STATISTICAL PREDICTION OF MINIMUM AND MAXIMUM AIR TEMPERATURE IN CALIFORNIA AND WESTERN NORTH AMERICA

A Report for:

California's Fourth Climate Change Assessment

Prepared By:
Daniela F. Dias[1], Daniel R. Cayan[1], Alexander Gershunov[1]

1 Scripps Institution of Oceanography, University of California San Diego

Edmund G. Brown, Jr., *Governor*

August 2018
CCCA4-CEC-2018-011

ACKNOWLEDGEMENTS

We thank Ben Livneh for producing and supplying the gridded historical minimum and maximum temperature dataset, Dennis Lettenmaier for the VIC historical data of soil moisture, and David Pierce for supplying those datasets. We also thank the Hadley Center Sea Ice and Sea Surface Temperature for the SST reanalysis. Finally, we would like to thank also Art Miller, Amato Evan and David Pierce for useful discussions. This research was supported by the California Energy Commission under Agreement Number PIR-15-005.

PREFACE

California's Climate Change Assessments provide a scientific foundation for understanding climate-related vulnerability at the local scale and informing resilience actions. These Assessments contribute to the advancement of science-based policies, plans, and programs to promote effective climate leadership in California. In 2006, California released its First Climate Change Assessment, which shed light on the impacts of climate change on specific sectors in California and was instrumental in supporting the passage of the landmark legislation Assembly Bill 32 (Núñez, Chapter 488, Statutes of 2006), California's Global Warming Solutions Act. The Second Assessment concluded that adaptation is a crucial complement to reducing greenhouse gas emissions (2009), given that some changes to the climate are ongoing and inevitable, motivating and informing California's first Climate Adaptation Strategy released the same year. In 2012, California's Third Climate Change Assessment made substantial progress in projecting local impacts of climate change, investigating consequences to human and natural systems, and exploring barriers to adaptation.

Under the leadership of Governor Edmund G. Brown, Jr., a trio of state agencies jointly managed and supported California's Fourth Climate Change Assessment: California's Natural Resources Agency (CNRA), the Governor's Office of Planning and Research (OPR), and the California Energy Commission (Energy Commission). The Climate Action Team Research Working Group, through which more than 20 state agencies coordinate climate-related research, served as the steering committee, providing input for a multisector call for proposals, participating in selection of research teams, and offering technical guidance throughout the process.

California's Fourth Climate Change Assessment (Fourth Assessment) advances actionable science that serves the growing needs of state and local-level decision-makers from a variety of sectors. It includes research to develop rigorous, comprehensive climate change scenarios at a scale suitable for illuminating regional vulnerabilities and localized adaptation strategies in California; datasets and tools that improve integration of observed and projected knowledge about climate change into decision-making; and recommendations and information to directly inform vulnerability assessments and adaptation strategies for California's energy sector, water resources and management, oceans and coasts, forests, wildfires, agriculture, biodiversity and habitat, and public health.

The Fourth Assessment includes 44 technical reports to advance the scientific foundation for understanding climate-related risks and resilience options, nine regional reports plus an oceans and coast report to outline climate risks and adaptation options, reports on tribal and indigenous issues as well as climate justice, and a comprehensive statewide summary report. All research contributing to the Fourth Assessment was peer-reviewed to ensure scientific rigor and relevance to practitioners and stakeholders.

For the full suite of Fourth Assessment research products, please visit www.climateassessment.ca.gov. This report contributes to energy sector resilience by improving our understanding of probabilistic forecasts, which have the potential to improve energy system management in California's variable and changing climate.

ABSTRACT

Probabilistic forecasts have the potential to improve energy system management by anticipating climate fluctuations and possible weather extremes that influence energy consumption. Such forecasts must contend with a high level of natural variability as well as challenges posed by climate change. However, such forecasts are constrained by limited understanding of local and regional atmospheric predictability. This work bridges the gap between country-wide studies and specific, local studies of atmospheric predictability. These results indicate that statistical methods can provide predictability for seasonal anomalies of air temperature over much of California. The level of forecast skill may provide useful guidance for utility industry applications, but is not universally applicable because skill was found to vary over several elements, including forecasts of nighttime and daytime temperatures, location, season, lead time, and choice of predictor variables. The statistical forecast skill of seasonal air temperatures obtained for some regions and seasons can be superior to that provided by dynamical forecast models, according to results provided by a companion study (Zhang et al. 2018) submitted to California's Fourth Climate Change Assessment. We developed and explored a statistical prediction model of minimum and maximum air temperature anomalies (Tmin and Tmax) over California using remote and local predictors. The first predictor variable field is sea surface temperature (SST) anomalies across the tropical and northern Pacific basin, representing the influence of large-scale climate variability patterns, which in turn affect local surface air temperature. The second predictor variable is soil moisture (SM), which is thought to exert a local or regional influence on temperatures near the surface through the effects on the surface energy balance. The evaluation of the predicted air temperature using historical observations indicates that both local and remote influences contribute to the prediction skill for air temperature. These influences vary with the season and vary between Tmin or Tmax predictands. At time leads of one season, SST has a strong effect on the summer temperatures, but SM provides little prediction skill. On the other hand, in the winter, both SM and SST anomalies provide significant skill in one season in advance predictions of Tmin and Tmax. However, those two predictands also present some important differences: while SST influence both Tmin and Tmax, the influence is stronger for Tmin for both summer and winter. On the other hand, SM influences Tmax more strongly during winter, especially for lags of two and three months. These results exhibit positive forecast skill, but demonstrate considerable variability across seasons in model predictors and forecast performance of seasonal air temperature for California and the overall far western North America region. The results also demonstrate the importance of careful analyses that consider season, variable being predicted, predictors, and time lags in forming statistical forecast models that will be used for decision making.

Keywords: seasonal predictability, air temperatures, California, sea surface temperature, soil moisture, canonical correlation.

Please use the following citation for this paper:

Dias, D. F., Cayan, D. R., Gershunov, A. (Scripps Institution of Oceanography, University of California, San Diego). 2018. ***Statistical prediction of minimum and maximum air temperature in California and western North America.*** California's Fourth Climate Change Assessment, California Energy Commission. Publication Number: CCCA4-CEC-2018-011.

HIGHLIGHTS

- Development of reliable probabilistic forecasts has the potential to improve energy sector management in California's highly variable and changing climate. For instance, simple statistical models that relate patterns of variability of different variables can be used with certain confidence to develop the aforementioned forecasts. The advantage of such models is that they are easy and economical to implement when compared to more complex dynamical models.
- A linear statistical model was developed to test the seasonal predictive skill of air temperatures over California. Local and remote predictors were used and the results demonstrate that both influence this predictive skill. However, there is considerable variability in the forecast skill across seasons in models predictors and forecast performance of seasonal air temperature.
- Pacific sea surface temperature (SST) and western North America soil moisture (SM) exhibit positive skills from linear statistical forecasts of winter (December, January and February; DJF) and summer (June, July and August: JJA) seasonal average minimum temperature (Tmin) and maximum temperature (Tmax) for California.
- In summer (JJA), SST has a strong effect on the temperatures, but SM did not yield much predictive skill, whereas in winter, both SM and SST have some ability to predict Tmin and Tmax.
- Pacific SST influences both Tmin and Tmax, but the influence is stronger for Tmin for both summer and winter, whereas SM influences Tmax more strongly during wintertime.
- A simple model such as the one explored here may offer predictive skill for air temperatures over California at certain seasons, and may be useful for energy utility applications. Over parts of California, the forecast skill of these linear statistical models is comparable or better than that obtained by the North American Multi-Model Ensemble (NMME). For example, the summer (JJA) and winter (DJF) forecast skill obtained by the linear model using SST anomalies as a predictor is greater than the forecast skill obtained by NMME over parts Southern and coastal California regions.
- If the forecasts' uncertainties associated with the variability across seasons and model predictors can be better understood, there may be opportunities to perform real-time air temperature forecasts for California at one season lead times using the linear statistical model developed here, in order to inform decision makers, especially in the energy sector.

TABLE OF CONTENTS

1: Introduction

Skillful predictions of climate fluctuations at seasonal or longer time scales have many economical and societal applications. Anticipating the climate fluctuations, and possible associated weather extremes, at one or more season in advance would have measurable benefits for decision-making relating to hydrology, agriculture, health, energy, and other sectors. Surface air temperature anomalies exert great influence on agriculture, where both warm and cool anomalous temperatures can affect crop production, especially for spring and early summer. And, concerning the energy sector, the variability of regional surface air temperature plays an important role. For example, a summer when temperatures are much higher than the average would increase the demand for cooling, therefore increasing energy consumption through the heavy load of electrical usage from air conditioners. Likewise, a winter cooler than the average will bring the energy consumption up by increasing the demand for heating.

Although the dynamics of the climate system have important nonlinear features, seasonal predictions are often made with linear statistical methods. In fact, it has been found that the traditional linear methods are comparable in skill to the nonlinear statistical ones (Tang et al. 2000; van den Dool 2007) and can perform better than dynamical ones (Gershunov and Cayan 2003). Most of those linear approaches seek to estimate related patterns of variability in the predictor and the predictand fields, such as by canonical correlation analysis (CCA; e.g, Gershunov and Cayan, 2003; Wilks 2008). The advantage of those statistical forecasts is that they can be formulated and implemented easily and they perform economically, and the forecast skill they produce can be evaluated in a straightforward manner.

The scientific basis for seasonal atmospheric climate predictability arises from the fact that variations in slowly changing influences, such as oceanic and terrestrial conditions at Earth's surface, in particular those having time scales of several days to several months, can affect atmospheric circulation and accordingly, the climate at the surface. Ocean thermal anomalies often persist for several months, and a large body of studies have shown that slowly evolving sea surface temperature (SST) anomaly patterns can be used to make useful forecast of seasonal weather and climate anomalies a season or more in advance (e.g, Namias, 1965; Barnett and Preisendorfer 1987; Barnston and Smith 1996; Gershunov 1998, Gershunov and Cayan, 2003; Neelin et al., 1998; Rasmusson and Wallace, 1983; Xoplaki et al., 2003). This is particularly important for North America, where the low-frequency variability in Pacific SST climate modes influences atmospheric circulation over the Pacific-North America sector and consequently, the United States climate (Hartman, 2015; Horel and Wallace, 1981; Mantua et al. 1997).

Those oceanic influences on surface temperatures over North America often involve large scale variability patterns and operate at relatively long distances. However, there may also be regional influences that come to play, including the state of the land surface. In particular, the variability of soil moisture has been shown to affect the surface energy balance (Koster et al. 2000) and thereby add predictability for seasonal time scales (Alfaro et al., 2006; Seneviratne et al., 2010; Whan et al., 2015). This may be the case when the impact of SST fields on the atmospheric variability is modest, which has been demonstrated during the summer season in mid-latitudes (Koster et al. 2000; Shukla 1998; Trenberth et al. 1998).

Additionally, antecedent surface moisture deficits have been statistically linked to the frequency of hot summer days and to the duration of summer heatwaves over a large fraction of the global

land surface, including most areas of South and North America, Europe, Australia, and parts of China (Mueller and Seneviratne, 2012, Gershunov and Douville 2006). Within these results, the predictability of hot extremes was shown to be enhanced by sequences of days following wetter conditions because the moisture in the soil leads to a decrease in the number of hot days (Quesada et al., 2012). Therefore, soil moisture may increase the skill for seasonal predictions at some specific locations and seasons (Koster and Suarez, 2003).

Some studies have explored the role of soil moisture in the predictability of summer air temperature over the United States. Initially, Barnett and Preisendorfer (1987) suggested that some portion of air temperature variability may be explained by forcings other than large scale ocean influences when studying the predictability of U.S. summer air temperatures using Pacific SST patterns. From studies of potential predictability, they found that the average SST predictive skill for summer air temperature was lower than expected. In fact, Georgakakos et al. (1995) found significant cross-correlation for the upper soil moisture leading maximum temperature values with 5 to 10 days of lag in two locations in the interior of the U.S. Additionally, Huang and Van den Dool (1993) found that soil moisture has an impact on the next month's surface air temperature in summer, and Huang et al. (1996) found an inverse correlation between soil moisture and air temperature in the interior of the U.S. This enhances the temperature persistence during warm season months, therefore providing extra skill in predicting air temperature in large interior areas, particularly for the daily maximum values.

Alfaro et al. (2006), hereafter A06, studied associations of Pacific SST and the Palmer Drought Severity Index (PDSI) with summer maximum and minimum temperatures over the central and western United States during 1950-2001. A06 found that large-scale and remote climate conditions exert their influence more strongly on minimum than maximum temperature. Interestingly, summer surface maximum air temperatures were more predictable than those in winter, despite summer having less spatially coherent and lower amplitude climate patterns than those in winter. This was attributed to the additive effect of soil moisture, presumably augmenting effects of large-scale patterns, represented by SST. In other words, during the warm season, regional or local soil moisture processes provide another source of predictability.

In this work, we investigate predictability of seasonal average minimum and maximum air temperatures over the western United States in the winter and summer. As in A06, we employ linear statistical methods, using Pacific SST and regional soil moisture predictors. This work is similar to A06, but includes more years, is more regionally focused, and provides a broader seasonal consideration. Here, using 66 years of record, we focus on the western conterminous U.S. and compare seasonal forecasts of summer air temperature with those of winter. In particular, we ask to what degree does this approach offer skill in predicting seasonal air temperature over the California region.

2: Data and Methods

Our objective was to explore the predictability of air temperature over California at seasonal time-scales. Encouraged by earlier results demonstrated in A06, we adopt a linear statistical approach using canonical correlation analysis. The variables and the data used as predictors and predictands, as well as the details of the implemented model, are described in this section.

2.1 Predictands

The variability of daytime temperatures, represented by time averages of the daily maximum temperature (Tmax), differ from that of nighttime temperatures, represented by time averages of daily minimum temperature (Tmin). From a statistical perspective, daytime temperature anomalies have higher amplitude (larger daily, monthly and seasonal fluctuations) than do nighttime temperatures, especially in summer months (A06). Additionally, seasonal anomalies of Tmax and Tmin over California are at best (in spring) correlated moderately, and at worst (late summer-fall) correlated poorly (***Table 1***).

Table 1: Correlation coefficients between seasonal averages of Tmin and Tmax anomalies averaged over the California region.

Season	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
Correlation	0.41	0.43	0.60	0.65	0.58	0.50	0.38	0.35	0.35	0.33	0.37	0.44

This lack of strong correlation evidently reflects the different mechanisms that affect daytime and nighttime temperatures and their predictability. From an energy utility perspective, daytime and nighttime temperature anomalies may have different effects on the supply of energy and on the demand for energy. For example, in summer extreme hot daytime temperatures increase demand for electricity because of the use of air conditioners but usually occur under clear skies, which results in high solar energy production. Warmer summer nighttime temperatures would diminish beneficial diurnal cooling relief from hot days and thus might increase the magnitude and duration of air conditioning electrical load, which could increase system failures. In winter, other kinds of impacts may occur. In cooler parts of California, warm nighttime anomalies may lead to less space heating and thus lower electrical and natural gas demand, whereas cool daytime anomalies may provoke higher indoor energy requirements from increased space heating and other indoor utility use.

In view of the above, in our study of air temperature predictability over the California region, we investigate a separate set of forecast models for Tmax and Tmin in summer and winter. The 1950-2013 period is used in the model development and validation. The predictands variables considered are the seasonally averaged Tmin and Tmax, respectively from the Livneh CONUS near-surface gridded meteorological data (Livneh et al. 2015), provided by the Earth System Research Laboratory of the National Oceanic & Atmospheric Administration (ESRL/NOAA). These datasets are gridded at a spatial resolution of 1/16° by 1/16° (approximately 6km square) and are derived from daily temperature observations from approximately 20,000 NOAA Cooperative Observer (COOP) stations (Livneh et al., 2015). We use seasonal averages of Tmax and Tmin anomalies because they represent short period climate fluctuations over multi-day time scales. The predictand dataset are obtained by first calculating the monthly anomalies for each variable, and then calculating seasonal means of the Tmin and Tmax anomalies. While the models are developed for Tmax and Tmin over a broader domain of the far western part of North America, this study present results for Tmax and Tmin predictions with focus on the California region by providing validation maps and statistics exclusively over California. Forecasts for winter (December, January and February – DJF season) and summer (June, July

and August – JJA season) are presented. The similarity of the model experiments with those of A06 provide an interesting comparison of Tmax and Tmin predictive skill obtained by that prior study, when the regional domain is narrowed to California and the far western United States and the model testing period is extended by several years.

2.2 Predictors

Two variables were used as predictor fields for the air temperature. The first predictor variable is Sea Surface Temperatures (SST) anomalies for the Pacific Basin (15°S to 60°N and 135°E to 110°W), which are obtained from the Hadley Center Sea Ice and Sea Surface Temperature (HadISST, Rayner et al., 2003). The HadISST data has a resolution of 1° by 1°and it can be obtained at the following website: https://www.metoffice.gov.uk/hadobs/hadisst/data/download.html.

