

**The influence of sea surface temperature and soil moisture in seasonal
predictions of air temperature over Western North America**

Daniela Faggiani Dias*, Daniel R. Cayan, Alexander Gershunov and Arthur J. Miller

Scripps Institution of Oceanography, La Jolla, California, USA

⁵ *Corresponding author: Daniela Faggiani Dias, dfaggian@ucsd.edu, Scripps Institution of
⁶ Oceanography, 9500 Gilman Drive 0206, La Jolla, CA, 92093, USA

ABSTRACT

Seasonal predictions have the potential to improve the management of different sectors of the society by anticipating climate fluctuations and possible weather extremes. Such forecasts must contend with a high level of natural variability as well as challenges posed by climate change. However, they are constrained by limited understanding of local and regional atmospheric predictability. Here, a canonical correlation analysis (CCA) prediction model of minimum and maximum air temperature anomalies (T_{min} and T_{max}) over Western North America (WNA) is developed. Remote and local predictors are used: sea surface temperature (SST) across the Pacific and local soil moisture (SM). The evaluation of the skill of predicted air temperature using historical observations indicates that CCA can provide skillful predictions for seasonal anomalies of air temperature over the region. However, skill is found to vary over seasons, location and combination of predictor and predictand variables. SST yields the best predictive skill for T_{max} and T_{min} during wintertime, but for spring and early-summer its influence is mostly on T_{min} . Remote large-scale patterns, in the form of climate indices, are captured by the CCA canonical modes and it is shown that they can be responsible for this predictive ability. On the other hand, the influence of SM is restricted to T_{max} and only during the winter, when it is shown that SM has the highest autocorrelation for the region. The results demonstrate the importance of careful analyses that consider season, variable being predicted, and predictors in forming statistical forecast models to be used for decision making.

²⁴ **1. Introduction**

²⁵ Skillful predictions of climate fluctuations at seasonal time scales have many economical and
²⁶ societal applications. Anticipating the climate fluctuations, and possible associated weather ex-
²⁷ tremes, at one or more month in advance would benefit decision-making relating to hydrology,
²⁸ agriculture, health, energy, and other sectors. Although the dynamics of the climate system display
²⁹ important nonlinear features, seasonal predictions can be made with linear statistical methods with
³⁰ skill comparable to that obtained from nonlinear statistical methods (Tang et al. 2000; Van den
³¹ Dool 2007). The linear approaches typically seek to estimate linearly related patterns of variability
³² in the predictor and the predictand fields, such as by canonical correlation analysis (CCA, Barnston
³³ 1994; Gershunov and Cayan 2003; Wilks 2008). The advantage of such statistical forecasts is that
³⁴ they can be formulated and implemented easily and they perform economically, and the forecast
³⁵ skill they produce can be evaluated in a straightforward manner. Moreover, such techniques are not
³⁶ "black boxes", in the sense that the sources of predictability can be easily identified and examined.
³⁷ The scientific basis for seasonal atmospheric climate predictability arises from the fact that vari-
³⁸ ations in slowly changing influences, such as oceanic and terrestrial conditions manifested at the
³⁹ Earth's surface, in particular those varying on time scales of several days to several months, can af-
⁴⁰ fect atmospheric circulation and, accordingly, the climate at the surface. Ocean thermal anomalies
⁴¹ often persist for or evolve systematically over several months, and many studies have shown that
⁴² slowly evolving sea surface temperature (SST) anomaly patterns can be a source of predictability
⁴³ for seasonal weather and climate anomalies a months in advance (e.g., Barnett and Preisendorfer
⁴⁴ 1987; Barnston 1994; Gershunov and Barnett 1998; Gershunov and Cayan 2003; Xoplaki et al.
⁴⁵ 2003). This is particularly important for North America, where low-frequency variability in Pa-
⁴⁶ cific SST climate modes influences atmospheric circulation over the Pacific-North America sector

⁴⁷ (Hartmann 2015; Horel and Wallace 1981; Mantua et al. 1997).

⁴⁸ Regional influences may also play a role in influencing the variability of air temperature over
⁴⁹ land. In particular, the variability of soil moisture has been shown to affect the surface energy
⁵⁰ balance (Koster et al. 2000) and thereby enhance predictability for seasonal time scales (Alfaro
⁵¹ et al. 2006; Seneviratne et al. 2010). This may be the case when the impact of SST fields on
⁵² the atmospheric variability is modest, which has been demonstrated during the summer season in
⁵³ mid-latitudes (Koster et al. 2000; Shukla 1998; Trenberth et al. 1998). Moreover, the predictability
⁵⁴ of hot extremes over most areas of South and North America, Europe, Australia, and parts of China
⁵⁵ was enhanced when antecedent precipitation was introduced (Quesada et al. 2012). Therefore, soil
⁵⁶ moisture may increase the skill for seasonal predictions at some specific locations and seasons.

⁵⁷ Alfaro et al. (2006) studied associations of Pacific SST and the Palmer Drought Severity Index
⁵⁸ (PDSI) with summer maximum and minimum temperatures over the central and western United
⁵⁹ States during 1950-2001. They found that large-scale and remote climate conditions exert their
⁶⁰ influence more strongly on minimum rather than maximum temperature. Interestingly, summer
⁶¹ surface maximum air temperatures were more predictable than those in winter, despite summer
⁶² temperatures having less spatially coherent and lower amplitude anomaly patterns than those in
⁶³ winter. This was attributed to the additive effect of soil moisture, presumably augmenting effects
⁶⁴ of large-scale patterns, represented by SST.

⁶⁵ The goal of this paper is to use a linear statistical model to evaluate the predictive skill of seasonal
⁶⁶ average maximum and minimum land surface temperature over WNA using up-to-date finely re-
⁶⁷ solved temperature and land surface moisture datasets. Additionally, the role of local and remote
⁶⁸ predictors in the predictive skill is investigated. The paper is organized as follows. Section 2
⁶⁹ describes the predictive approaches used in this study, including the description of the statistical
⁷⁰ model, the data used and the skill evaluation and comparison. Section 3 shows the results of the

71 optimized models that were used to make the seasonal predictions of Tmax and Tmin. Section 4
72 explores the climate patterns and their relationships that are responsible for the predictive skill and
73 Section 5 shows the results for the statistical forecasts. Finally, Section 6 presents a summary of
74 the main findings and some concluding remarks.

75 **2. Data and Methods**

76 The domain includes the southwestern Canada and the western United States (including the
77 Rocky Mountains), and northwestern Mexico. Detail of the domain with the topography is found
78 in Figure 1. Following results demonstrated in Alfaro et al. (2006), we adopt a linear statistical
79 approach using canonical correlation analysis. The variables and the data used as predictors and
80 predictands, as well as the details of the implemented model, are described in this section.

81 *a. Data*

82 1) PREDICTANDS

83 The variability of daytime temperatures, represented by time averages of the daily maximum
84 temperature (Tmax), differ from that of nighttime temperatures, represented by time averages of
85 daily minimum temperature (Tmin). From a statistical perspective, Tmax anomalies have higher
86 amplitude (larger daily, monthly and seasonal fluctuations) than do Tmin, especially in summer
87 months (Alfaro et al. 2006). Additionally, seasonal anomalies of Tmax and Tmin over WNA are
88 at best (in winter months) correlated moderately, and at worst (late summer-fall) correlated poorly
89 (Figure 2). This lack of strong correlation and high variability throughout the year evidently reflects
90 the different mechanisms that affect Tmax and Tmin and, quite possibly, differences in the pattern
91 and degree of their predictability in different seasons. Additionally, Tmax and Tmin anomalies
92 may have different societal consequences. From an energy utility perspective, Tmax and Tmin

93 anomalies may have different effects on the supply and on the demand for energy. For example, in
94 summer, extreme hot Tmax increase demand for electricity because of the use of air conditioners,
95 but usually occur under clear skies, which results in high solar energy production. Warmer summer
96 Tmin would diminish beneficial diurnal cooling relief from hot days and thus might increase the
97 magnitude and duration of air conditioning electrical load, which could increase system failures.
98 In winter, other kinds of impacts may occur. In cooler parts of the domain, warm Tmin anomalies
99 may lead to less space heating and thus lower electrical and natural gas demand, whereas cool
100 Tmax anomalies may provoke higher indoor energy requirements from increased space heating and
101 other indoor utility use. From a public health perspective, the expression of the impacts may differ
102 between Tmin and Tmax. Over the Southwest of North America, heat waves are becoming not
103 only stronger, longer and more extensive, but also more humid. High humidity, which exacerbates
104 the nighttime expression of heat waves, can have greater and specific impacts on public health that
105 may require different intervention strategies to manage (Gershunov et al. 2009).

106 In view of the above, in our study of air temperature predictability over the WNA (latitudes 25°N
107 to 53°N and longitudes 105°W to 125°W), we investigate a separate set of forecast models for
108 Tmax and Tmin. The 1950-2013 period is used in the model development and validation. The
109 predictands variables considered are the seasonally averaged Tmin and Tmax, respectively from
110 the Livneh CONUS near-surface gridded meteorological data (Livneh et al. 2015), provided by
111 the Earth System Research Laboratory of the National Oceanic and Atmospheric Administration
112 (ESRL/NOAA). These datasets are gridded at a spatial resolution of 1/16° by 1/16° (approximately
113 6km square) and are derived from daily temperature observations from approximately 20,000
114 NOAA Cooperative Observer (COOP) stations (Livneh et al. 2015). We use seasonal averages of
115 Tmax and Tmin anomalies because they represent short period climate fluctuations over multi-day
116 time scales. The predictand datasets are obtained by first calculating the monthly anomalies for

¹¹⁷ each variable, and then calculating seasonal means of the Tmin and Tmax anomalies for 12 different
¹¹⁸ seasons, each of those represented as a 3-months average of temperature anomalies.

¹¹⁹ 2) PREDICTORS

¹²⁰ Two fields of predictor variables are used. The first predictor field is Sea Surface Temperatures
¹²¹ (SST) anomalies for the Pacific Basin (15°S to 60°N and 135°E to 110°W), which are obtained
¹²² from the Hadley Center Sea Ice and Sea Surface Temperature (HadISST, Rayner et al. 2003).

¹²³ The HadISST data has a resolution of 1° by 1° and it can be obtained at the following website:
¹²⁴ <https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html>.

¹²⁵ The second predictor field is soil moisture (SM) anomalies, which is applied individually and also
¹²⁶ in combination with SST (SM+SST). In contrast to prior studies, that have used soil moisture
¹²⁷ indices such as the Palmer Drought Severity Index (PDSI), SM employed here is estimated from
¹²⁸ an existing hydrological model reanalysis produced by the Variable Infiltration Capacity (VIC)
¹²⁹ model, forced by observed daily precipitation and temperature. VIC is a macroscale hydrological
¹³⁰ water and energy balance accounting model (Liang et al. 1994; Cherkauer 2003) that has been
¹³¹ used in numerous studies of climate and hydrological variability and changes (e.g., Hamlet et al.
¹³² 2007; Das et al. 2009). The forcing data is the Livneh et al. (2015) gridded daily precipitation
¹³³ and temperature historical dataset, which covers the conterminous United States and adjoining
¹³⁴ portions of Canada and Mexico. This version of VIC is forced at the same resolution as the Livneh
¹³⁵ predictand fields, resulting in SM data that is distributed over the identical $1/16^{\circ}$ grid. The VIC
¹³⁶ SM domain is also identical to that of the Tmax and Tmin predictands, covering the land area from
¹³⁷ 25°N to 53°N and 105°W to 125°W .

