Arnold Schwarzenegger
Governor

PREDICTIVE RELATIONSHIP BETWEEN PACIFIC SST AND SUMMER CALIFORNIA AIR SURFACE TEMPERATURE

PIER PROJECT REPORT

Prepared For:
California Energy Commission
Public Interest Energy Research Program

Prepared By:
Scripps Institution of Oceanography
Universidad de Costa Rica
U.S. Geological Survey
University of Washington

February 2005
CEC-500-2005-015

California Climate Change Center
Report Series Number 2005-005

Prepared By:
Eric Alfaro[1,2], Alexander Gershunov[1], Daniel Cayan[1,3], Anne Steinemann[1,4], David Pierce[1], and Tim Barnett[1]
[1] Climate Research Division, Scripps Institution of Oceanography, University of California, San Diego
[2] Escuela de Fisica-Centro de Investigaciones Geofisicas, Universidad de Costa Rica
[3] U.S. Geological Survey
[4] University of Washington

Contract No. 500-02-004
Work Authorization No. MR-004

Prepared For:
California Energy Commission
Public Interest Energy Research (PIER) Program

Guido Franco,
Contract Manager

Kelly Birkinshaw,
Program Area Team Lead
Energy-Related Environmental Research

Ron Kukulka,
Acting Deputy Director
ENERGY RESEARCH AND DEVELOPMENT DIVISION

Robert L. Therkelsen
Executive Director

Acknowledgements

This work was done during the first author's visit at the Climate Research Division, Scripps Institution of Oceanography, University of California, San Diego, supported by the California Climate Change Center, sponsored by the California Energy Commission's Public Interest Energy Research Program. Additional support was provided by the NOAA/NA17RJ1231 grant funded by the National Oceanic and Atmospheric Administration's Office of Oceanic and Atmospheric Research, under the California Energy Security Project**.**

Please cite this report as follows:

Alfaro, E., A. Gershunov, D. Cayan, A. Steinemann, D. Pierce, and T. Barnett. 2005. *Predictive Relationship Between Pacific SST and Summer California Air Surface Temperature.* California Energy Commission, PIER Energy-Related Environmental Research. CEC-500-2005-015.

Preface

The Public Interest Energy Research (PIER) Program supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The PIER Program, managed by the California Energy Commission (Energy Commission), annually awards up to $62 million to conduct the most promising public interest energy research by partnering with Research, Development, and Demonstration (RD&D) organizations, including individuals, businesses, utilities, and public or private research institutions.

PIER funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency
- Energy Innovations Small Grant Program
- Energy-Related Environmental Research
- Energy Systems Integration Environmentally Preferred Advanced Generation
- Industrial/Agricultural/Water End-Use Energy Efficiency
- Renewable Energy Technologies

The California Climate Change Center (CCCC) is sponsored by the PIER program and coordinated by its Energy-Related Environmental Research area. The Center is managed by the California Energy Commission, Scripps Institution of Oceanography at the University of California at San Diego, and the University of California at Berkeley. The Scripps Institution of Oceanography conducts and administers research on climate change detection, analysis, and modeling; and the University of California at Berkeley conducts and administers research on economic analyses and policy issues. The Center also supports the Global Climate Change Grant Program, which offers competitive solicitations for climate research.

The California Climate Change Center Report Series details ongoing Center-sponsored research. As interim project results, these reports receive minimal editing, and the information contained in these reports may change; authors should be contacted for the most recent project results. By providing ready access to this timely research, the Center seeks to inform the public and expand dissemination of climate change information; thereby leveraging collaborative efforts and increasing the benefits of this research to California's citizens, environment, and economy.

The work described in this report was conducted under the Preliminary Climatic Data Collection, Analyses, and Modeling contract, contract number 500-02-004, Work Authorization MR-004, by the Climate Research Division, Scripps Institution of Oceanography, University of California, San Diego.

For more information on the PIER Program, please visit the Energy Commission's Web site www.energy.ca.gov/pier/ or contract the Energy Commission at (916) 654-4628.

Abstract

Climate prediction for summertime tropospheric variables has been recognized as one of the most challenging issues in middle latitudes. However, advances in summertime temperature forecasts are important for planning activities of different economic sectors, such as the energy industry. In this study, seasonal predictability of summer temperature in California was explored using Canonical Correlation Analysis (CCA). The diagnostic and prediction scheme used Pacific Ocean Sea Surface Temperatures (PSST) as predictors and surface air temperature as predictands over 1950–2001. About 220 stations reporting daily data over California and Nevada were used. The analysis focused on three aspects: average temperature, extreme warm events, and cooling degree days. Results show a strong variability in air surface temperatures over California when a SST contrast between the North West-Central and East Pacific is found during the spring, suggesting the influence of the Pacific Decadal Oscillation (PDO). Maximum predictive skill for the June, July, and August (JJA) season was given by March, April, and May (MAM), but significant skill was found for PSST predictors leading summer air temperature by several seasons. Contingency analysis showed that under Below (Above) Normal spring PDO conditions, the most probable scenario for summer air surface temperature is Below (Above) Normal in California. Other important SST modes identified in the CCA analysis are the El Niño/Southern Oscillation (ENSO) variability and long term trends, but their influence on summer air surface temperatures was smaller than for the PDO. The strength and simplicity of the relationship between spring PDO and summer temperature suggests a simple yet useful approach for predicting summer temperatures in California.