The second predictor variable is soil moisture (SM) anomalies. In contrast to prior studies, including A06, that have used soil moisture indices such as the Palmer Drought Severity Index (PDSI), SM employed here is estimated from an existing (developed by Dennis Lettenmaier and colleagues) hydrological model reanalysis produced by the Variable Infiltration Capacity (VIC) model, forced by observed daily precipitation and temperature. VIC is a macroscale hydrological water and energy balance accounting model (Liang et al., 1994; Cherkauer and Lettenmaier, 2003) that has been used in numerous studies of climate and hydrological variability and changes (e.g. Hamlet et al. 2007; Das et al. 2009). The forcing data is the gridded daily precipitation and temperature historical dataset from Livneh et al. 2015, which covers the conterminous United States and adjoining portions of Canada and Mexico, and is forced at the same resolution as the Livneh predictand fields, resulting in SM data that is distributed over the identical 1/16° grid. The VIC SM domain is also identical to that of the Tmax and Tmin predictands, covering the land area from 25°N to 53°N and 105°W to 125°W.

SST and SM were selected as predictors because they are broadly representative of large scale and regional and local scale climate measures that vary over time scales that are relevant to seasonal air temperature fluctuations. While Tmax and Tmin are affected by several other variables such as winds, clouds, and topographic influences, we have not included those here since our purpose is to explore linear prediction using a plausible, readily available, and manageable predictor set. Both predictor datasets consist of monthly means from 1950 to 2013, from which monthly anomalies are calculated by removing the climatological (1950-2013) monthly mean for each grid cell and each month. Importantly, trends in the predictors and predictands were not removed.

2.3 Model Details

2.3.1 Canonical Correlation Analysis and Prediction

The model framework implemented to make the seasonal predictions of air temperature over California uses Canonical Correlation Analysis (CCA), following an approach that has been used in several prior climate prediction and analysis studies (e.g. Barnett and Preisendorfer 1987; Gershunov et al. 2000; Gershunov and Cayan 2003). This method explores and identifies the linear combinations between two sets of variables that have the greatest correlation with each other, seeking to match patterns in the predictor fields (SST and SM) with patterns in the predictand fields (Tmin and Tmax). For many climate-related phenomena, the CCA approach condenses much of the spatial and temporal variability and it is a simple and cheap way to

perform the predictions. These spatial patterns and their time variations provide a set of modes for which we can interpret climate-scale associations and test for possible association with commonly known climate variation patterns at different timescales. Moreover, the predictions are made by fitting a linear model that relates antecedent observation of the monthly-averaged predictor fields to the time-lagged seasonally-averaged predictand fields. This methodology follows the hypotheses addressed by A06: on one side, the anomalous Pacific SST causes changes in the large-scale atmospheric circulation, which in turn influences the land surface temperatures over a broad sector; on the other side, soil moisture modulates the local air temperature via the interaction between latent and sensible heating. With those hypotheses addressed, a CCA prediction model for the winter (DJF) and summer (JJA) is constructed considering SST and SM lagged by one to three months. Therefore, for the winter, "Lag 1" considers the monthly-averaged predictor field from November, "Lag 2" from October, and "Lag 3" from September; for the summer, "Lag 1" includes the monthly-averaged predictor field from May, "Lag 2" from April, and "Lag 3" from March.

The statistical model is constructed first by selecting the antecedent observations of the predictor fields and the contemporaneous observations of the predictand field. For example, to make the Tmin one-month lag winter prediction using SST as predictor, we select all the Novembers of SST and all the DJFs of Tmin, but excluding the year that is being predicted. The predictor and predictand fields are separately pre-filtered with the same number of p principal components (PCs), which are statistically orthogonal patterns of spatial and temporal variability, ordered by amount of variance explained, in the predictor and predictand fields. Those patterns are then related to each other using the q canonical correlates (CCs) extracted from the CCA analysis. To avoid artificial skill (e.g, Davis 1976) from over-fitting, the models are optimized (Section 2.3.2) wherein skill over a range of predictor and predictand membership is evaluated under a cross-validation scheme – the number of PCs and CCs are thus determined as the model complexity which maximizes cross-validated prediction skill (Gershunov and Cayan 2003). Therefore, $p \leq q \leq T$ (in fact p and q are much fewer than T), where T is the number of years available for the training period.

Importantly, all results pertaining to model performance are cross-validated, wherein the year of a given prediction has been left out of the model development. That is, separate models are made for each year, wherein the training period of each year's model is made using a period that excludes the season and adjacent months being predicted. This is implemented by excluding the year predicted from the training period (lagged month for the predictor and season for the predictand).

2.3.2 Skill Optimization and Model Validation

Given the cross-validated forecasts, an evaluation is then conducted of the sequence of resulting yearly forecasts compared to that of the observed data to provide a realistic estimate of forecast skill. The first measure of skill we applied is the correlation between the observed and the predicted air temperature (Tmin and Tmax) at each grid point. The squared correlation is the amount of Tmax or Tmin variance explained. From this measure maps of correlation values can be investigated that represent the correlation skill of the statistical models, using a different set of predictors (SST and SM), predictands (Tmin and Tmax), seasons (summer and winter), and lags (one, two and three months).

Another skill measure applied is the "Heidke Skill Score" (HSS), which measures the fraction of corrected forecasts after eliminating those forecasts that would be correct due to random chance. HSS is applied to the forecast time series of a region near Sacramento, in central California, and for the eight different models described here. Sacramento is chosen because of its central location geographically, because it represents a major population who depend on natural gas and energy utilities, and because it is a well-known location. The formulation of this skill score is given by:

$$HSS = \frac{(NC - E)}{T - E}$$

Where *NC* is the number of correct forecasts, hence the number of forecasts that falls in the same category as the observed values, *T* is the total number of forecasts, and *E* is the number of forecasts expected to verify based on chance. To calculate those values, the observations time series are divided in three categories (above normal, normal and below normal). *NC* represents the number of the predicted values that fall into one of those categories (x_{ii}), whereas *E* represents the predicted values that do not correspond to those categories (x_{ip} or x_{pi}):

$$NC = \sum_{i=1}^{m} x_{ii}$$

$$E = \sum_{i=1}^{m} x_{ip} x_{pi} / T$$

The resultant HSS is a number between Infinity and 1, where 1 represents the perfect model, 0 indicates no skill, and negative values indicate that the forecasts are less skillful than randomly generated ones.

For each of the forecast model cases (season, lag, predictor, and predictand), a certain number of p PCs and q CCs is determined by finding the optimum model, i.e., the combination of PCs and CCs that maximize the forecast skill. The model is calculated using all different combinations of PCs and CCs, from 1 to 18 and with PCs ≤ CCs. For each option of model complexity (i.e., combination of the number of PCs and CCs), the cross-validated skill is estimated for each grid point and then summarized in one value by averaging the correlation of all the points. In practice, we use the field averaged correlation between the predicted and observed air temperature as a measure of forecast skill when determining the optimum model. Therefore, for each model combination (i.e., predictor-predictand-season), we have a different model complexity (combination of PCs and CCs numbers) that maximizes the forecast skill. Finally, the optimum model is determined by selecting the model complexity that presents the better skill. Details of the model optimization process are presented below in Section 3.1.

3: Results

In this section, results from the CCA prediction models are presented. We first describe the skill optimization and show the different model complexity for the different experiments. Then we briefly explain the canonical correlation patterns, that represent the physical sources of predictability for Tmin and Tmax. Finally, we show the forecast skill for Tmin and Tmax using all the model combinations described in the previous section.

3.1 Optimum Model

The model complexity (numbers of PCs and CCs, i.e. p and q) used for each experiment is presented in **Table *2*** and the variance explained by each CC mode is presented in **Table 3** for the predictands (Tmin and Tmax) and in **Table 4** for the predictors (SST and SM).

Table 2: Optimum number of Principal Components (PCs) and Canonical Correlates (CCs) for all one month lag summer (JJA) and winter (DJF) forecast experiments, the field-averaged skill for the optimum model (Skill), and the lower (rl) and upper (ru) limits of the 95th confidence interval of skill obtained from 100 randomized experiments.

Season Predicted	Predictor	Predictand	PC	CC	Skill	rl	ru
JJA	SST	Tmin	16	16	0.21	-0.003	0.013
DJF	SST	Tmin	10	9	0.20	-0.011	0.011
JJA	SST	Tmax	16	16	0.24	-0.001	0.017
DJF	SST	Tmax	14	12	0.26	-0.008	0.026
JJA	SM	Tmin	11	10	0.13	-0.007	0.010
DJF	SM	Tmin	10	6	0.15	-0.008	0.012
JJA	SM	Tmax	2	2	0.22	0.001	0.026
DJF	SM	Tmax	13	5	0.10	-0.001	0.016

Table 3: Variance explained by each predictand Canonical Correlate (rows, CC mode), for each experiment (columns). The number of CC modes included in each experiment was determined by the optimization process (see *Table 2* and Section 3.1 for details).

	SST				SM			
CC mode	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**
1	34.7%	14.8%	4.9%	8.1%	10.3%	17.4%	50.1%	57.0%
2	3.9%	1.0%	4.7%	7.5%	1.3%	10.8%	49.9%	41.3%
3	7.9%	0.5%	1.8%	5.7%	0.4%	0.0%	--	0.0%
4	0.3%	7.4%	19.8%	7.9%	20.1%	70.9%	--	1.1%
5	9.4%	6.4%	11.9%	0.0%	23.3%	0.7%	--	0.6%
6	0.1%	0.6%	0.5%	28.0%	17.8%	0.2%	--	--
7	0.2%	8.6%	0.0%	9.6%	4.8%	--	--	--
8	12.0%	44.1%	12.2%	2.3%	3.2%	--	--	--
9	4.4%	16.7%	5.1%	1.1%	1.8%	--	--	--
10	3.3%	--	2.5%	6.1%	17.1%	--	--	--
11	4.5%	--	10.7%	22.7%	--	--	--	--
12	0.5%	--	9.7%	0.9%	--	--	--	--
13	8.2%	--	0.0%	--	--	--	--	--
14	2.6%	--	9.0%	--	--	--	--	--
15	4.4%	--	1.5%	--	--	--	--	--
16	3.8%	--	5.6%	--	--	--	--	--

Table 4: Variance explained by each predictor Canonical Correlate (rows, CC mode), for each experiment (columns). The number of CC modes included in each experiment was determined by the optimization process (see *Table 2* and Section 3.1 for details).

	SST				SM			
CC mode	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**
1	14.1%	3.3%	7.9%	0.1%	20.8%	35.1%	50.0%	33.2%
2	0.9%	23.2%	5.4%	17.5%	30.0%	15.9%	50.0%	11.6%
3	0.0%	16.3%	8.0%	2.1%	5.1%	28.7%	--	0.9%
4	2.0%	13.0%	0.2%	0.1%	14.6%	1.1%	--	14.4%
5	6.1%	12.5%	3.0%	14.3%	0.0%	18.8%	--	40.0%
6	16.7%	2.0%	5.8%	0.8%	0.0%	0.4%	--	--
7	0.0%	9.8%	2.7%	4.8%	0.2%	--	--	--
8	21.9%	2.7%	2.7%	19.8%	1.8%	--	--	--
9	21.2%	17.1%	17.7%	31.0%	26.6%	--	--	--
10	2.3%	--	1.5%	1.3%	0.9%	--	--	--
11	1.6%	--	5.3%	2.4%	--	--	--	--
12	2.8%	--	6.1%	5.8%	--	--	--	--
13	5.2%	--	18.4%	--	--	--	--	--
14	1.4%	--	3.7%	--	--	--	--	--
15	1.0%	--	5.2%	--	--	--	--	--
16	2.8%	--	6.5%	--	--	--	--	--

Those choices were made based on the skill optimization matrix, where the field-averaged skill is calculated for each different combination of PCs and CCs. To reduce the number of calculations required in this exercise, the number of PCs representing the predictor and predictand fields was fixed as the same number of modes in producing these optimization estimates.

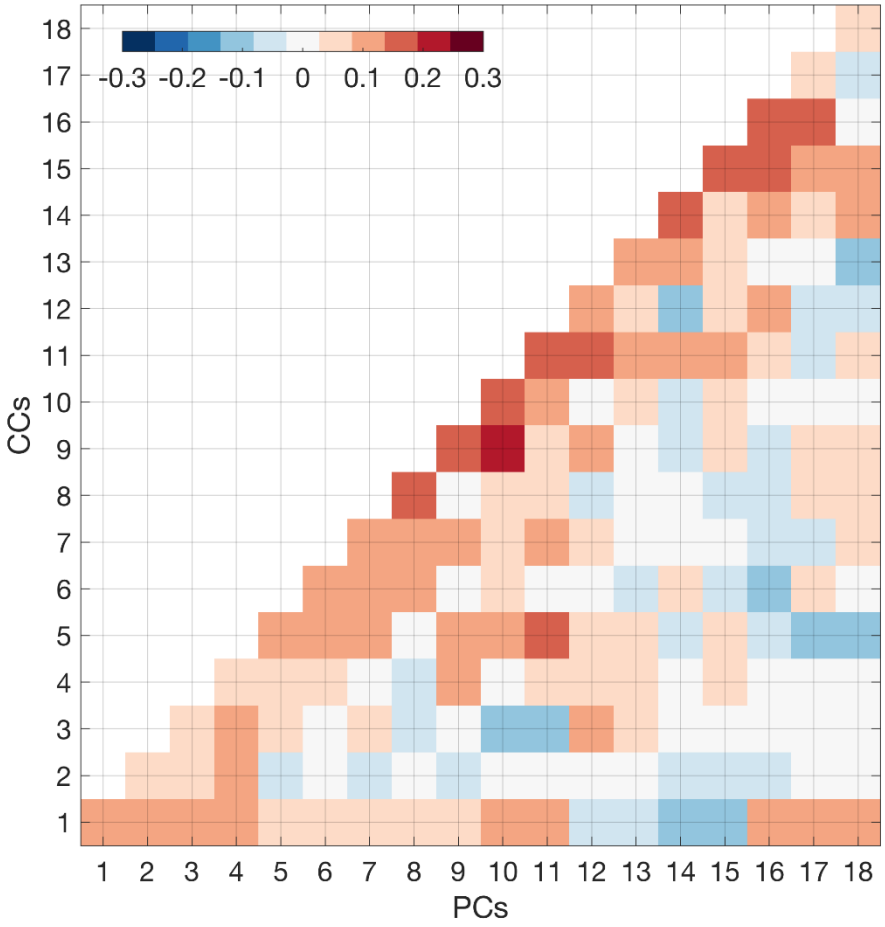

Figure 1: Skill optimization for DJF Tmin prediction using SST as predictor. Figure shows the cross-validated field-averaged skill, expressed as the correlation between the predicted and the observed temperatures, using different combination of the Principal Components (PCs) and the Canonical Correlates (CCs). The optimum is defined the number of PC and CC modes that yielded the highest cross validated skill (PC=10 and CC=9).

As an example, **Figure *1*** shows the plot of this matrix for the prediction of winter Tmin using one-month lagged SST as a predictor. In this case, the resulting skill matrix exhibits not one, but a set of optimal regions where the skill is equally highest (for example, number of CCs between 7 and 11, and the number of PCs between 8 and 12). In proceeding with the model experiments, in the spirit of parsimony, i.e. "simpler is better", we chose the simpler model complexity (the fewer PCs/CCs number) that can have the best predictability. In ***Figure 1***, this number is 10 PCs and 9 CCs, and this procedure was repeated for all the other experiments, as shown in ***Table 2***. This table also presents the field-averaged skill for that chosen number of CCs and PCs and compares it to the field-averaged correlation of a randomized experiment: the time series of the predictor was randomized and used to generate the CCA model to forecast air temperature. This procedure was repeated 100 times and the lower and upper limits of the 95th confidence intervals of those repetitions were calculated. The field-averaged correlation is always bigger

than both of those limits, which indicates that the predictability derived from the model is not due to a stochastic forcing results from a deterministic process between the predictors and the predictand. Therefore, in all of these experiments, the skill achieved is highly likely to be real and not simply artifacts from the data or the model.

Surveying the model results, the optimization produces different combinations of predictors/predictands and canonical correlations produced for different seasons. The model complexity ranges from very simple models, where few PCs and CCs results in the maximum predictability, to more complex models, where a greater number of modes were required to obtain the highest forecast skill. For example, the JJA SM-Tmax model gives the greatest skill when using only two PCs and CCs modes, while the JJA SST-Tmax is optimized at 16 of each. This is a hint that the sources of predictability for Tmin and Tmax vary according to the predictor and the season that is being analyzed. Therefore, it appears to be important to perform this skill optimization analysis across each case to maximize the potential air temperature forecast skill.

3.2 Canonical Patterns

Before exploring the prediction skill with CCA, it is important to elucidate the spatial patterns and their time variations to provide insight into the physical sources of predictability. To do so, the canonical correlation patterns that relate SST and SM to Tmax and Tmin can be explored. We show only the three leading CC modes because the spatial structure associated with the higher-order modes typically becomes continuously less spatially coherent and they contribute progressively less to the field-averaged forecast skill, although they may be important for predicting sub-regional anomalies in specific years.

