 performance of varres-CESM within each region, the state has been divided into five regional
239 zones, including the Central Valley (CV), Mountain Region (MR), North Coast (NC), South Coast
240 (SC), and Desert Region (DR). The spatial extent of these regions is depicted in Figure 2. The
241 division of these five zones is loosely based on the results of Abatzoglou et al. (2009) and the
242 climate zones used by the California Energy Commission. To restrict the analysis in each zone,
243 simulations and datasets have been masked to restrict climate variables to each region.

244 Standard statistical measures have been used to quantify the performances of the models in
245 comparison with the reference datasets. These statistical variables include the Root-mean-square
246 deviation (RMSD), mean absolute difference (MAD) **When you use MAD in the tables, you're**
247 **using mean difference not mean absolute difference**, mean relative difference (MRD) and correla-
248 tion, and sample standard deviation. **We should include a short description of each of these, such**
249 **as in the LaTeX comments below.**

250 Grid point differences are calculated by remapping the reference datasets to the model's output
251 grid using bilinear interpolation. Remapping using patch-based interpolation has also been tested
252 and does not exhibit notable differences.

253 Student's t-test has been used at the 0.05 significance level to determine if two sets of yearly or
254 seasonally averaged data show statistically significant differences. We need to point out that this is
255 just a approximate test to further support our results analysis since the this test assumes normality
256 of the sample population. (add F-test here)

257 **What about the supplement?**

258 *c. Gridded and Reanalysis Datasets*

259 For validation purpose, reanalysis and gridded observational datasets of the highest available
260 quality are employed (see Table 1). In particular, gridded datasets at resolutions finer than 10km

represent the best foundation for assessing the variable-resolution simulation results. Differences between gridded observations can be due to choice of meteorological stations, interpolation techniques, elevation models and processing algorithms. Consequently, the use of multiple reference datasets is necessary to understand the uncertainty underlying the observational data. Moreover, in this study, our purpose of using these products is to serve as realistic proxies to allow for a comparison of the model results. We acknowledge that reanalysis products are particularly sensitive to model choice and choice of assimilated observations and so cannot be treated as truth. Detailed descriptions of these datasets are as follows.

(i) *NARR*: The North American Regional Reanalysis (NARR) provides dynamically downscaled data over North America at ~ 32 km resolution and 3 hourly intervals from 1979 through present (Mesinger et al. 2006). It is the National Centers for Environmental Prediction (NCEP)'s high resolution reanalysis product. All major climatological variables are present in NARR, making it an excellent candidate for assessment of regional climate. Nonetheless, some inaccuracies have been identified in NARR that must be accounted for, including deficiencies in precipitation fields away from the continental US (Bukovsky and Karoly 2007).

(ii) *NCEP CPC*: This data set is CPC unified gauge-based analysis of daily precipitation provided by the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC). It is a suite of unified precipitation products with consistent and improved quality by combining all information available at CPC and by taking advantage of the optimal interpolation (OI) objective analysis technique. The gauge analysis covers the Conterminous United States with a fine-resolution at 0.25° from 1948/01/01 to 2006/12/3.

(iii) *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).

284 UW incorporates topographic corrections by forcing the long-term average precipitation to match
285 that of the PRISM dataset. The temperature dataset is produced in a similar fashion as precip-
286 itation, but uses a simple 6.1 K/km lapse rate for topographic effect. The dataset is at 0.125°
287 horizontal resolution and is provided from year 1949 to 2010.

288 (iv) *PRISM*: The Parameter-elevation Regressions on Independent Slopes Model (PRISM) (Daly
289 et al. 2008) supports a 4km gridded dataset obtained by taking a wide range of point measurements
290 and applying a weighted regression scheme that accounts for many factors affecting the local cli-
291 matology. The datasets include total precipitation and minimum/maximum, (derived) mean tem-
292 peratures and dewpoints, based on sophisticated quality control measures. Monthly climatological
293 variables are available for 1895 through 2014 provided by the PRISM Climate Group. PRISM is
294 U.S. Department of Agriculture (USDA)'s official climatological data. We will use this product as
295 the main reference dataset for model assessment.

296 (v) *Daymet*: Daymet is an extremely high resolution (1 km) gridded dataset with daily outputs
297 of total precipitation, humidity, and minimum/maximum temperature covering the years of 1980
298 through 2013 (Thornton et al. 1997; Thornton and Running 1999; Thornton et al. 2000). The
299 dataset is produced using an algorithmic technique that ingests point station measurements in
300 conjunction with a truncated Gaussian weighting filter. Some adjustments are made to account for
301 topography. Daymet is available through the Oak Ridge National Laboratory Distributed Active
302 Archive Center (ORNL DAAC).

303 To assess differences in these data products, we have used the Student's t-test to see if the JJA
304 Tmax, Tmin and Tavg from PRISM, UW and Daymet are statistically different from one another
305 **comparing seasonally averaged values**. For these quantities, all products exhibited similarity at
306 the 0.05 significance level over most regions of the study area, except in the two coastal regions

307 [I'm not sure this means you can assume uncertainty is negligible], Huang: Actually, I did not
308 assume uncertainty is negligible. A similar assessment was performed for mean DJF precipitation
309 climatology and, in this case, all products exhibited similarity at the 0.05 significance level over
310 the entire simulation domain even north coast.

311 A figure or table showing these results should be included here. Can we estimate the uncertainty
312 in these datasets somehow? Huang: uncertainty is usually like mean+/-uncertainty, this is not
313 applicable to our case. Here, uncertainty tends to mean the statistical difference. This can be done
314 by combining the result of the t-test and absolute difference. I will plot the t-test results, and the
315 absolute differences for these and those mentioned in the results part. How about put them in the
316 supplement?

317 3. Results

318 a. Temperature

319 The mean JJA Tmax, Tmin and Tavg climatology over the simulation period, together with
320 NARR and PRISM data is shown in Figure 4. Statistical measures over CA are tabulated in Table
321 2. All simulations have captured the spatial climate patterns exhibited by the PRISM, with high
322 spatial correlations (>0.95), especially for Tmax and Tavg.

323 For Tmax, when compared with the reference datasets, varres-CESM showed a warm bias of
324 about 2 to 3°C. Uniform CESM showed a similar outcome to varres-CESM, but with a larger
325 RMSD value (~4°C). However, WRF produced an overall colder climate, especially the WRF
326 9km simulation, which were about 2 to 3°C cooler than PRISM. Tmax over CV has been overes-
327 timated by all the simulations. Possible reasons for this overestimation are discussed at the end of
328 this section.

329 For Tmin, varres-CESM still showed a strong warm bias (~ 3 to 4°C), with a particularly egre-
330 gious overestimation over Nevada ($> 5^{\circ}\text{C}$). WRF also exhibited a warm bias, but of a much
331 smaller magnitude (~ 2 to 3°C). However, the pattern of Tmin presented in Figure 4 in both WRF
332 simulations suggests a cooler interior to the Central Valley and warmer perimeter, which is not
333 supported by observations.

334 For Tavg, the warm bias of Tmin and Tmax by varres-CESM leads a similar overestimation for
335 Tavg. For WRF, underestimation of Tmax and overestimation of Tmin lead to an overall closer
336 match to Tavg over most of the domain, but is indicative of suppressed variability.

337 The spatial standard deviation of JJA Tmax, Tmin and Tavg by models and PRISM are depicted
338 in Figure 5. It can be seen that the variability is largely consistent across different sub-zones, and
339 the values are around 0.5 to 1.5°C for all the datasets, except for the high Sierras in the WRF 9km
340 simulation which show enhanced variability ($\sim 2^{\circ}\text{C}$). **It would be interesting to calculate RMSE**
341 **of the standard deviation. A visual inspection suggests varres-0.125d has the best representation**
342 **of variability.**

343 Compared with the reference datasets over CA, varres-CESM 0.125 degree produced the lowest
344 RMSD for Tmax, whereas WRF produced the lowest RMSD for Tmin. In both cases the RMSD
345 was still around 2°C . Notably, Tmin from varres-CESM matched much more closely with NARR,
346 although this may be indicative of a related warm bias in NARR. In fact, closer examination of
347 the differences between varres-CESM, WRF and NARR marine 2m temperature patterns indicates
348 that CESM and NARR exhibit Tmin values that are approximately 2°C larger than WRF. Since
349 marine 2m temperature is strongly correlated with ocean SSTs, this suggests a possible source of
350 the warm bias in CESM. The sea breeze effect, associated with cooler temperatures near the San
351 Francisco Bay, is apparent in all runs. It is especially encouraging that differences in the varres-

352 CESM simulations, which only used prescribed SSTs, closely match those of WRF, which were
353 also forced at the lateral domain boundaries with reanalysis data.

354 The seasonal cycle of Tavg is shown in Figure 6 for simulations and reference data from PRISM
355 and NARR. Modeled results match closely with reference data with no larger than a 2°C difference,
356 with the largest errors occurring in the summer and winter seasons showing statistical significance
357 of differences over most regions. Compared with PRISM, Varres-CESM overpredicted summer
358 season temperatures in all sub-zones except coastal regions, and underpredicted winter season
359 temperatures in all zones, corresponding to a larger annual temperature range. Uniform CESM
360 was similar to, but statistically different from varres-CESM, with a larger deviation of around 3°C.
361 WRF has **better performance** in presenting the monthly trend than CESM, which is further proved
362 by the t-test, with about 1°C underestimation over all seasons. There is no clear improvement in
363 the seasonal cycle across resolutions.

364 Variability in monthly average Tavg is expressed by the sample standard deviation showed in
365 Figure 7. The variability is the interannual standard deviation of monthly Tavg over 26 years.
366 Generally, standard deviation is between 1 to 2°C. Among the simulations, WRF 27km is **most**
367 **consistent to PRISM with smallest difference**. WRF 9km is **also close to PRISM**, but has ~1°C
368 larger variability over January and February with statistical significance of difference. Varres-
369 CESM basically showed about 0.5°C more scattered values (either above or lower) comparing to
370 reference datasets, and uniform CESM has about 0.5°C lower variability than PRISM. **The F-test**
371 **results showed that most of the variability values among these models do not have statistically**
372 **significant differences.**

373 Due to the impact of summer heat waves, we now focus on Tmax over summer season. In Figure
374 8, the frequency distribution of Tmax using all JJA daily values over 26 years is depicted. Proper-
375 ties of the frequency distribution, including average, variability, skewness and Kurtosis are tabu-

376 lated in Table 3. As exemplified by the similarity in the moments of the distribution, varres-CESM
377 clearly captures the general distribution of Tmax correctly. Outside of the central valley, skewness
378 and Kurtosis measures match closely between varres-CESM and the UW dataset. In the north and
379 south coastal regions, Daymet overestimates the frequency of cold days leading to deviation in the
380 moments from UW. Consistent with the observations in Figure 4, outside of the CV, WRF tends
381 to be cooler in general and varres-CESM tends to be warmer. In coastal regions all models show
382 better performance with higher Tmax values than with lower Tmax values. Enhanced frequency
383 of cool Tmax values appears to be the primary driver in overestimation of sample variance in these
384 regions. For both varres-CESM and WRF there is no apparent improvement in statistics at higher
385 resolutions.

386 In the Central Valley, models show a clear warm bias and underestimated skewness, associated
387 with a long forward tail and temperatures reaching near 50°C. As discussed earlier, all models do
388 overestimate Tmax over CV. In order to further assess the accuracy of the gridded observations,
389 we examine the Tmax data directly from recorded weather station observations over the CV. The
390 results validate that Tmax values above 45°C are rare (although station observations suggest these
391 days may be slightly more frequent than suggested by UW and Daymet). The warm bias associated
392 with the aforementioned extreme hot days in both varres-CESM and WRF is likely correlated
393 with overly dry summertime soil moisture, as discussed in Caldwell et al. (2009). This could be
394 caused by the lack of accurate land surface treatment in climate models. areas. Bonfils and Lobell
395 (2007) found that irrigation in Central Valley has significantly decreased summertime maximum
396 temperatures especially in heavily-irrigated areas. Other studies have also found the cooling effects
397 of irrigation, such as (Kueppers et al. 2007).