Table of Contents

1. Introduction

In recent years significant progress has been made in seasonal climate prediction, yet summertime midlatitude climate prediction remains problematic (e.g., Gershunov and Cayan 2003). A series of previous studies have explored the skill of Pacific Sea Surface Temperatures (PSST) in seasonal prediction of various atmospheric variables (e.g., Barnett and Preisendorfer, 1987), but few have focused on its value in forecasting summer conditions (e.g., Douville, 2003). Advances in summertime temperature forecasts are important for planning in different economic sectors, such as the energy industry. This issue is especially important in California, where the summer peak energy demand is about 50% higher than that in winter, in response to heavy loads from air conditioners, pumping water, and other seasonal issues. The California Energy Security Project was designed to address this issue by developing summertime climate forecasts in a format, and containing information, that is tailored to the needs of decision makers in the energy sector (both public and private).

A statistical model based on Canonical Correlation Analysis (CCA) was used for seasonal diagnosing and predicting summer temperatures in California. This statistical technique has proven skillful when used recently by Gershunov and Cayan (2003) for the prediction of winter precipitation statistics over the contiguous United States and by Westerling et al. (2002) in predicting wild fires in the West. The CCA is a statistical technique developed to identify and quantify associations between two sets of variables, matching mainly large-scale patterns in the predictor with patterns in the predictand fields; in this case coupled patterns in spring PSST and California summer temperatures were identified. As will be described, results from the CCA analysis provide the basis for a simpler statistical procedure using the state of the dominant mode of the North Pacific SST variability in spring to forecast summer temperature variability over the California region.

2. Predicting California Temperatures

2.1 Data

We used 220 stations with daily temperature data covering 52 years (1950–2001) over California and Nevada, a subset of the National Climatic Data Center (NCDC) first order and cooperative observer data (Groisman et al., 2004; NCDC, 2003). Seasonal SST anomaly data from the Pacific basin over the 1950–2001 period were used as predictors (Kaplan et al., 1998). The analysis focused on three variables important for energy demand during June, July, and August (JJA): (1) the frequency of daily maximum temperature extremes (daily maximum temperatures above the 90th percentile of their summertime climatology (Tmax-T90), see next section for details), (2) seasonal average temperature (Tmean), and (3) Cooling Degree Days or CDD (Pierce 2004).

2.2 Extreme events, relevant climate aspects, and background

As a first approach, the same threshold for all stations in a particular season was used as a definition of extreme event. Fig. 2.2.1 shows the spatial distribution of the probability of days with maximum temperature equal to or greater than 95°F (35°C) during the summer. From Fig. 2.2.1, we see that this definition is inconvenient, because some stations present maximum temperature (Tmax) values greater than the threshold almost during the whole summer and some of them hardly ever. This problem could also be identified for different days in a row (see for example Fig. 2.2.2).

Taking this problem into account, the extreme warm events were defined as those values greater or equal to the 90th percentile during the summer (see for example Fig. 2.2.3). Then, annual time series, expressed as the percentage of days with extreme values for the season, were calculated.

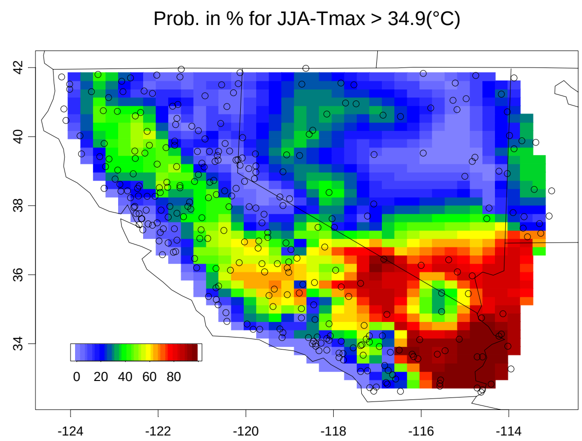

Figure 2.2.1. Probability of having a day with maximum temperature equal to or greater than 95°F (35°C) during summer.

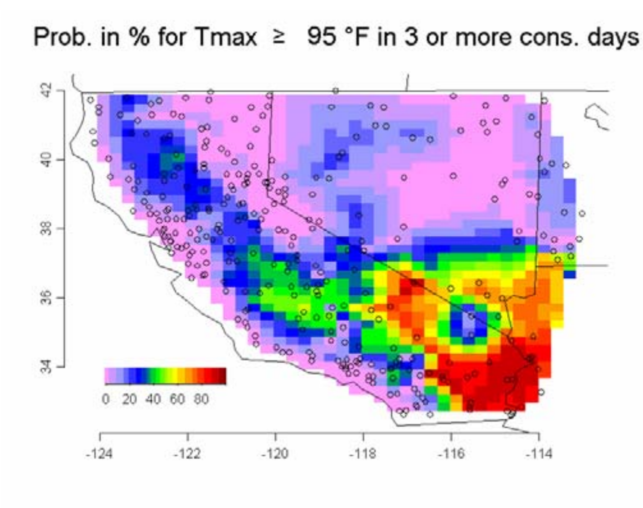

Figure 2.2.2. Probability of having three days in a row with maximum temperature equal to or greater than 95°(35°C) during summer for California and Nevada.