¹³⁸ SST and SM were selected as predictors because they contain some measure of climate memory
¹³⁹ and they are broadly representative of large, regional and local scales climate measures that vary

140 over time scales that are relevant to seasonal air temperature fluctuations. While Tmax and Tmin
141 are affected by several other variables such as winds, clouds, and topographic influences, we have
142 not included those here since our purpose is to explore linear prediction using a plausible, readily
143 available, and manageable predictor set. Therefore, it is necessary to include a set of predictors
144 that contain a measure of climate memory, so that it makes sense to use antecedent observations
145 of these fields as predictors. Both predictor datasets consist of monthly means from 1950 to
146 2013, from which monthly anomalies are calculated by removing the climatological (1950–2013)
147 monthly mean for each grid cell and each month.

148 *b. Model details*

149 The statistical model framework implemented to make the seasonal predictions of air temperature
150 over the WNA uses Canonical Correlation Analysis (CCA), following an approach that has been
151 used in several prior climate prediction and analysis studies (e.g., Barnett and Preisendorfer 1987;
152 Gershunov et al. 2000; Gershunov and Cayan 2003; Alfaro et al. 2006). This method explores and
153 identifies the linear combinations between two sets of variables that have the greatest correlation
154 with each other, seeking to match patterns in the predictor fields (SST and SM) with patterns in
155 the predictand fields (Tmin and Tmax) whose temporal evolutions are optimally correlated. For
156 many climate-related phenomena, the CCA approach condenses much of the spatial and temporal
157 co-variability into a few modes and it provides a simple and cheap way to predict one field of
158 variables from another.

159 This methodology follows the hypotheses addressed by Alfaro et al. (2006): on one side, the
160 anomalous Pacific SST causes changes in the large-scale atmospheric circulation, which in turn
161 influences the land surface temperatures over a broad sector; on the other side, soil moisture
162 modulates the local air temperature via the interaction between latent and sensible heating. With

¹⁶³ those hypotheses addressed, a CCA prediction model is constructed for each season, employing
¹⁶⁴ predictors formed by one month antecedent SST and SM. For example, to make the Tmin one-
¹⁶⁵ month lag winter prediction using SST as predictor, we select all Novembers of SST and all DJFs
¹⁶⁶ of Tmin.

¹⁶⁷ The predictor and predictand fields are pre-filtered separately with the same number of p principal
¹⁶⁸ components (PCs), which are statistically orthogonal patterns of spatial and temporal variability,
¹⁶⁹ ordered by amount of variance explained. Those patterns are then related to each other using
¹⁷⁰ the q canonical correlates (CCs) extracted from the CCA analysis. The number of p PCs and
¹⁷¹ q CCs are determined using a skill optimization scheme (Section 2c). For the predictor using
¹⁷² both SST and SM combined (SST+SM), we separately calculate their p PCs and then we join
¹⁷³ those PCs together ($PC_{SST+SM} = PC_{SST} + PC_{SM}$). After that, the procedure is similar to that for
¹⁷⁴ single-variable predictors with p and q for SST+SM taken to be the sum of p and q obtained for
¹⁷⁵ SST and SM separately.

¹⁷⁶ Importantly, to avoid artificial skill from over-fitting, all results pertaining to model performance are
¹⁷⁷ cross-validated, wherein the year of a given prediction has been left out of the model development.

¹⁷⁸ Separate models are made for each year that is being predicted, in which the training period of
¹⁷⁹ each year's model is made using a period that excludes the season and adjacent months being
¹⁸⁰ predicted. This is implemented by excluding the year predicted from the training period (month
¹⁸¹ for the predictor and lagged-season for the predictand). To calculate the confidence intervals for
¹⁸² our CCA forecasts, we perform some randomized CCA experiments: first, the time series of the
¹⁸³ predictors are randomized and used to build the CCA model to forecast air temperature. This
¹⁸⁴ procedure is repeated 1,000 times and the forecast skill for each of those repetitions is calculated.
¹⁸⁵ Finally, using the skill obtained with this randomized experiments, for all these repetitions, the
¹⁸⁶ confidence intervals are calculated.

187 *c. Skill optimization*

188 Given the cross-validated forecasts, an evaluation is conducted of the yearly sequence of resulting
189 forecasts compared to the observed seasonal temperatures to estimate forecast skill. To avoid
190 artificial skill from over-fitting, the models are optimized, wherein skill over a range of predictor
191 and predictand membership is evaluated under a cross-validation scheme. For each of the forecast
192 model cases (season, predictor, and predictand), a certain number of p PCs and q CCs is determined
193 by finding the optimum model, i.e., the model complexity (the combination of PCs and CCs) that
194 maximize the forecast skill (Gershunov and Cayan 2003). The model is calculated using all
195 different combinations of PCs and CCs, from 2 to 18 and with $\text{PCs} \leq \text{CCs}$. Therefore, $p \leq q \leq T$,
196 where T is the number of years available for the training period.

197 For each option of model complexity, the cross-validated skill is estimated for each grid point and
198 then summarized in one value by averaging the correlation of all the points. In practice, we use
199 the field averaged correlation between the predicted and observed air temperature as a measure of
200 forecast skill when determining the optimum model. Therefore, for each model combination (i.e.,
201 predictor-predictand-season), a different p and q are selected to maximize the forecast skill.

202 *d. Skill evaluation*

203 We first evaluate the skill of all the CCA forecast models by calculating the anomaly correlation
204 coefficient (ACC) between each model's forecast result and the correspondent observation. We
205 analyze the seasonal variations of the predictive skill for each combination of predictor, predictand
206 and season by calculating the field-averaged ACC. The same is done for the randomized experiments
207 to verify the significance of the skill values obtained with the CCA forecasts.

208 To explore the independence of SST and SM in producing predictive skill, i.e., how much skill can
209 be added for each predictor separately, we perform some CCA forecast experiments for the residuals:

210 first, the difference between Tmin and Tmax predicted and observed fields is calculated (residuals);
211 secondly, those residuals fields are used as a set of new predictands, wherein the residuals obtained
212 with the model using SST are predicted with the SM field (SST_{res}) and likewise the residuals
213 obtained with SM are predicted with the SST field (SM_{res}); finally, the skill is calculated for each
214 of those residuals experiment and it is contrasted with the original CCA forecast skill.
215 Once we verify how the skill varies seasonally in field-averaged sense, we also explore regional
216 differences in the skill by calculating maps of ACC for the CCA models.

217 **3. Identification of the optimum model**

218 The optimum model complexity (i.e., numbers of PCs and CCs) used for each experiment is
219 presented in Table A1. Those choices were made based on the skill optimization matrices, in
220 which the field-averaged skill is calculated for each combination PCs and CCs. To reduce the
221 number of calculations required in this exercise, the number of PCs representing the predictor
222 and predictand fields is fixed as the same. As an example, Figure 3 shows the skill optimization
223 matrices for JFM using all the combinations of predictors and predictands. The skill values vary
224 smoothly as the number of PCs (x-axis) and CCs (y-axis) changes, therefore, similar predictive
225 skill is achieved by different model complexities (p PCs and q CCs). In proceeding with the model
226 experiments, in the spirit of parsimony, i.e., “simpler is better”, we chose the simplest model
227 complexity (the fewer PCs/CCs number) that yields near-optimal predictive skill.

228 Surveying the model results, the optimization produces different combinations of predic-
229 tors/predictands and canonical correlations produced for different seasons. The model complexity
230 ranges from very simple models, where few PCs and CCs results in the maximum predictive skill,
231 to more complex models, where a greater number of modes is required. For example, the FMA
232 SST-Tmin model gives the greatest skill when using only four PCs and CCs modes, while the

233 NDJ SST-Tmax is optimized when $p = 18$ and $q = 10$ (Table A1). This is a hint that the sources
234 of predictability for Tmin and Tmax vary according to the predictor and the season that is being
235 analyzed. Therefore, it appears to be important to perform this skill optimization analysis across
236 each case to maximize the potential air temperature forecast skill.

237

238 **4. Relationships between climate patterns**

239 Before exploring the predictive skill, it is important to elucidate the spatial patterns and their
240 time variations as described by the diagnostic application of CCA. We aim to provide some insight
241 into the physical sources of predictability that arises from this analysis. As shown in Section 2b,
242 CCA condenses the spatial and temporal variability of the variables that are being used to build
243 the model based on the salient patterns. As a result, we have a set of spatial patterns and their time
244 variations, known as canonical correlates or canonical modes (CCs), which provide a set of modes
245 ordered by the correlation between the CCs of the predictor (SST and SM) and the predictands
246 (Tmax and Tmin).

247 Those CC modes can be interpreted as climate-scale associations between predictors and
248 predictand and they can be tested for possible associations with commonly known climate
249 variation patterns at different timescales. To do so, we compare each CC mode from all the CCA
250 models with several well-known oceanic and atmospheric climate indices in the Pacific (Table
251 1). This comparison is made by first calculating the correlation between each climate mode
252 with the temporal variation of each CC mode for all the combination of the three predictors, two
253 predictands and 12 seasons. After that, the absolute value is calculated (since both high positive or
254 negative correlations are significant and may represent some physical association). Additionally,
255 this correlation value is weighted by the amount of variance explained in each of those CC modes.

256 With that, it is possible to get some insight about the physical patterns that may give rise to
257 seasonal predictability skill for air temperature in the WNA.

258 The climate modes considered here have been identified in numerous previous studies, capturing
259 regional to global scale variability at the ocean surface in the Pacific basin and in the atmosphere
260 near Earth's surface and aloft in the mid-troposphere. Figure 4a-f shows the maximum absolute
261 correlation value among all the canonical modes used for each CCA experiment (the number of
262 canonical modes included is shown in Figures 3). High values indicate that the time series of the
263 canonical mode, and as consequence its spatial pattern, resembles the climate mode that is being
264 analyzed.

265 Figures 4a and 4b show that throughout the year several canonical modes are closely related to
266 large-scale features in the Tropical and Extra-tropical Pacific ocean. The modes derived from
267 SST exhibit significant and substantial correlation with oceanic climate indices over a range of
268 time scales: seasonal, including the the Pacific Meridional Mode (PMM); interannual, such as
269 several indices used to represent El Niño Southern Oscillation (Niño 1.2, Niño 3.4 and Niño 4);
270 and decadal, such as the Pacific Decadal Oscillation (PDO). For both Tmin and Tmax, there are
271 significant correlation for all the seasons (black dots). However, the canonical modes derived for
272 Tmin have stronger correlation throughout the year, particularly for the winter (December and
273 January) and for the spring (April and May). In both of these seasons, the correlation is higher than
274 0.6 for all the modes cited above (PMM, ENSO and PDO). The correlations with the atmospheric
275 modes (NPI and PNA) are weaker ($\rho < 0.45$), but still significant for the late fall and winter. For
276 Tmax, significant correlations are also found for those Tropical and Extratropical SST-related
277 modes, but the strongest correlations are found for late fall and winter (November to February).

278 The order of the canonical modes that represent those known-climate variations provides insight
279 into the common variability between the predictor and the lagged-predictand and thus provide

280 a better understanding of the nature of the predictive skill. The modes in CCA are ordered
281 by the correlation between the predictor and the lagged-predictand canonical correlates. If the
282 climate patterns are represented only in lower-order canonical correlates, the variability shared
283 between predictor and predictand is probably smaller and it may decrease the ability to predict the
284 variations of air temperature for that season. To investigate that, we show in Figures 5a and 5b in
285 which canonical modes of the SST-Tmin and SST-Tmax experiments the majority of the climate
286 indices are represented. This figure presents the number of times that the correlation between
287 a canonical mode and a climate mode is larger than 0.6, for each month and canonical mode
288 number. For example, in Figure 5a, in January the second canonical mode presents correlation
289 higher than 0.6 with three climate modes (the three ENSO indices), whereas the third canonical
290 mode has correlation higher than 0.6 with one climate climate mode (PDO). In June, only the 12th
291 canonical mode has substantial correlations with two climate modes (Niño 1.2 and Niño 3.4). In
292 general, for both Tmin and Tmax, the higher number of those strong correlations occur within
293 the first five CC modes. More occurrences are concentrated in the second canonical mode during
294 the late fall and winter (for Tmax) and during the winter and the spring (for Tmin). In the spring,
295 association between those climate patterns and the Tmin occur between the second and the fifth
296 canonical modes.