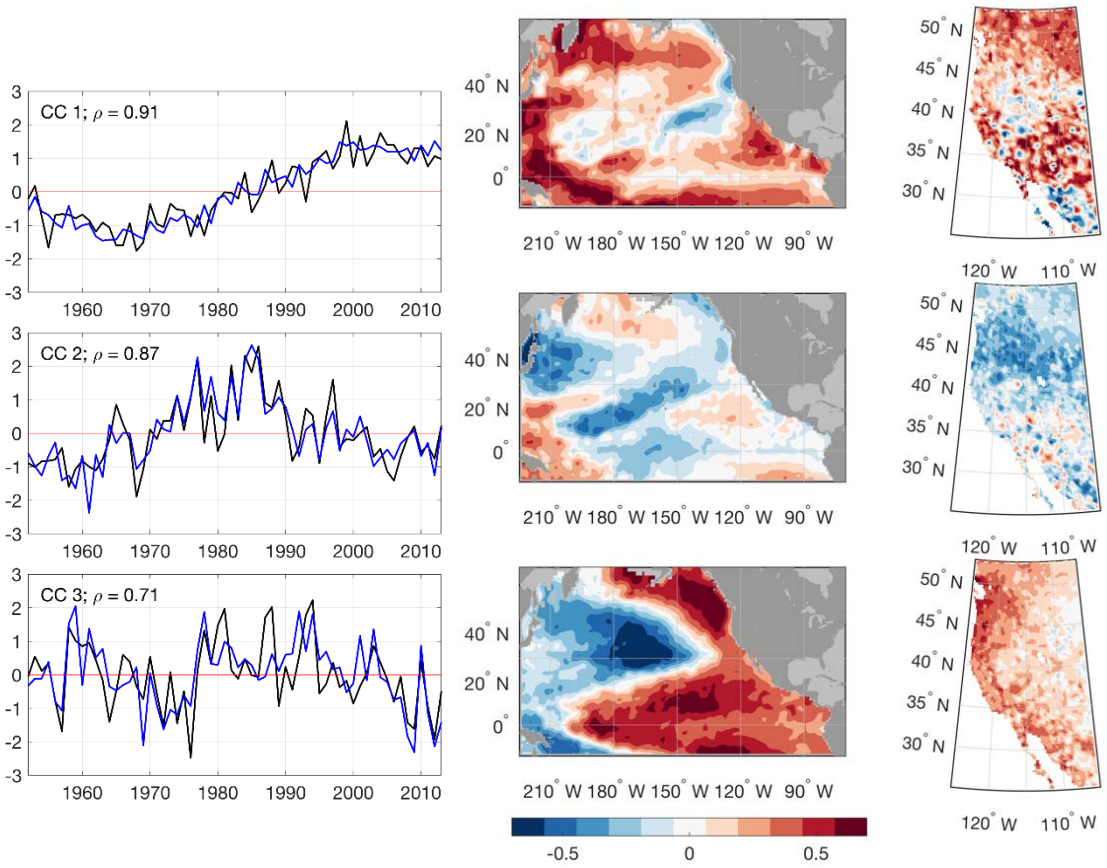

Figure 2: First three coupled CCA modes of DJF Tmin related to November SST. Left panel displays the time series of the CCA modes (CC1, CC2 and CC3), where predictor (SST) is in black and predictand (Tmin) is in blue; ρ is the correlation between the two time-series (CC1=0.91; CC2=0.87 and CC3=0.71). Middle and right panels show the spatial pattern associated with the SST and Tmin fields, respectively, expressed as correlation between the CCA mode time series and their respective variable field for each location.

The spatial patterns of the canonical correlates are explored. Those spatial patterns are maps of the correlations, calculated at each point, between the canonical correlates (CC) time series and the original data time series. First, concerning winter forecasts using antecedent monthly SST, **Figure 2** presents the spatial patterns that relate November SST to DJF Tmin, as well as the time-series associated with those patterns. The leading mode (CC1) of the SST-Tmin pattern captures a warming effect for the air temperature over land, and its SST expression as a warming in the Western Pacific. The pattern exhibited by the third mode (CC3) is an expression of the Pacific Decadal Oscillation (PDO, Mantua et al., 1997). The second mode (CC2) presents weaker correlations than the other two, although the SST pattern resembles the shape of the Pacific Meridional Mode (PMM, Vimont et al., 2001). For the case where November SM was used to predict winter temperatures, interesting regional scale patterns also emerge.

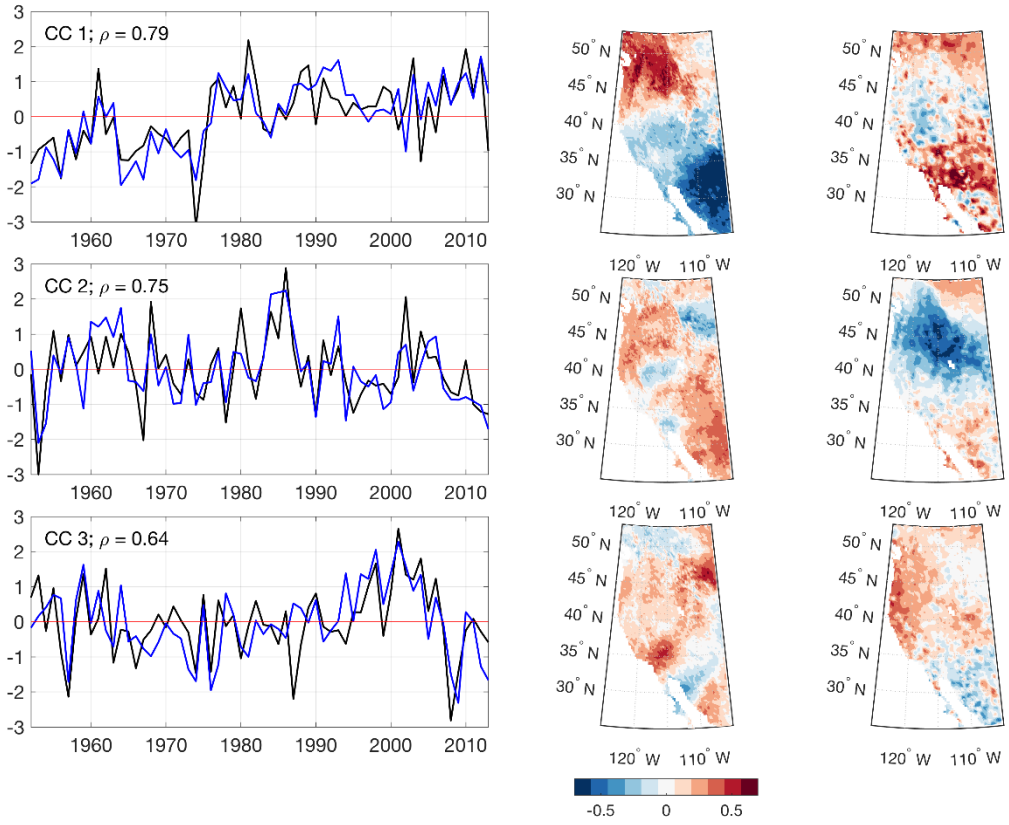

Figure 3: First three coupled CCA modes of DJF Tmin related to November SM. Left panel displays the time series of the CCA modes (CC1, CC2 and CC3), where predictor (SM) is in black and predictand (Tmin) is in blue; ρ is the correlation between the two time-series (CC1=0.79; CC2=0.75 and CC3=0.64). Middle and right panels show the spatial pattern associated with the SM and Tmin fields, respectively, expressed as correlation between the CCA mode time series and their respective variable field for each location.

Figure 3 shows the November SM spatial pattern that is related to DJF Tmin and the time series to which those patterns are associated. Although we cannot name the SM modes because they have not been so broadly studied as the SST modes, we can see that the patterns are spatially coherent and the CC time series appear to have interannual to decadal temporal variation. Within the SM-CC correlation maps, some of the regions of strongest correlations are opposite of those produced by the Tmin-CC correlation maps, meaning that in these regions, positive SM anomalies are associated with negative Tmin anomalies and vice versa. The first SM canonical mode (CC1) shows a clear north-south gradient, and it is associated with an opposite gradient for Tmin, especially in the southernmost part of the domain. For the second mode (CC2), this opposite pattern between SM and Tmin is more evident in the Northwestern part of the domain, although the correlations are weaker. Finally, the third mode (CC3) has also weaker correlations (i.e. weaker coupling overall by definition, and also less local variance), and the patterns are alike, with positive SM anomalies linked to positive Tmin anomalies (and vice versa).

In general, the Tmin and Tmax CC modes consist of spatial patterns that have few to several hundred km scales. This indicates that the forecasted temperature anomalies are regional, not local, so the value to utilities would often cover major parts of their service areas.

Several of the CC time series are quite closely related to known climate modes. ***Figure 4*** shows the correlation of the leading three CC modes from each of the experiments with 12 different climate variability patterns. The climate patterns are ones that have been identified in numerous previous studies, capturing regional to global scale variability at the ocean surface in the Atlantic and Pacific basins, and in the atmosphere near Earth's surface and aloft in the mid-troposphere. Some studies also have explored the effect of those modes in the predictability of California air temperature. For example, Alfaro et al. (2004) demonstrated the predictive capability of the spring PDO for the following summer temperatures over California, and Schwartz et al. (2014) explored the mechanism behind this forcing, through links between the coastal low cloudiness and PDO. Schwartz et al. (2014) found that the North Pacific SST variability, in the form of PDO, is well correlated with the coastal cloud variability, helping to organize those summer patterns along the coast of North America, which in turn may accomplish in part the PDO association with California summer temperature anomalies.

The time series associated with these patterns vary over a range of time scales from seasonal to multi-decadal. Several of the CC modes derived from the SST predictor present high correlations with those large-scale patterns. For example, as shown in **Figure 2**i, the third mode of the winter SST-tmin model resembles the pattern of PDO and indeed it presents a correlation of 0.7 with that time series (***Figure 4***b). Moreover, the time series representing ENSO (Niño1.2, Niño3.4, and Niño4) also exhibit strong correlations with certain CC modes, for example the second mode of the winter tmax-SST model. Therefore, those CC patterns indicate that the forecast models developed here, especially using SST as a predictor, rely on the effect of the large-scale climate variations over the regional variability. In contrast to the SST CC modes, the SM CC modes do not correlate strongly with any of the climate modes.

3.3 Forecast Skill

Since a primary goal of this work is aimed at developing temperature forecasts related to California energy supply and demand, forecast skill over the state of California is presented here. It was expected that forecast skill would diminish as the forecast lead is increased – e.g., the greatest forecast skill is obtained for forecast lead time of one month, and least for lead of three months. Considering results from A06, it was expected that SM forecasts would yield more skillful forecasts during the warm season, and SM forecasts for Tmax would be superior to those for Tmin. Furthermore, A06 found that summer SM forecast skill was greater than the skill of Pacific SST forecasts.

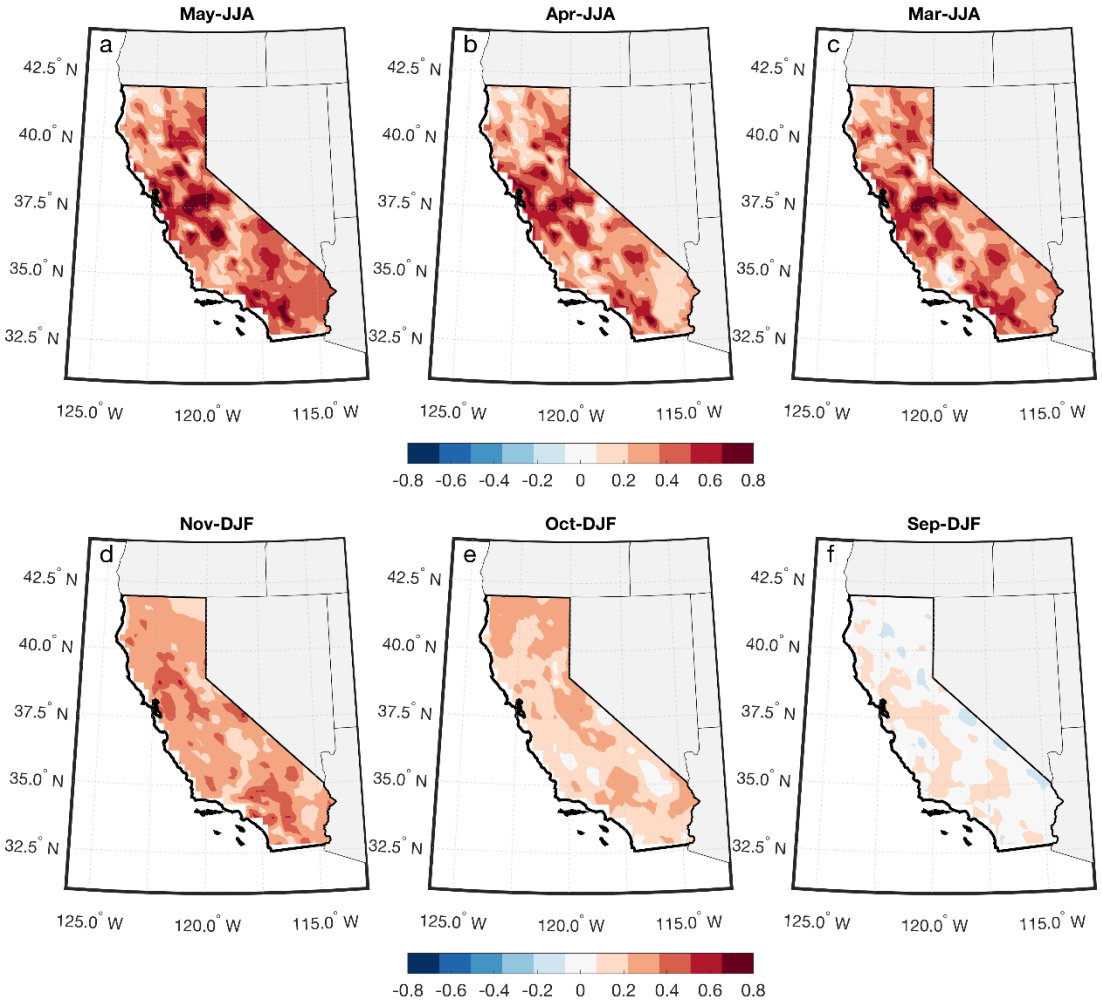

Figure 4: Tmin forecast skill using SST as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SST predicting JJA Tmin (a) and Nov SST predicting DJF Tmin (d)); two months (Apr SST predicting JJA Tmin (b) and Oct SST predicting DJF Tmin (e)); and three months (Mar SST predicting JJA Tmin (c) and Sep SST predicting DJF Tmin (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

Forecast skill for the CCA models is presented in ***Figure 5*** to ***Figure 8,*** which are maps of cross-validated correlations between the time series of forecasts at each grid point and that of the attendant Tmin or Tmax observations. The maps are presented for the array of different predictors, predictands, seasons, and lags. The correlation values indicate the portion of the variability of Tmin and Tmax that was explained by the predictor for the season that is being predicted. As noted earlier, squared correlations represent the amount of variance explained. Those maps show that there are significant differences in forecast skill for both Tmax and Tmin, with considerable differences across the season that is being predicted and by the predictor that is being used. Furthermore, the ability to predict air temperature for higher lags (two and three months) is also dependent on those factors, with some remarkable differences.

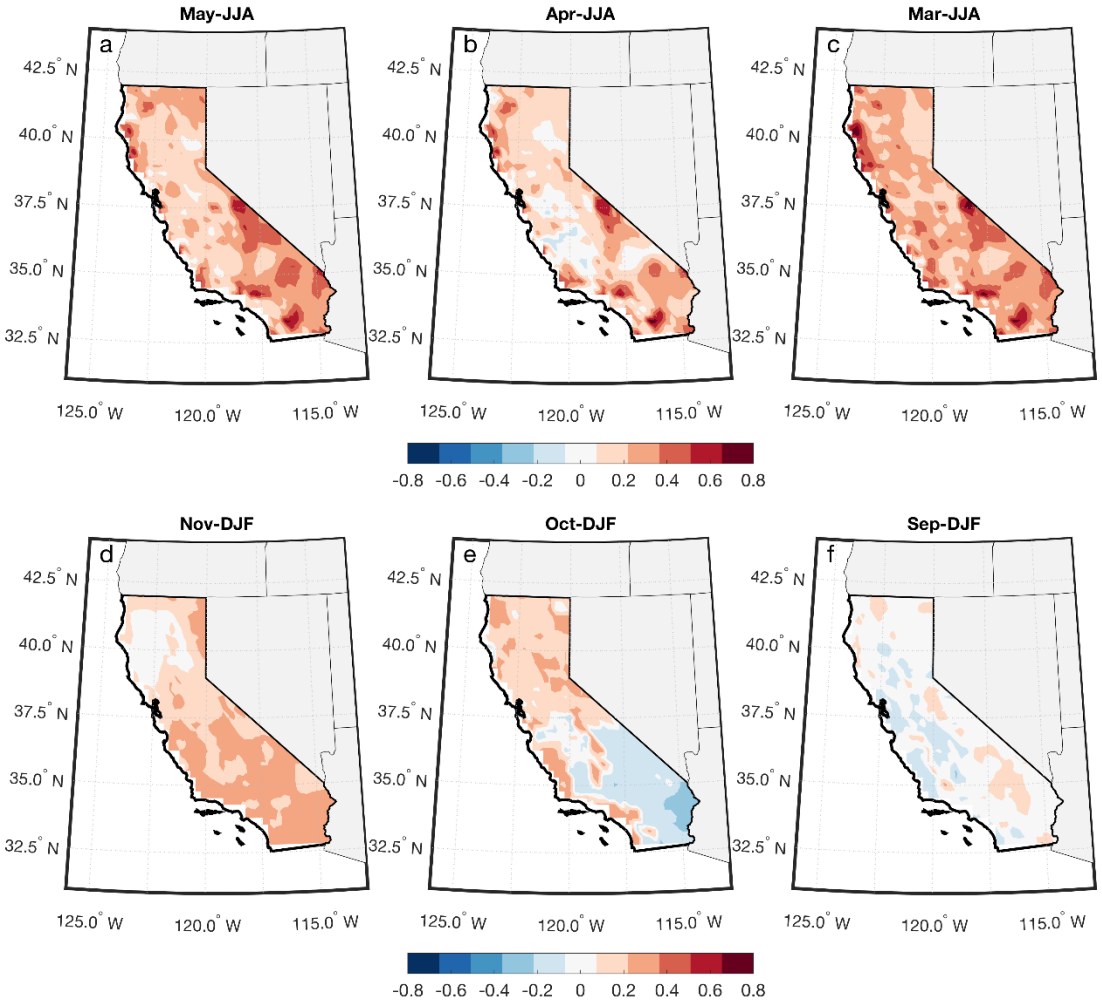

Figure 5: Tmax forecast skill using SST as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SST predicting JJA Tmax (a) and Nov SST predicting DJF Tmax (d)); two months (Apr SST predicting JJA Tmax (b) and Oct SST predicting DJF Tmax (e)); and three months (Mar SST predicting JJA Tmax (c) and Sep SST predicting DJF Tmax (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

For summer (JJA) forecasts, SST provides significant (positive) skill in predicting Tmax and Tmin variability, but SM offers little. The skill of predicting summer Tmin is somewhat higher than forecasts of Tmax. Summer Tmin skill over California using SST as a predictor is higher than 0.5 over nearly the entire State (***Figure 5***a); when using SM as a predictor the skill decays to zero and to negative values (***Figure 7***a). Likewise, the model using SST as a predictor for Tmax during the summer provides positive forecast skill although not quite as high as for Tmin, but higher than that of the model using SM (compare ***Figure 6***a with ***Figure 8***a). Moreover, the skill of SST predicting Tmax and Tmin persists with little decline throughout the three months of lead time (May, April and March leading JJA) (***Figure 5***b,c and ***Figure 7***b,c). Forecasts of summer Tmin and Tmax from SM predictors (***Figure 7***a,b,c and ***Figure 8a,***b,c) have mostly insignificant or negative skill but interestingly March SM leading JJA Tmin produces positive skill over much of the state, perhaps reflecting the fact that March SM is greater and has greater influence on subsequent air temperature anomalies than April or May SM, when soils are rapidly drying out in California.