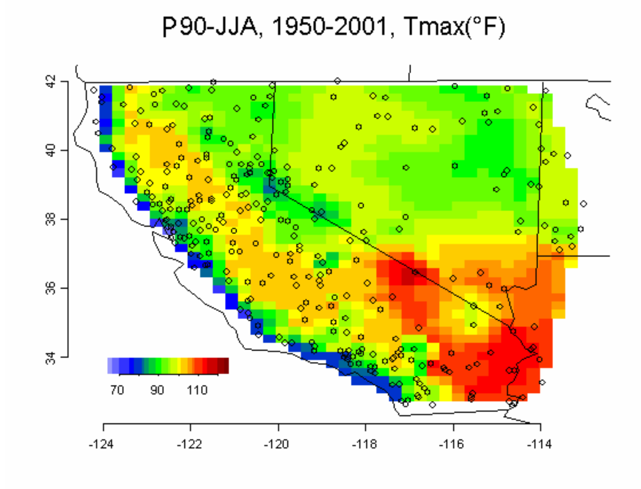

Figure 2.2.3. The spatial distribution of the 90th percentile for Tmax during the summer.

For the summer, the lower 90^{th} percentile values tend to occur in the coastal and mountain regions and most of upper values tends to occur inland, around the border of California, Nevada, and Arizona, in which the North American Monsoon influence (Higgins et al., 2003) is appreciable.

Because the analysis does not have the annual cycle removed, the monthly distribution of extreme events is illustrated for summer in Figs. 2.2.4 (a), (b), and (c). These analyses show that in general July is the month with most of the warm extreme events. This could be explained by the annual air surface temperature cycle (Fig. 2.2.5) that shows a maximum in July and a minimum during January. Notice also that some locations, like those around the San Francisco Bay area, presents a distribution of extreme events during the summer with almost thirty percent for every month and the region around the coast in Southern California presents more extreme events in August. Going a step forward, the daily frequency distribution of extreme events was calculated in Fig. 2.2.6, shows a relatively smooth distribution of extreme events during the month of the maximum temperatures (July), but with a maximum on August 7.

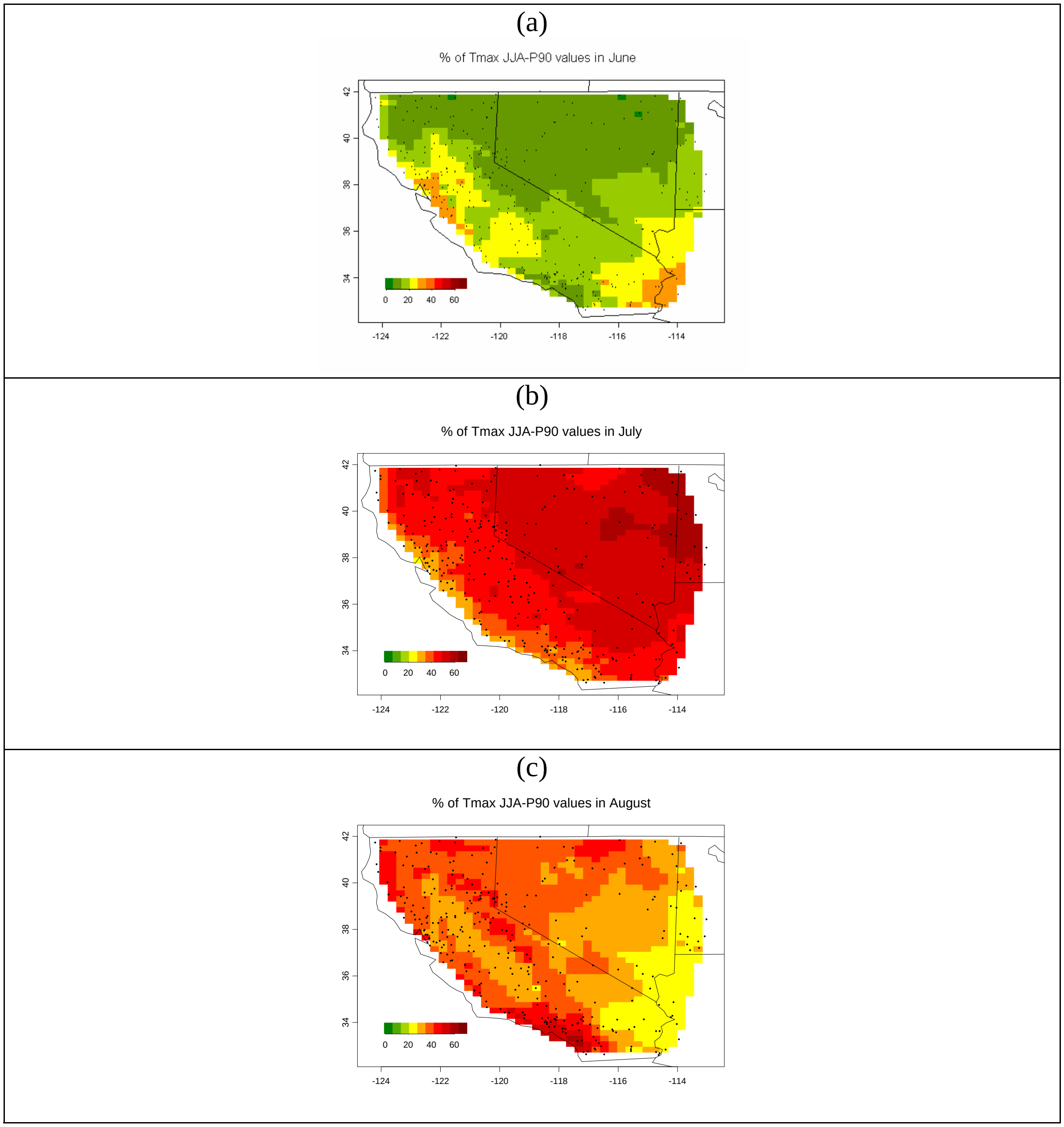

Figure 2.2.4. Monthly distribution of extreme values for Tmax during the summer for California and Nevada, (a) June, (b) July, and (c) August.