398 *b. Precipitation*

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

403 The long-term average climatologies of DJF and annual daily precipitation (Pr) over 26 years
404 from simulations and reference datasets are depicted in Figure 9. Statistical quantities over CA are
405 given in Table 4. Precipitation is heavily influenced by orography, leading to most accumulation
406 occurring along the North Coast and Sierra Nevada mountains. As with temperature, the model
407 results match the spatial patterns of the PRISM, with high correlation coefficients (>0.94).

408 Along the western edge of the Sierra Nevada and into the CV, varres-CESM overestimates total
409 precipitation relative to PRISM, especially the coarser resolution (28 km) simulation (about 40%-
410 50%) **with statistically significant difference from PRISM**. However, varres-CESM 0.125° is still
411 statistically the same as PRISM **with p value around 0.2 to 0.3 over the aforementioned region**.
412 On the other hand, precipitation is slightly underestimated relative to PRISM along the North
413 Coast, particularly near the Oregon border. Uniform CESM has slightly better results than varres-
414 CESM 0.25deg, especially over the western edge of the Sierra Nevada, with about 10% reduced
415 MRD values. There are also notable differences between WRF 27km and WRF 9km. WRF 27km
416 underestimates precipitation along the North Coast (by about 30%) but fairly accurately captures
417 precipitation in the CV, whereas WRF 9km greatly overestimates precipitation (by about 60%-
418 80%) along the North coast and the Sierra Nevada. **However, considering the variability shown**
419 **in the Figure 10, WRF 9km and WRF 27km are both significantly the same at the significance**
420 **level of 0.05 as PRISM except over the mountain region. [I find this hard to believe – how did you**

421 compute? if using t-test you should use sample variance from PRISM. Huang: That what I used.
422 I will plot the t test result explicitly. Notably, variability has a similar pattern to the precipitation
423 intensity, and increases as the precipitation magnitude increases. Models capture the variation of
424 precipitation well with close values to PRISM Huang: I will also calculate the RMSD, particularly
425 looking at the varres-CESM 0.125deg and WRF 27 km, however, variability is ~50% higher for
426 WRF9km.

427 Using RMSD values from Table 4 as a guide, varres-CESM 0.125° performs slightly better than
428 CESM 0.25° and WRF 27km [Is this all you have to say about the table?].

429 The annual cycle of precipitation averaged over each sub-zone over 26 years is presented in
430 Figure 11. It can be seen that the simulations exhibit similar trends to the reference datasets with
431 highest precipitation over winter and lowest values over summer. The main deviation occurred
432 during the rainy seasons, especially in winter. WRF 27km is drier than PRISM and UW with
433 relative differences ranging from ~10%-40%, whereas WRF 9km is far wetter with relative dif-
434 ferences reaching up to 40%-80% over these five climate sub-zones. Varres-CESM tracks well
435 [vague - t-test needed] Huang: I will add t-test for this. The t-test showed varres-CESM, uniform-
436 CESM and WRF 27 is same as reference datasets especially over rainy seasons (except: WRF
437 27 over desert). but WRF 27 is not the same as CESM. with observed precipitation with ~10%-
438 20% relative difference everywhere except in the CV, where precipitation is overestimated at rainy
439 seasons with about 70%-80%. Nonetheless, the strong seasonal dependence on precipitation is
440 apparent with extremely dry conditions during summer months. A slight increase in summertime
441 precipitation is apparent in the Desert region, indicating the North American monsoon. However,
442 we also observe that the peak month for precipitation tends to occur earlier in varres-CESM than in
443 observations [comment on jaggedness of varres-CESM as well]. It is not surprising that a seasonal

444 time drift occurred with the varres-CESM simulations as it was not forced by a reanalysis dataset
445 like the WRF simulations. [Why?]

446 The monthly cycle sample standard deviation is depicted in Figure 12. The variability has a
447 similar monthly trend compared to precipitation rate, with overall values from 0 to 4 mm/day.
448 Generally higher inter-annual variability occurs over locations of higher mean precipitation 11.
449 Comparing with observations, varres-CESM exhibited basically no more than 1mm/day larger
450 variability in the rainy season except over the CV. WRF 27km is also close to observations with
451 generally no more than 1mm/day over all regions. WRF 9km again showed larger variability
452 (~1.5 mm/day more) during rainy seasons over most regions. Such higher variability within
453 higher magnitude of precipitation has also been found in previous studies; for example,Duffy et al.
454 (2006) observe higher variability associated with higher spatial resolution in RCMs, attributed to
455 a more accurate representation of topography. The main cause of the interannual variability of
456 precipitation over CA is the El NiñoSouthern Oscillation (ENSO), which varies the amount of
457 moisture flux transported to this region.

458 [Show statistical significance of differences in variability via f-test] WRF 9km is different from
459 others over most region except north coast. Others are the same over most region especially over
460 rainy seasons (except CV, especially for varres-CESM; except desert for WRF 27km, for CESM
461 over summer, though there is rare rain; except uniform CESM over south coast).

462 The frequency distribution of DJF Pr has been constructed from rainy days in winter
463 ($Pr \geq 0.1 \text{ mm/d}$) and depicted in Figure 13. Generally, varres-CESM matches more closely with
464 observations everywhere except in the CV. In CV region, WRF 27km appears to better capture
465 high-intensity precipitation events, but performs poorly on low-intensity events ($Pr < 20 \text{ mm/d}$).
466 The underestimation of rainfall frequency in WRF 27km appears consistent across regions. WRF
467 9km produces a significantly better treatment of low-intensity events, but greatly overestimates the

frequency of high-intensity events ($\text{Pr} > 20 \text{ mm/d}$). For strong precipitation events, varres-CESM
is more close to UW dataset than WRF except at the CV.

The overestimation of precipitation for WRF at high resolution has also been found in previous studies. Although not as pronounced, Caldwell et al. (2009) demonstrated that WRF at 12km largely overestimated the precipitation over the California's mountainous regions (however, this paper did employ a different set of 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. Further discussion can be found in former studies that employ different microphysics schemes (and so produce a wide range of precipitation magnitudes) including (Leung et al. 2003b; Jankov et al. 2005; Gallus Jr and Bresch 2006; Chin et al. 2010; Caldwell 2010).

A concise summary of model performance over CA 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 reference data. It can be seen that the models correlate well with the PRISM reference dataset. Normalized standard deviation and bias are larger for precipitation, especially for WRF 9km. Overall, varres-CESM has demonstrated that it can competitively compare to WRF in capturing the regional climatology of California.

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. varres-CESM) as an alternative approach for two-way dynamically down-scaled climate modeling. The performance of varres-CESM was evaluated in simulating Cali-

491 fornia's unique regional climatology. This relatively new technique has been evaluated against
492 gridded reference datasets, regional reanalysis data and the Weather Research and Forecasting
493 model.

494 Based on 26 years of high-resolution historical climate simulations, we analyzed the mean cli-
495 matology of California and across its climate divisions from both temperature and precipitation.
496 Generally, when compared with gridded observational datasets, both varres-CESM and WRF do
497 a good job of capturing regional climatological patterns with high spatial correlations (>0.94).
498 Uncertainty between reference datasets exists, but is relatively small and not statistically signifi-
499 cant over most regions. We found that varres-CESM showed comparable performance as WRF in
500 regional climate study. Even compared with a uniform high-resolution GCM (CESM-FV), varres-
501 CESM also performed competitively.

502 Deviations from reference datasets do exist in these simulations, but they have different features.
503 During summer, varres-CESM model possessed about 2 to 3°C warmer climate, especially in
504 the Central Valley. WRF exhibited a colder ($\sim 2^{\circ}\text{C}$) Tmax over most regions except the Central
505 Valley, but a little warmer in Tmin. Overall, varres-CESM showed better ability in reproducing
506 mean climatology of Tmax, but WRF was better at modeling Tmin and Tavg. The variability
507 of JJA mean temperature is largely within the range of 0.5 to 1.5°C . WRF presents the annual
508 cycle of Tavg better than CESM with about 1°C underestimation. CESMs showed about 2°C
509 overestimation of Tavg over the summer season and similar magnitude of underestimation over
510 winter season, indicating larger temperature range over most regions.

511 When assessing the frequency distribution of JJA Tmax, both varres-CESM and WRF 27km
512 match closely to UW dataset over all study area except in Central Valley (CV). The failure to
513 correctly capture CV Tmax is likely caused by the lack of irrigation cooling effect over this region.
514 Future work will address this issue by adding an irrigation parameterization to varres-CESM so

515 as to figure out the role irrigation played in regulating Tmax, and the overestimation and longer
516 upbounded tail of frequency distribution for Tmax,

517 As for precipitation representation, varres-CESM matches closely with PRISM everywhere ex-
518 cept for an overestimation of winter or annual precipitation (about 40%-50%) along the western
519 side of Sierra Nevada and into the CV. Increasing the resolution produces a slight reduction in
520 this overestimation (10%) likely due to improved treatment of orographic effects. WRF 27km
521 underestimates precipitation (about 30%) along the North coast and Sierra Nevada mountains,
522 where almost all the precipitation comes from, whereas WRF 9km shows a large overestimation
523 (about 70%–80%). Variability of precipitation ranges from 0 to 6 mm/day, with generally higher
524 inter-annual variability over locations of higher mean precipitation. For strong precipitation events
525 probability, varres-CESM is more close to UW dataset than WRF except at the CV.

526 Higher resolution (0.125°) simulation of varres-CESM do show better results in capturing sum-
527 mer Tmax, precipitation and their variability, than the coarser resolution run. However, the im-
528 provements are not statistically significant. For WRF, when resolution increased, the model pro-
529 duces obviously overestimated precipitation as previous studies have also found when using RCMs
530 for fine-scale simulations as aforementioned. The use convection scheme is perhaps not needed
531 when grid spacing is near 10km. However, it turned out that almost all of the precipitation comes
532 from resolved (large-scale) processes for all these models. In this way, model deviation is mainly
533 related with resolved-scale processes and microphysics scheme plays a major role, which makes
534 it necessary to develop more scale-aware parameterizations.

535 The importance and necessity of high resolution for regional climate studies has been widely
536 stressed by previous studies. However, whether the current regional climate models can fulfill this
537 demand when resolution is pushed to local scales is questionable. It is clear that further work is
538 urgently needed to solve the scale limitation of current regional climate models at fine horizontal

539 resolutions. The possible causes of the scale limitation may include a lack of accurate scale-aware
540 physical parameterizations near or below 10 km horizontal resolution, the treatment of dynamics
541 at fine scales, and the interactions among different components of RCMs or VRGCMs (e.g., land-
542 atmosphere interactions).

543 In summary, varres-CESM demonstrated competitive utility for studying high-resolution re-
544 gional climatology when compared to a regional climate model (WRF) and a uniform high-
545 resolution GCM (CESM-FV). Deviations, showed within these models, are not indicative of deep
546 underlying problems with the model formulation, but one should be aware of these differences
547 when using these models for assessing future climate change. This study suggests that variable-
548 resolution GCMs are useful tools for assessing climate change over the coming century. As the
549 need for assessments of regional climate change is increasing, alternative modeling strategies, in-
550 cluding variable-resolution global climate models will be needed to improve our understanding of
551 the effects of fine-scale processes representation in regional climate regulation.