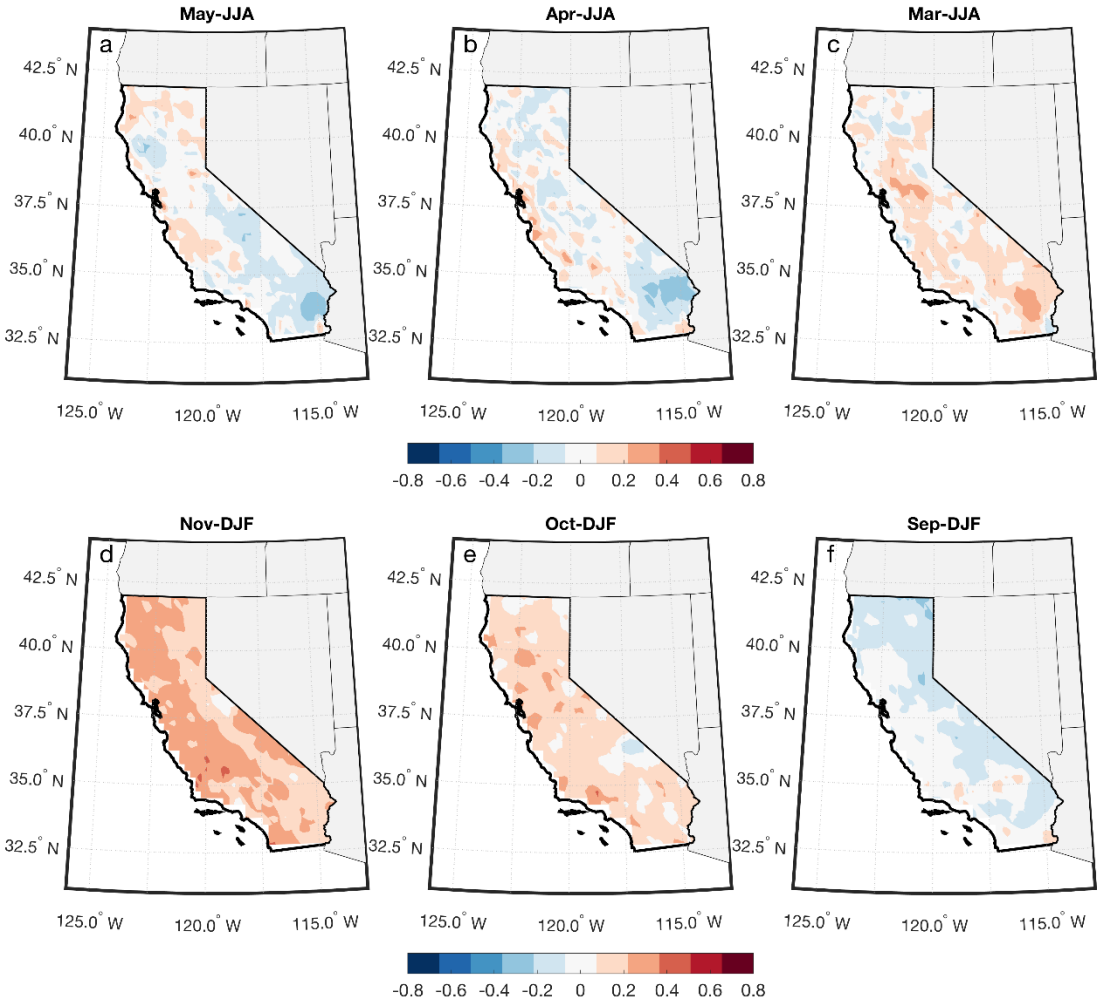

Figure 6: Tmin forecast skill using SM as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SM predicting JJA Tmin (a) and Nov SM predicting DJF Tmin (d)); two months (Apr SM predicting JJA Tmin (b) and Oct SM predicting DJF Tmin (e)); and three months (Mar SM predicting JJA Tmin (c) and Sep SM predicting DJF Tmin (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

For winter forecasts, Tmax and Tmin predictions for the winter (DJF) present some noteworthy differences when comparing to the predictions for JJA. First, SST seems to explain less of the variance for both Tmax and Tmin, given the lower correlation coefficients over California. Moreover, with increasing forecast lags, the skill decreases faster than in summer: whereas for the summer skill persists at relatively high levels up to three months of lag, when predicting the winter the skill decreases to less than 0.3 for lag 2 and it is near zero for lag 3 (compare ***Figure 5***a,b,c with ***Figure 5***d,e,f and ***Figure 6***a,b,c with ***Figure 6***d,e,f). On the other hand, the predictions of winter Tmin and Tmax made using SM as a predictor have much higher skill than the ones made for the summer (compare ***Figure 7***a,b,c with ***Figure 7***d,e,f and ***Figure 8***a,b,c with ***Figure 8***d,e,f). There is skill comparable to the one obtained using SST and sometimes higher for longer lead times. For example, the DJF Tmax forecast using SM as predictor has positive skill for the three lead times (***Figure 8***d,e,f), while the skill is partially lost for the second lag and is completely gone in the third lag when using SST as predictor (***Figure 6***d,e,f).

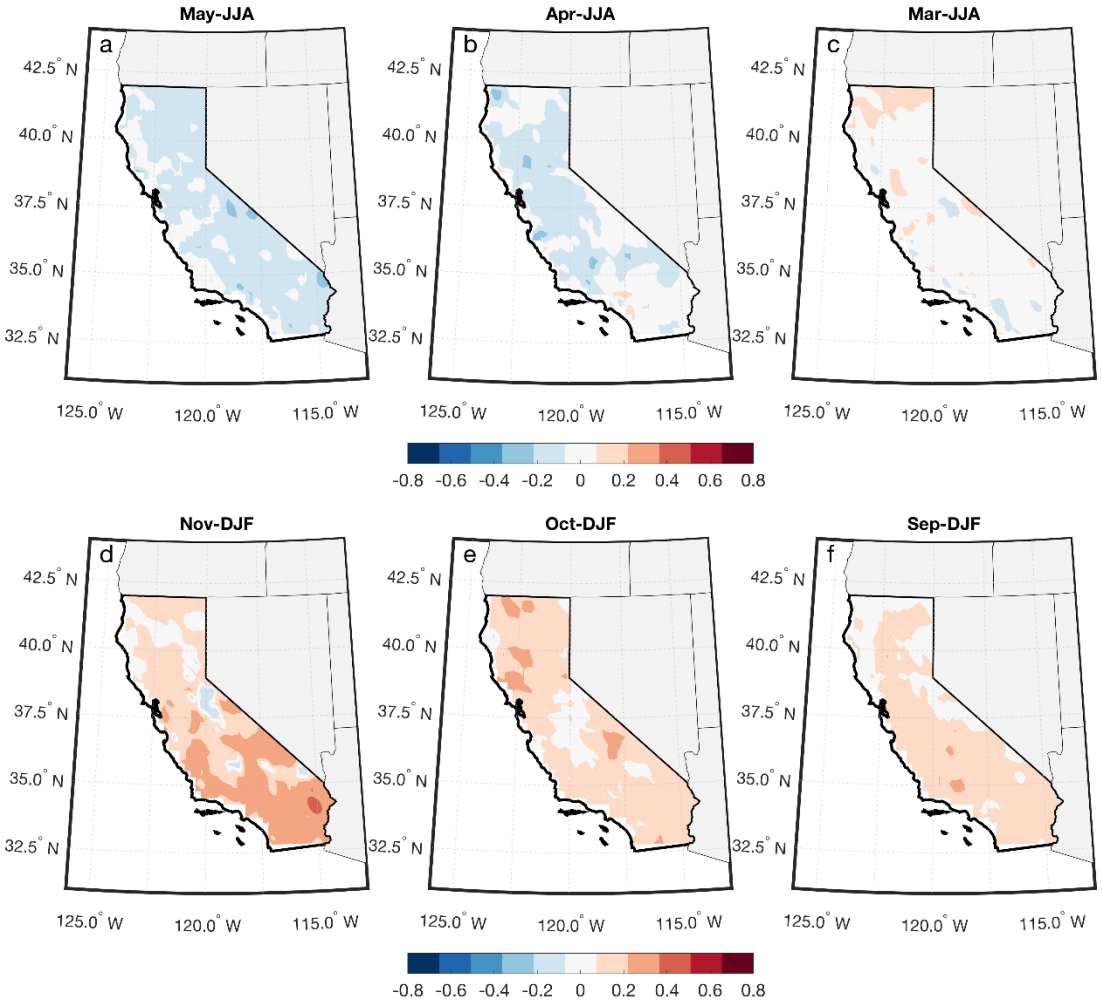

Figure 7: Tmax forecast skill using SM as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SM predicting JJA Tmax (a) and Nov SM predicting DJF Tmax (d)); two months (Apr SM predicting JJA Tmax (b) and Oct SM predicting DJF Tmax (e)); and three months (Mar SM predicting JJA Tmax (c) and Sep SM predicting DJF Tmax (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

Some observations can be made from the skill maps, of the predictability of Tmax and Tmin for different seasons, and of their relationship with SST and SM. Tmin seems to be more strongly related to the SST for both summer and winter seasons, although during the summer SST can explain more of the Tmin variability. The greatest association of SST with subsequent air temperature over California is found in summer, given the high correlation values for the Lag 3 predictions; this is true both for Tmax and especially for Tmin. On the contrary, SM exhibits a relatively strong influence on winter Tmax, but has less effect on winter Tmin, especially for higher lags. During summer (JJA), SM plays a negligible role in controlling either Tmax and Tmin over California. Therefore, Tmin over California seems to be more influenced by large-scale climate patterns that are expressed as SST variability, and this effect is stronger during summer. For the winter, those large-scale climate fluctuations also influence both Tmin and Tmax, although local modulation by SM variability may also operate, as indicated by positive forecast skill using SM predictors. Furthermore, the forecast skill for Tmin and Tmax obtained

by these linear statistical models provides comparable and sometimes better than the forecast skill for mean air temperatures over California obtained by the North American Multi-Model Ensemble (NMME). As described by Zhang et al. (2018), NMME forecast skill for California for summer (JJA) and winter (DJF) are very low or insignificant in areas where the present linear models are relatively high, namely for Southern California and throughout the coastal California region.

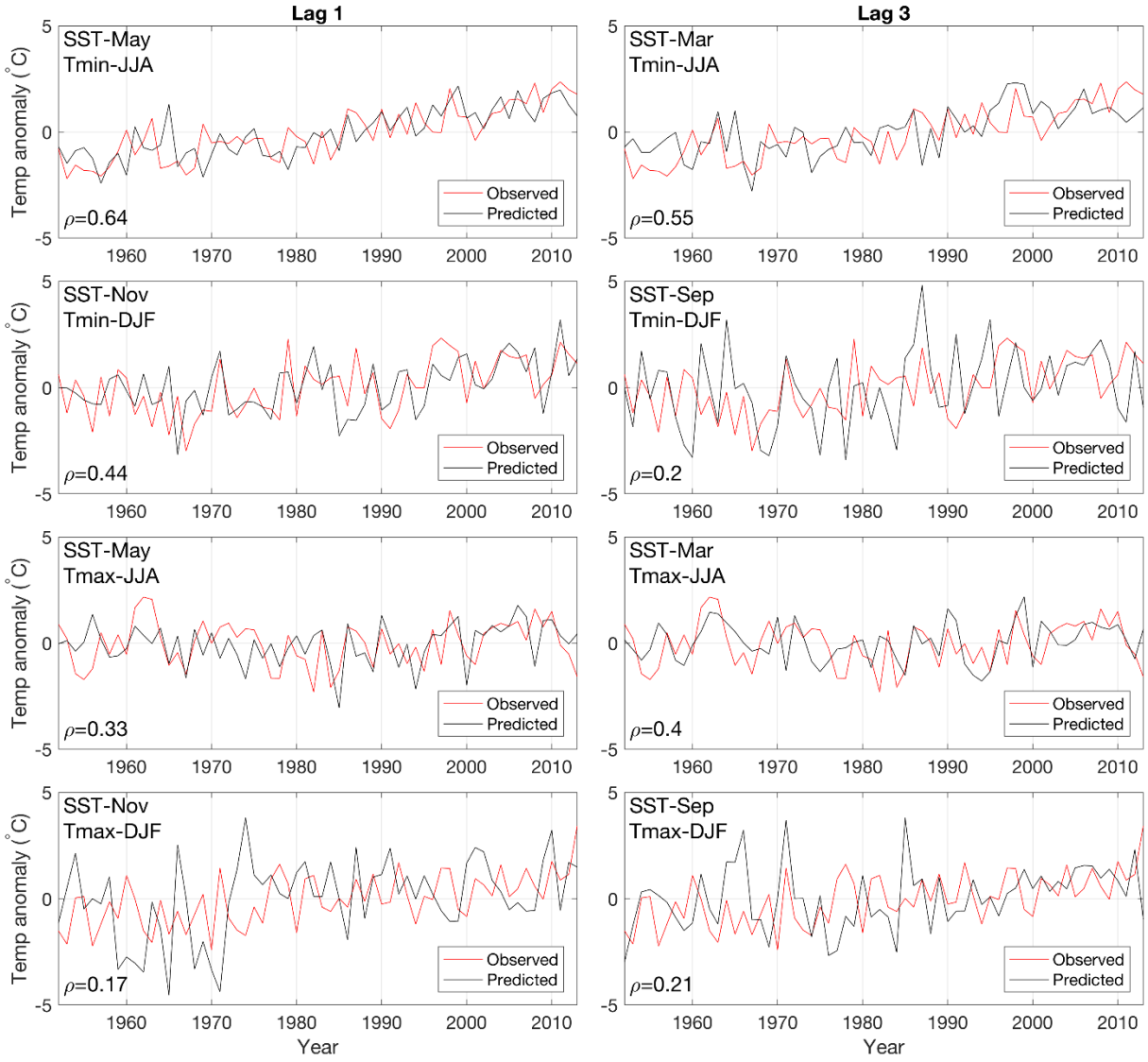

Figure 8: Time series for the predicted (black) and observed (red) air temperature for Sacramento (38.58° N and 121.49° W), for all the experiments using SST as a predictor, and for lags 1 (left panels) and lag 3 (right panels). ρ shows the correlation between the two time-series.