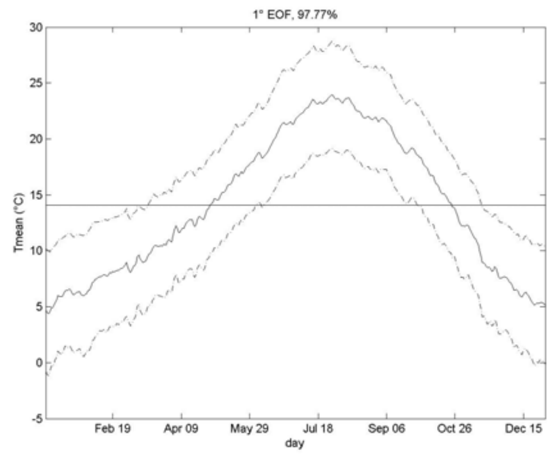

Figure 2.2.5. Annual cycle of the first principal component for mean temperature stations in California and Nevada (solid line). This component explains around 98% of the explained variance. Dashed lines represent one standard deviation.

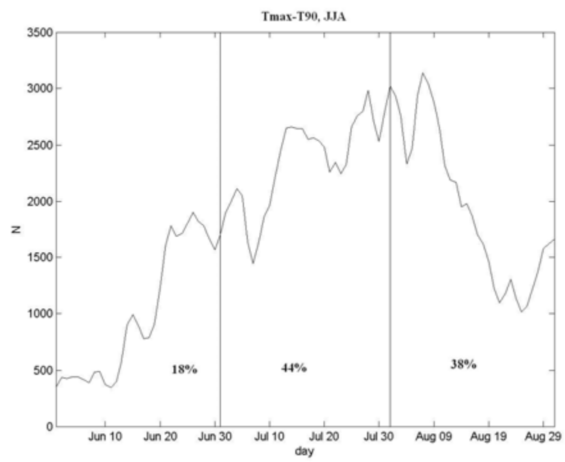

Figure 2.2.6. Seasonal daily frequency distribution of summer warm extreme events (Tmax-T90, 1950–2001).

3. CCA Statistical Model

To begin with, we used CCA as a diagnostic to identify coupled patterns between California summertime temperature and simultaneous or antecedent PSST. Fig. 3.1 shows that using March, April, and May (MAM) SST and Tmean fields, the leading coupled mode (r = 0.91, Fig. 3.1a) captures an east-west contrast in the PSST correlations in the North Pacific resembling the Pacific Decadal Oscillation or PDO[1] (Mantua et al. 1997; Gershunov and Barnett 1998; Pierce 2002, Fig. 3.1b). Coupled to this PSST structure is a swath of anomalous temperature along the coast of California (Fig. 3.1c). The sense of the PSST-air temperature linkage is that positive (negative) values of the MAM-PDO index tend to be associated with warm (cool) JJA coastal temperature. This suggests that summer temperature may be seasonally predictable along the California coast, where much of the State's population is concentrated.

[1] Depending on the author and definition, it could be associated with the term: North Pacific Oscillation, or NPO.