552 *Acknowledgments.* The authors would like to thank Michael Wehner and Daíthí Stone for pro-
553 viding the uniform resolution CESM dataset. We would further like to thank Travis O'Brien for
554 his many useful conversations on climate model assessment. We also want to thank Xuemeng
555 Chen for her assistance in WRF running. This project is supported in part by the University of
556 California, Davis and by the Department of Energy “Multiscale Methods for Accurate, Efficient,
557 and Scale-Aware Models of the Earth System” program.

558 References

- 559 Abatzoglou, J. T., K. T. Redmond, and L. M. Edwards, 2009: Classification of regional climate
560 variability in the state of California. *Journal of Applied Meteorology and Climatology*, **48** (8),
561 1527–1541.

- 562 Bacmeister, J. T., M. F. Wehner, R. B. Neale, A. Gettelman, C. Hannay, P. H. Lauritzen, J. M.
563 Caron, and J. E. Truesdale, 2014: Exploratory high-resolution climate simulations using the
564 Community Atmosphere Model (CAM). *Journal of Climate*, **27** (9), 3073–3099.
- 565 Bonfils, C., and D. Lobell, 2007: Empirical evidence for a recent slowdown in irrigation-induced
566 cooling. *Proceedings of the National Academy of Sciences*, **104** (34), 13 582–13 587.
- 567 Bukovsky, M. S., and D. J. Karoly, 2007: A brief evaluation of precipitation from the North
568 American Regional Reanalysis. *Journal of Hydrometeorology*, **8** (4), 837–846.
- 569 Bukovsky, M. S., and D. J. Karoly, 2009: Precipitation simulations using WRF as a nested regional
570 climate model. *Journal of Applied Meteorology and Climatology*, **48** (10), 2152–2159.
- 571 Caldwell, P., 2010: California wintertime precipitation bias in regional and global climate models.
572 *Journal of Applied Meteorology and Climatology*, **49** (10), 2147–2158.
- 573 Caldwell, P., H.-N. S. Chin, D. C. Bader, and G. Bala, 2009: Evaluation of a WRF dynamical
574 downscaling simulation over California. *Climatic Change*, **95** (3-4), 499–521.
- 575 Cayan, D. R., A. L. Luers, G. Franco, M. Hanemann, B. Croes, and E. Vine, 2008: Overview of
576 the California climate change scenarios project. *Climatic Change*, **87** (1), 1–6.
- 577 Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn
578 State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly
579 Weather Review*, **129** (4), 569–585.
- 580 Chin, H.-N. S., P. M. Caldwell, and D. C. Bader, 2010: Preliminary study of California wintertime
581 model wet bias. *Monthly Weather Review*, **138** (9), 3556–3571.
- 582 Christensen, J. H., and Coauthors, 2007: Regional climate projections. *Climate Change, 2007:
583 The Physical Science Basis. Contribution of Working group I to the Fourth Assessment Report*

584 of the Intergovernmental Panel on Climate Change, University Press, Cambridge, Chapter 11,
585 847–940.

586 Collins, W. D., and Coauthors, 2004: Description of the NCAR community atmosphere model
587 (CAM 3.0). Tech. rep., Tech. Rep. NCAR/TN-464+ STR.

588 Daly, C., M. Halbleib, J. I. Smith, W. P. Gibson, M. K. Doggett, G. H. Taylor, J. Curtis, and P. P.
589 Pasteris, 2008: Physiographically sensitive mapping of climatological temperature and precipi-
590 tation across the conterminous United States. *International Journal of Climatology*, **28** (15),
591 2031–2064.

592 Dee, D., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of
593 the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137** (656),
594 553–597.

595 Dennis, J., and Coauthors, 2011: Cam-se: A scalable spectral element dynamical core for the com-
596 munity atmosphere model. *International Journal of High Performance Computing Applications*,
597 1094342011428142.

598 Dinse, K., 2009: Climate variability and climate change: what is the difference. *Book Climate
599 Variability and Climate Change: What is the Difference*.

600 Duffy, P., and Coauthors, 2006: Simulations of present and future climates in the western United
601 States with four nested regional climate models. *Journal of Climate*, **19** (6), 873–895.

602 Fox-Rabinovitz, M., J. Côté, B. Dugas, M. Déqué, and J. L. McGregor, 2006: Variable resolution
603 general circulation models: Stretched-grid model intercomparison project (SGMIP). *Journal of
604 Geophysical Research: Atmospheres (1984–2012)*, **111** (D16).

- 605 Fox-Rabinovitz, M. S., G. L. Stenchikov, M. J. Suarez, and L. L. Takacs, 1997: A finite-
606 difference GCM dynamical core with a variable-resolution stretched grid. *Monthly Weather*
607 **Review**, **125** (11), 2943–2968.
- 608 Fox-Rabinovitz, M. S., G. L. Stenchikov, M. J. Suarez, L. L. Takacs, and R. C. Govindaraju, 2000:
609 A uniform-and variable-resolution stretched-grid GCM dynamical core with realistic orography.
610 *Monthly Weather Review*, **128** (6), 1883–1898.
- 611 Fox-Rabinovitz, M. S., L. L. Takacs, R. C. Govindaraju, and M. J. Suarez, 2001: A variable-
612 resolution stretched-grid general circulation model: Regional climate simulation. *Monthly*
613 *Weather Review*, **129** (3), 453–469.
- 614 Gallus Jr, W. A., and J. F. Bresch, 2006: Comparison of impacts of WRF dynamic core, physics
615 package, and initial conditions on warm season rainfall forecasts. *Monthly Weather Review*,
616 **134** (9), 2632–2641.
- 617 Gates, W. L., 1992: AMIP: The Atmospheric Model Intercomparison Project. *Bulletin of the*
618 *American Meteorological Society*, **73**, 1962–1970.
- 619 Hamlet, A. F., and D. P. Lettenmaier, 2005: Production of Temporally Consistent Gridded Precipi-
620 tation and Temperature Fields for the Continental United States*. *Journal of Hydrometeorology*,
621 **6** (3), 330–336.
- 622 Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on Califor-
623 nia. *Proceedings of the National Academy of Sciences of the United States of America*, **101** (34),
624 12 422–12 427.
- 625 Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme
626 (WSM6). *Asia-Pacific Journal of Atmospheric Sciences*, **42** (2), 129–151.

- 627 Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit
628 treatment of entrainment processes. *Monthly Weather Review*, **134** (9), 2318–2341.
- 629 Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski, 2008: A new sea surface temper-
630 ature and sea ice boundary dataset for the Community Atmosphere Model. *Journal of Climate*,
631 **21** (19), 5145–5153.
- 632 Hurrell, J. W., and Coauthors, 2013: The community earth system model: A framework for col-
633 laborative research. *Bulletin of the American Meteorological Society*, **94** (9), 1339–1360.
- 634 Jankov, I., W. A. Gallus Jr, M. Segal, B. Shaw, and S. E. Koch, 2005: The impact of different WRF
635 model physical parameterizations and their interactions on warm season MCS rainfall. *Weather*
636 and Forecasting, **20** (6), 1048–1060.
- 637 Kain, J. S., 2004: The Kain-Fritsch convective parameterization: An update. *Journal of Applied*
638 *Meteorology*, **43** (1), 170–181.
- 639 Kanamitsu, M., and H. Kanamaru, 2007: Fifty-seven-year California Reanalysis Downscaling at
640 10 km (CaRD10). Part I: System detail and validation with observations. *Journal of Climate*,
641 **20** (22), 5553–5571.
- 642 Kueppers, L. M., M. A. Snyder, and L. C. Sloan, 2007: Irrigation cooling effect: Regional climate
643 forcing by land-use change. *Geophysical Research Letters*, **34** (3).
- 644 Laprise, R., 2008: Regional climate modelling. *Journal of Computational Physics*, **227** (7), 3641–
645 3666.
- 646 Laprise, R., and Coauthors, 2008: Challenging some tenets of regional climate modelling. *Meteo-*
647 *rology and Atmospheric Physics*, **100** (1-4), 3–22.

- 648 Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby, 2003a: Regional climate research: needs
649 and opportunities. *Bulletin of the American Meteorological Society*, **84** (1), 89–95.
- 650 Leung, L. R., and Y. Qian, 2009: Atmospheric rivers induced heavy precipitation and flooding in
651 the western US simulated by the WRF regional climate model. *Geophysical research letters*,
652 **36** (3).
- 653 Leung, L. R., Y. Qian, and X. Bian, 2003b: Hydroclimate of the western United States based on
654 observations and regional climate simulation of 1981-2000. Part I: Seasonal statistics. *Journal*
655 *of Climate*, **16** (12), 1892–1911.
- 656 Leung, L. R., Y. Qian, X. Bian, W. M. Washington, J. Han, and J. O. Roads, 2004: Mid-century
657 ensemble regional climate change scenarios for the western United States. *Climatic Change*,
658 **62** (1-3), 75–113.
- 659 Lo, J. C.-F., Z.-L. Yang, and R. A. Pielke, 2008: Assessment of three dynamical climate downscal-
660 ing methods using the Weather Research and Forecasting (WRF) model. *Journal of Geophysical*
661 *Research: Atmospheres (1984–2012)*, **113** (D9).
- 662 Maurer, E., A. Wood, J. Adam, D. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically
663 based dataset of land surface fluxes and states for the conterminous United States*. *Journal of*
664 *climate*, **15** (22), 3237–3251.
- 665 McDonald, A., 2003: Transparent boundary conditions for the shallow-water equations: testing in
666 a nested environment. *Monthly Weather Review*, **131** (4), 698–705.
- 667 Mearns, L. O., and Coauthors, 2012: The North American regional climate change assessment
668 program: overview of phase I results. *Bulletin of the American Meteorological Society*, **93** (9),
669 1337–1362.

- 670 Mesinger, F., and K. Veljovic, 2013: Limited area NWP and regional climate modeling: a test
671 of the relaxation vs Eta lateral boundary conditions. *Meteorology and Atmospheric Physics*,
672 **119 (1-2)**, 1–16.
- 673 Mesinger, F., and Coauthors, 2006: North American regional reanalysis. *Bulletin of the American*
674 *Meteorological Society*, **87**, 343–360.
- 675 Neale, R. B., and Coauthors, 2010a: Description of the NCAR community atmosphere model
676 (CAM 5.0). *NCAR Tech. Note NCAR/TN-486+STR*.
- 677 Neale, R. B., and Coauthors, 2010b: Description of the NCAR Community Atmosphere Model
678 (CAM 5.0). NCAR Technical Note NCAR/TN-486+STR, National Center for Atmospheric Re-
679 search, Boulder, Colorado, 268 pp.
- 680 Oleson, K., and Coauthors, 2010: Technical description of version 4.0 of the Community Land
681 Model (CLM). NCAR Technical Note NCAR/TN-478+STR, National Center for Atmospheric
682 Research, Boulder, Colorado, 257 pp. doi:10.5065/D6FB50WZ.
- 683 Pan, L.-L., S.-H. Chen, D. Cayan, M.-Y. Lin, Q. Hart, M.-H. Zhang, Y. Liu, and J. Wang, 2011:
684 Influences of climate change on California and Nevada regions revealed by a high-resolution
685 dynamical downscaling study. *Climate Dynamics*, **37 (9-10)**, 2005–2020.
- 686 Pierce, D. W., and Coauthors, 2013: Probabilistic estimates of future changes in California temper-
687 ature and precipitation using statistical and dynamical downscaling. *Climate Dynamics*, **40 (3-**
688 **4)**, 839–856.
- 689 Rauscher, S. A., E. Coppola, C. Piani, and F. Giorgi, 2010: Resolution effects on regional climate
690 model simulations of seasonal precipitation over Europe. *Climate Dynamics*, **35 (4)**, 685–711.