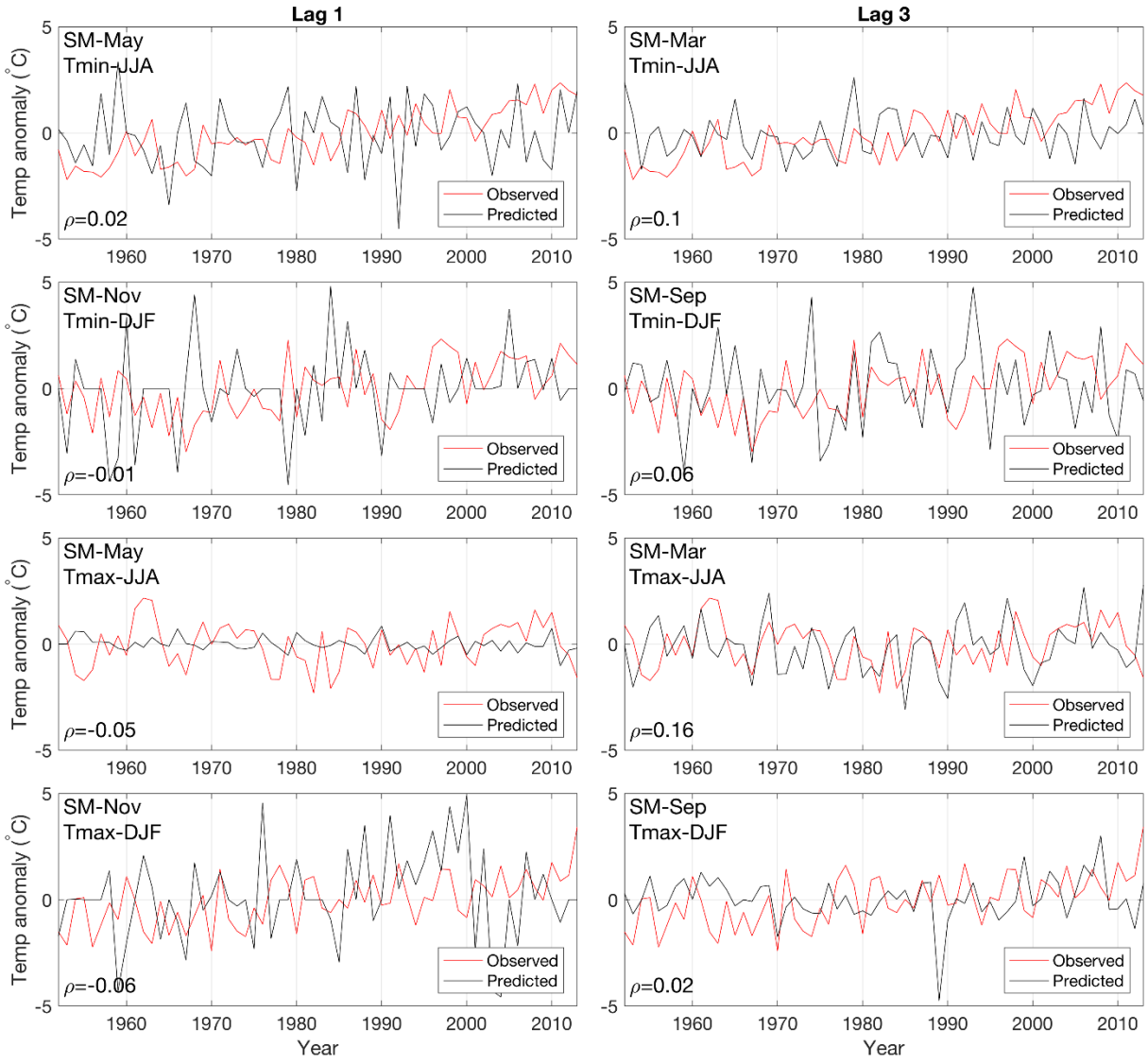

Figure 9: Same as Figure 9, but showing the time-series for the experiments using SM as predictor.

To illustrate the predictions made for a specific region over the period of record, we show the time series of the predicted and the observed Tmin and Tmax for the city of Sacramento (38.58° N and 121.49° W), for Lag 1 and Lag 3, for both summer and winter season, and both SST (**Figure 9**) and SM predictors (**Figure 10**). Those time series reinforce the description provided above: SST variations modulate more strongly Tmin, given the higher correlation values between observed and predicted; this influence is stronger in the summer; Tmax is also influenced by SST, although it is weaker. For this specific region, the correlation values for the models that used SM as a predictor are very low, with the highest value being 0.16 for the prediction of Tmax for the summer.

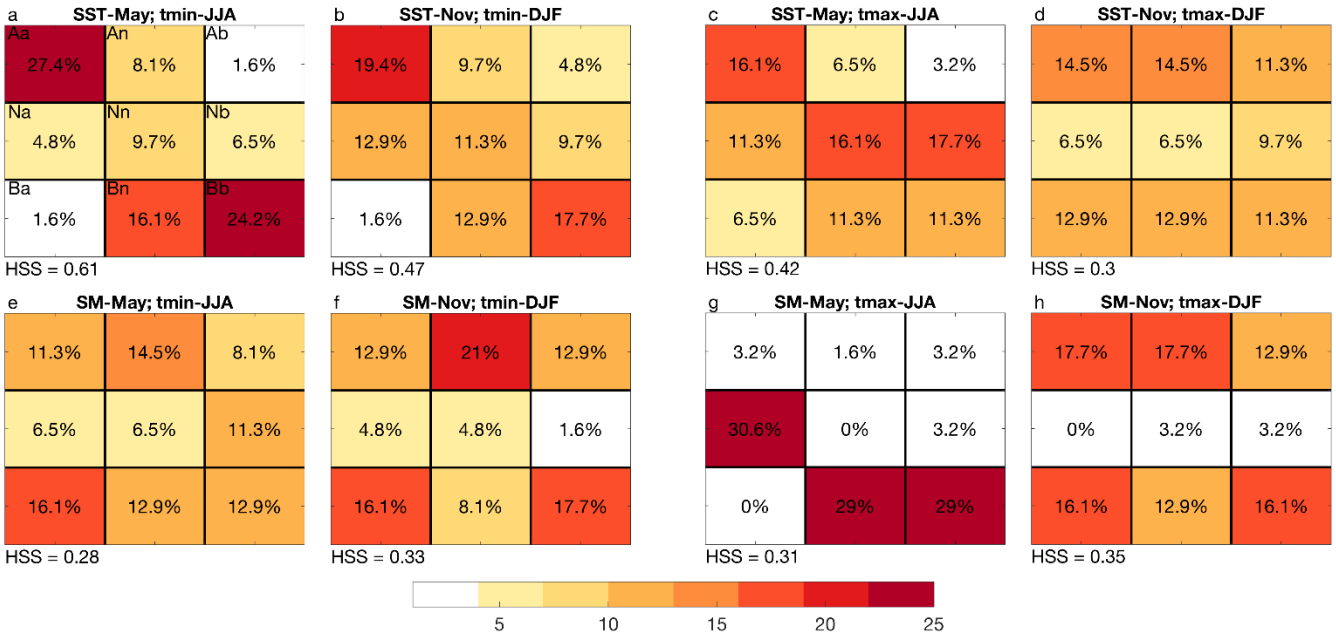

Figure 10: Performance of lag 1 forecasts evaluated for three categories (Above, Normal, Below) for Sacramento (38.58° N and 121.49° W) Tmin (leftmost panels) and Tmax (rightmost panels). Forecasts using SST (upper panels) and SM (lower panels) predictors are shown for May predicting JJA (columns 1 and 3) and November forecasting DJF (columns 2 and 4). Errors are calculated by first splitting the observations and forecast values into three equally occurring classes: above (A), with values larger than the 66th percentile); normal (N), with values between the 33rd and 66th percentile; and below (B), with values below the 33rd percentile. The percentages represent the number of the predicted values that falls into one of those categories: (1) Diagonal (Aa, Nn, Bb), where the predicted values are in the same category as the observed; (2) Class one error (An, Nb, Na, Bn), where the predicted values are one category below or above the observed; (3) Class two errors (Ab and Ba), where the predicted values are two categories below (Ab) or above (Ba) the observed. Capital letters represent the observations and lowercase letters the forecast. The ideal model would have 33% along the diagonal, indicating that the terciles of the predicted values match with the terciles of the observations. The values in the lower left corner of each figure represent the Heidke Skill Score (HSS), in which 1 represents a perfect model, 0 indicates no skill, and a negative value indicates that the forecasts are worse than a randomly generated forecast.

Finally, **Figure 11** shows the errors associated with the predictions for Sacramento for each model. To calculate those errors, the observations were divided in three equally occurring tercile categories, above normal, normal and below normal and then the percentage of the predicted values that fall into one of those categories were calculated. A high-performing forecast would have most occurrences along upper-left to lower-right diagonal, where both observations and predictions are in the same category (Aa, Nn, Bb), while and the worst model is the one that have greater values in the furthermost off-diagonal (class two errors), when the observed values are above the normal and the predicted values are below the normal (Ab) or vice-versa (Ba). **Figure 11** also reports the Heidke Skill Score (HSS) (see Section 2.3.2) for each of the models performed.

When using SST as a predictor, especially to predict Tmin, most of the predicted values are in the same tercile category as the observed values; the highest percentages lie along the diagonal. This is also reflected in the HSS values, which are the highest for the experiments using SST as a

predictor and Tmin as predictands (0.61 for the summer predictions and 0.47 for the winter). Two-class errors (when the above normal is forecast and below normal is observed or below normal is forecast and above normal is observed) are relatively low. The exception is the experiment predicting Tmax for the winter (**Figure 11**d): the values among the categories are similar and the two-class errors have relatively large values (11.3% for Ab and 12.9% for Ba) when compared to the other experiments using SST as a predictor (**Figure 11**a,b,c); also HSS is lower (0.3 against 0.61, 0.47 and 0.42). Those experiments present low two-class errors, with the maximum being 6.5% for the summer Tmax predictions. On the other hand, the experiments using SM as a predictor present bigger errors and lower HSS in most of the cases. With the exception of Tmax predictions for the summer, when two-class errors are smaller than 3.2% (**Figure 11**g), the other experiments present those errors varying between 12.9% and 16.1%, which are also similar to the values along the diagonal (**Figure 11**e,f,h). Moreover, HSS is generally lower (HSS<0.35) than most of the experiment using SST as predictor, except for the Tmax forecast for the winter (HSS=0.3). As was shown in the maps for California (***Figure 7*** and ***Figure 8***) using SM predictors for summer Tmin and Tmax forecasts for the Sacramento region yields very little skill.

4: Conclusions and Future Directions

We developed and explored a statistical prediction model for minimum and maximum air temperature anomalies (Tmax and Tmin) over California using remote and local predictors. The first predictor variable is sea surface temperature anomalies (SST) across the tropical and northern Pacific basin, representing the influence of large-scale climate variability patterns, which in turn affect local surface air temperature. The second predictor variable is antecedent soil moisture (SM), which is thought to represent the local effects on surface air temperature that affect the atmospheric energy balance through the interaction between latent and sensible heating.

These experiments indicate that both local and remote predictors influence the predictability of air temperature, and that this influence is dependent on the season that is being predicted as well as if the prediction were made for Tmin or Tmax. For summer (JJA) air temperature, antecedent SST exhibited a strong effect, but SM did not. On the other hand, for winter (DJF) air temperatures, antecedent SST and antecedent SM exhibited skill in predicting Tmin and Tmax. However, those two predictands also present some important differences: while SST influence both Tmin and Tmax, the influence is stronger for Tmin for both summer and winter. On the other hand, for winter Tmax, SM influences more strongly Tmax during winter, especially for lags of two and three months.

The results of this study yielded some agreement and some interesting differences from those of a prior study by Alfaro et al. 2006 (A06), which also used Pacific SST and an index of soil moisture as predictors of seasonal air temperatures (applied to a larger region than explored here). As in A06, the present study demonstrates a degree of skill from both predictor variables. However, while A06 found greater predictive skill of summer air temperatures from SM than SST, the present study finds greater skill from SST predictors, and very little skill from SM predictors. And while A06 found more skill from SM in predicting Tmax than Tmin, the present study does not (if anything, Tmin skill from SM is higher). These differences are likely a result

of the larger Pacific coast to the Mississippi basin domain of the A06 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). The processes involved in the positive skill of winter Tmin and Tmax forecasts from SM found here for the California region are not clearly understood and will require further study.

Although other studies have applied a similar statistical approach to predict Tmin and Tmax over the United States, this work bridges a gap between generalized country-wide studies and specific, local studies. These results indicate that statistical methods can provide modest predictability for summer and winter seasonal mean temperature anomalies over much of California. Moreover, the forecast skill for Tmin and Tmax obtained by our statistical model is comparable and sometimes better than the forecast skill for mean air temperatures over California obtained by the North American Multi-Model Ensemble (NMME). As shown by Zhang et al. (2018), NMME forecast skill for California for summer (JJA) and winter (DJF) is very low or insignificant, especially for the Southern and coastal California region.

While the results contain positive forecast skills, they produced considerable variability across seasons in model predictors and forecast performance of seasonal air temperature for California and the overall far western North America region. Several of the predictor patterns relate to large scale climate variability patterns. Somewhat surprisingly, higher skill was obtained for summer (JJA) forecasts than for winter (DJF), when many of the climate variability patterns have stronger expressions. The results demonstrate the importance of careful analysis when forming a statistical forecast model to determine the mix of predictor variables, and the number of modes of predictor and predictand variables that should be considered.

Given updated data representing SST and SM predictors, the forecast tool presented here can be applied to perform real-time forecasts. With that, linear models can estimate seasonal minimum and maximum air temperature anomalies in the following season. Although the temperature is expected to increase on average, the frequency of extreme conditions is also expected to increase in California with climate change in the coming decades. As a result, season-ahead forecasts should include confidence intervals to help planners understand the level of uncertainty and potential for unexpected extreme events to occur during the season.

Looking forward, the skill obtained from the set of linear statistical forecasts considered here suggests that more insight can be gained from additional experiments. Importantly, some of the present results produced positive skill with little drop off out to three months, so it will be useful to investigate forecast skills from similar models that have predictors at extended lead times. Additionally, it will be useful to investigate how such models perform in predicting other variables, such as warm extreme (heat waves or “heat storms”) occurrence in California and across the western U.S. The present study only considered forecasts of summer and winter mean Tmax and Tmin, but it would be of interest to explore predictability of other three month seasons. The results here were confined to forecasts constructed from single variable predictors (i.e. only SST or only SM), but it would be interesting to explore models that combine both predictors (SST and SM) to determine if skill provided by the two predictors is additive. This seems especially interesting potentially for the winter, when both SST and SM produced significant forecast skill.

This report provides more insight into seasonal air temperature variability over California and advances the development of a linear statistical technique for seasonal forecasting. To be ready for use by the energy sector, further investigation is needed to better understand the forecast uncertainties associated with the variability across seasons and model predictors and it is necessary the use of observational real-time data to perform real-time forecasts that can inform decision makers. We hope that further efforts by others who are well placed in the utility industry and other applications will be able to contribute in taking next steps toward practical seasonal forecasting.

Table 5: Correlation coefficients between seasonal averages of Tmin and Tmax anomalies averaged over the California region.

Season	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
Correlation	0.41	0.43	0.60	0.65	0.58	0.50	0.38	0.35	0.35	0.33	0.37	0.44

Table 6: Optimum number of Principal Components (PCs) and Canonical Correlates (CCs) for all one month lag summer (JJA) and winter (DJF) forecast experiments, the field-averaged skill for the optimum model (Skill), and the lower (rl) and upper (ru) limits of the 95th confidence interval of skill obtained from 100 randomized experiments.

Season Predicted	**Predictor**	**Predictand**	**PC**	**CC**	**Skill**	**rl**	**ru**
JJA	SST	Tmin	16	16	0.21	-0.003	0.013
DJF	SST	Tmin	10	9	0.20	-0.011	0.011
JJA	SST	Tmax	16	16	0.24	-0.001	0.017
DJF	SST	Tmax	14	12	0.26	-0.008	0.026
JJA	SM	Tmin	11	10	0.13	-0.007	0.010
DJF	SM	Tmin	10	6	0.15	-0.008	0.012
JJA	SM	Tmax	2	2	0.22	0.001	0.026
DJF	SM	Tmax	13	5	0.10	-0.001	0.016

Table 7: Variance explained by each predictand Canonical Correlate (rows, CC mode), for each experiment (columns). The number of CC modes included in each experiment was determined by the optimization process (see *Table 2* and Section 3.1 for details).

	SST				SM			
CC mode	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**
1	34.7%	14.8%	4.9%	8.1%	10.3%	17.4%	50.1%	57.0%
2	3.9%	1.0%	4.7%	7.5%	1.3%	10.8%	49.9%	41.3%
3	7.9%	0.5%	1.8%	5.7%	0.4%	0.0%	--	0.0%
4	0.3%	7.4%	19.8%	7.9%	20.1%	70.9%	--	1.1%
5	9.4%	6.4%	11.9%	0.0%	23.3%	0.7%	--	0.6%
6	0.1%	0.6%	0.5%	28.0%	17.8%	0.2%	--	--
7	0.2%	8.6%	0.0%	9.6%	4.8%	--	--	--
8	12.0%	44.1%	12.2%	2.3%	3.2%	--	--	--
9	4.4%	16.7%	5.1%	1.1%	1.8%	--	--	--
10	3.3%	--	2.5%	6.1%	17.1%	--	--	--
11	4.5%	--	10.7%	22.7%	--	--	--	--
12	0.5%	--	9.7%	0.9%	--	--	--	--
13	8.2%	--	0.0%	--	--	--	--	--
14	2.6%	--	9.0%	--	--	--	--	--
15	4.4%	--	1.5%	--	--	--	--	--
16	3.8%	--	5.6%	--	--	--	--	--

Table 8: Variance explained by each predictor Canonical Correlate (rows, CC mode), for each experiment (columns). The number of CC modes included in each experiment was determined by the optimization process (see *Table 2* and Section 3.1 for details).