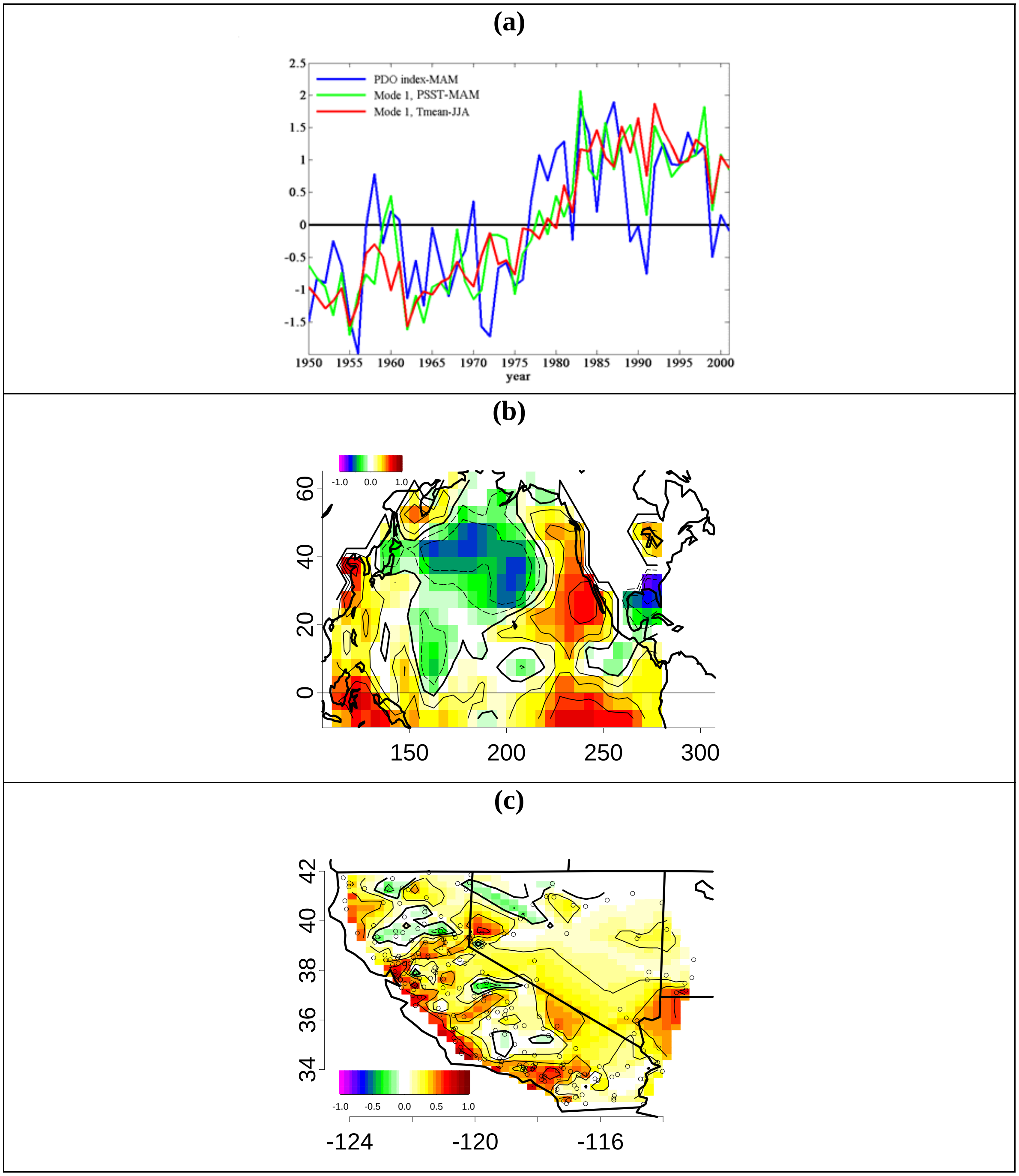

Figure 3.1. (a) CCA mode 1 of JJA-Tmean (red line) related with MAM-PSST (green line). MAM-PDO index is also plotted for comparison (blue line). Spatial patterns of predictor fields (MAM-PSST) and JJA-Tmean are displayed in (b) and (c) as correlations of the temporal evolutions with their respective variable fields at each location (i.e., grid cell for predictors and station for the predictand). These correlation patterns are displayed in color for PSST and Tmean fields on the common color scale (-1: violet, 1: dark red). Correlations are displayed as contours at 0.2 intervals. Negative contours are dashed; positive are solid. Small circles in (c) represent station locations. Correlation values greater than 0.32 are significant at the 0.01 level.