- 691 Rauscher, S. A., T. D. Ringler, W. C. Skamarock, and A. A. Mirin, 2013: Exploring a Global
692 Multiresolution Modeling Approach Using Aquaplanet Simulations. *Journal of Climate*, **26** (8),
693 2432–2452.
- 694 Rhoades, A. M., X. Huang, P. A. Ullrich, and C. M. Zarzycki, 2015: Characterizing Sierra Nevada
695 Snowpack Using Variable-Resolution CESM. *J. Appl. Meteor. Clim.*, Under minor revision.
- 696 Ringler, T., L. Ju, and M. Gunzburger, 2008: A multiresolution method for climate system model-
697 ing: application of spherical centroidal voronoi tessellations. *Ocean Dynamics*, **58** (5-6), 475–
698 498.
- 699 Skamarock, W., J. Klemp, J. Dudhia, D. Gill, and D. Barker, 2005: Coauthors, 2008: A Descrip-
700 tion of the Advanced Research WRF Version 3. NCAR Technical Note. Tech. rep., NCAR/TN–
701 475+ STR.
- 702 Skamarock, W. C., and J. B. Klemp, 2008: A time-split nonhydrostatic atmospheric model for
703 weather research and forecasting applications. *Journal of Computational Physics*, **227** (7),
704 3465–3485.
- 705 Skamarock, W. C., J. B. Klemp, M. G. Duda, L. D. Fowler, S.-H. Park, and T. D. Ringler, 2012: A
706 multiscale nonhydrostatic atmospheric model using centroidal voronoi tesselations and c-grid
707 staggering. *Monthly Weather Review*, **140** (9), 3090–3105.
- 708 Soares, P. M., R. M. Cardoso, P. M. Miranda, J. de Medeiros, M. Belo-Pereira, and F. Espírito-
709 Santo, 2012: WRF high resolution dynamical downscaling of ERA-Interim for Portugal. *Cli-
710 mate dynamics*, **39** (9-10), 2497–2522.
- 711 Solomon, S., 2007: *Climate change 2007—the physical science basis: Working group I contribution
712 to the fourth assessment report of the IPCC*, Vol. 4. Cambridge University Press.

- 713 Staniforth, A. N., and H. L. Mitchell, 1978: A variable-resolution finite-element technique for
714 regional forecasting with the primitive equations. *Monthly Weather Review*, **106** (4), 439–447.
- 715 Stauffer, D. R., and N. L. Seaman, 1990: Use of four-dimensional data assimilation in a limited-
716 area mesoscale model. Part I: Experiments with synoptic-scale data. *Monthly Weather Review*,
717 **118** (6), 1250–1277.
- 718 Taylor, M. A., 2011: Conservation of mass and energy for the moist atmospheric primitive equa-
719 tions on unstructured grids. *Numerical Techniques for Global Atmospheric Models*, Springer,
720 357–380.
- 721 Taylor, M. A., and A. Fournier, 2010: A compatible and conservative spectral element method on
722 unstructured grids. *Journal of Computational Physics*, **229** (17), 5879–5895.
- 723 Thornton, P. E., H. Hasenauer, and M. A. White, 2000: Simultaneous estimation of daily solar ra-
724 diation and humidity from observed temperature and precipitation: an application over complex
725 terrain in Austria. *Agricultural and Forest Meteorology*, **104** (4), 255–271.
- 726 Thornton, P. E., and S. W. Running, 1999: An improved algorithm for estimating incident daily
727 solar radiation from measurements of temperature, humidity, and precipitation. *Agricultural and*
728 *Forest Meteorology*, **93** (4), 211–228.
- 729 Thornton, P. E., S. W. Running, and M. A. White, 1997: Generating surfaces of daily meteorolog-
730 ical variables over large regions of complex terrain. *Journal of Hydrology*, **190** (3), 214–251.
- 731 Ullrich, P. A., 2014: SQuadGen: Spherical Quadrilateral Grid Generator. Accessed: 2015-02-11,
732 <http://climate.ucdavis.edu/squadgen.php>.

- 733 Warner, T. T., R. A. Peterson, and R. E. Treadon, 1997: A tutorial on lateral boundary conditions
734 as a basic and potentially serious limitation to regional numerical weather prediction. *Bulletin*
735 *of the American Meteorological Society*, **78** (11), 2599–2617.
- 736 Wehner, M., Prabhat, K. Reed, D. Stone, W. D. Collins, J. Bacmeister, and A. Gettleman, 2014a:
737 Resolution dependence of future tropical cyclone projections of CAM5.1 in the US CLIVAR
738 Hurricane Working Group idealized configurations. *J. Climate*, doi:10.1175/JCLI-D-14-00311.
739 1, in press.
- 740 Wehner, M. F., and Coauthors, 2014b: The effect of horizontal resolution on simulation qual-
741 ity in the Community Atmospheric Model, CAM5.1. *J. Model. Earth. Sys.*, doi:10.1002/
742 2013MS000276.
- 743 Zarzycki, C. M., C. Jablonowski, and M. A. Taylor, 2014: Using Variable-Resolution Meshes
744 to Model Tropical Cyclones in the Community Atmosphere Model. *Monthly Weather Review*,
745 **142** (3), 1221–1239.
- 746 Zarzycki, C. M., C. Jablonowski, D. R. Thatcher, and M. A. Taylor, 2015: Effects of localized
747 grid refinement on the general circulation and climatology in the community atmosphere model.
748 *Journal of Climate*, **28** (7), 2777–2803.

749 LIST OF TABLES

750	Table 1. Reanalysis and statistically downscaled observational datasets used in this	36
751	study.	
752	Table 2. RMSD, MAD and Correlation (Corr) for seasonally-averaged JJA temperature	37
753	metrics over California	
754	Table 3. The first four moments of the JJA Tmax frequency in each sub-zone. Column	38
755	titles refer to Average (Avg), Variance (Var), Skewness (Skew) and Kurtosis	
756	(Kurt).	
757	Table 4. RMSD, MAD, MRD, Correlation (Corr) for precipitation over California	39

TABLE 1. Reanalysis and statistically downscaled observational datasets used in this study.

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

TABLE 2. RMSD, MAD and Correlation (Corr) for **seasonally-averaged** JJA temperature metrics over California

RMSD	UW		PRISM			Daymet	
	T _{max}	T _{min}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}
varres-CESM 0.25d	2.322	3.745	2.924	3.121	2.604	2.810	3.934
varres-CESM 0.125d	1.900	3.631	2.447	2.944	2.184	2.475	3.701
WRF 27km	2.310	2.738	2.933	2.254	2.169	2.511	2.992
WRF 9km	3.319	2.937	3.492	1.837	1.769	3.203	2.942
uniform CESM 0.25d	3.885	4.088	4.265	3.614	3.536	4.315	4.274
MAD	UW		PRISM			Daymet	
	T _{max}	T _{min}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}
varres-CESM 0.25d	0.981	2.907	0.606	1.731	0.823	1.177	2.877
varres-CESM 0.125d	0.645	2.848	0.203	1.660	0.579	0.818	2.744
WRF 27km	-0.577	0.819	-0.952	-0.357	-0.771	-0.386	0.789
WRF 9km	-2.277	1.862	-2.720	0.674	-1.142	-2.103	1.757
uniform CESM 0.25d	1.812	2.993	1.449	1.815	1.280	2.013	2.961
Corr	UW		PRISM			Daymet	
	T _{max}	T _{min}	T _{max}	T _{min}	T _{avg}	T _{max}	T _{min}
varres-CESM 0.25d	0.998	0.982	0.996	0.986	0.994	0.997	0.979
varres-CESM 0.125d	0.998	0.985	0.997	0.988	0.996	0.997	0.983
WRF 27km	0.997	0.982	0.996	0.989	0.996	0.997	0.978
WRF 9km	0.996	0.985	0.997	0.993	0.998	0.996	0.984
uniform CESM 0.25d	0.994	0.980	0.992	0.981	0.991	0.993	0.977

758 TABLE 3. The first four moments of the JJA Tmax frequency in each sub-zone. Column titles refer to Average
 759 (Avg), Variance (Var), Skewness (Skew) and Kurtosis (Kurt).

	Central valley				Mountain				North coast				South coast				Desert			
	Avg	Var	Skew	Kurt	Avg	Var	Skew	Kurt	Avg	Var	Skew	Kurt	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	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	32.7	23.5	-0.9	1.5	25.9	39.3	-0.5	0.5	26.5	30.1	-0.3	0.4	26.5	30.1	-0.3	0.4	37.0	24.3	-0.6	0.6
CESM 0.25d	34.1	26.2	-0.4	0.2	28.1	27.6	-0.4	0.3	26.4	37.4	0.1	-0.7	26.4	37.4	0.1	-0.7	37.6	19.0	-0.5	0.8
CESM 0.125d	34.3	28.5	-0.5	0.4	27.2	30.0	-0.4	0.3	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	33.9	34.8	-0.5	0.2	24.9	34.8	-0.3	0.0	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	32.4	33.1	-0.7	0.6	22.4	38.5	-0.5	0.6	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], it is usually more sharply peaked [flatter] than the normal distribution (leptokurtic and platykurtic, respectively).

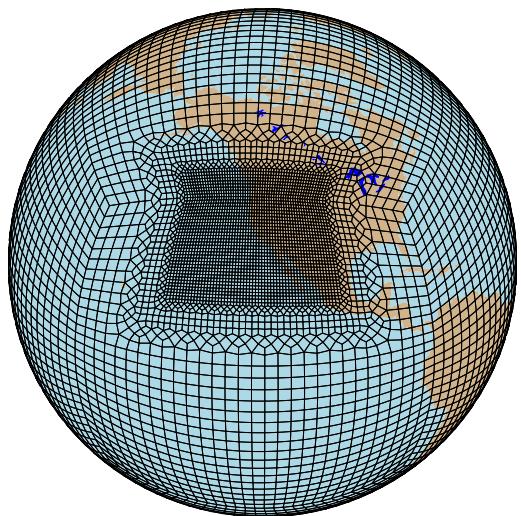
TABLE 4. RMSD, MAD, MRD, Correlation (Corr) for precipitation over California

Annual	CPC				UW				PRISM				DAYMET				
	RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr	
varres-CESM 0.25d	0.607	0.394	0.413	0.981	0.616	0.292	0.434	0.968	0.727	0.203	0.429	0.952	0.567	0.191	0.375	0.972	
varres-CESM 0.125d	0.469	0.207	0.321	0.980	0.526	0.115	0.339	0.970	0.624	0.045	0.328	0.961	0.504	0.027	0.310	0.973	
WRF 27km	0.419	-0.205	0.269	0.977	0.580	-0.308	0.274	0.971	0.765	-0.396	0.296	0.965	0.647	-0.409	0.312	0.970	
WRF 9km	2.226	1.485	0.950	0.957	2.052	1.393	0.864	0.964	1.889	1.322	0.815	0.970	2.005	1.306	0.773	0.961	
uniform CESM 0.25d	0.555	0.134	0.277	0.969	0.600	0.031	0.302	0.961	0.700	-0.057	0.290	0.953	0.600	-0.069	0.284	0.962	
DJF		CPC				UW				PRISM				DAYMET			
		RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr	RMSD	MAD	MRD	Corr
varres-CESM 0.25d	1.486	0.986	0.532	0.972	1.445	0.673	0.531	0.959	1.654	0.577	0.547	0.943	1.346	0.514	0.435	0.964	
varres-CESM 0.125d	1.194	0.638	0.396	0.976	1.234	0.346	0.398	0.965	1.395	0.287	0.400	0.955	1.170	0.212	0.337	0.969	
WRF 27km	0.888	-0.376	0.269	0.975	1.289	-0.688	0.289	0.967	1.552	-0.785	0.298	0.962	1.351	-0.848	0.324	0.966	
WRF 9km	4.264	2.607	0.742	0.950	3.835	2.315	0.616	0.955	3.570	2.256	0.604	0.964	3.804	2.183	0.554	0.955	
uniform CESM 0.25d	1.392	0.377	0.300	0.960	1.431	0.064	0.316	0.951	1.544	-0.033	0.314	0.946	1.406	-0.095	0.288	0.953	