297 It is important to consider how much of the variance is explained by each of the canonical modes
298 and how this is related to the climate patterns they represent. We show in Figures 4g-l the same
299 correlations of Figures 4a-f weighted by the variance explained for each canonical mode. Despite
300 the fact that the canonical modes present significant and substantial correlations with the climate
301 modes throughout the year, when considering the variance explained the significant correlations
302 only hold for the late fall and winter (for Tmax) and for winter and spring (for Tmin). This can
303 also be seen in Figures 5c and 5d, showing that high values are restricted mostly to the leading

304 CCA modes during the winter and fall. The only exception is for Tmin in April and May, when
305 the second canonical mode displays significant correlation and variance explained (Figure 5c),
306 represented exclusively by the Niño 3.4 index (Figure 4g).

307 These results are consistent with many previous studies that have explored the effect of those
308 climate patterns in the WNA. For the winter, the ENSO-PDO-related variability has been shown
309 to affect circulation anomalies that, in turn, influence synoptic weather patterns over the WNA
310 (e.g., Gershunov and Barnett 1998; Hidalgo and Dracup 2003; Thompson and Wallace 2001;
311 Guzman-Morales et al. 2016; Gershunov et al. 2017; Guirguis et al. 2019) and have implications
312 for the predictability of its hydroclimate (Gershunov and Cayan 2003). However, the influence of
313 these large-scale features in the weather of Western U.S. is not restricted to winter. Alfaro et al.
314 (2006) showed that the PDO is responsible in part for the coupled variability between the climate
315 in May and summer air temperatures over Western U.S., as well as for its predictive capability.
316 Our results for the models built with SST as predictor show that the leading CCA modes reflect the
317 influence of the Tropical and Extra Tropical Pacific ocean-atmospheric climate throughout most of
318 the year and over a range of time scales. Moreover, we show the greatest SST and air temperature
319 related variability is present in the winter, given that this is when the CCA modes present the
320 highest amount variance explained. Those CCA patterns have a physical meaning and are not
321 simply statistical artifacts and this indicates that the forecast models developed here using SST as
322 predictor rely on the effect of the large-scale climate variations over the regional variability.

323 In contrast to the SST canonical modes, those derived from soil moisture do not correlate strongly
324 with any of the climate modes (Figures 4c and 4d). As consequence, when weighting by the
325 variance explained by each mode, the values drop significantly (Figures 4i and 4j). However,
326 SM canonical modes are more strongly correlated locally with the surface air temperature
327 variability, particularly with Tmax for fall and winter, as it is shown in Figure A1. This

328 figure presents the pairs of SM and Tmin and SM and Tmax patterns that constitute the first
329 two SM-Tmin and SM-Tmax canonical modes. SM and Tmax patterns for fall and winter
330 are co-located, suggesting that the SM influence on subsequent surface temperature occurs
331 locally. This influence is quantified in Figure A2, showing the SM-Tmin and SM-Tmax patterns
332 weighted by the variance explained by each of the SM canonical modes. We observe that
333 the only cases when there is still significant values occur for Tmax during fall and winter
334 (Figure A2b). When both SST and SM are used to build the CCA model, the correlations
335 between the canonical modes and the climate indices are similar to those obtained using only
336 SST. However, the variance explained is significantly reduced. The inclusion of SM as a
337 second predictor reduces the variance explained by each of those modes (Figures 6a to 6f). Im-
338 plications for predictability and predictive skill of the CCA models are discussed in the next section.

339

340 **5. Predictive relationships and skill**

341 In this section we present an overview of the annual cycle of the cross-validated seasonal forecast
342 skill, calculated as the correlation between observed and predicted air temperature anomalies,
343 obtained with the CCA models for the different combinations of predictors (SST, SM and SST+SM)
344 and predictands (Tmin and Tmax).

345 Figure 7 shows annual cycle of the field-averaged skill of all CCA models and the field-averaged
346 skill for the randomized experiments. Skill values that are outside of the shaded area in Figure 7
347 are considered significant with 95% confidence, representing a result of deterministic processes
348 between the predictors and the predictands.

349 Remote SSTs yield more seasonal skill at one month lead time than does local soil moisture, for
350 both Tmax and Tmin (Figures 7a and 7b). Additionally, there is a noteworthy difference in the

351 annual cycle of the field-averaged skill between the two predictors: while for SM significant skill
352 occurs only during late fall and winter, SST yields significant skill throughout the year for both set
353 of predictands, with the the highest skill values occurring during late winter/early spring (FMA
354 and MAM) and summer (JJA and JAS).

355 In comparison to predictions from SST alone, the combination of SM and SST predictors add
356 only little to seasonal forecast skill (Figure 7c). However, some interesting features are found.

357 In location and seasons where the skill for the SM-CCA model is the lowest, the inclusion of
358 this predictor with SST tends to deteriorate the predictive skill. For example, during spring and
359 summer (MAM to JAS), the skill to predict Tmax and Tmin with SM is at its lowest; those seasons
360 correspond to a decrease in the skill for the SST+SM model compared to the SST-alone model
361 (Figure 7c). It may seen surprising that the inclusion of an additional predictor deteriorates the
362 skill. However, when there is not much influence, complicating the predictor field provides more
363 opportunity for CCA to find optimally correlated modes that do not explain much variance, e.g.,
364 spurious or non-physical modes. This low amount of variance explained by the SST+SM model
365 combined with the low skill achieved by the SM-alone model during spring and summer impacts
366 the ability to predict land surface temperature for those seasons. On the other hand, during the
367 winter (DJF, JFM and FMA), there is a slight improvement in the Tmax skill for the SST+SM
368 model. This season corresponds to the best skill for the SM model and it is also when the leading
369 CCA modes for SST+SM explain the greatest amount of variance (Figures 6e and 6f). This is
370 not seen in the Tmin results, suggesting that the wintertime influence of soil moisture in Tmin
371 variability is very low compared to that from Tmax.

372 The improvement in the SST+SM skill is very modest and it is much smaller than the sum of
373 the skill for SST and SM exclusive models. This raises the question whether the skill achieved
374 individually by each of those predictors is actually independent. To investigate that, we run the

375 experiments with the residuals, when SM is used to predict the residuals of the SST CCA model
376 and vice versa (see Section 2d). The results for the skill for the prediction of the residuals are
377 presented in Figures 7a and 7b, as the dashed lines. We observe that the domain-aggregated skill
378 for predicting the Tair residuals obtained from the SST-only models with SM (for both Tmax and
379 Tmin) is not statistically significant, explaining less than 1% of the variance ($\rho^2 < 0.01$) and its
380 annual cycle is very noisy (dashed lines in Figure 7a). On the other hand, the soil moisture residuals
381 predicted with SST present some noteworthy patterns. Although its skill is, in general, below the
382 significance levels, the annual cycle is less noisy than that from the SST residuals predicted with
383 SM. Moreover, the SM residuals for Tmin during the late winter and early spring can be skillfully
384 predicted by SST (dashed blue line in Figure 7c). Those results show that SM cannot predict the
385 residuals from the SST model, but SST can sometimes predict the residuals from the SM models,
386 particularly when the original skill is at its lowest. This indicates that the skill obtained from SM
387 is not completely independent of SST, contributing to the fact that combining SST and SM into a
388 single predictor does not considerably improve the forecast skill.

389 The field-averaged skill analysis shows patterns in the annual variability of the predictive skill for
390 the WNA. However, there are some regional variations in the skill pattern throughout the domain.
391 These variations are explored in the maps of the cross-validated skill obtained with the three
392 combinations of predictor fields for selected seasons (JFM, MAM, JJA and SON), and for Tmin
393 (Figure 8a) and Tmax (Figure 9a).

394 In general, these maps reflect what is seen in field-averaged skill analysis: SST is the best predictor
395 for both Tmin and Tmax for all seasons, the inclusion of SM degrades the skill for Tmin throughout
396 the year and for Tmax during Spring and summer, but it improves the predictive skill for Tmax
397 during the winter. Specifically, SM yields the best skill in predicting Tmax in the central-east and
398 south region during JFM (Figure 9a). The spatial pattern for the residuals of Tmax emphasize this

399 result: for JFM, SM can significantly predict the Tmax residuals obtained with the SST model,
400 particularly for the central-east portion of the domain (Figure 9b, frame a). This region correspond
401 to that where SST fails in predicting Tmax, leading to an improvement in the skill when SST and
402 SM are combined (Figure 9a, frame i).

403 For the seasons other than JFM (Tmax), SM does not explain a lot of neither Tmin nor Tmax
404 variability. Although there are statistically significant values, the skill is smaller than that of JFM
405 ($\rho < 0.25$) and, differently from JFM, the inclusion of SM does not improve the skill. Moreover,
406 the spatial pattern of the skill of SST residuals predicted with SM is not coherent and the values
407 are small and in general not significant (Figures 9b and 8b, frames a-d). On the other hand, for
408 Tmin, the SM residuals can be skillfully predicted with SST, particularly for the spring (MAM)
409 (Figure 8b, frame f). In fact, MAM is the season with the highest predictive skill and this skill
410 seems to be almost entirely derived from relationships with SST patterns. Among those patterns,
411 the leading canonical mode corresponds to a warming trend (Figure A4) and it accounts for almost
412 30% of the variance, the highest among all the 12 seasons (Figure A3). Indeed, MAM is the season
413 with the largest warming trend observed over WNA. Therefore, the warming trend accounts for an
414 important part of the predictive skill for the spring, particularly for Tmin.

415 Compared to models using SST predictors, those using SST + SM for spring, summer and fall
416 show that there is a decrease in the skill in predicting Tmin and Tmax, especially in regions where
417 SM skill is the lowest. This result contrasts with previous results from Alfaro et al. (2006), which
418 also used Pacific SST and an index of soil moisture as predictors of seasonal air temperatures. The
419 authors found that summertime air temperatures are more predictable than those in the winter. The
420 main reason for this discrepancy is that, during the summer, both SST, in the form of the large-scale
421 patterns, and local soil moisture provide a source for temperature variability, especially for Tmax.
422 For the winter, Alfaro et al. (2006) found that the effects of soil moisture are less important and

SST plays a more important role in the temperature variability, especially for Tmin. Here we found that SM does not improve the predictions for Tmin and only has an impact in the Tmax predictions during the winter. This disagreement with the results from Alfaro et al. (2006) is likely a result of the larger Pacific coast to the Mississippi basin domain of the their study, whose Great Plains region has been shown to harbor relatively strong soil moisture feedback to the overlying atmosphere during the summer period, while the West Coast United States does not (Huang et al. 1996; Zhang et al. 2008; Gershunov and Douville 2008; Wing et al. 2016).
The significant skill of winter Tmax forecasts from SM, to our best knowledge, has not been reported before. To understand the process that yielded those positive skills, we explore the annual variability of the SM autocorrelation at different lags over different regional domains compared to the annual cycle of the regional-averaged skill of Tmin and Tmax. Those regional domains are shown as the red boxes in Figure 1 and they were chosen based on regions where SM presented strong predictive skill for Tmax during wintertime: Central Rockies (region A), Coastal California (region B) and Inland Southwest (region C). The annual cycle of the regional-averaged skill for each of those regions is presented in Figures 10a-c for Tmin and in Figures 10d-f for Tmax. For Tmin the SST-alone model has a better predictive skill than SM and the inclusion of SM in general does not improve the skill. Moreover, the SM skill does not vary much throughout the year and the SST+SM skill variability is closely related to that from SST-alone. This is consistent with findings shown above that SM does not influence Tmin variability and therefore the predictability in any season. On the other hand, SM did exhibit association with Tmax variability. During wintertime, SM yields a significant predictive skill of Tmax, very close to the skill achieved by SST-alone. As a consequence, the combination of two predictors accounts for more skill than the individual predictors, particularly for Coastal California and the Inland Southwest during JFM. This annual variability in the SM predictive skill for those three regions is very similar to the annual cycle of

447 the SM autocorrelation in each of these regions (Figure 10g-i). These figures show the correlation
448 between the monthly and the seasonally-averaged SM at different lags. For example, January soil
449 moisture is correlated with FMA soil moisture and it corresponds to lag 1; January with MAM
450 corresponds to lag 2 and so forth. The figure shows that this autocorrelation is large and more
451 persistent during the winter months (November to January). Over WNA, soil moisture has a longer
452 memory during the winter than during the other warmer periods of the year. This is reflected in
453 the seasonal predictions, since those are the months that the Tmax predictions from SM yielded
454 significant skill.