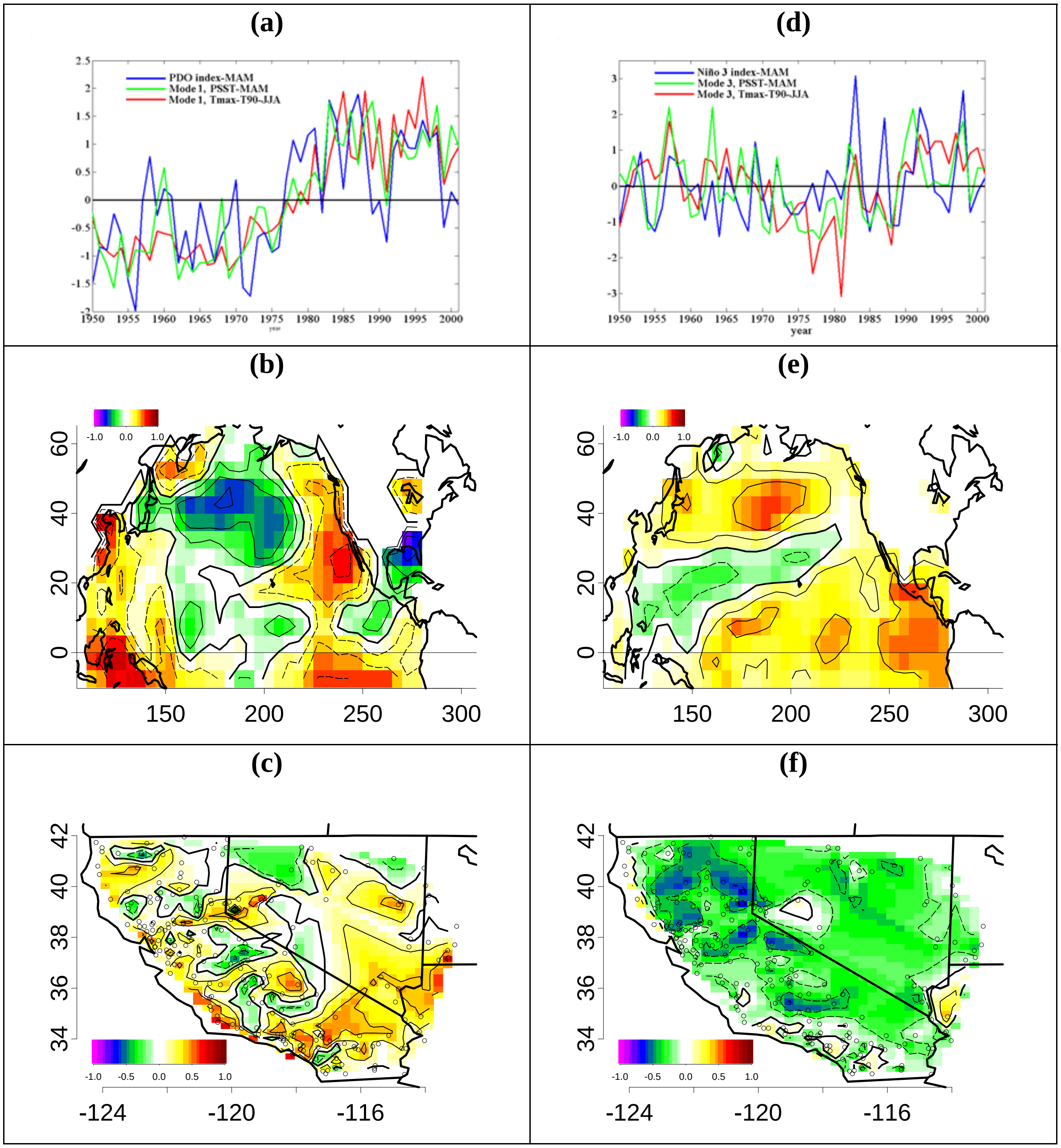

Figure 3.2. CCA modes (a) 1 and (d) 3 of JJA-Tmax-T90 (red lines) related with MAM-PSST (green lines). The MAM PDO and MAM-Niño 3 indices are also plotted for comparison in (a) and (b) respectively (blue lines). Spatial patterns of predictor fields (MAM-PSST) and JJA-Tmean are displayed in (b), (c), (e) and (f) as correlations of the temporal evolutions with their respective variable fields at each location (i.e., grid cell for predictors and station for the predictand). These correlation patterns are displayed in color for PSST and Tmax-T90 fields on the common color scale (-1: violet, 1: dark red). Correlations are displayed as contours at 0.2 intervals. Negative contours are dashed, positive are solid. Small circles in (c) and (f) represent station locations. Correlation values greater than 0.32 are significant at the 0.01 level.

Comparing the leading modes of summer Tmean and Tmax-T90, we found that the first leading coupled mode (r = 0.88, Fig. 3.2a) also resembles the PDO structure (Fig. 3.2b) but with a more scattered response in the stations correlation pattern over California (Fig. 3c). Additionally the third leading coupled mode (r = 0.55, Fig. 3.2d) is associated to interannual variability in the tropical eastern Pacific (Fig. 3.2e) with a broad but weak negative correlation pattern over almost all the State (Fig. 3.2f).

We then used this coupled mode along with several others (not shown) in a CCA model estimated at different lead times from January-February-March (JFM) to MAM for prediction and from April-May-June (AMJ) to JJA for specification. Fig. 3.3a. shows the optimum skill values for JJA-Tmean, CDD and Tmax-T90 using different leading PSST seasons. These measures were obtained from the "optimum" CCA model, which was determined from optimizing the number of predictand and predictor modes to maximize the average skill, using cross-validation techniques to avoid overfitting (see Gershunov and Cayan (2003) for optimization details). Cross-Validation is a resampling technique that operates in a way similar to the bootstrap and permutation tests, dividing the total group of data repeatedly into a subgroup for control and another for verification. The most common point, like in this case, is that the size of the first group is *n-1* and that of the second is *1*, with *n* different data partitions. As expected, the skill values tend to decrease as the predictor lead time increases. The SST season that produces the maximum non-overlapping predictive skill for JJA Temperature season is MAM. The skill maps (Figs. 3.3b, c, and d) illustrate the correlation between the observed value and the predicted value in a cross-validated framework. The skill of Tmean and CDD prediction (Figs. 3.3b and c) shows values greater than 0.4 (more than 20% of the variability predicted by model) in a large part of the region, while the prediction of warm extreme values gives acceptable values only for some locations in the California central valley (Fig. 3.3d). A set of statistical models were constructed to investigate the predictability of PSST in forecasting three summer variables: (1) cooling degree days (CDD), (2) average temperature, (3) and the number of maximum temperature extremes. For each of these variables there appears to be significant model skill. Model skills are maximum for JJA PSST predicting JJA temperature measure and decline successively with increasing PSST time lead. The model for Tmean has greater skill than those for CDD and CDD models have greater skill than those for Tmax-T90 measure. Additionally, Fig. 3.4 presents the prediction skill maps for the summer Tmax-T90 using MJJ, AMJ, and February-March-April (FMA) PSST as predictors. Although, the first two maps of this figure implied some simultaneous months (Fig. 3.4a, b) with summer, these maps are useful to update the forecast or in the use of a perfect prognosis scheme. Notice also that in spite that the skill map using FMA predictors (Fig. 3.4c) has a lower correlation average value than the one using MAM predictors (Figs. 3.3a and d), it still maintains an important area with values having statistical significance greater than 99%.