760 LIST OF FIGURES

761	Fig. 1. Grid meshes for the two varres-CESM simulations.	41
762	Fig. 2. Domains of WRF simulations (left) and five climate divisions in California (right) with topography in meters (m).	42
764	Fig. 3. Topography in meters (m) for (top left to bottom right) varres-CESM 0.25° , varres-CESM 0.125° , uniform CESM-FV 0.25° , WRF 27km, WRF 9km and ERA-Interim (~ 80 km).	43
766	Fig. 4. JJA average daily Tmax, Tmin and Tavg from models and reference datasets, and differences between models and PRISM ($^\circ\text{C}$). Maybe remove Uniform CESM and move PRISM to the first row at the end. Place difference color bar somewhere more in line with plots.	44
769	Fig. 5. Sample standard deviation of JJA average daily Tmax, Tmin and Tavg from models and PRISM ($^\circ\text{C}$).	45
771	Fig. 6. Seasonal cycle of monthly-average Tavg for each subzone ($^\circ\text{C}$). The shading refers to the 0.95 confidence interval of PRISM and NARR.	46
773	Fig. 7. Seasonal standard deviation (s) values of monthly-average Tavg for each subzone ($^\circ\text{C}$).	47
774	Fig. 8. Frequency distribution of summer Tmax ($^\circ\text{C}$).	48
775	Fig. 9. Annual and DJF precipitation from models and reference datasets, and differences between models and PRISM (mm/d).	49
777	Fig. 10. Sample standard deviation of Annual and DJF precipitation from models and PRISM (mm/d).	50
779	Fig. 11. As Figure 6, but for monthly-average total precipitation (mm/d). Can you add the 0.95 confidence interval to PRISM and UW?	51
781	Fig. 12. As Figure 7, but for monthly-average total precipitation (mm/d).	52
782	Fig. 13. Frequency distribution of winter Pr, the unit of x-axis is mm/d (note that the vertical scale is logarithmic).	53
784	Fig. 14. Taylor diagram of annual climatology for the entire California region, using the PRISM dataset as reference.	54

1 degree -> 0.25 degree



1 degree -> 0.125 degree

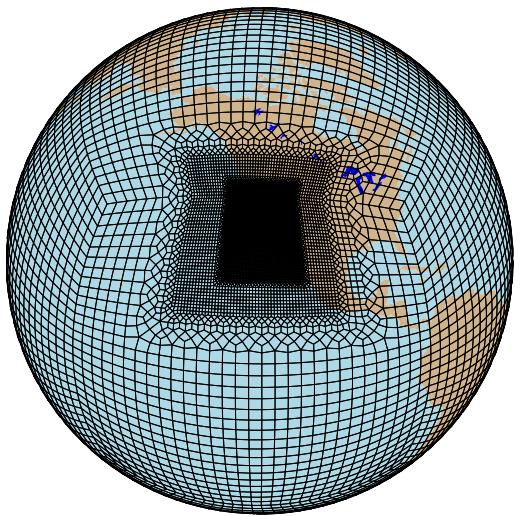
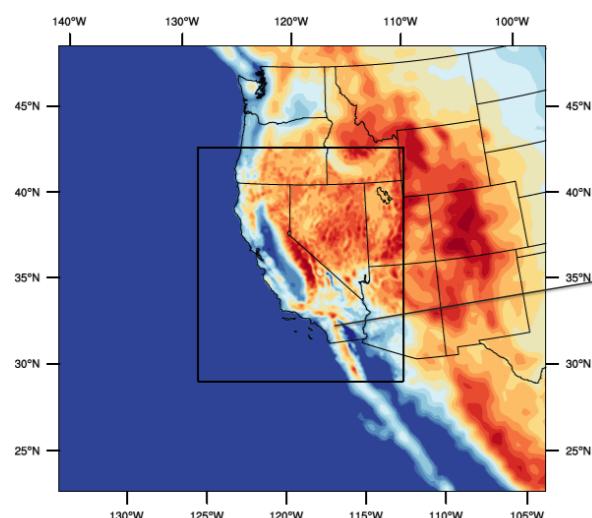
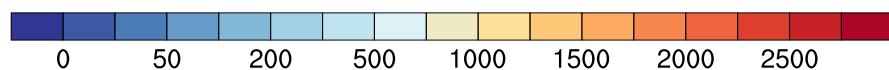
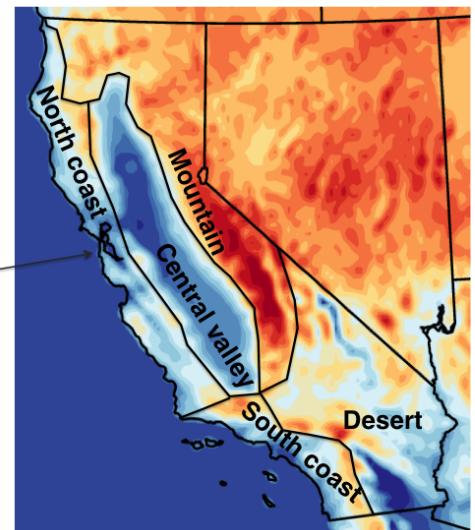


FIG. 1. Grid meshes for the two varres-CESM simulations.

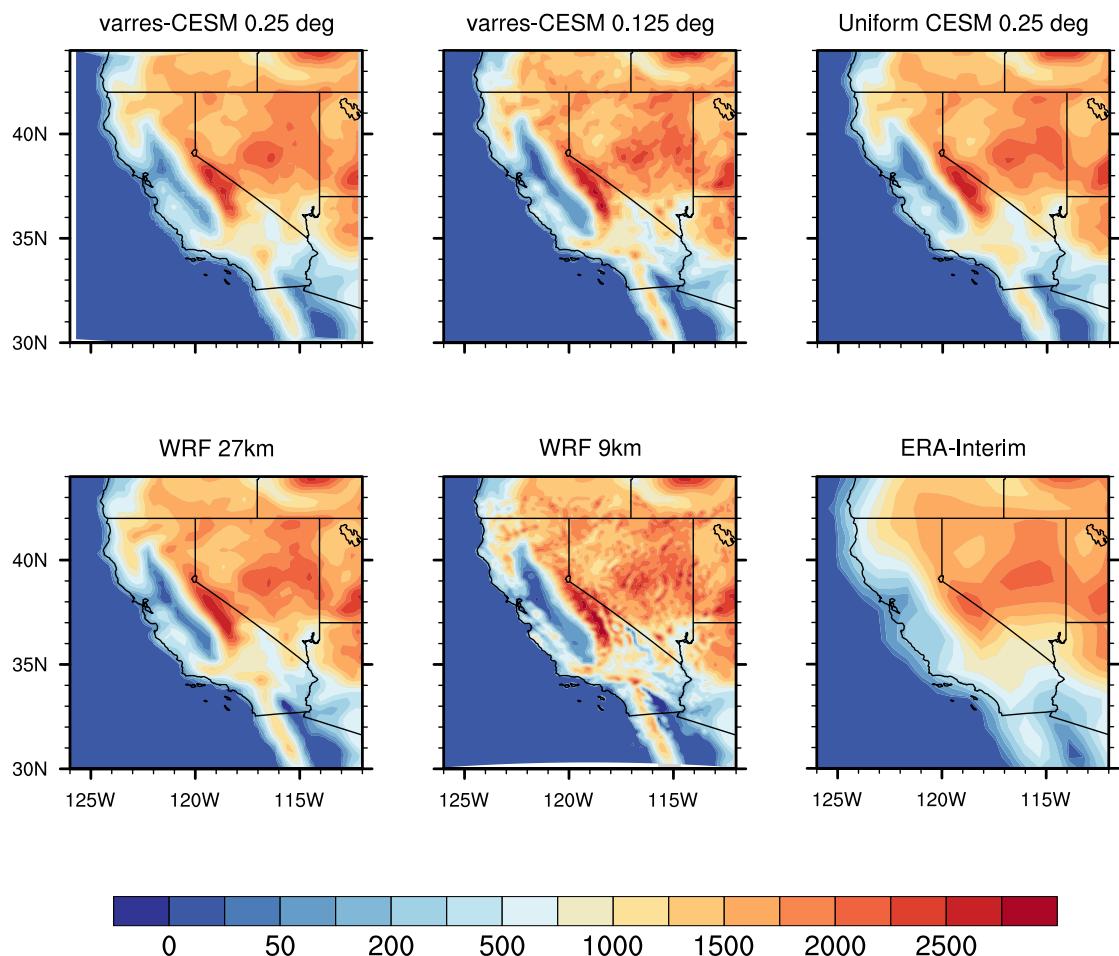
WRF 9km : Outer and inner domain



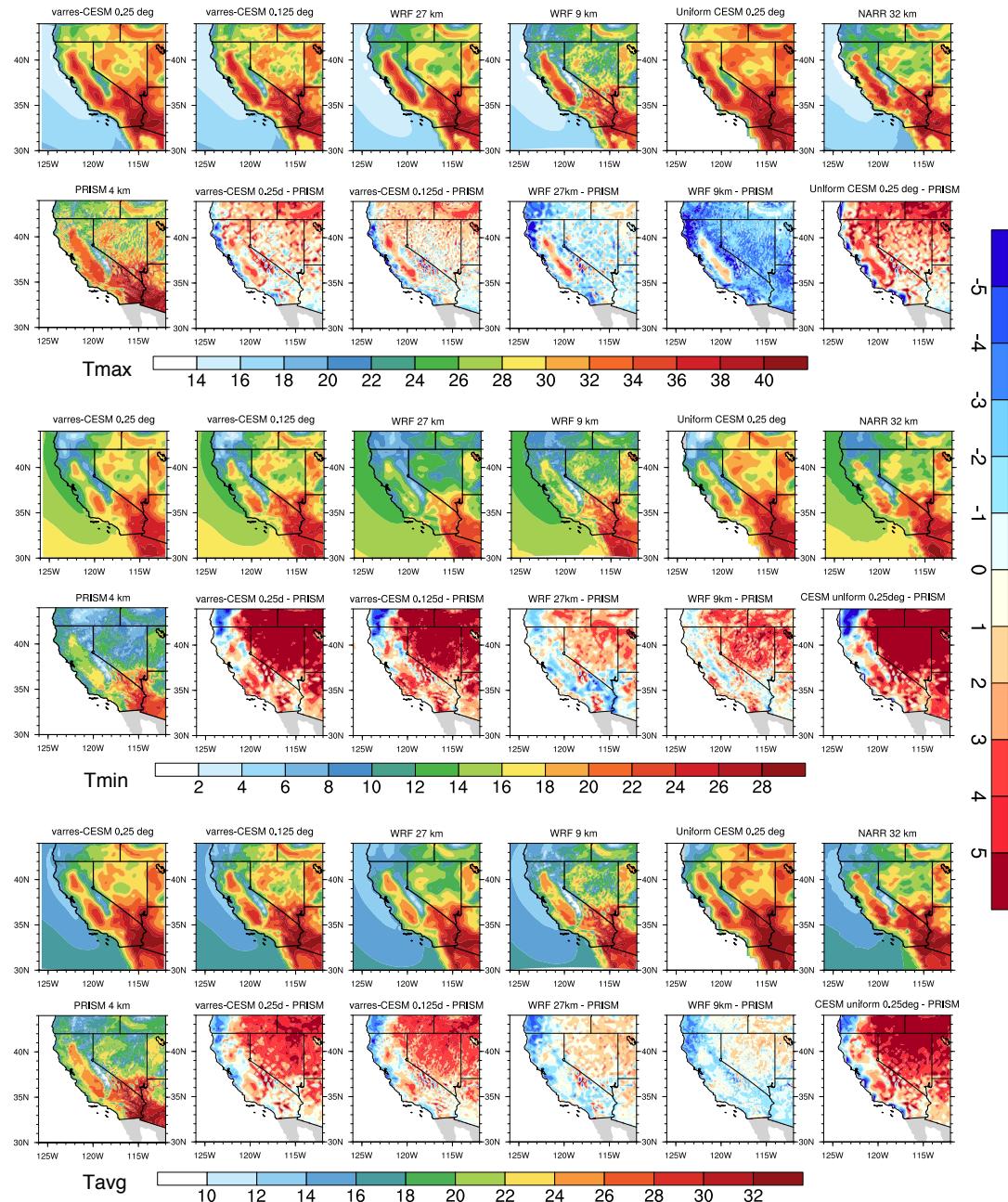
Climate divisions across CA



786 FIG. 2. Domains of WRF simulations (left) and five climate divisions in California (right) with topography in
787 meters (m).