455 **6. Summary and Conclusions**

456 We developed a statistical prediction model for seasonal Tmax and Tmin anomalies over WNA
457 using remote and local predictors. The first and most skillful predictor is sea surface temperature
458 (SST) across the tropical and northern Pacific basin, representing the influence of large-scale
459 climate variability patterns, which affect surface air temperature over land via association with
460 atmospheric circulation patterns. The second and least skillful predictor is soil moisture (SM),
461 which is thought to represent the local effects on surface air temperature that affect the surface
462 energy balance through partitioning of latent and sensible heating. We showed that both remote
463 and local predictors influence the variability and, as consequence, the predictability of local Umax
464 and Tmin over WNA. Skills shown are cross-validated and are also significant when compared to
465 a randomized experiment, indicating that the skill is real.

466 We explored the annual cycle of Tmax and Tmin predictive skill from SST, SM and from those
467 two predictors combined. The general pattern in the annual cycle of the predictive skill obtained
468 with SST is similar for both Tmin and Tmax in which the highest skill occurs during the late
469 winter and in spring, in particular for the MAM season, that is shown to be the season with the

470 highest predictive skill, in part because of the enhanced warming trend during that season. The
471 spatial variation in the skill, shown as anomaly correlation coefficient maps, is also very similar
472 for Tmin and Tmax, in which the highest skill occurs along the coast and in the northernmost
473 and southernmost parts of the domain. For both predictands, the Central region, corresponding to
474 Nevada, Utah and Colorado, presents very low predictive skill during the wintertime. During the
475 summer season, SST yields higher skill for Tmin than Tmax.

476 The leading SST-WNA temperature canonical modes reflect distinct forms of remote climate
477 variability operating via teleconnections, explored as correlations between the canonical modes
478 and several known climate indices in the Pacific ocean and atmosphere. Such patterns are strongly
479 correlated with the canonical modes throughout the year, but there is more variance explained by
480 each individual mode during the wintertime, which accounts for more predictability for both Tmin
481 and Tmax during winter. However, ENSO indices are highly correlated with the Tmin canonical
482 patterns during late spring and early summer, leading to a higher predictive skill for Tmin compared
483 to Tmax during summertime. Therefore, those large-scale climate conditions influence both Tmax
484 and Tmin during the winter, but their influence in the summer mainly affects Tmin.

485 When considering soil moisture (SM), there are some noteworthy differences between the predictive
486 ability for Tmin and Tmax. Over the WNA domain employed here, soil moisture seems to
487 not influence Tmin in any season, but it improved the predictability skill for Tmax during the
488 wintertime, particularly for JFM. This result agreed with previous findings showing that soil
489 moisture exerts higher influence on temperature variability mainly through Tmax. This result is
490 not consistent with Huang et al. (1996) and Alfaro et al. (2006) results showing that the strongest
491 influence and greatest predictability of Tmax from SM occurs in the summer. Those studies,
492 however, were made over a larger domain that included the Great Plains region in the United States
493 interior, where soil moisture has been shown to have stronger influence on surface air temperature.

494 Additionally, we found that, for the central and southern parts of WNA, soil moisture displays a
495 strong autocorrelation up to one season ahead during the winter months, therefore having more
496 potential to influence the overlying atmosphere one season ahead, which likely explains the larger
497 influence exerted on the predictive skill by soil moisture in late winter.

498 These results provide more insight into seasonal air temperature variability over WNA and advances
499 the development of a linear statistical technique for seasonal forecasting, which could be easily
500 operationalized. Evaluation of statistical methods relative to state-of-art operational dynamical
501 approaches was not the goal of this paper, but the statistical model skills are comparable and in
502 some cases superior to those from dynamical models. Exploring these comparative benefits will
503 be the goal of a future study. Finally, prediction of Tmax and Tmin separately can provide useful
504 information for stakeholders, given the distinct consequences that Tmin and Tmax have for different
505 sectors of the society.

506 *Data availability statement.* The SST dataset used during the current study is available in the
507 the Met Office Hadley Center website: <https://www.metoffice.gov.uk/hadobs/hadisst/>.
508 Livneh dataset is available in the Earth System Research Laboratory website: <https://www.esrl.noaa.gov/psd/data/gridded/data.livneh.html>. VIC dataset is available from the
509 corresponding author under reasonable request.

511 *Acknowledgments.* We thank Ben Livneh for producing and supplying the gridded historical
512 minimum and maximum temperature dataset, Dennis Lettenmaier for the VIC historical data
513 of soil moisture, and David Pierce for supplying those datasets. We also thank the Hadley
514 Center Sea Ice and Sea Surface Temperature for the SST reanalysis. Finally, we would like to
515 thank also Amato Evan and David Pierce for useful discussions. This study forms a part of
516 the Ph.D. dissertation of DFG, who was partially supported by the Brazilian National Council

517 for Scientific and Technological Development (CNPq) under the Grant 221222/2014-6 and by
518 the California Energy Commission under Agreement Number PIR-15-005. Additional funding
519 from the NSF (OCE1419306) and the National Oceanic and Atmospheric Administration (NOAA;
520 NA17OAR4310106) is greatly appreciated.

521 **APPENDIX**

522 **Supplementary tables and figures**

523 Figures A1 to A4 and Table A1 provide further details of the analyses and results presented in
524 the text.

525 **References**

- 526 Alfaro, E. J., A. Gershunov, D. Cayan, E. J. Alfaro, A. Gershunov, and D. Cayan, 2006: Prediction
527 of Summer Maximum and Minimum Temperature over the Central and Western United States:
528 The Roles of Soil Moisture and Sea Surface Temperature. *Journal of Climate*, **19** (8), 1407–1421,
529 doi:10.1175/JCLI3665.1, URL <http://journals.ametsoc.org/doi/abs/10.1175/JCLI3665.1>.
- 530 Baldwin, M. P., and Coauthors, 2001: The Quasi-Biennial Oscillation. *Reviews of Geophysics*,
531 **29** (2), 179–229.
- 532 Barnett, T. P., and R. Preisendorfer, 1987: Origins and Levels of Monthly and Seasonal Forecast
533 Skill for United States Surface Air Temperatures Determined by Canonical Correlation Anal-
534 ysis. *Monthly Weather Review*, **115** (9), 1825–1850, doi:10.1175/1520-0493(1987)115<1825:
535 OALOMA>2.0.CO;2.
- 536 Barnston, A. G., 1994: Linear Statistical Short-Term Climate Predictive Skill in the Northern
537 Hemisphere. *Journal of Climate*, **7** (10), 1513–1564.

- 538 Barnston, A. G., and R. E. Livezey, 1987: Classification, Seasonality and Persistence of Low-
539 Frequency Atmospheric Circulation Patterns. *Monthly Weather Review*, **115** (6), 1083–1126.
- 540 Cherkauer, K. A., 2003: Simulation of spatial variability in snow and frozen soil. *Journal of*
541 *Geophysical Research*, **108** (D22), 8858, doi:10.1029/2003JD003575.
- 542 Chiang, J. C. H., and D. J. Vimont, 2004: Analogous Pacific and Atlantic Meridional Modes
543 of Tropical Atmosphere – Ocean Variability. *Journal of Climate*, **17**, 4143–4158, doi:10.1175/
544 JCLI4953.1.
- 545 Das, T., and Coauthors, 2009: Structure and Detectability of Trends in Hydrological Measures
546 over the Western United States. *Journal of Hydrometeorology*, **10** (4), 871–892, doi:10.1175/
547 2009JHM1095.1.
- 548 Di Lorenzo, E., and Coauthors, 2008: North Pacific Gyre Oscillation links ocean climate and
549 ecosystem change. *Geophysical Research Letters*, **35** (8), L08 607, doi:10.1029/2007GL032838.
- 550 Gershunov, A., and T. P. Barnett, 1998: Interdecadal Modulation of ENSO Teleconnec-
551 tions. *Bulletin of the American Meteorological Society*, **79** (12), 2715–2725, doi:10.1175/
552 1520-0477(1998)079<2715:IMOET>2.0.CO;2.
- 553 Gershunov, A., T. P. Barnett, D. R. Cayan, T. Tubbs, and L. Goddard, 2000: Predicting and
554 downscaling ENSO impacts on intraseasonal precipitation statistics in California: The 1997/98
555 event. *Journal of Hydrometeorology*, **1** (3), 201–210, doi:10.1175/1525-7541(2000)001<0201:
556 PADEIO>2.0.CO;2.
- 557 Gershunov, A., and D. R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous
558 United States: Sources of climatic variability and seasonal predictability. *Journal of Climate*,
559 **16** (16), 2752–2765, doi:10.1175/1520-0442(2003)016<2752:HDPFOT>2.0.CO;2.

- 560 Gershunov, A., D. R. Cayan, and S. F. Iacobellis, 2009: The Great 2006 Heat Wave over California
561 and Nevada : Signal of an Increasing Trend. *Journal of Climate*, **22** (2008), 6181–6203, doi:
562 10.1175/2009JCLI2465.1.
- 563 Gershunov, A., and H. Douville, 2008: Extensive summer hot and cold extremes under current and
564 possible future climatic conditions : Europe and North America. *Climate Extremes and Society*,
565 74–98.
- 566 Gershunov, A., T. Shulgina, M. Ralph, D. A. Lavers, and J. J. Rutz, 2017: Assessing the climate-
567 scale variability of atmospheric rivers affecting western North America. *Geophysical Research
Letters*, **(44)**, 7900–7908, doi:10.1002/2017GL074175.
- 568 Guirguis, K., A. Gershunov, T. Shulgina, R. E. S. Clemesha, and F. M. Ralph, 2019: Atmospheric
569 rivers impacting Northern California and their modulation by a variable climate. *Climate Dy-
570 namics*, **52** (11), 6569–6583, doi:10.1007/s00382-018-4532-5.
- 571 Guzman-Morales, J., A. Gershunov, J. Theiss, H. Li, and D. Cayan, 2016: Santa Ana Winds of
572 Southern California: Their climatology, extremes, and behavior spanning six and a half decades.
573 *Geophysical Research Letters*, **43** (6), 2827–2834, doi:10.1002/2016GL067887.
- 574 Hamlet, A. F., P. W. Mote, M. P. Clark, and D. P. Lettenmaier, 2007: Twentieth-century trends in
575 runoff, evapotranspiration, and soil moisture in the western United States. *Journal of Climate*,
576 **20** (8), 1468–1486, doi:10.1175/JCLI4051.1.
- 577 Hartmann, D. L., 2015: Pacific sea surface temperature and the winter of 2014. *Geophysical
Research Letters*, **42**, 1894–1902, doi:10.1002/2015GL063083.