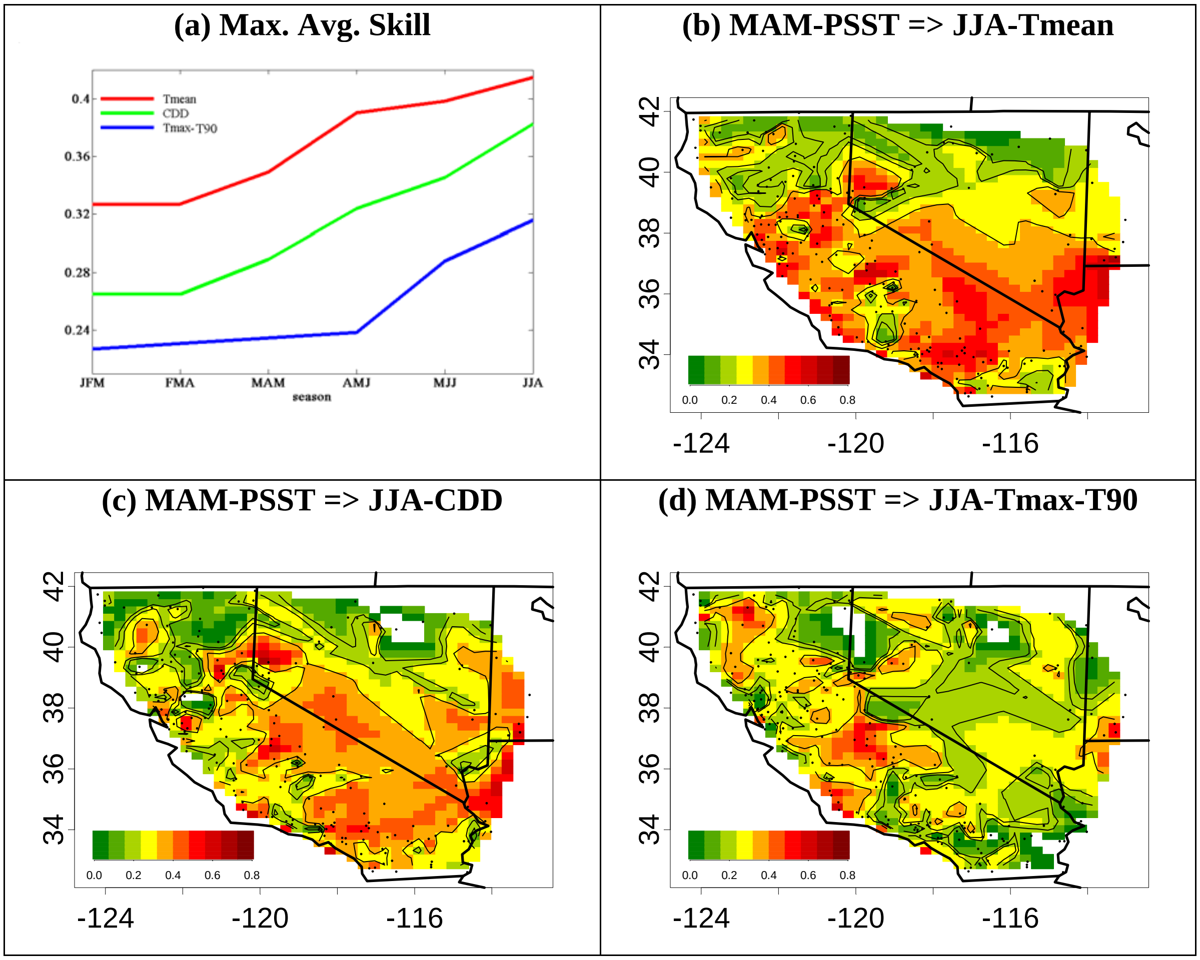

Figure 3.3. (a) Optimized maximum average skill for Tmean (red line), CDD (green line), Tmax-T90 (blue line), and Tmax-T95 (black line). (b), (c), and (d) are the respectively JJA Tmean, CDD and Tmax-T90 optimized skill maps, expressed as correlations between the cross-validated forecast and observations at stations using MAM-PSST as predictors. All values are displayed on the same range. Uncolored areas are regions of insignificant negative correlations. The three contours represent the 90th, 95th, and 99th percent levels of significance in order of increasing correlations.