788 FIG. 3. Topography in meters (m) for (top left to bottom right) varres-CESM 0.25° , varres-CESM 0.125° ,
789 uniform CESM-FV 0.25° , WRF 27km, WRF 9km and ERA-Interim (~ 80 km).



790 FIG. 4. JJA average daily Tmax, Tmin and Tavg from models and reference datasets, and differences between
 791 models and PRISM ($^{\circ}\text{C}$). **Maybe remove Uniform CESM and move PRISM to the first row at the end. Place**
 792 **difference color bar somewhere more in line with plots.**

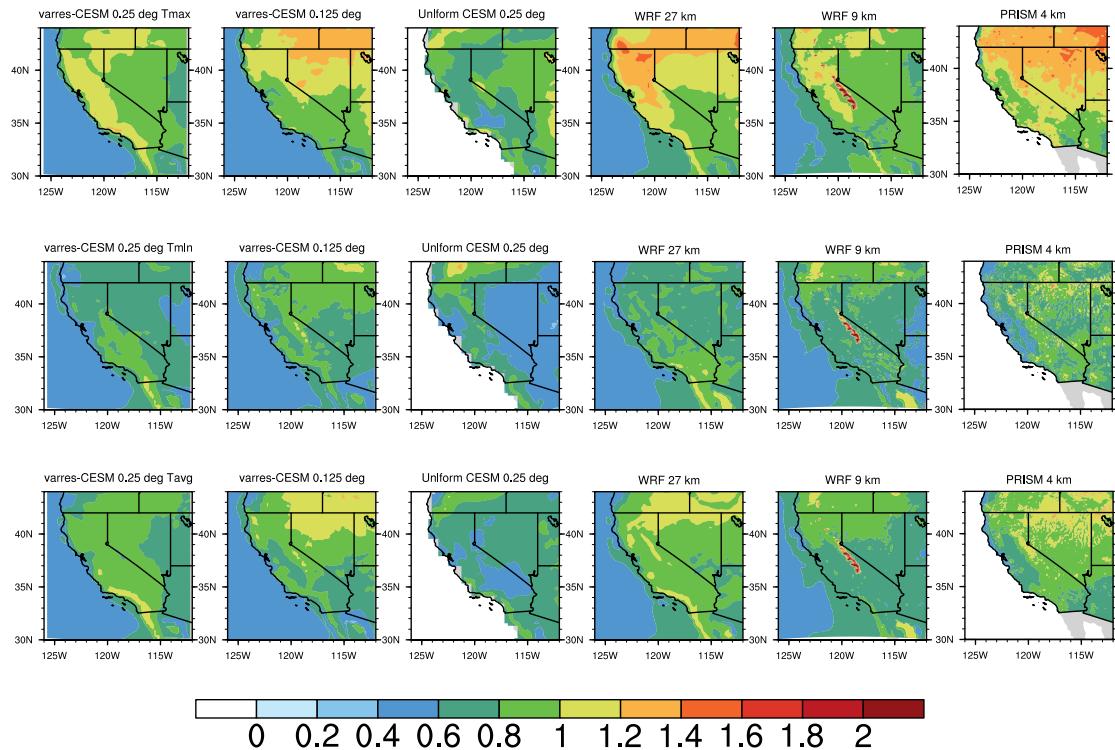
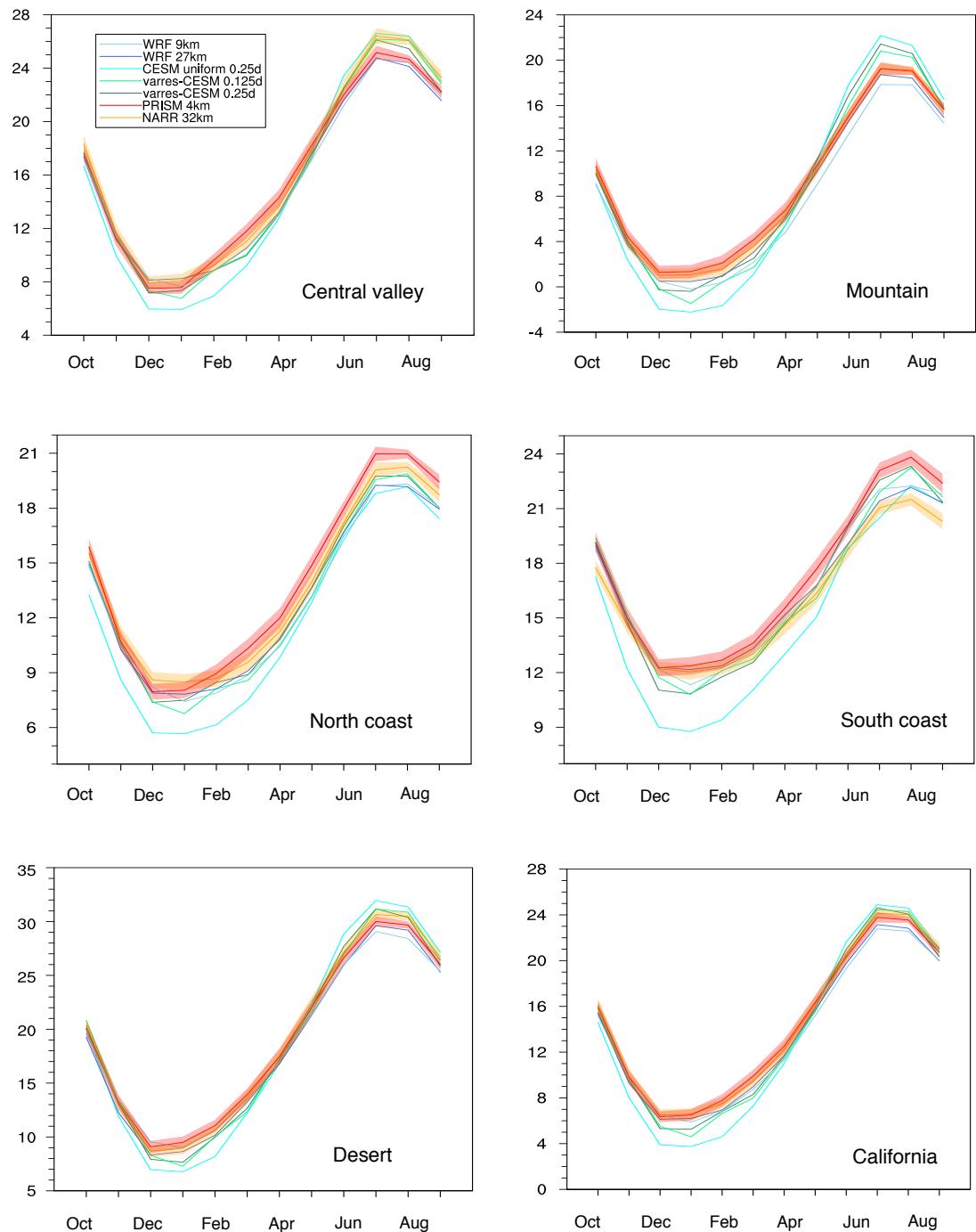


FIG. 5. Sample standard deviation of JJA average daily Tmax, Tmin and Tavg from models and PRISM ($^{\circ}\text{C}$).



793 FIG. 6. Seasonal cycle of monthly-average Tavg for each subzone ($^{\circ}\text{C}$). The shading refers to the 0.95
794 confidence interval of PRISM and NARR.

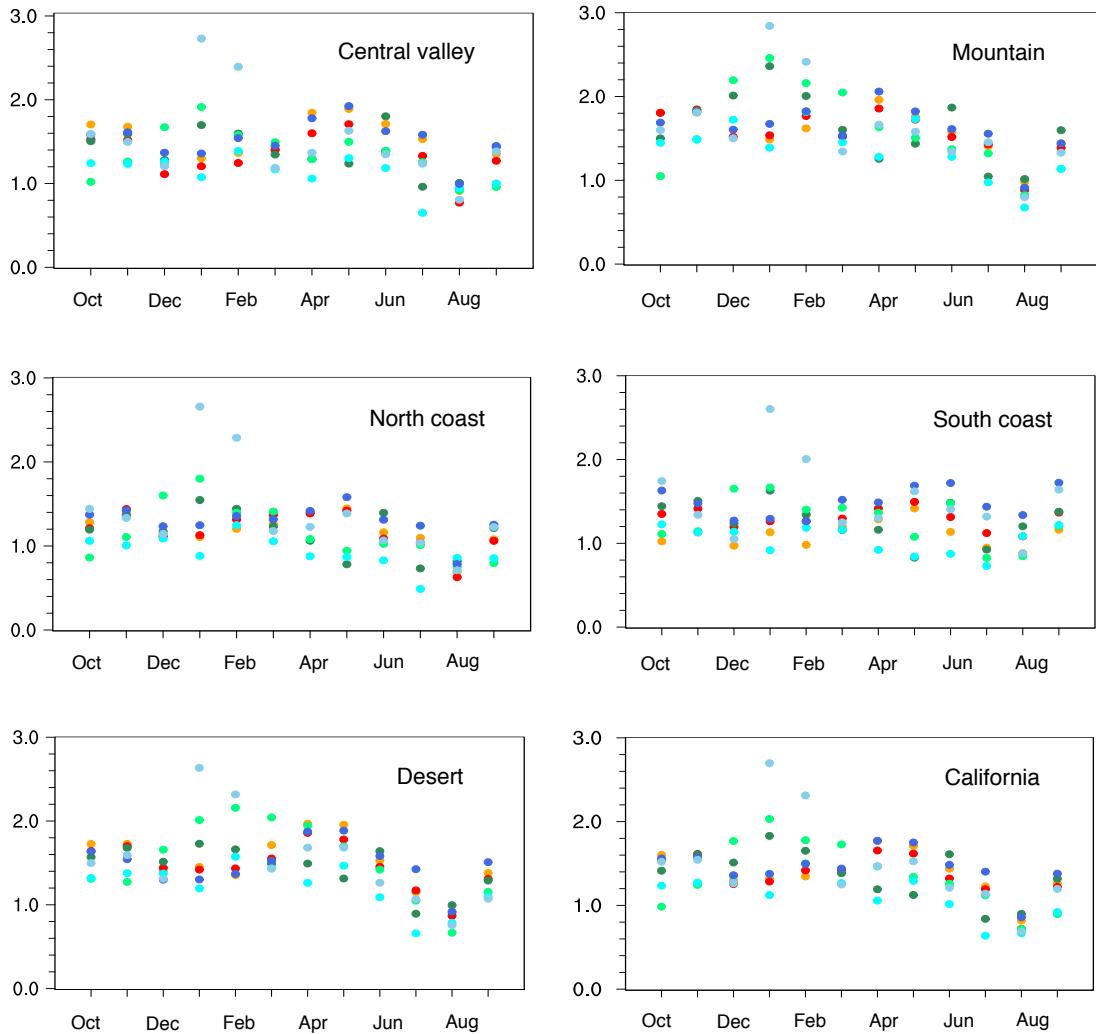


FIG. 7. Seasonal standard deviation (s) values of monthly-average T_{avg} for each subzone ($^{\circ}C$).