	SST				SM			
CC mode	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**	**Tmin JJA**	**Tmin DJF**	**Tmax JJA**	**Tmax DJF**
1	14.1%	3.3%	7.9%	0.1%	20.8%	35.1%	50.0%	33.2%
2	0.9%	23.2%	5.4%	17.5%	30.0%	15.9%	50.0%	11.6%
3	0.0%	16.3%	8.0%	2.1%	5.1%	28.7%	--	0.9%
4	2.0%	13.0%	0.2%	0.1%	14.6%	1.1%	--	14.4%
5	6.1%	12.5%	3.0%	14.3%	0.0%	18.8%	--	40.0%
6	16.7%	2.0%	5.8%	0.8%	0.0%	0.4%	--	--
7	0.0%	9.8%	2.7%	4.8%	0.2%	--	--	--
8	21.9%	2.7%	2.7%	19.8%	1.8%	--	--	--
9	21.2%	17.1%	17.7%	31.0%	26.6%	--	--	--
10	2.3%	--	1.5%	1.3%	0.9%	--	--	--
11	1.6%	--	5.3%	2.4%	--	--	--	--
12	2.8%	--	6.1%	5.8%	--	--	--	--
13	5.2%	--	18.4%	--	--	--	--	--
14	1.4%	--	3.7%	--	--	--	--	--
15	1.0%	--	5.2%	--	--	--	--	--
16	2.8%	--	6.5%	--	--	--	--	--

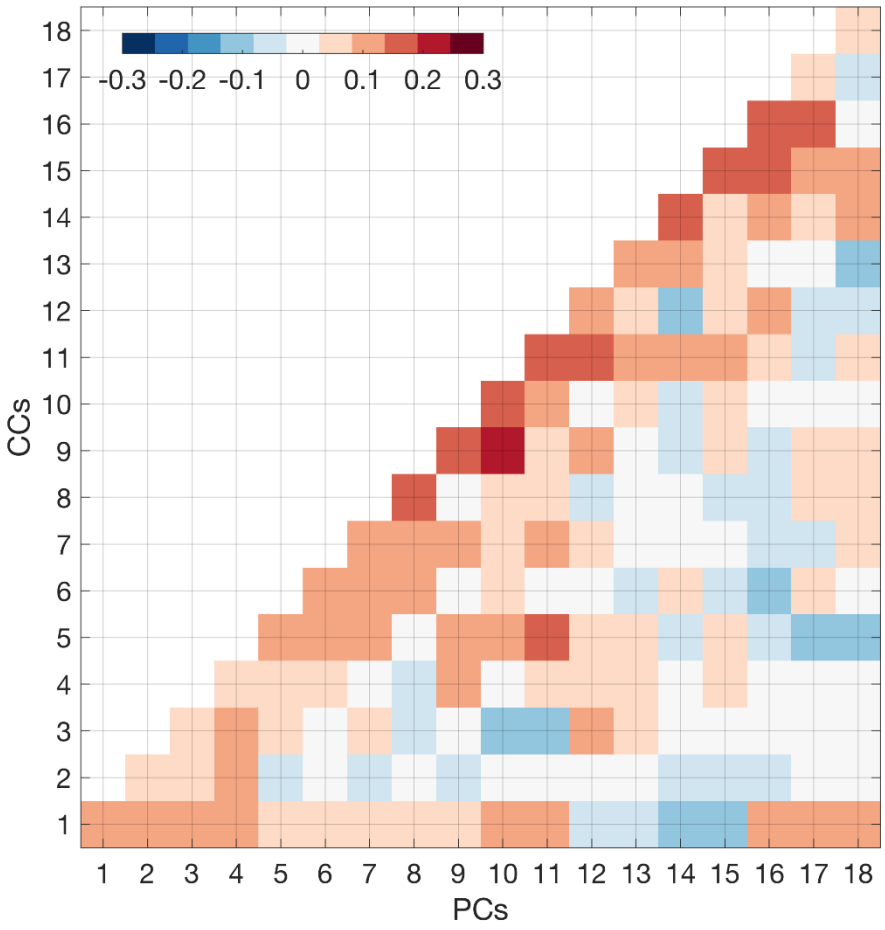

Figure 11: Skill optimization for DJF Tmin prediction using SST as predictor. Figure shows the cross-validated field-averaged skill, expressed as the correlation between the predicted and the observed temperatures, using different combination of the Principal Components (PCs) and the Canonical Correlates (CCs). The optimum is defined the number of PC and CC modes that yielded the highest cross validated skill (PC=10 and CC=9).

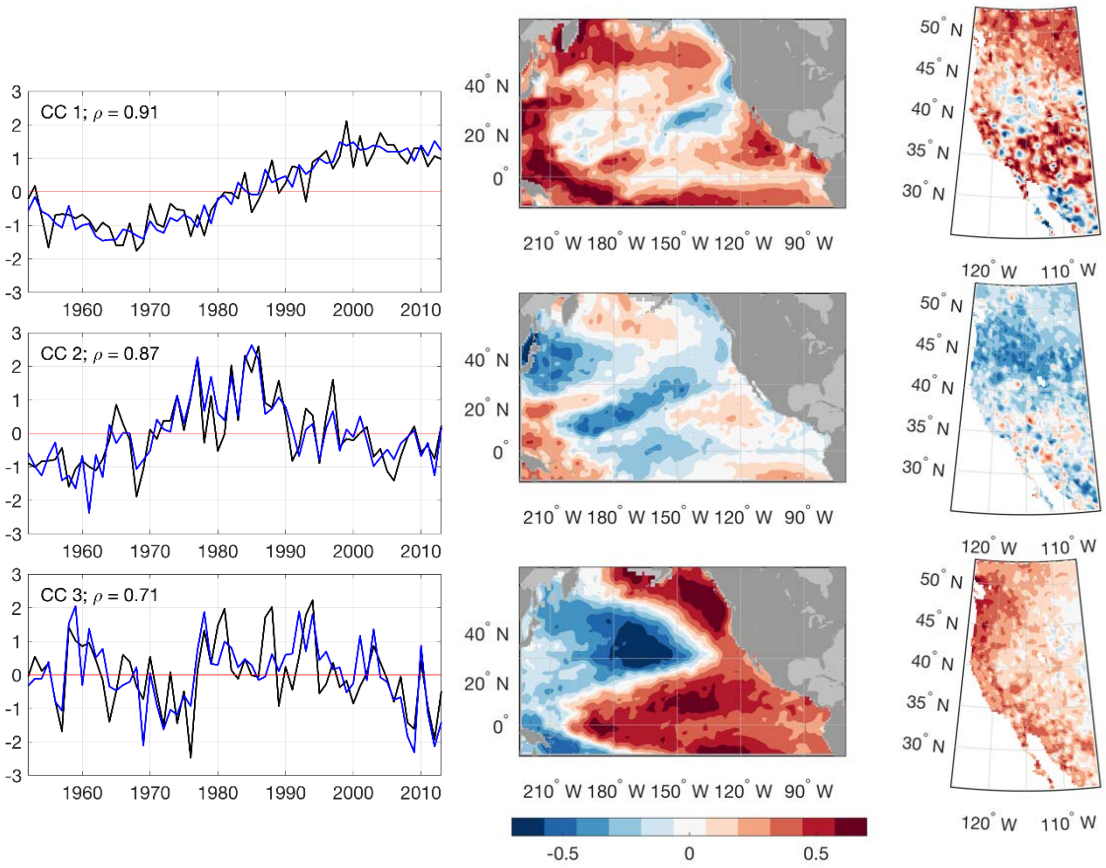

Figure 12: First three coupled CCA modes of DJF Tmin related to November SST. Left panel displays the time series of the CCA modes (CC1, CC2 and CC3), where predictor (SST) is in black and predictand (Tmin) is in blue; ρ is the correlation between the two time-series (CC1=0.91; CC2=0.87 and CC3=0.71). Middle and right panels show the spatial pattern associated with the SST and Tmin fields, respectively, expressed as correlation between the CCA mode time series and their respective variable field for each location.

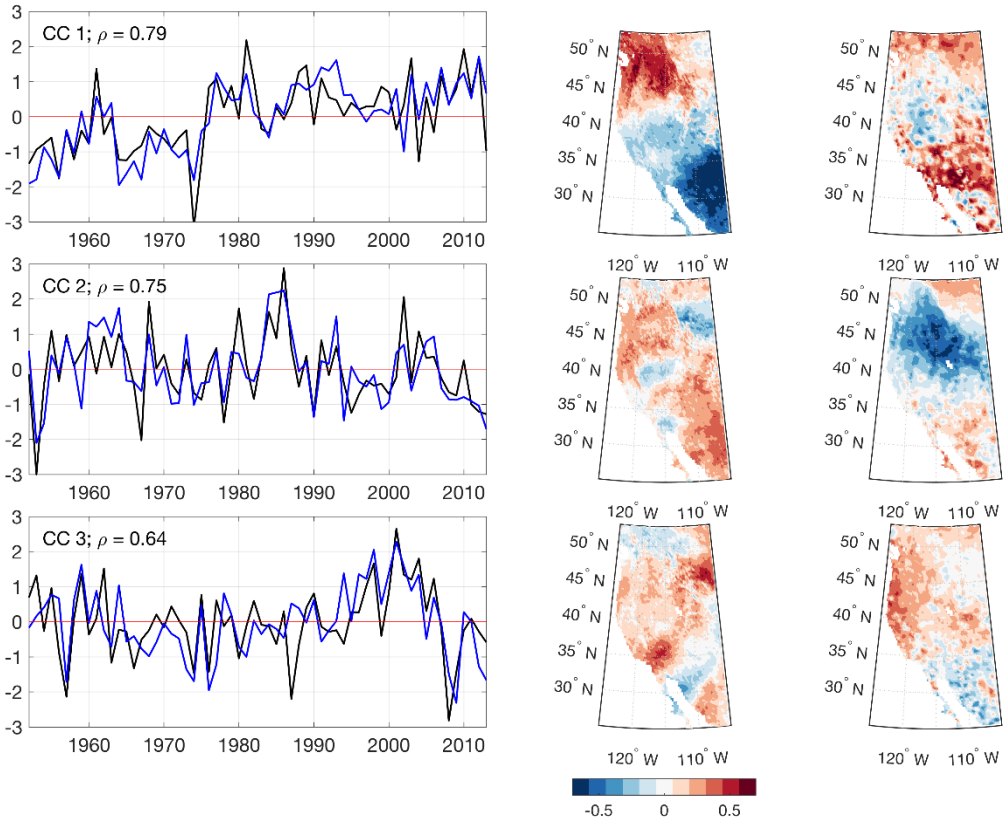

Figure 13: First three coupled CCA modes of DJF Tmin related to November SM. Left panel displays the time series of the CCA modes (CC1, CC2 and CC3), where predictor (SM) is in black and predictand (Tmin) is in blue; ρ is the correlation between the two time-series (CC1=0.79; CC2=0.75 and CC3=0.64). Middle and right panels show the spatial pattern associated with the SM and Tmin fields, respectively, expressed as correlation between the CCA mode time series and their respective variable field for each location.

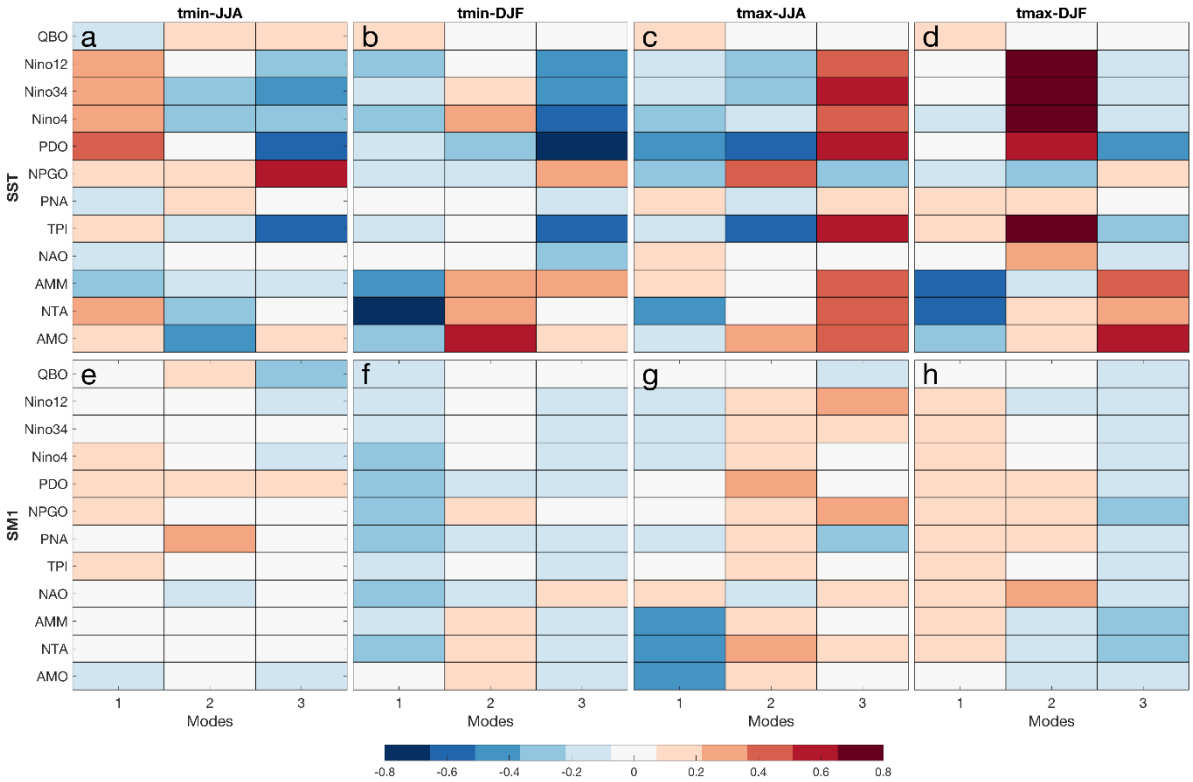

Figure 14: Correlation between the time series of selected climate modes (named along the y-axis) with the first three canonical correlates (CC) for all the lag 1 forecast experiments (x-axis). (a) SST predicting summer (JJA) Tmin; (b) SST predicting winter (DJF) Tmin; (c) SST predicting JJA Tmax; (d) SST predicting DJF Tmax; (e) SM predicting JJA Tmin; (f) SM predicting DJF Tmin; (g) SST predicting JJA Tmax; (h) SM predicting DJF Tmax. Climate modes are as follows: QBO (quasi biennial oscillation); Nino1.2, Nino3.4 and Nino4 are the different indices for El Niño Southern Oscillation (ENSO), correspondent to Niño 1+2 (Extreme Eastern Tropical Pacific SST), Niño 3.4 (East Central Tropical Pacific SST), and Niño 4 (East Central Tropical Pacific SST), respectively; PDO (Pacific Decadal Oscillation); NPGO (North Pacific Gyre Oscillation); PNA (Pacific North American Index); TPI (Tripole Index for Interdecadal Pacific Oscillation); NAO (North Atlantic Oscillation); AMM (Atlantic Meridional Mode); NTA (North Tropical Atlantic SST Index); and AMO (Atlantic Multidecadal Oscillation). The time series for those indices can be found at: https://www.esrl.noaa.gov/psd/data/climateindices/list/

Table 9: Description of each climate mode used in Figure 4

Index	Name	Description
QBO	Quasi-Biennale Oscillation	Oscillation of the equatorial wind between easterlies and westerlies in the tropical stratosphere. (Baldwin et al., 2001)
Nino1.2	ENSO	Extreme Eastern Tropical Pacific SST: SST anomalies between 0 – 10S, 90W – 80W
Nino3.4	ENSO	East Central Tropical Pacific SST: SST anomalies between 5N – 5S, 160E – 150W

Nino4	ENSO	Central Tropical Pacific SST: SST anomalies between 5N – 5S, 170W – 120W
PDO	Pacific Decadal Oscillation	The leading principal component of monthly SST anomalies in the North Pacific Ocean (Mantua et al., 1997).
NPGO	North Pacific Gyre Oscillation	The second leading principal component of sea surface height in the Northeast Pacific (Di Lorenzo et al., 2008)
AMM	Atlantic Meridional Mode	Leading mode of the maximum covariance analysis between SST and 10 m wind fields between 21S – 32N and 74W – 15W (Chiang and Vimont, 2004).
NTA	North Tropical Atlantic Index	Timeseries of SST anomalies averaged over 10N – 20N and 79W – 20W (Penland and Matrosova, 1998)
AMO	Atlantic Multidecadal Oscillation	Average anomalies of SST in the North Atlantic basin, typically over 0-80N (Enfield et al., 2001).