- 580 Hidalgo, H. G., and J. a. Dracup, 2003: ENSO and PDO Effects on Hydroclimatic Variations
581 of the Upper Colorado River Basin. *Journal of Hydrometeorology*, **4** (1), 5–23, doi:10.1175/
582 1525-7541(2003)004<0005:EAPEOH>2.0.CO;2.
- 583 Horel, J. D., and J. >. Wallace, 1981: Planetary-Scale Atmospheric Phenomena Associated with
584 the Southern Oscillation. *Monthly Weather Review*, **109**, 813–829.
- 585 Huang, J., H. Van Den Dool, and K. P. Georgakakos, 1996: Analysis of Model-Calculated Soil
586 Moisture over the United States (1931 - 1993) and Applications to Long-Range Temperature
587 Forecasts. *Journal of Climate*, **9**, 1350–1362.
- 588 Koster, R. D., M. J. Suarez, and M. Heiser, 2000: Variance and Predictability of Precipitation
589 at Seasonal-to-Interannual Timescales. *Journal of Hydrometeorology*, **1**, 26–46, doi:10.1175/
590 1525-7541%282000%29001.
- 591 Liang, X., D. P. Lettenmaier, E. F. Wood, and J. Burges, 1994: A simple hydrologically based model
592 of land surface water and energy fluxes for general circulation models. *Journal of Geophysical
593 Research*, **99** (D7), 14 415–14 428, doi:10.1029/94JD00483.
- 594 Livneh, B., T. J. Bohn, D. W. Pierce, F. Munoz-ariola, B. Nijssen, R. Vose, D. R. Cayan, and
595 L. Brekke, 2015: A spatially comprehensive , hydrometeorological data set for Mexico , the
596 U . S ., and Southern Canada 1950 – 2013. *Nature Scientific Data*, **2** (150042), 1–12, doi:
597 10.1038/sdata.2015.42.
- 598 Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific Interdecadal
599 Climate Oscillation with Impacts on Salmon Production. *Bulletin of the American Meteorological
600 Society*, **78** (January), 1069–1079.

- 601 Quesada, B., R. Vautard, P. Yiou, M. Hirschi, and S. I. Seneviratne, 2012: Asymmetric Eu-
602 ropean summer heat predictability from wet and dry southern winters and springs. *Nature*
603 *Climate Change*, **2 (10)**, 736–741, doi:10.1038/nclimate1536, URL <http://dx.doi.org/10.1038/nclimate1536>.
- 604
- 605 Rayner, N. A., D. E. Parker, E. B. Horton, C. K. Folland, L. V. Alexander, D. P. Rowell, E. C. Kent,
606 and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air
607 temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*,
608 **108 (D14)**, 1–37, doi:10.1029/2002JD002670, URL <http://dx.doi.org/10.1029/2002JD002670>.
- 609
- 610 Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J.
611 Teuling, 2010: Investigating soil moisture-climate interactions in a changing climate: A review.
Earth-Science Reviews, **99 (3-4)**, 125–161, doi:10.1016/j.earscirev.2010.02.004.
- 612
- 613 Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting.
Science, **282 (5389)**, 728–731, doi:10.1126/science.282.5389.728.
- 614
- 615 Tang, B., W. W. Hsieh, A. H. Monahan, and F. T. Tangang, 2000: Skill comparisons between neural
616 networks and canonical correlation analysis in predicting the equatorial Pacific sea surface
617 temperatures. *Journal of Climate*, **13 (1)**, 287–293, doi:10.1175/1520-0442(2000)013<0287:
SCBNNA>2.0.CO;2.
- 618
- 619 Thompson, D. W. J., and J. M. Wallace, 2001: Regional Climate Impacts of the Northern Hemisphere
Annular Mode. *Science*, **293 (July)**, 85–90.
- 620
- 621 Trenberth, K. E., G. W. Branstator, D. Karoly, A. Kumar, N.-C. Lau, and C. Ropelewski, 1998:
Progress during TOGA in understanding and modeling global teleconnections associated with

- 622 tropical sea surface temperatures. *Journal of Geophysical Research: Oceans*, **103 (C7)**, 14 291–
623 14 324, doi:10.1029/97JC01444.
- 624 Trenberth, K. E., and J. W. Hurrell, 1994: Decadal atmosphere-ocean variations in the Pacific.
625 *Climate Dynamics*, **9**, 303–319.
- 626 Van den Dool, H. M., 2007: *Empirical methods in short-term climate prediction*. Oxford University
627 Press, 215 pp., doi:10.1029/2005GL023422, arXiv:1011.1669v3.
- 628 Wilks, D. S., 2008: Improved statistical seasonal forecasts using extended training data. *Internation-
629 al Journal of Climatology*, **28**, 1589–1598, doi:10.1002/joc.1661.
- 630 Wing, G. Y. K., L. Sushama, and G. T. Diro, 2016: The intraannual variability of land-atmosphere
631 coupling over North America in the Canadian Regional Climate Model (CRCM5). *Journal of
632 Geophysical Research: Atmospheres*, **121 (23)**, 13–859, doi:10.1002/2016JD025423.
- 633 Xoplaki, E., J. F. González-Rouco, J. Luterbacher, and H. Wanner, 2003: Mediterranean summer
634 air temperature variability and its connection to the large-scale atmospheric circulation and SSTs.
635 *Climate Dynamics*, **20 (7-8)**, 723–739, doi:10.1007/s00382-003-0304-x.
- 636 Zhang, J., W.-c. Wang, and L. R. Leung, 2008: Contribution of land-atmosphere coupling to
637 summer climate variability over the contiguous United States. *Journal of Geophysical Research*,
638 **113 (August)**, 1–15, doi:10.1029/2008JD010136.

639 LIST OF TABLES

640	Table 1. Description of each climate index used in comparison with canonical correlates	31
641	in Figure 4	
642	Table A1. Optimum number of Principal Components (PCs) and Canonical Correlates	32
643	(CCs) for all combinations of predictors and predictands	

TABLE 1: Description of each climate index used in comparison with canonical correlates in Figure 4

Index	Name	Description
Nino1.2	ENSO	Extreme Eastern Tropical Pacific SST: SST anomalies between 0 – 10°S, 90°W – 80°W
Nino3.4	ENSO	East Central Tropical Pacific SST: SST anomalies between 5°N – 5°S, 160°E – 150°W
Nino4	ENSO	Central Tropical Pacific SST: SST anomalies between 5°N – 5°S, 170°W – 120°W
PDO	Pacific Decadal Oscillation	The leading principal component of SST anomalies in the North Pacific (Mantua et al. 1997)
NPGO	North Pacific Gyre Oscillation	The second dominant mode of sea surface height variability in the Northeast Pacific (Di Lorenzo et al. 2008)
PMM	Pacific Meridional Mode	Meridional variability in the tropical Pacific Ocean. It is defined as the first mode of a maximum covariance analysis (MCA) of SST and the zonal and meridional components of the 10m wind field, over the region defined between 21°S-32°N and 74°W-15°E (Chiang and Vimont 2004)
NPI	North Pacific Index	The area-weighted sea level pressure over the region 30°N-65°N,160°E-140°W (Trenberth and Hurrell 1994)
PNA	Pacific North American Pattern	Rotated Principal Component Analysis applied to monthly mean standardized 500-mb height anomalies over the region between 20°N-90°N (Barnston and Livezey 1987)
QBO	Quasi-Biennale Oscillation	Oscillation of the equatorial wind between easterlies and westerlies in the tropical stratosphere (Baldwin et al. 2001)

TABLE A1: Optimum number of Principal Components (PCs) and Canonical Correlates (CCs) for all combinations of predictors and predictands

SST-Tmin												
	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	JFM
PCs	4	7	5	5	7	14	7	12	15	17	10	7
CCs	4	4	5	5	4	11	3	5	14	7	9	4
SST-Tmax												
	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	JFM
PCs	6	6	11	7	9	13	9	15	11	18	9	7
CCs	4	4	8	7	8	5	3	11	9	10	4	6
SM-Tmin												
	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	JFM
PCs	13	8	12	13	10	7	6	3	2	9	12	7
CCs	6	4	5	11	6	5	5	3	2	7	5	2
SM-Tmax												
	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ	DJF	JFM
PCs	8	5	14	12	10	5	16	5	15	16	12	3
CCs	3	4	11	6	8	3	8	2	7	6	4	2

644 LIST OF FIGURES

- Fig. 1.** Region defined in this study as the WNA and the topography (in meters). Red boxes correspond to the three regional indexes used, named: A) Central Rockies; B) Coastal California; and C) Inland Southwest.

Fig. 2. Correlation between seasonal averages of Tmax and Tmin anomalies averaged over the WNA. Each season corresponds to a 3-months average of Tmax and Tmin anomalies, represented as: **FMA**: February, March and April; **MAM**: March, April and May; **AMJ**: April, May and June; **MJJ**: May, June and July; **JJA**: June, July and August; **JAS**: July, August and September; **ASO**: August, September and October; **SON**: September, October and November; **OND**: October, November and December; **NDJ**: November, December and January; **DJF**: December, January and February; **JFM**: January, February and March.

Fig. 3. Skill optimization matrices for JFM season for CCA forecasts using different combinations of predictors (SST and SM), predictands (Tmin and Tmax), and seasons (FMA to JFM). Each matrix represents one specific predictor-predictand combination (named in the top left of each individual matrix). Color scale represents the cross-validated field-averaged skill, expressed as the correlation between the predicted and the observed temperatures, using different combination of the Principal Components (PCs) and the Canonical Correlates (CCs). The optimum is defined the number of PC and CC modes that yielded the highest cross validated skill (black dots in each matrix).

Fig. 4. Maximum absolute correlation between all the canonical modes (CCs) from each CCA experiment and selected climate modes (**a - f**); maximum absolute correlation weighted by the variance explained by each canonical mode (**g - l**). Each matrix is a combination of a predictor (SST, SM and SST+SM) and a predictand (tmin and tmax) wherein the columns represent the months that each CCA model was initialized (January to December) and the lines represent the climate modes (named along the lines of each matrix and described in the Table 1). Black dots indicate values that have a statistical significance greater than 99%.

Fig. 5. Number of occurrences that the correlation between the SST canonical modes and the climate mode is bigger than a threshold, for each initialization month (y-axis), and each canonical mode number (x-axis). (**a**) and (**b**) bring the results for Tmin and Tmax, respectively, not considering the variance explained and which the threshold value is 0.6; (**c**) and (**d**) bring the results for Tmin and Tmax, respectively, weighted by the amount of variance explained for each canonical mode and which threshold value is 0.2.

Fig. 6. Variance explained by each canonical mode from the CCA models, averaged by seasons. The colors show the values referent for each season: winter includes those experiments initialized from December to February; spring, from March to May; Summer, from June to August; and fall, from September to November.

Fig. 7. Annual cycle of field-averaged skill (average skill among all the grid points) obtained from the CCA model run at one month lead time using (a) SST as predictor, (b) SM as predictor, and (c) SST and SM combined. Solid blue line is the skill for Tmin and solid red line is the skill for Tmax. Shaded areas are the confidence interval obtained from the randomized experiments: values that fall outside this shaded area have statistical significance greater than 95%. Dashed lines are the annual cycle for the field-averaged skill of the residuals, where in (a) are the residuals from SST predicted with SM and in (b) are the residuals from SM predicted with SST (see Section 2d for details).