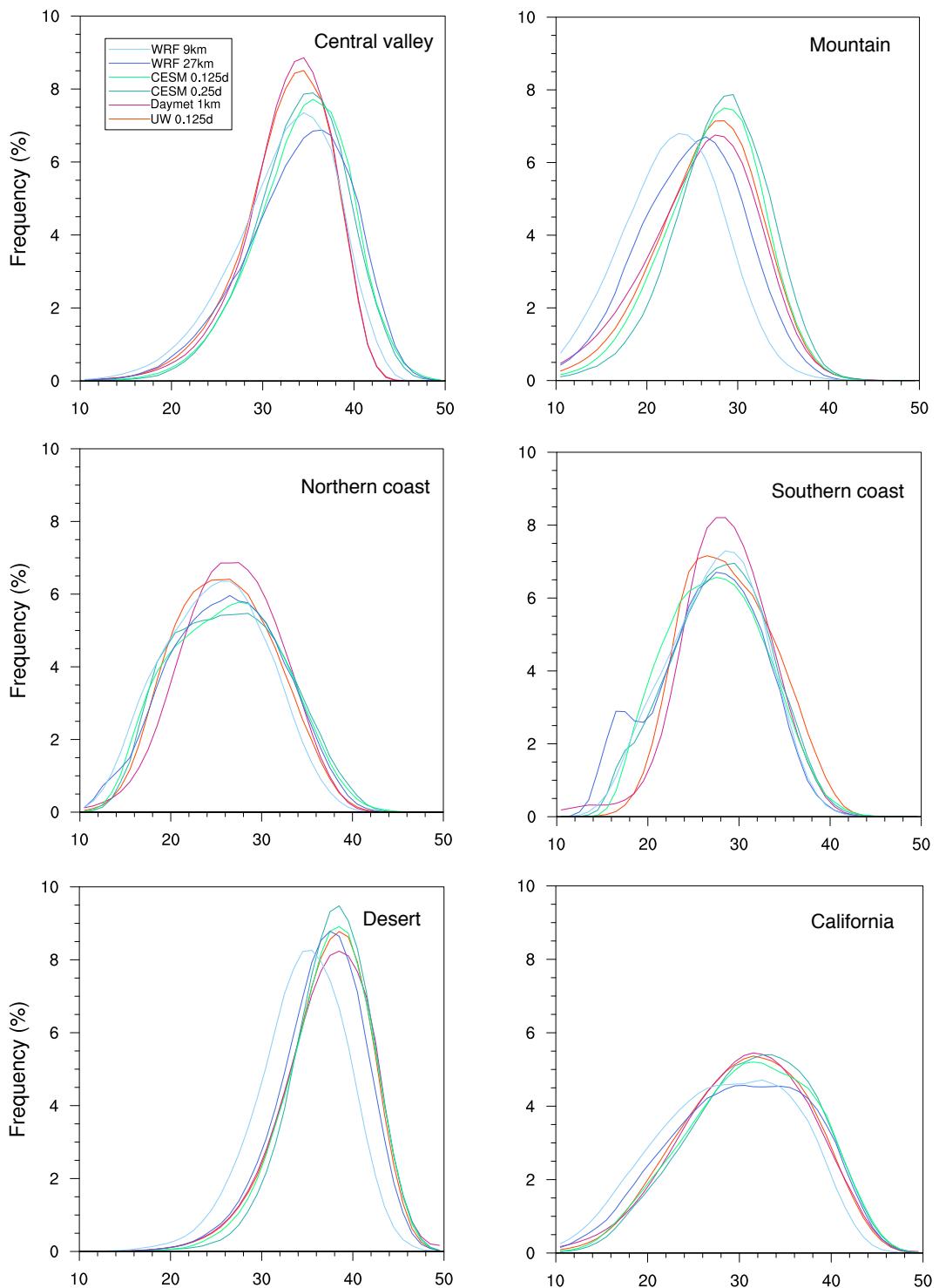
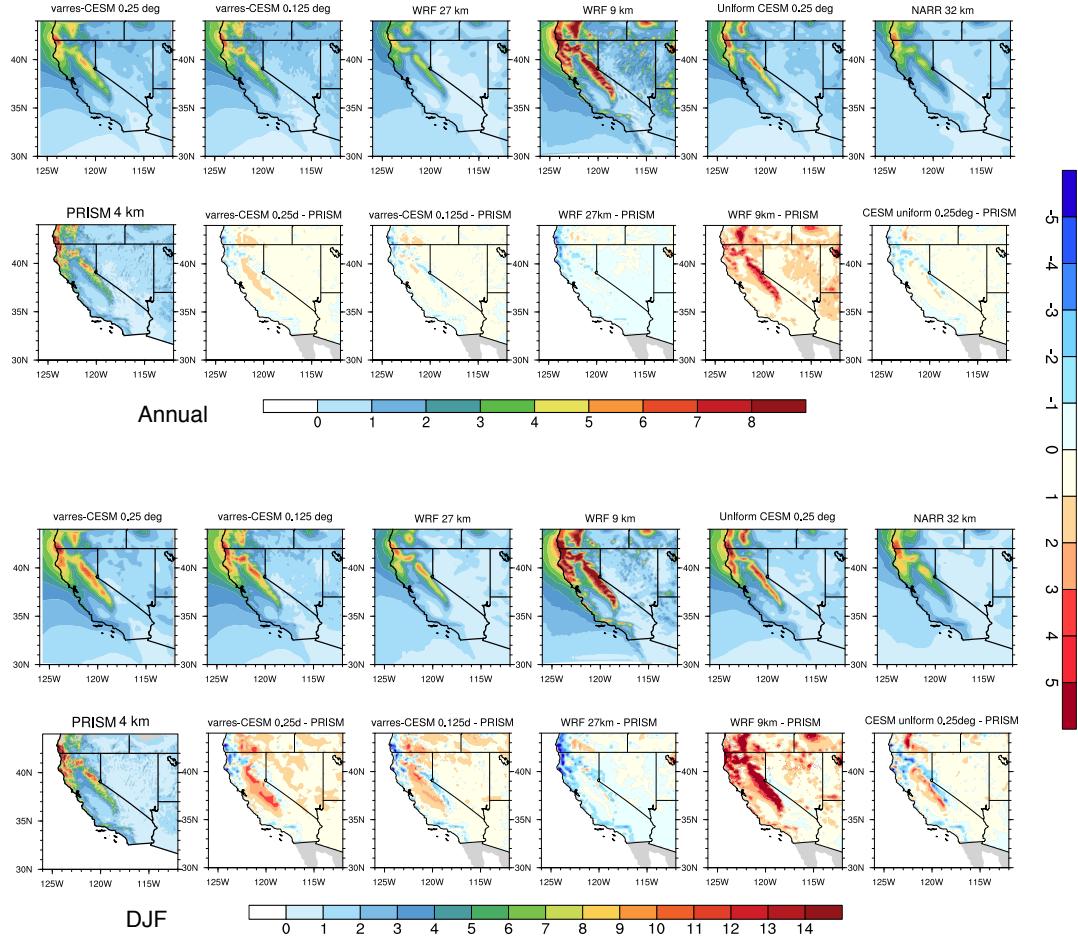


FIG. 8. Frequency distribution of summer Tmax ($^{\circ}\text{C}$).



795 FIG. 9. Annual and DJF precipitation from models and reference datasets, and differences between models
796 and PRISM (mm/d).

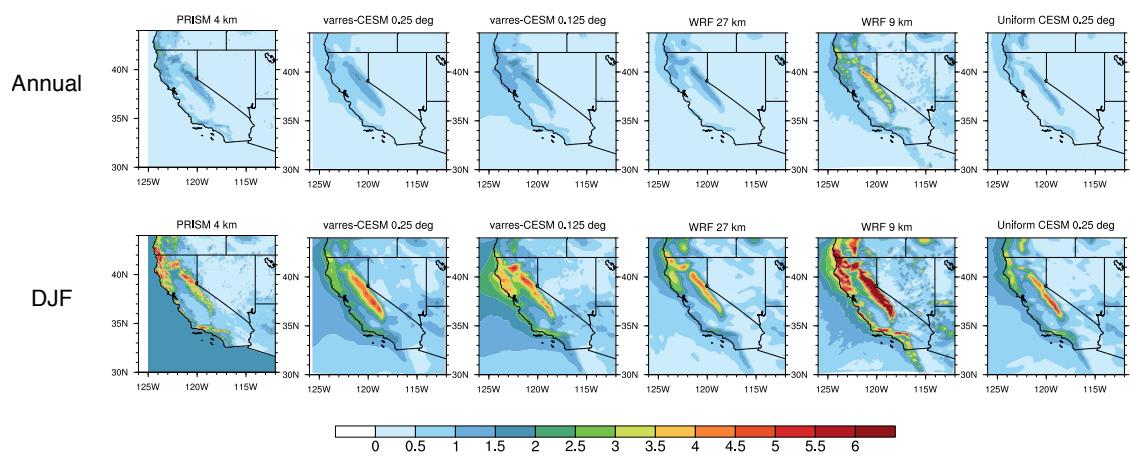
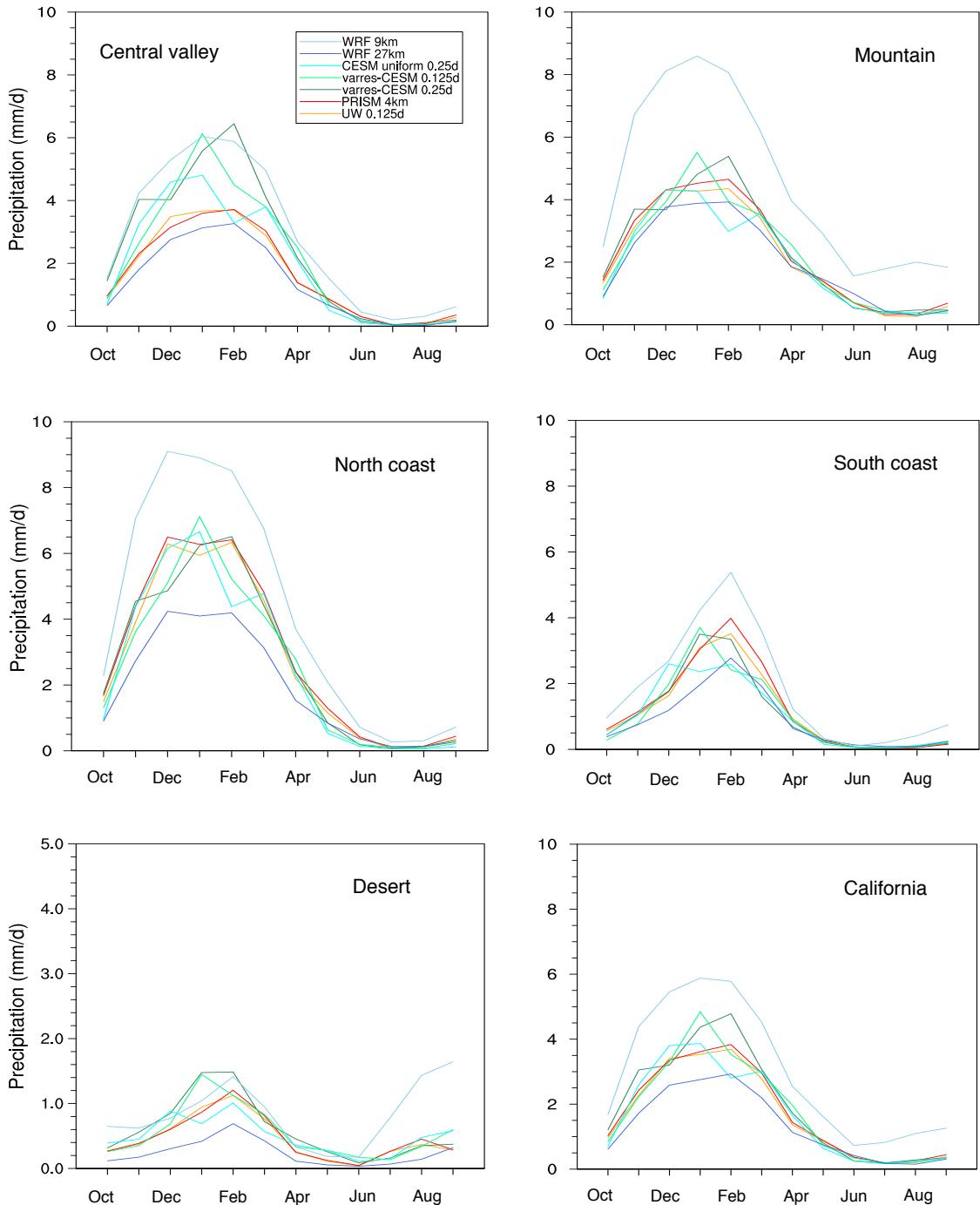