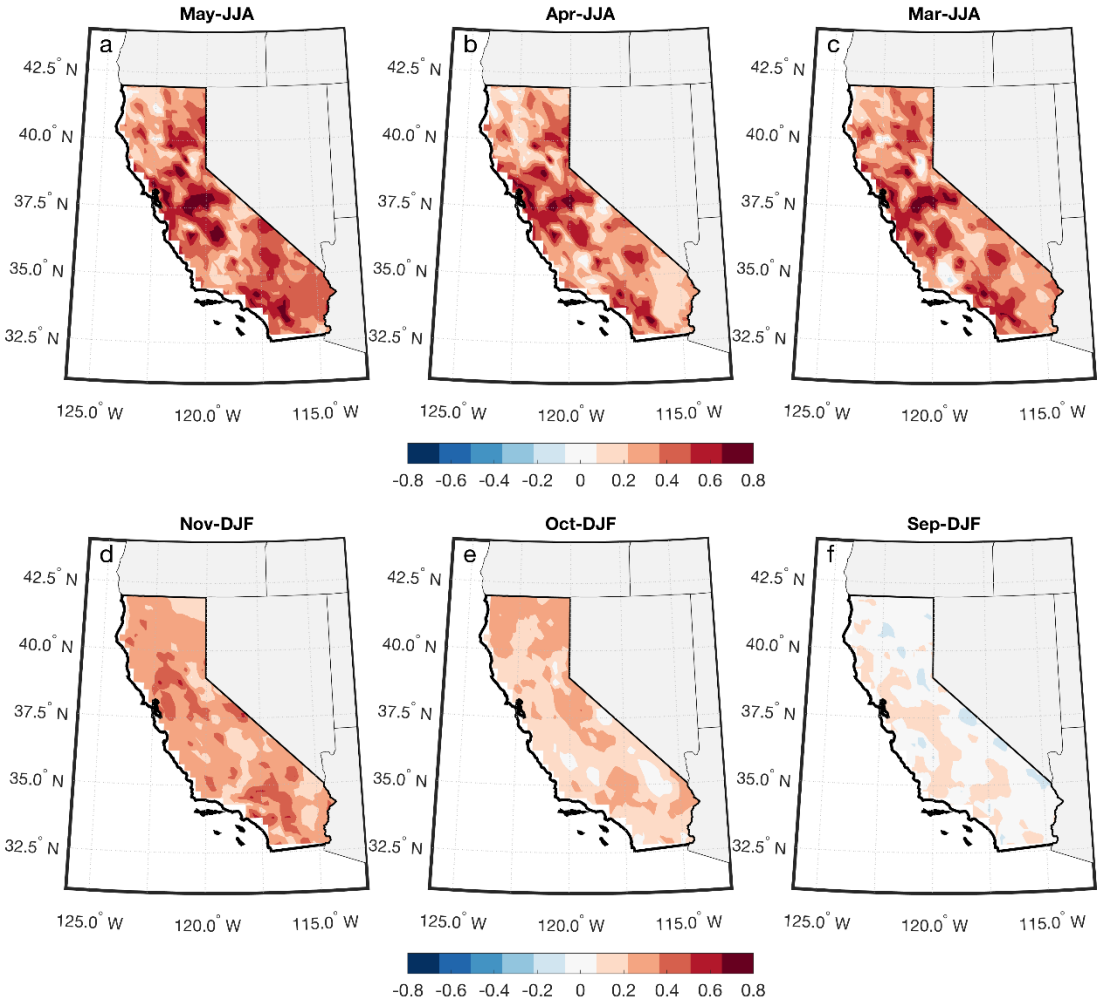

Figure 15: Tmin forecast skill using SST as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SST predicting JJA Tmin (a) and Nov SST predicting DJF Tmin (d)); two months (Apr SST predicting JJA Tmin (b) and Oct SST predicting DJF Tmin (e)); and three months (Mar SST predicting JJA Tmin (c) and Sep SST predicting DJF Tmin (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

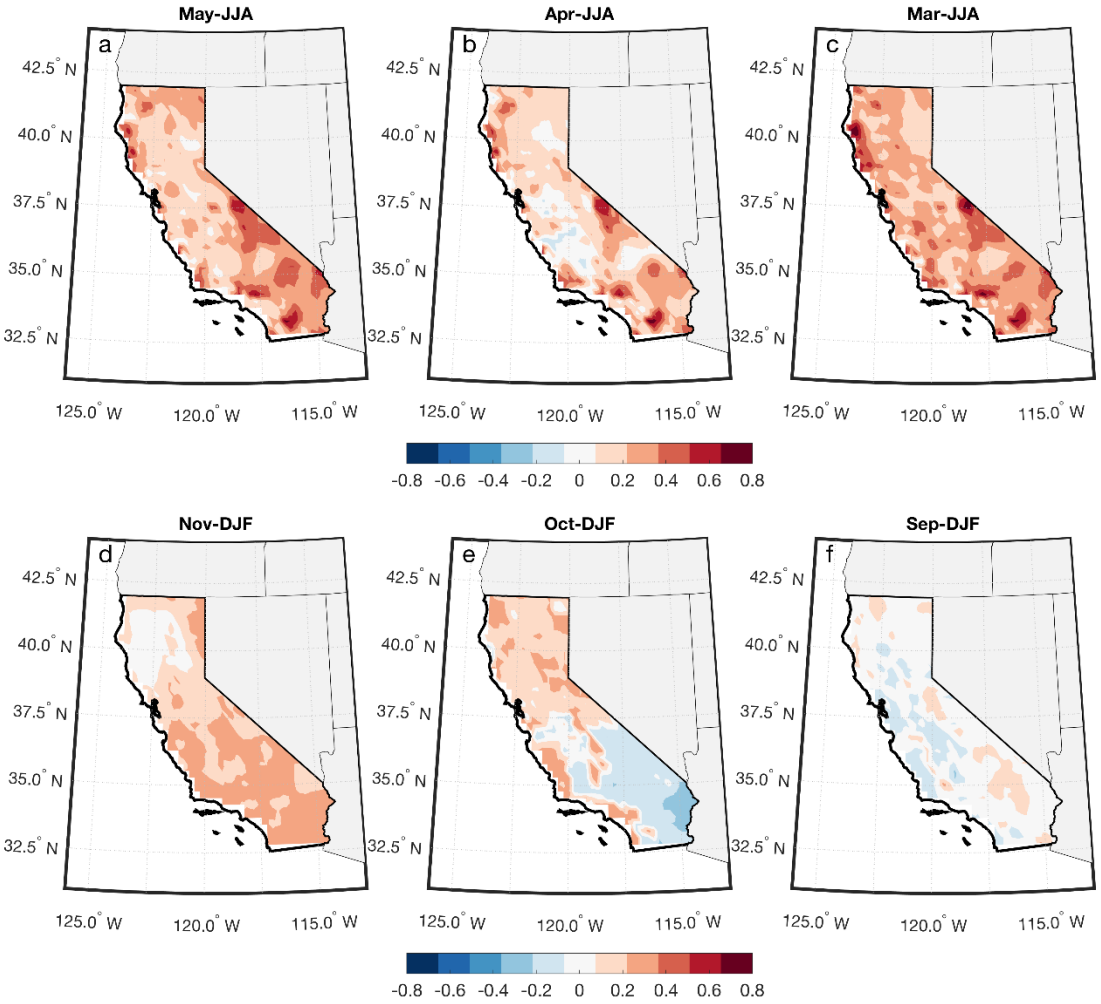

Figure 16: Tmax forecast skill using SST as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SST predicting JJA Tmax (a) and Nov SST predicting DJF Tmax (d)); two months (Apr SST predicting JJA Tmax (b) and Oct SST predicting DJF Tmax (e)); and three months (Mar SST predicting JJA Tmax (c) and Sep SST predicting DJF Tmax (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

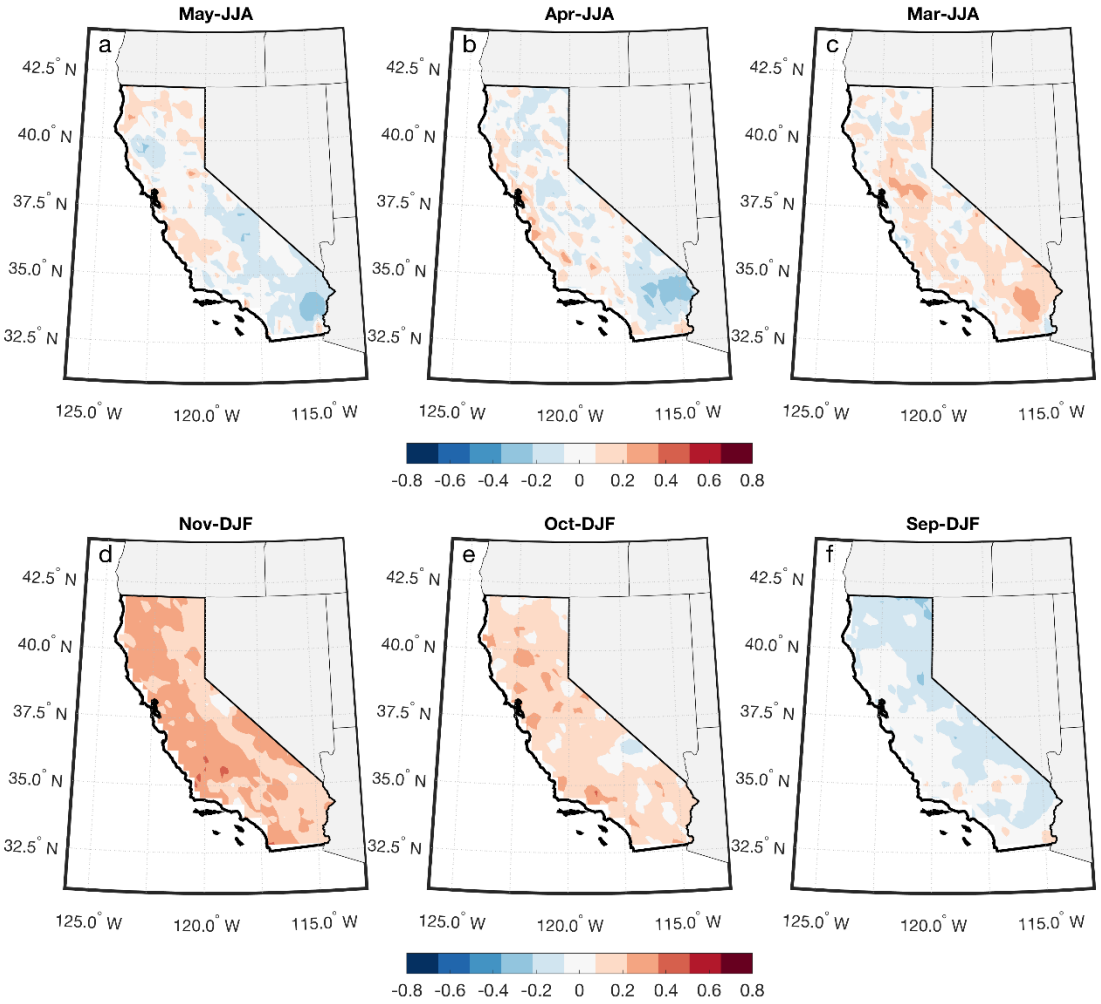

Figure 17: Tmin forecast skill using SM as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SM predicting JJA Tmin (a) and Nov SM predicting DJF Tmin (d)); two months (Apr SM predicting JJA Tmin (b) and Oct SM predicting DJF Tmin (e)); and three months (Mar SM predicting JJA Tmin (c) and Sep SM predicting DJF Tmin (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

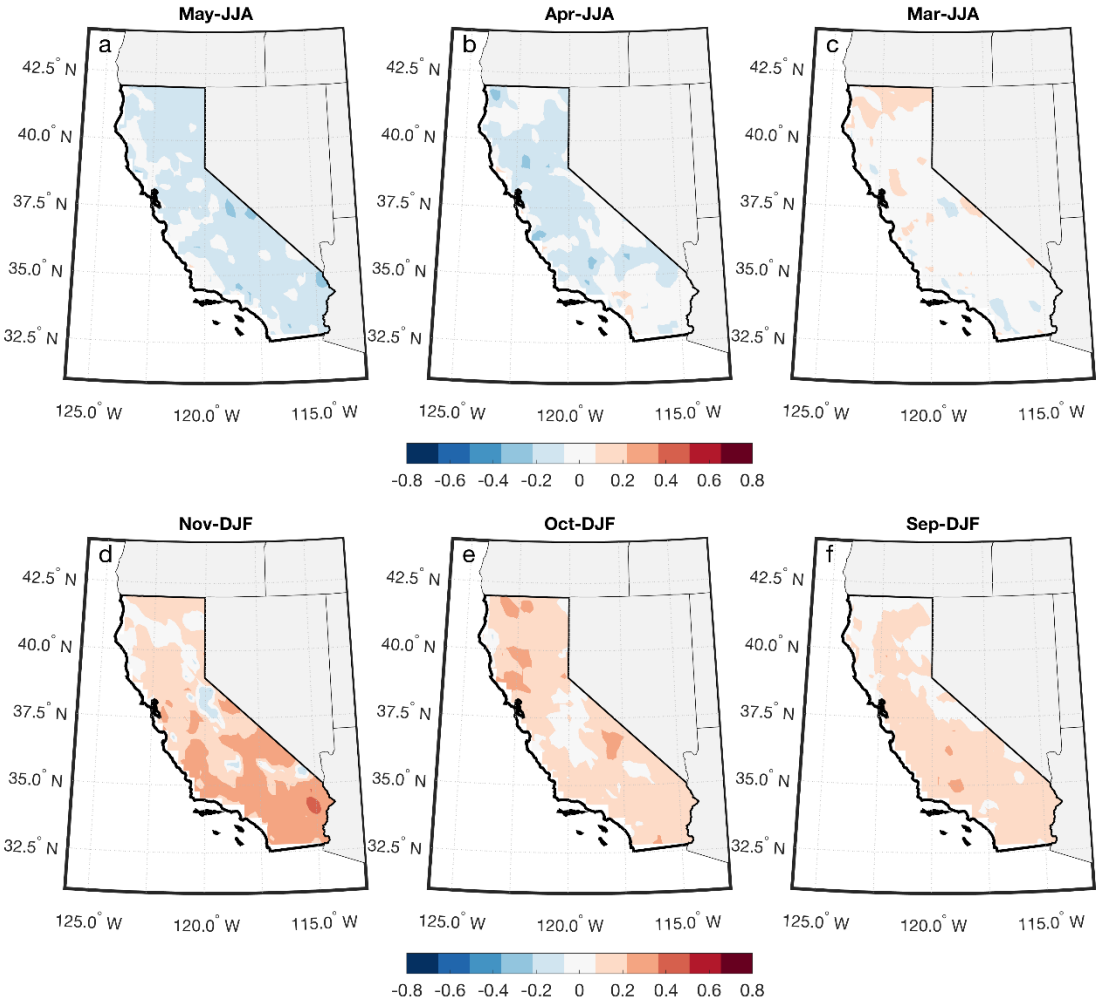

Figure 18: Tmax forecast skill using SM as predictor for JJA (top panel) and DJF (bottom panel). The forecast was made for different lags: one month (May SM predicting JJA Tmax (a) and Nov SM predicting DJF Tmax (d)); two months (Apr SM predicting JJA Tmax (b) and Oct SM predicting DJF Tmax (e)); and three months (Mar SM predicting JJA Tmax (c) and Sep SM predicting DJF Tmax (f)). Skill is expressed as correlation between the cross-validated CCA forecast and observation at each grid point in California.

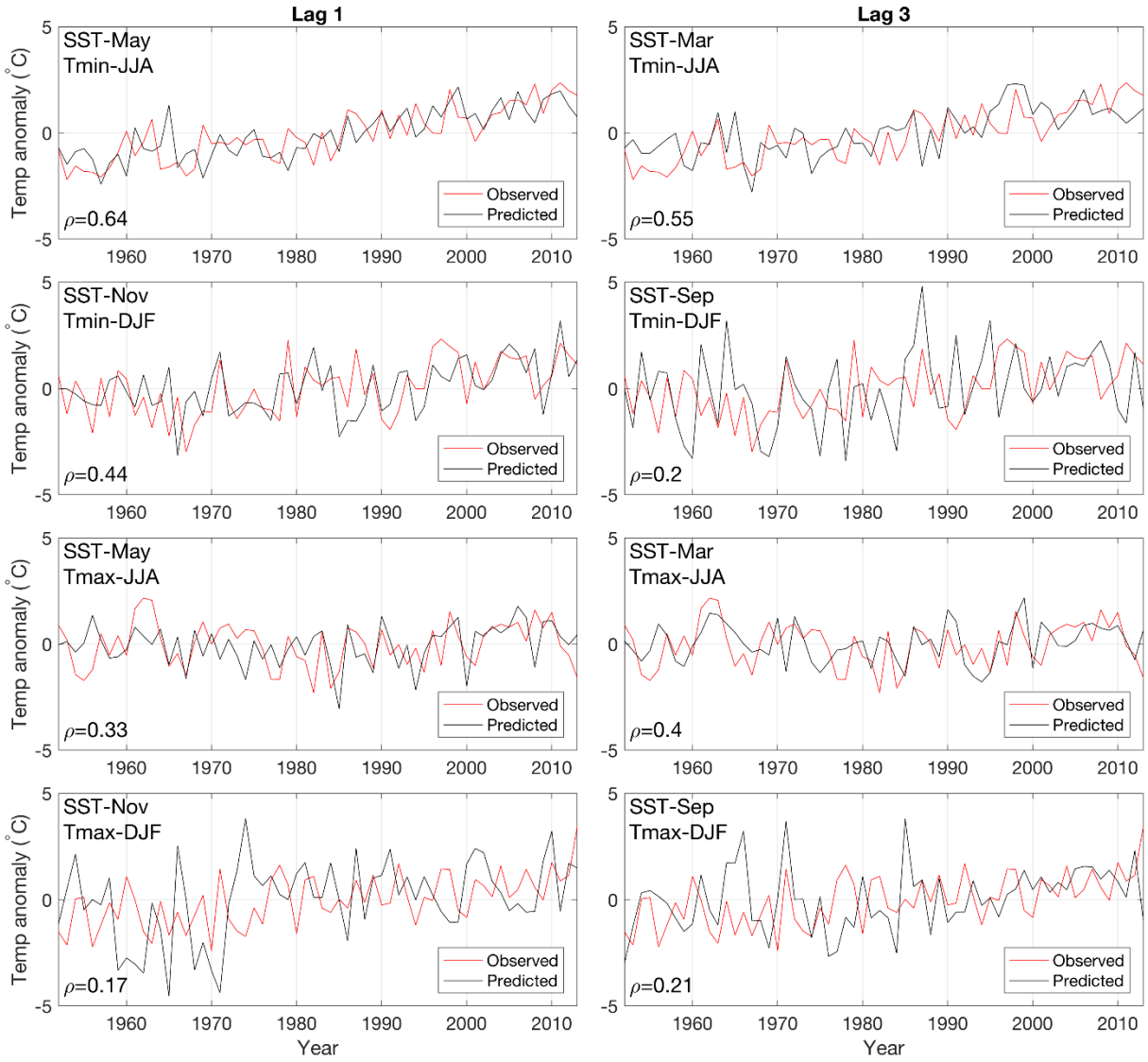

Figure 19: Time series for the predicted (black) and observed (red) air temperature for Sacramento (38.58° N and 121.49° W), for all the experiments using SST as a predictor, and for lags 1 (left panels) and lag 3 (right panels). ρ shows the correlation between the two time-series.