688	Fig. 8. Spatial pattern of Tmin forecast skill for selected seasons (shown above the maps in the upper panels). (a) Tmin seasonal forecast skill obtained with the model using SST (top panels), SM (middle panels) and SST+SM (lower panels) as predictors; (b) SST and SM Tmin residuals forecast skill. SST-Tmin residuals forecast using SM as predictor (upper panels) and SM-Tmin residuals forecast using SST as predictor (lower panels). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in the WNA. Black dots represent values that have statistical significance greater than 95%.	42
689		
690		
691		
692		
693		
694		
695	Fig. 9. Same as Figure 8, but for Tmax as predictand	43
696		
697	Fig. 10. Annual cycle of seasonal predictive skill for (a-c) Tmin and (d-f) Tmax, and (g-i) annual cycle of soil moisture memory for the three regional domains shown in Figure 1.	44
698		
699		
700	Fig. A1. First (upper panels) and second(lower panels) soil moisture canonical modes for selected seasons (JFM, MAM, JJA and SON) correlated with (a) Tmin and (b) Tmax time series. Black dots indicate values that have a statistical significance greater than 99%	45
701		
702	Fig. A2. Same as Figure A1, but with the projected value weighted by the variance explained by each of soil moisture canonical mode.	46
703		
704	Fig. A3. Variance explained by the leading mode of the CCA model using SST as a predictor and Tmin as a predictand, for the 12 seasons analyzed	47
705		
706		
707	Fig. A4. Leading canonical mode for the SST-Tmin model for MAM. (a) Time series related to this leading canonical mode. The value ρ corresponds to the correlation between these two time series; (b) corresponding spatial pattern for the predictor (SST) and (c) for the predictand (Tmin).	48
708		

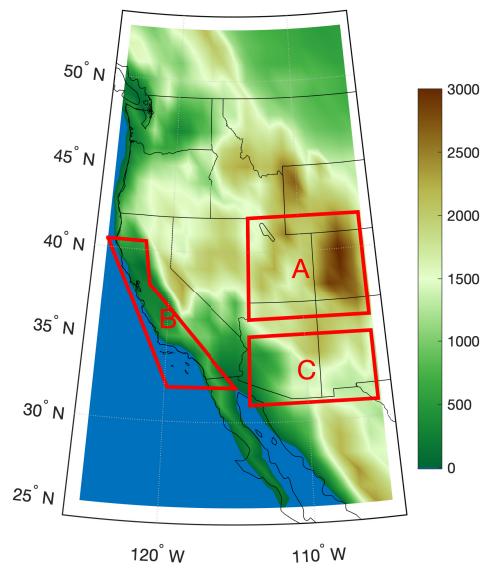


FIG. 1: Region defined in this study as the WNA and the topography (in meters). Red boxes correspond to the three regional indexes used, named: A) Central Rockies; B) Coastal California; and C) Inland Southwest.

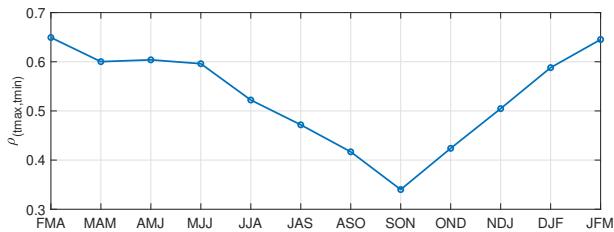


FIG. 2: Correlation between seasonal averages of Tmax and Tmin anomalies averaged over the WNA. Each season corresponds to a 3-months average of Tmax and Tmin anomalies, represented as: **FMA**: February, March and April; **MAM**: March, April and May; **AMJ**: April, May and June; **MJJ**: May, June and July; **JJA**: June, July and August; **JAS**: July, August and September; **ASO**: August, September and October; **SON**: September, October and November; **OND**: October, November and December; **NDJ**: November, December and January; **DJF**: December, January and February; **JFM**: January, February and March.

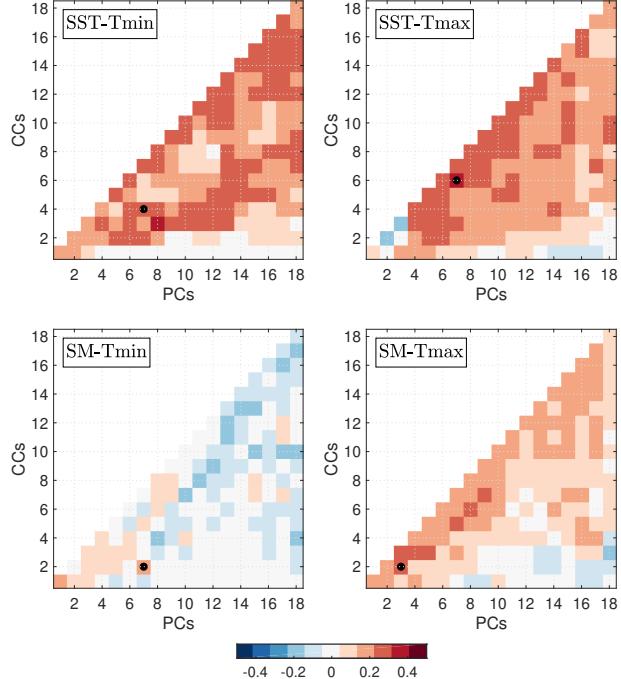


FIG. 3: Skill optimization matrices for JFM season for CCA forecasts using different combinations of predictors (SST and SM), predictands (Tmin and Tmax), and seasons (FMA to JFM). Each matrix represents one specific predictor-predictand combination (named in the top left of each individual matrix). Color scale represents the cross-validated field-averaged skill, expressed as the correlation between the predicted and the observed temperatures, using different combination of the Principal Components (PCs) and the Canonical Correlates (CCs). The optimum is defined the number of PC and CC modes that yielded the highest cross validated skill (black dots in each matrix).

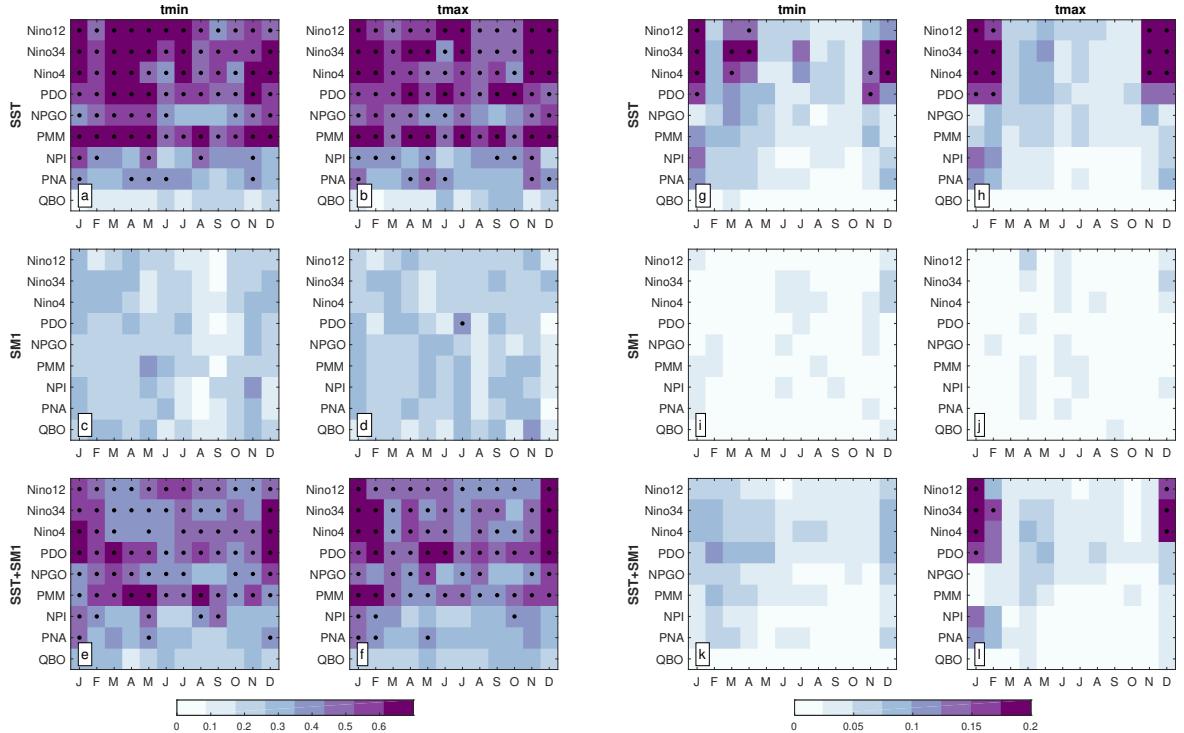


FIG. 4: Maximum absolute correlation between all the canonical modes (CCs) from each CCA experiment and selected climate modes (**a - f**); maximum absolute correlation weighted by the variance explained by each canonical mode (**g - l**). Each matrix is a combination of a predictor (SST, SM and SST+SM) and a predictand (tmin and tmax) wherein the columns represent the months that each CCA model was initialized (January to December) and the lines represent the climate modes (named along the lines of each matrix and described in the Table 1). Black dots indicate values that have a statistical significance greater than 99%.