FIG. 10. Sample standard deviation of Annual and DJF precipitation from models and PRISM (mm/d).



797 FIG. 11. As Figure 6, but for monthly-average total precipitation (mm/d). **Can you add the 0.95 confidence**
798 **interval to PRISM and UW?**

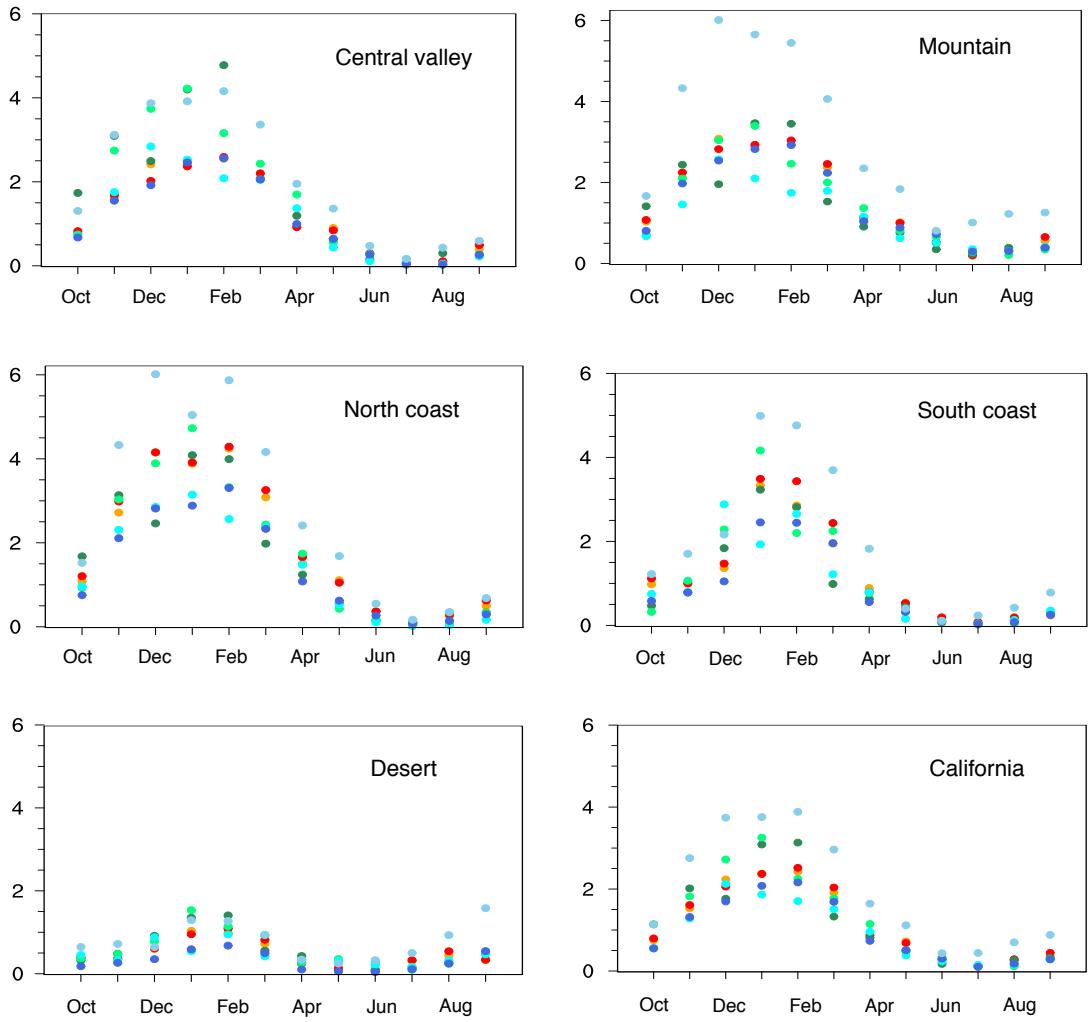
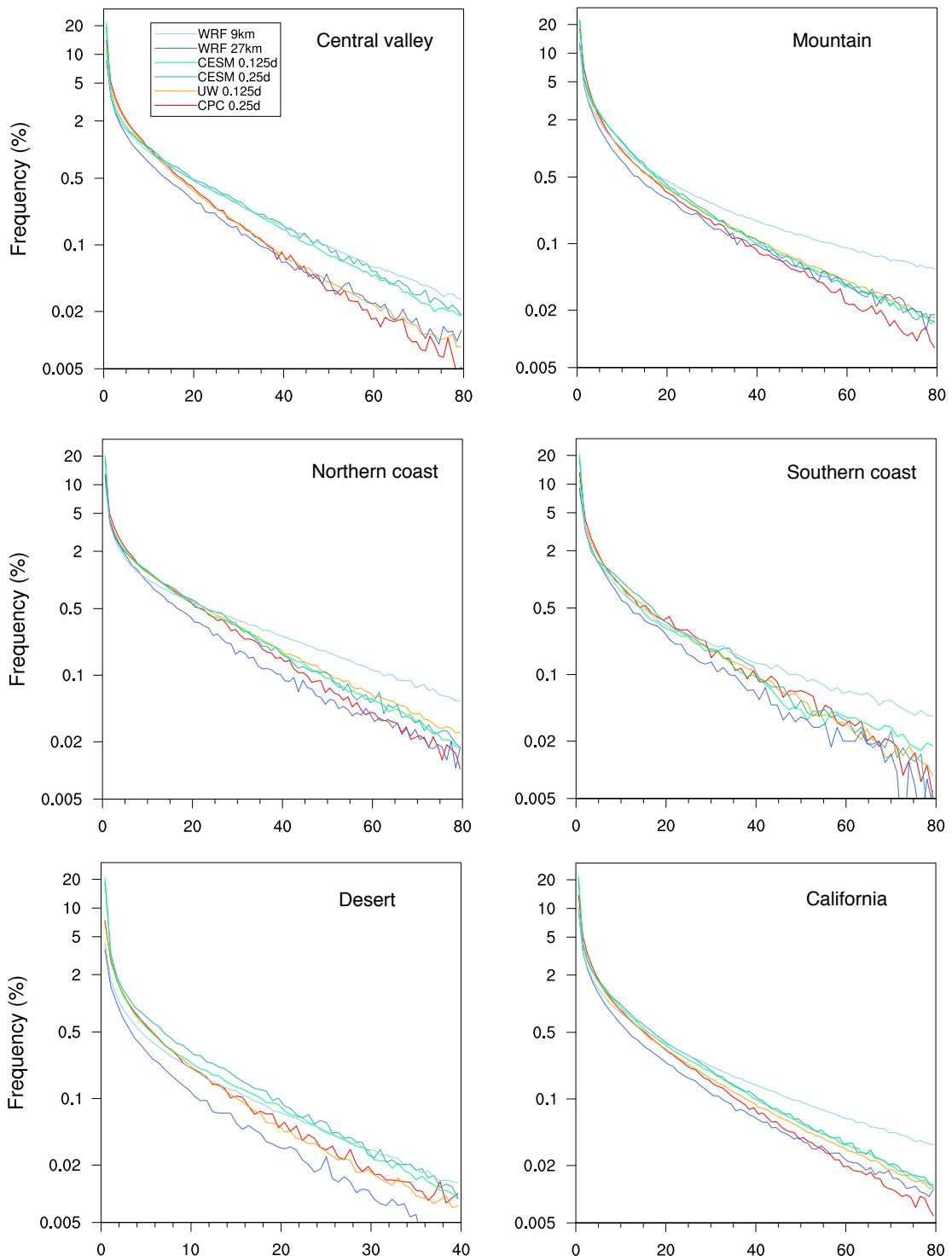
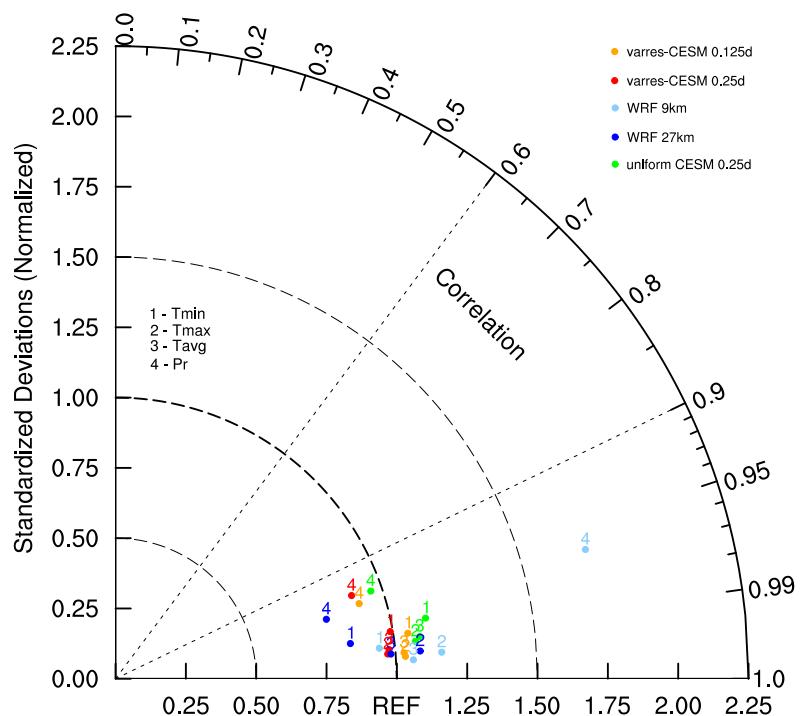


FIG. 12. As Figure 7, but for monthly-average total precipitation (mm/d).



799 FIG. 13. Frequency distribution of winter Pr, the unit of x-axis is mm/d (note that the vertical scale is loga-
800 rithmic).



801 FIG. 14. Taylor diagram of annual climatology for the entire California region, using the PRISM dataset as
 802 reference.