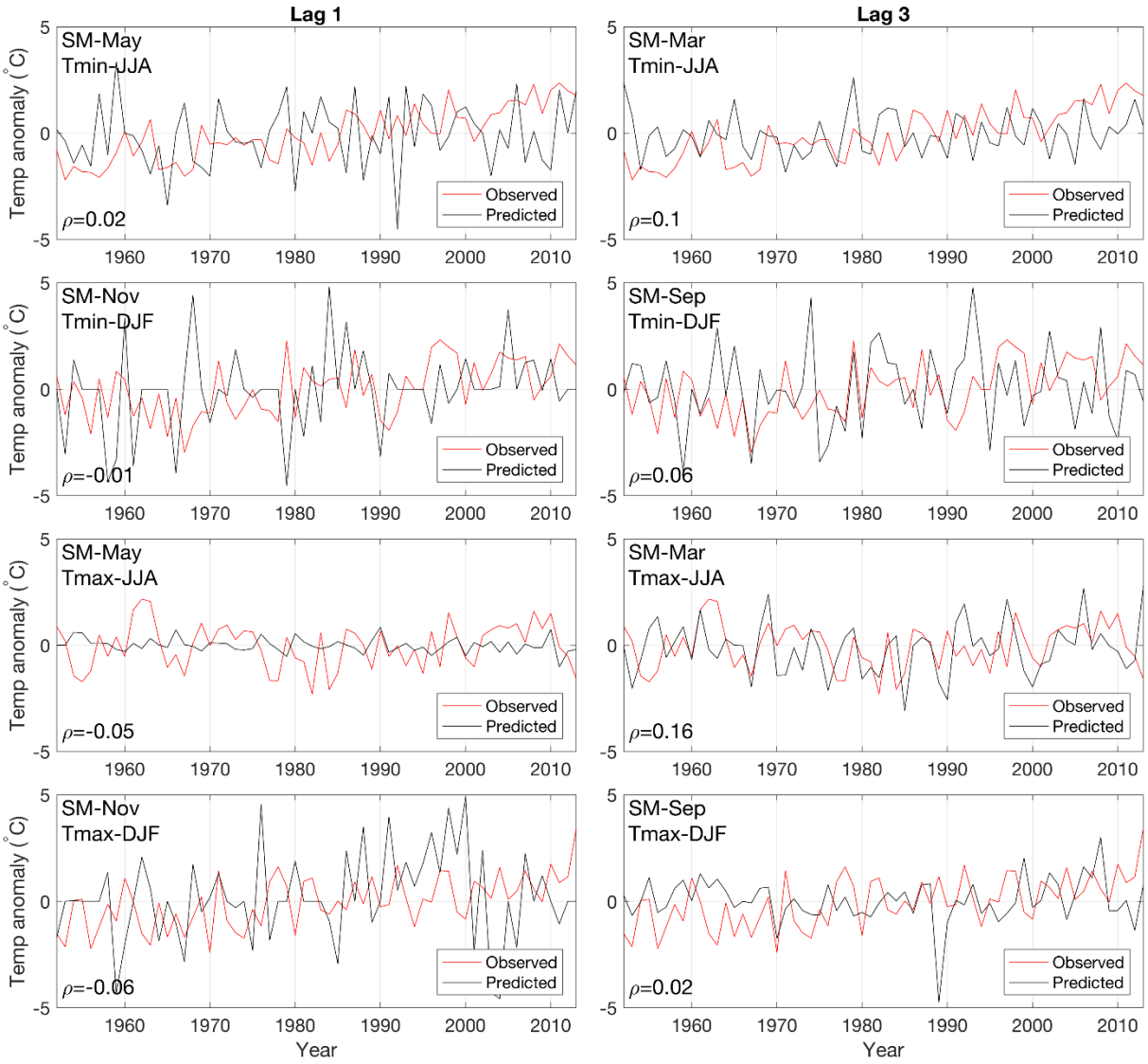

Figure 20: Same as Figure 9, but showing the time-series for the experiments using SM as predictor.

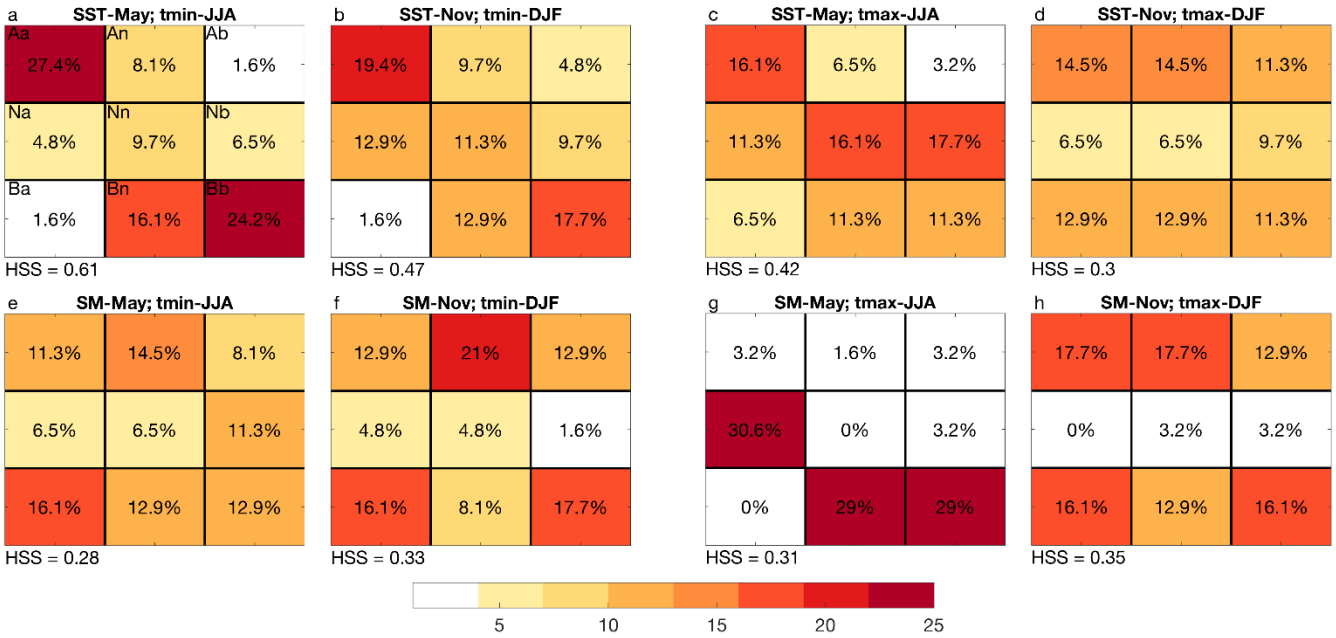

Figure 21: Performance of lag 1 forecasts evaluated for three categories (Above, Normal, Below) for Sacramento (38.58° N and 121.49° W) Tmin (leftmost panels) and Tmax (rightmost panels). Forecasts using SST (upper panels) and SM (lower panels) predictors are shown for May predicting JJA (columns 1 and 3) and November forecasting DJF (columns 2 and 4). Errors are calculated by first splitting the observations and forecast values into three equally occurring classes: above (A), with values larger than the 66th percentile); normal (N), with values between the 33rd and 66th percentile; and below (B), with values below the 33rd percentile. The percentages represent the number of the predicted values that falls into one of those categories: (1) Diagonal (Aa, Nn, Bb), where the predicted values are in the same category as the observed; (2) Class one error (An, Nb, Na, Bn), where the predicted values are one category below or above the observed; (3) Class two errors (Ab and Ba), where the predicted values are two categories below (Ab) or above (Ba) the observed. Capital letters represent the observations and lowercase letters the forecast. The ideal model would have 33% along the diagonal, indicating that the terciles of the predicted values match with the terciles of the observations. The values in the lower left corner of each figure represent the Heidke Skill Score (HSS), in which 1 represents a perfect model, 0 indicates no skill, and a negative value indicates that the forecasts are worse than a randomly generated forecast.

5: References

Alfaro, E., A. Gershunov, D.R. Cayan, A. Steinemann, D.W. Pierce and T.P. Barnett, 2004: A method for prediction of California summer air surface temperature. Eos. 85. 553, 557-558.

Alfaro, E.J., Gershunov, A., Cayan, D., 2006: Prediction of summer maximum and minimum temperature over the central and western United States: the roles of soil moisture and sea surface temperature. Journal of Climate, 19(8), pp.1407-1421.

Baldwin, M.P. et al., 2001. The Quasi-Biennial Oscillation. Reviews of Geophysics, 39, 179-229.

Barnett, T., R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. Monthly Weather Review, 115, 1825–1850.

Barnston, A., and T. Smith, 1996, Specification and prediction of global surface temperature and precipitation from global SST using CCA. Journal of Climate, 9, 2660–2697.

Cherkauer, K. A., and D. P. Lettenmaier, 2003: Simulation of spatial variability in snow and frozen soil. J. Geophys. Res., 108, 8858, doi:10.1029/2003JD003575.

Chiang, J. C. H., and D. J. Vimont, 2004. Analogous meridional modes of atmosphere-ocean variability in the tropical Pacific and tropical Atlantic. J. Climate,17(21), 4143-4158.

Davis, R. E. 1976. Predictability of Sea-Surface Temperature and Sea-Level Pressure Anomalies over North Pacific Ocean. Journal of Physical Oceanography. 6:249-266.

Di Lorenzo, E., Schneider, N., Cobb, K. M., Chhak, K., Franks, P. J. S., Miller, A. J., McWillians, J. C., Bograd, S, J., Arango, H., Curchister, E., Powell, T. M. and Rivere, P. 2008. North Pacific Gyre Oscillation links ocean climate and ecosystem change. Geophysical. Research Letters, 35, L08607, doi:10.1029/2007GL032838.

Enfield, D.B., A. M. Mestas-Nunez and P.J. Trimble, 2001: The Atlantic multidecadal oscillation and it's relation to rainfall and river flows in the continental U.S. Geophysical Research Letters, 28, 2077-2080.

Georgakakos, K. P., D.H., Bae, D. R. Cayan, 1995: Hydroclimatology of continental watersheds 1: Temporal analysis, Water Resources Research, 31(3), 655-678.

Gershunov, A., 1998: ENSO influence on intraseasonal extreme rainfall and temperature frequencies in the contiguous US: Implications for long-range predictability. Journal of Climate, 11, 3192-3203.

Gershunov, A, Barnett TP, Cayan DR, Tubbs T, Goddard L. 2000. Predicting and downscaling ENSOimpacts on intraseasonal precipitation statistics in California: The 1997/98 event. Journal of Hydrometeorology. 1:201-210.

Gershunov, A., D.R. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climatic variability and seasonal predictability. Journal of Climate, 16(16), 2752-2765.

Hamlet, A.F., Mote, P.W., Clark, M.P., Lettenmaier, D.P., 2007: 20th Century Trends in Runoff, Evapotranspiration, and Soil Moisture in the Western U.S., J. Climate, 20 (8): 1468-1486.

Hartmann, D. L., 2015: Pacific sea surface temperature and the winter of 2014, Geophysical Research Letters, 42, 1894–1902, doi:10.1002/2015GL063083.

Horel, J. D., and J. M. Wallace, 1981: Planetary-scale atmospheric phenomena associated with the Southern Oscillation, Monthly Weather Review, 109, 813–829.

Koster, R. D., M. J. Suarez, M. Heiser, 2000: Variance and predictability of precipitation at seasonal-to-interannual timescales. Journal of Hydrometeorology, 1, 26–46.

Huang, J., H. Van den Dool, 1993: Monthly precipitation-temperature relations and temperature prediction over the United States, Journal of Climate, 6, 1111–1132.

Huang, J., H. Van den Dool, K. Georgakakos, 1996: Analysis of model calculated soil moisture over the United States (1931–1993) and applications to long-range temperature forecasts. Journal of Climate, 9, 1350–1362.

Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. Journal of Geophysical Research: Atmospheres, 99(D7), pp.14415-14428.

Livneh, B., T. J. Bohn, D. W. Pierce, F. Munoz-Arriola, B. Nijssen, R. Vose, D. R. Cayan, and L. Brekke, 2015: A spatially comprehensive, hydrometeorological data set for Mexico, the U.S., and Southern Canada 1950-2013. Scientific Data, v. 2, article 150042 (2015). doi:10.1038/sdata.2015.42.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production, Bulletin of American Meteorological Society, 78(6), 1069–1079.

Mueller, B., Seneviratne, S. I., 2012: Hot days induced by precipitation deficits at the global scale. Proc. Natl. Acad. Sci. 109(31), 12398–12403. http://dx.doi.org/ 10.1073/pnas.1204330109.

Namias, J., 1965: Macroscopic association between mean monthly sea-surface temperature and the overlying winds. J. Geophys. Res., 70, 2307-2318.

Neelin, J. D., D. S. Battisti, A. C. Hirst, F.-J. Fei, Y. Wakata, T. Yamagata, and S. E. Zebiak, 1998: ENSO theory, Journal of Geophysical Research, 103(C7), 14, 261–14, 290, doi:10.1029/97JC03424.

Penland, C., and L. Matrosova, 1998: "Prediction of tropical Atlantic sea surface temperatures using Linear Inverse Modeling," J. Climate, March, 483-496 pp.

Pierce, D.W., Cayan, D.R., Thrasher, B.L., 2014: Statistical downscaling using localized constructed analogs (LOCA). Journal of Hydrometeorology, 15(6), pp.2558-2585.

Quesada, B., Vautard, R., Yiou, P., Hirschi, M., Seneviratne, S. I., 2012: Asymmetric European summer heat predictability from wet and dry southern winters and springs. Nature Climate Change, 2(10), 736–741. http://dx.doi.org/10.1038/ nclimate1536.

Rasmusson, E. M., and J. M. Wallace, 1983: Meteorological aspects of the El Niño/Southern Oscillation, Science, 222, 1195–1202.

Rayner, N.A., Parker, D.E., Horton, E.B., Folland, C.K., Alexander, L.V., Rowell, D.P., Kent, E.C. and Kaplan, A., 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. Journal of Geophysical Research: Atmospheres, 108(D14).

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Teuling, A. J., 2010: Investigating soil moisture–climate interactions in a changing climate: a review. Earth-Sciences Review 99(3), 125–161. http://dx.doi.org/10.1016/j.earscirev.2010.02.004.

Schwartz, R.E., A. Gershunov, S.F. Iacobellis and D.R. Cayan, 2014: North American West Coast Summer Low Cloudiness: Broad Scale Variability Associated with Sea Surface Temperature. Geophysical Research Letters. 41, 3307–3314, DOI: 10.1002/2014GL059825.

Shukla, J., 1998: Predictability in the midst of chaos: A scientific basis for climate forecasting. Science, 282, 728–731.

Tang, B., W. W. Hsieh, A. H. Monahan, and F. T. Tangang, 2000: Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures. Journal of Climate, 13, 287–293.

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 tropical sea surface temperatures. J. Geophys. Res., 103, 14 291–14 324.

Van den Dool, H., 2007: Empirical Methods in Short-Term Climate Prediction. Oxford University Press, 215 pp.

Vimont, D.J., Battisti, D.S., Hirst, A.C., 2001: Footprinting: A seasonal connection between the tropics and mid-latitudes. Geophysical research letters, 28(20), pp.3923-3926.

Wilks, D. S., 2008: Improved statistical seasonal forecasts using extended training data. International Journal of Climatology, 28, 1589–1598.

Xoplaki, Elena & González Rouco, J. Fidel & Luterbacher, Jürg & Wanner, Heinz. (2003). Mediterranean summer air temperature variability and its connection to the large-scale atmospheric circulation and SSTs. Climate Dynamics. 20.723-739.10.1007/s00382-003-0304-x.

Zhang, Zhenhai Daniel R. Cayan, David W. Pierce. (Scripps Institution of Oceanography). 2018. ***Seasonal Temperature Forecast Skill over the California Region in NMME.*** California's Fourth Climate Change Assessment, California Energy Commission. Publication number: CCCA4-CEC-2018-010.