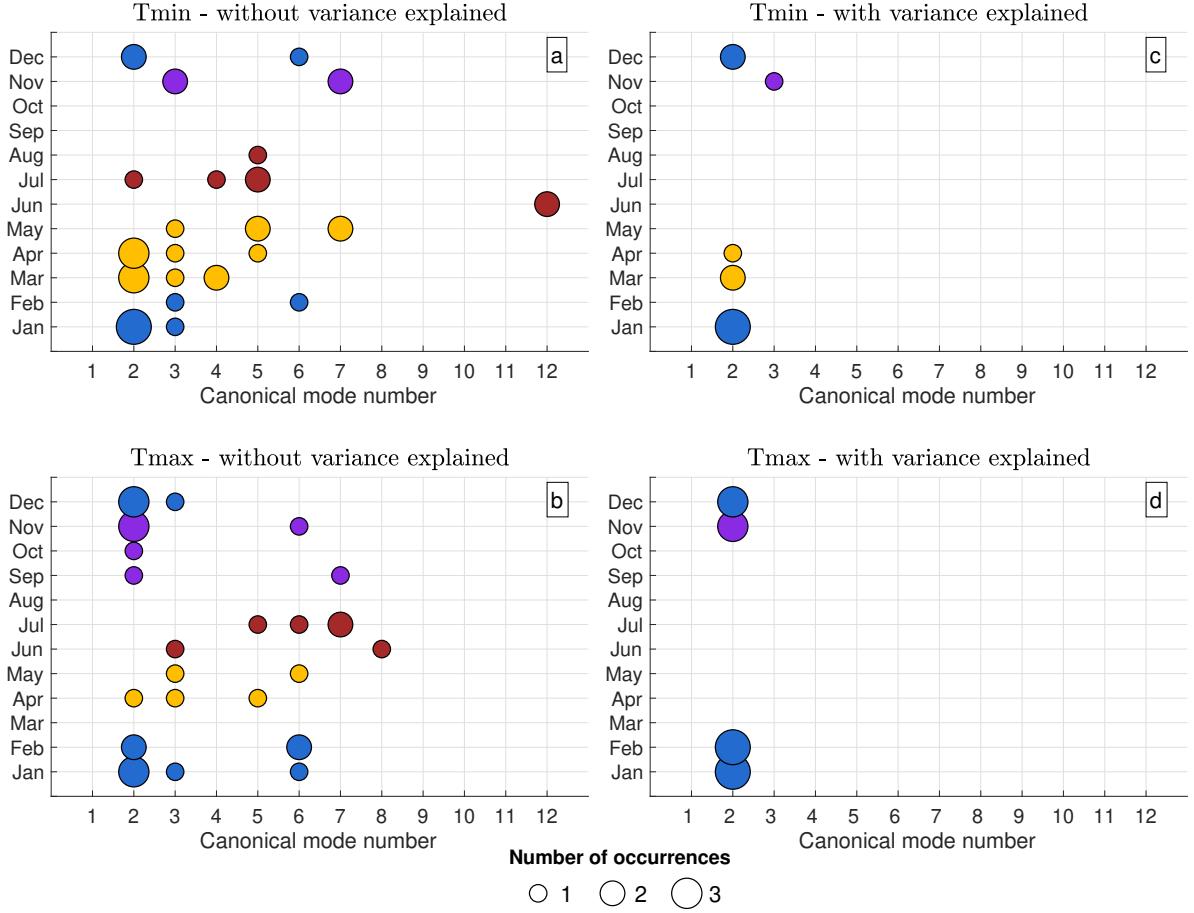


FIG. 5: Number of occurrences that the correlation between the SST canonical modes and the climate mode is bigger than a threshold, for each initialization month (y-axis), and each canonical mode number (x-axis). (a) and (b) bring the results for T_{\min} and T_{\max} , respectively, not considering the variance explained and which the threshold value is 0.6; (c) and (d) bring the results for T_{\min} and T_{\max} , respectively, weighted by the amount of variance explained for each canonical mode and which threshold value is 0.2.

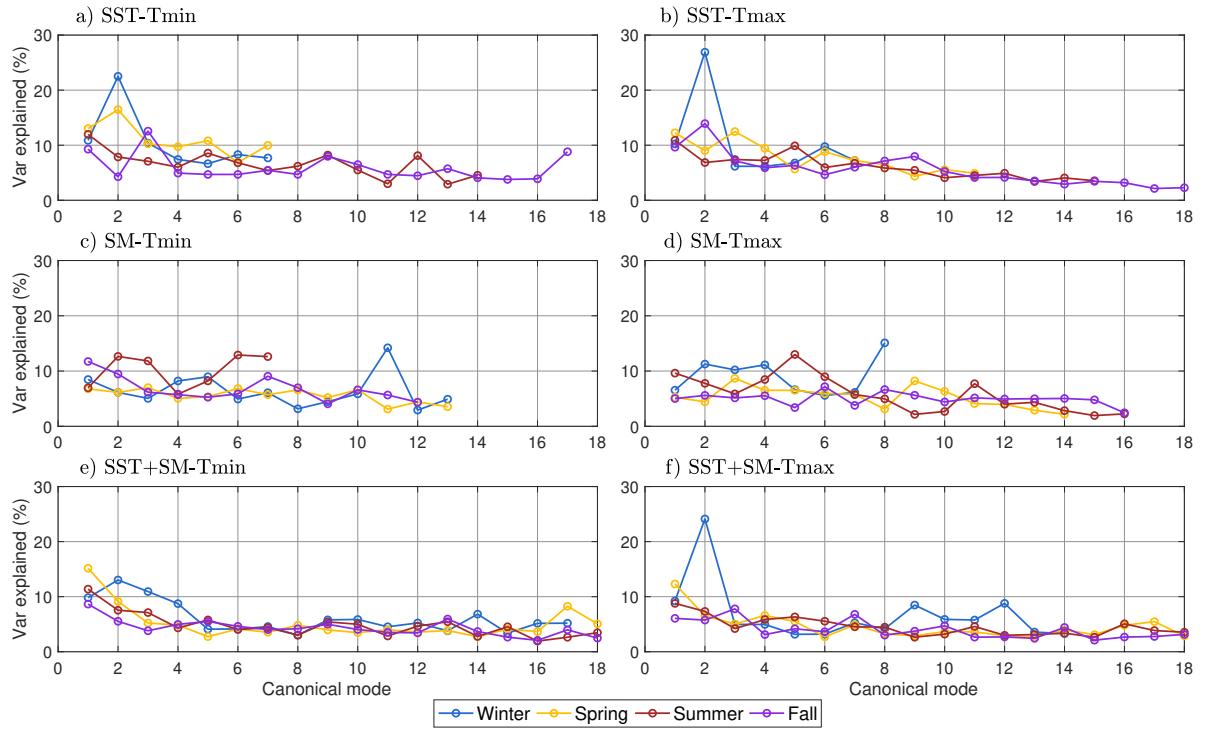


FIG. 6: Variance explained by each canonical mode from the CCA models, averaged by seasons. The colors show the values referent for each season: winter includes those experiments initialized from December to February; spring, from March to May; Summer, from June to August; and fall, from September to November.

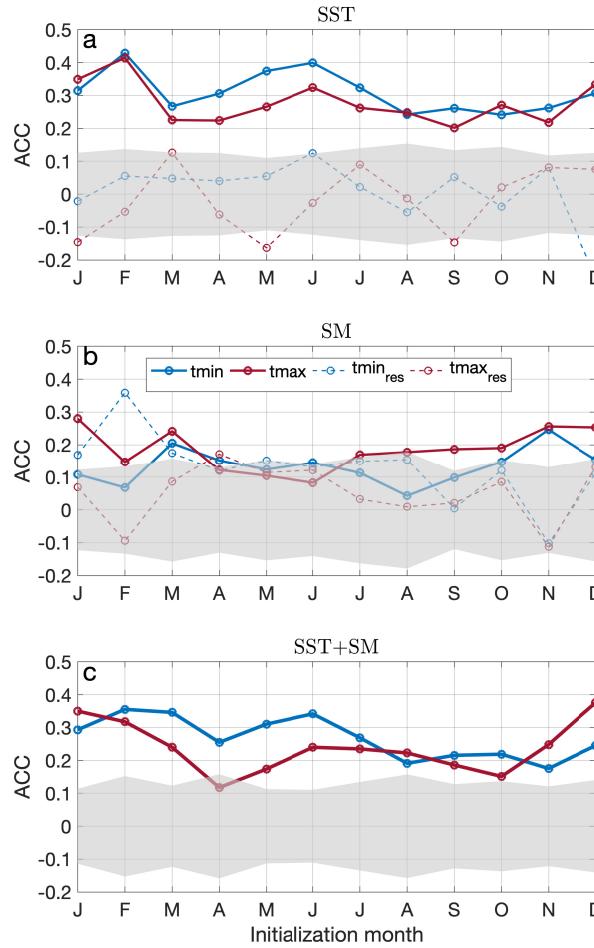


FIG. 7: Annual cycle of field-averaged skill (average skill among all the grid points) obtained from the CCA model run at one month lead time using (a) SST as predictor, (b) SM as predictor, and (c) SST and SM combined. Solid blue line is the skill for T_{\min} and solid red line is the skill for T_{\max} . Shaded areas are the confidence interval obtained from the randomized experiments: values that fall outside this shaded area have statistical significance greater than 95%. Dashed lines are the annual cycle for the field-averaged skill of the residuals, where in (a) are the residuals from SST predicted with SM and in (b) are the residuals from SM predicted with SST (see Section 2d for details).

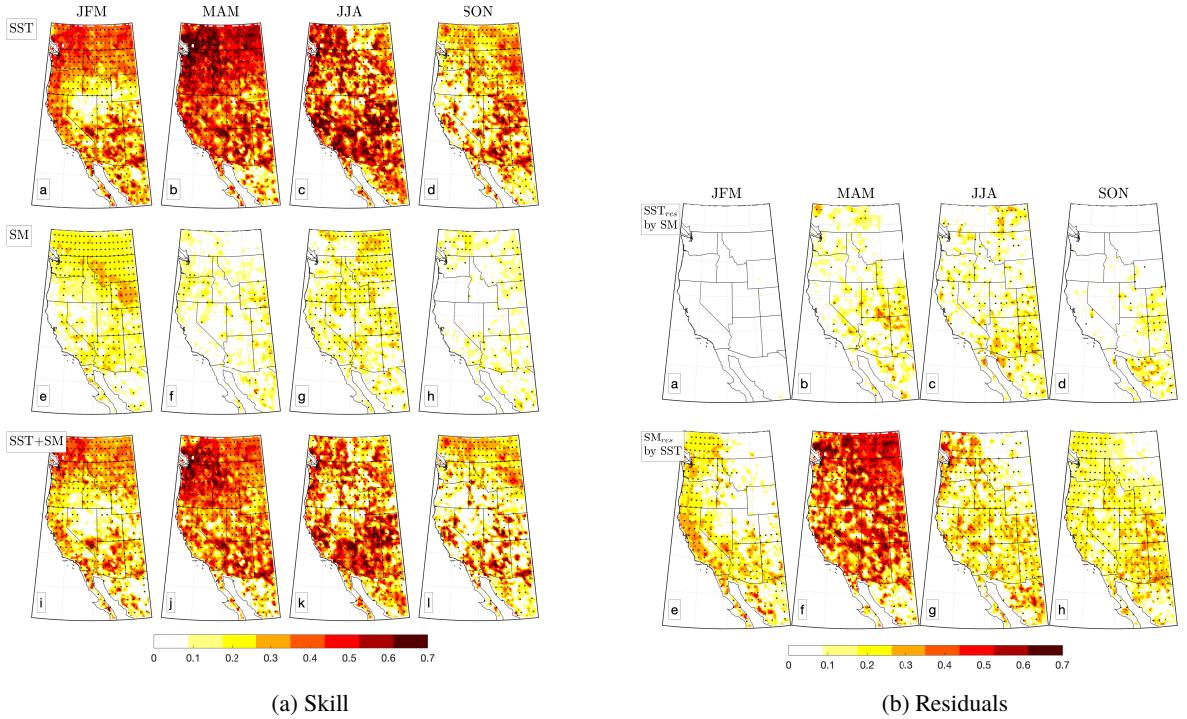


FIG. 8: Spatial pattern of Tmin forecast skill for selected seasons (shown above the maps in the upper panels). **(a)** Tmin seasonal forecast skill obtained with the model using SST (top panels), SM (middle panels) and SST+SM (lower panels) as predictors; **(b)** SST and SM Tmin residuals forecast skill. SST-Tmin residuals forecast using SM as predictor (upper panels) and SM-Tmin residuals forecast using SST as predictor (lower panels). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in the WNA. Black dots represent values that have statistical significance greater than 95%.

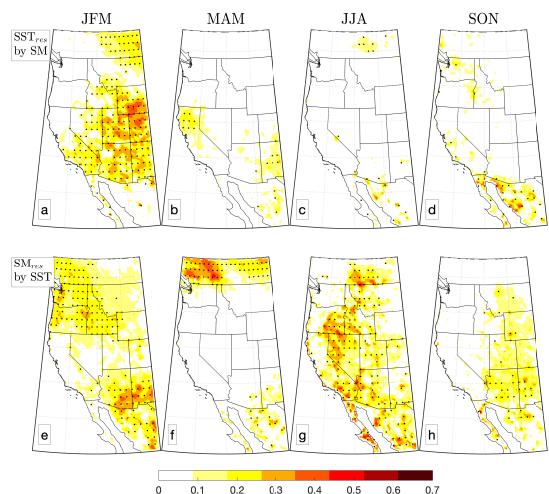
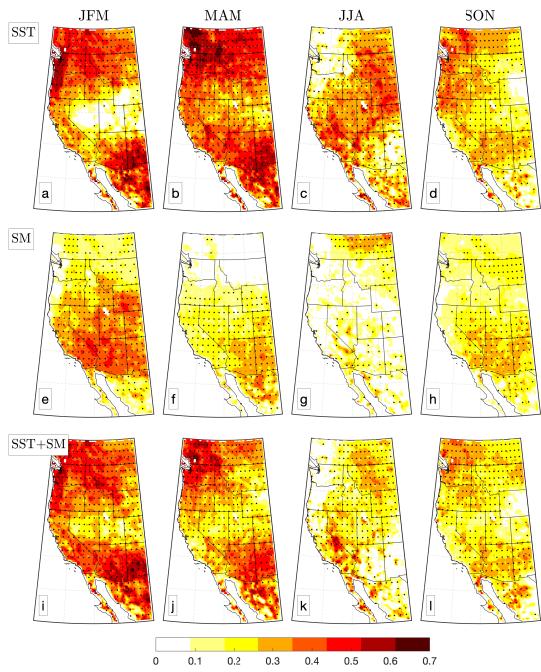


FIG. 9: Same as Figure 8, but for Tmax as predictand

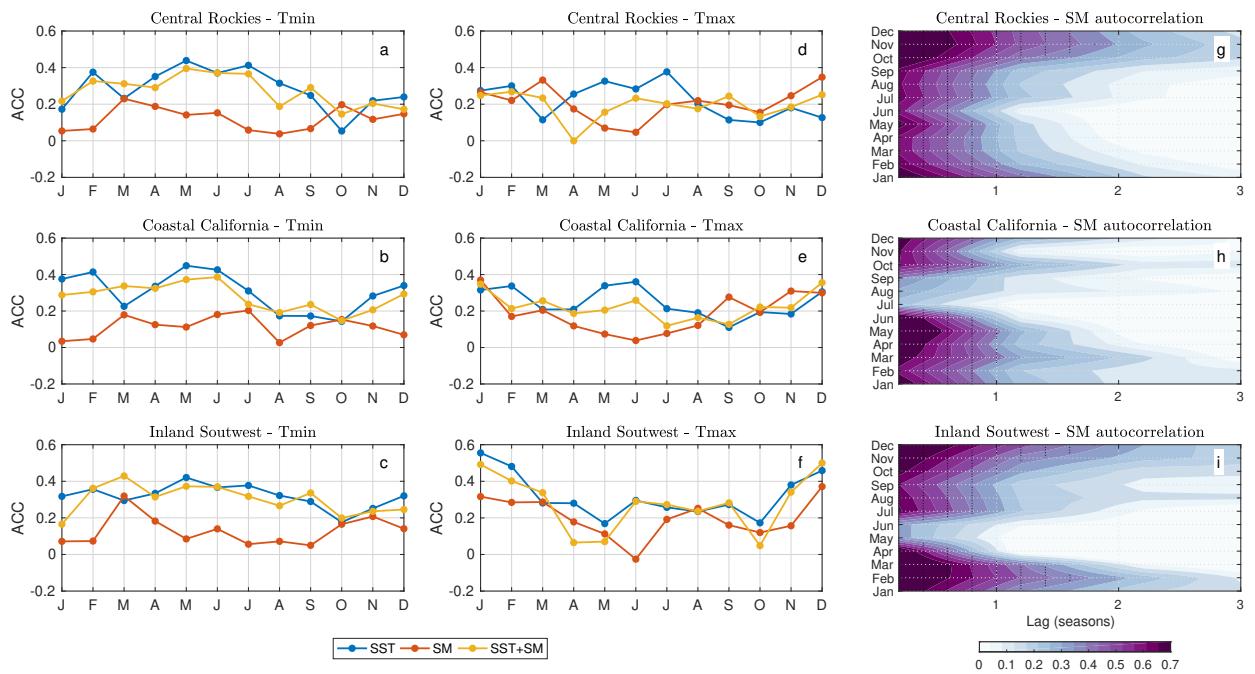


FIG. 10: Annual cycle of seasonal predictive skill for (a-c) Tmin and (d-f) Tmax, and (g-i) annual cycle of soil moisture memory for the three regional domains shown in Figure 1.

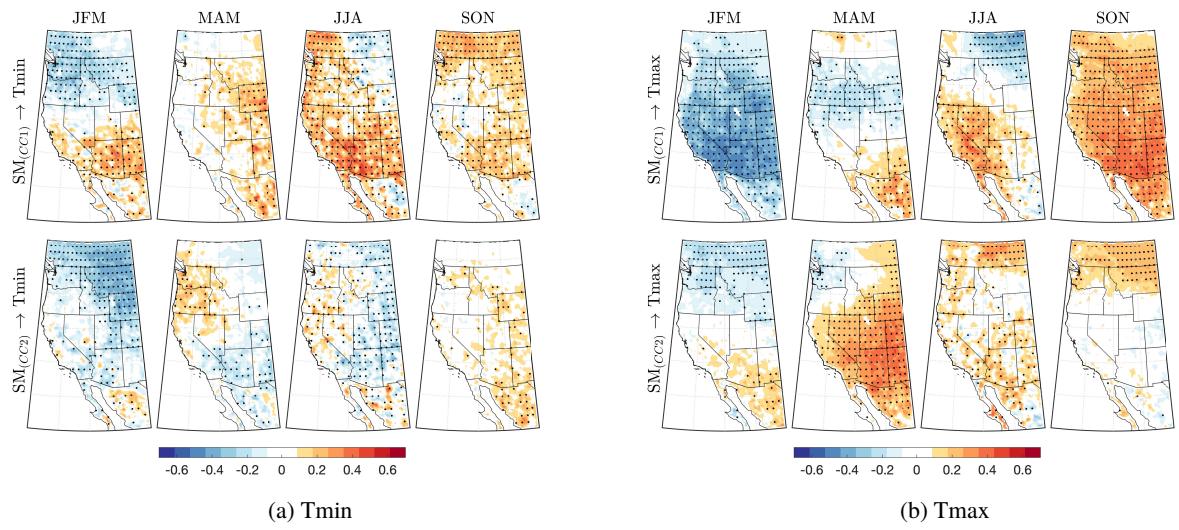


FIG. A1: First (upper panels) and second(lower panels) soil moisture canonical modes for selected seasons (JFM, MAM, JJA and SON) correlated with (a) Tmin and (b) Tmax time series. Black dots indicate values that have a statistical significance greater than 99%

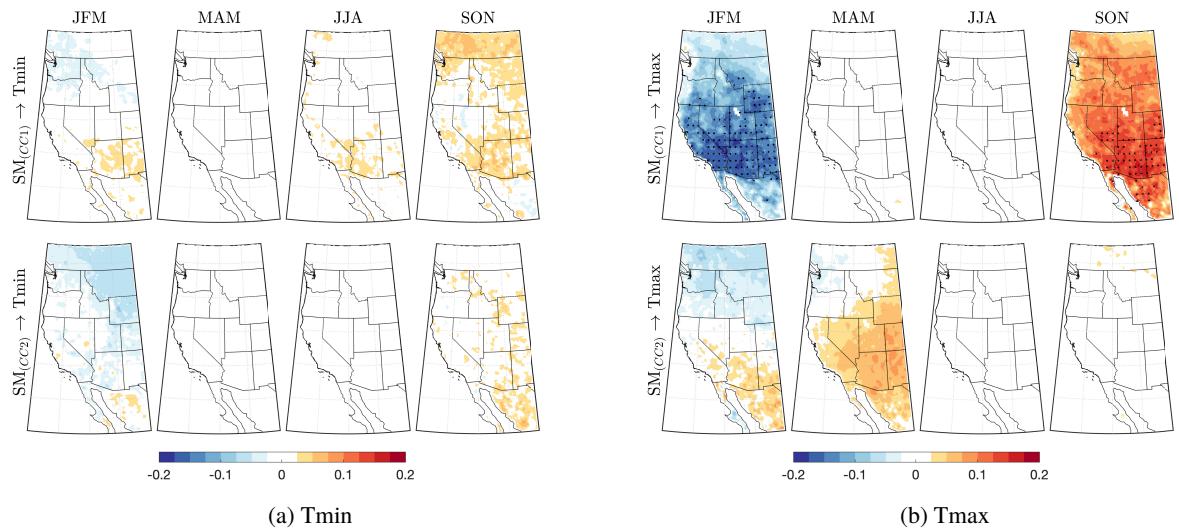


FIG. A2: Same as Figure A1, but with the projected value weighted by the variance explained by each of soil moisture canonical mode.

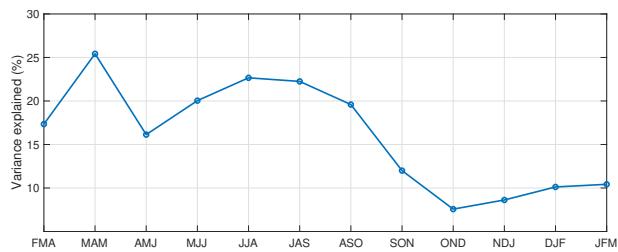


FIG. A3: Variance explained by the leading mode of the CCA model using SST as a predictor and Tmin as a predictand, for the 12 seasons analyzed

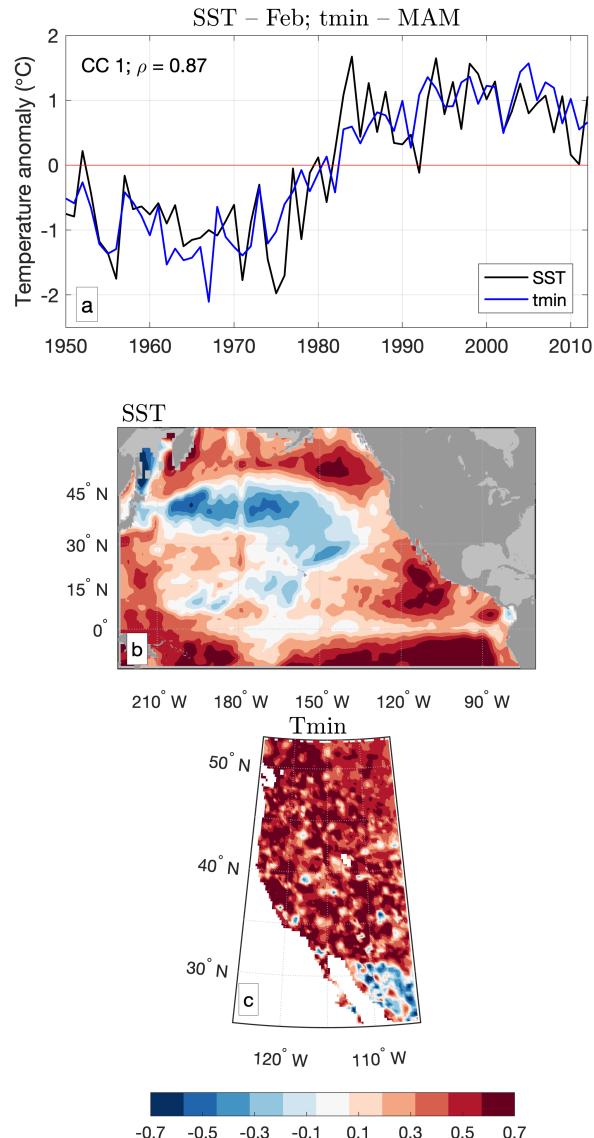


FIG. A4: Leading canonical mode for the SST-Tmin model for MAM. (a) Time series related to this leading canonical mode. The value ρ corresponds to the correlation between these two time series; (b) corresponding spatial pattern for the predictor (SST) and (c) for the predictand (Tmin).