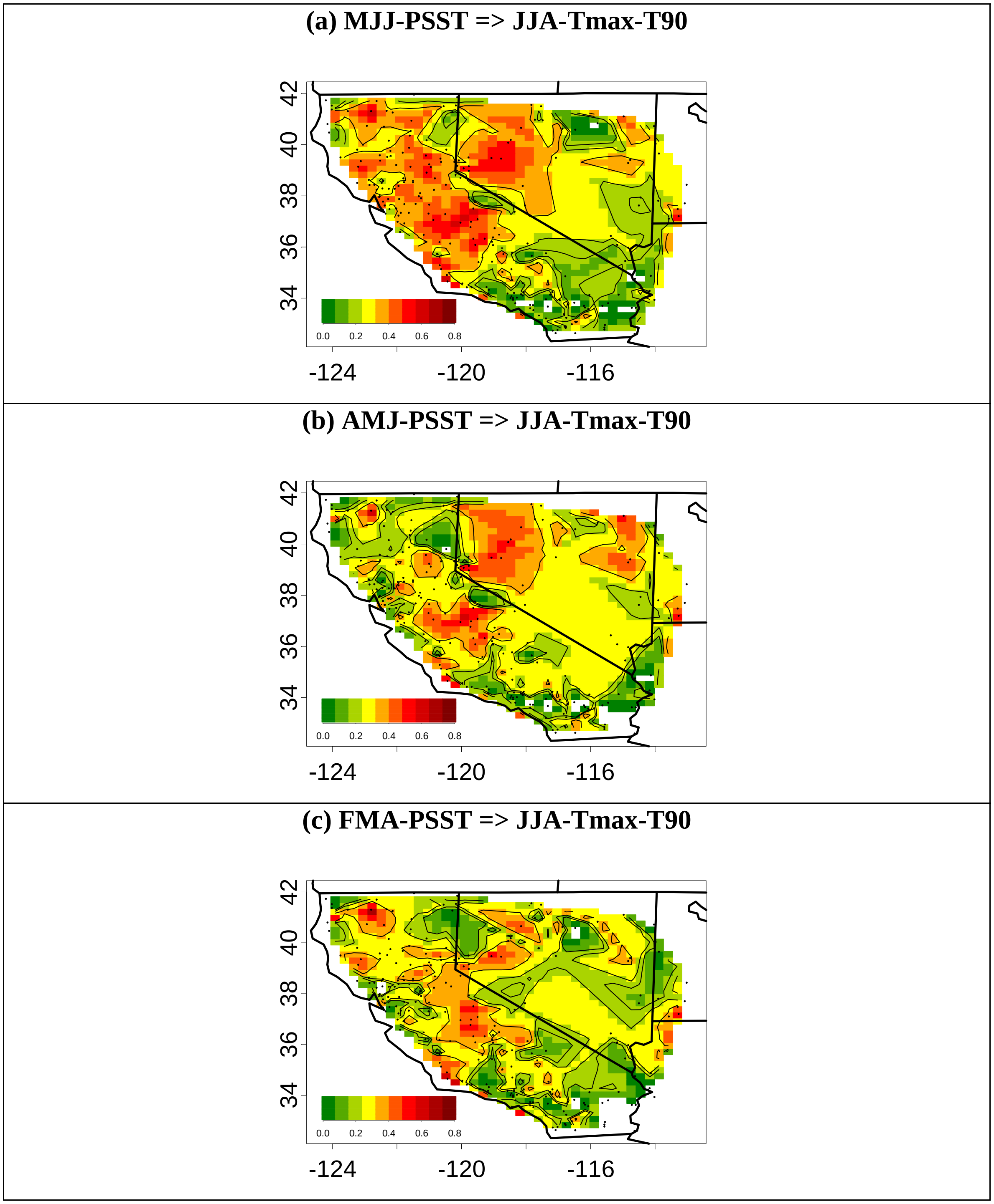

Figure 3.4. JJA-Tmax-T90 optimized skill maps expressed as correlations between the cross-validated forecast and observations at stations using (a) MJJ-PSST, (b) AMJ-PSST, and (c) FMA-PSST as predictors. All values are displayed on the same range. Uncolored areas are regions of insignificant negative correlations. The three contours represent the 90th, 95th, and 99th percent levels of significance, in order of increasing correlations.

4. A Simplified Statistical Model Using Spring PDO

The CCA methodology as used here amounts to a rather sophisticated statistical model based on many coupled patterns in Pacific SST and California temperature. However, Fig. 3.1 identifies a well-known climate pattern, the PDO, which is a dominant mode of anomalous SST variability in the extratropical North Pacific (Mantua et al. 1997), preceding coastal California summertime temperature anomalies. The PDO index is available at: http://jisao.washington.edu/pdo/PDO.latest. It is possible, therefore, to produce a simple rule-of-thumb California coastal summertime temperature-forecasting scheme based on the springtime state of the PDO. The strength of the spring PDO—summer temperature relationship suggests that it might be possible to exploit a rather simple contingency analysis. Such an analysis could be formatted to transform these predictions in a product tailored explicitly for energy producers. For this purpose we used the PDO index and related it to climate at stations that California Energy Commission has defined as representative of the different California Climate zones and important summertime energy demand sectors (Pierce 2004). These stations are: Eureka, Ukiah, Sacramento, Fresno, San Francisco, Long Beach, Los Angeles, San Diego, Burbank-Glendale-Pasadena, Blythe, San Bernardino, and San Jose.

Fig. 4.1 shows that under Below (Above) Normal spring PDO conditions, the most probable scenario for summer air surface temperature is also Below (Above) Normal in much California and mainly in those stations along the coast. Notice also that PDO observed conditions during previous winter and spring tend to persists trough summer. Not only is the occurrence of Below (Above) normal spring PDO biased in favor of the occurrence of Below (Above) normal summer Tmean or CDD at most of the stations, it also disfavors the occurrence of the opposite class anomalies of summer Tmean or CDD. For example, the expected value of 33% Below (Above) normal CDD is elevated to greater than 50% at several of the stations that were selected. The occurrence of the opposite class Above (Below) anomalies is likewise reduced to less than 20% and in some cases less than 7% of the sample considered. These statistical relationships are generally stronger along the coastal and in the southern and central portion of the station network.

5. Implications for Utility Industry

From a decision-making perspective, improved forecasts of summer California temperatures could provide several benefits to the electricity industry, including increased reliability, decreased risk of outages, and more efficient use of resources. Electrical energy use in California is highly dependent on temperature; for instance, on hot afternoons, air conditioning can account for over 40% of the total electricity use in California. California electrical providers are concerned not only with warm temperature anomalies, but also the spatial distribution of those temperature anomalies. Partly, this is because the interconnected electricity system can handle localized increased temperatures and corresponding demands, but becomes stressed when high demands spread throughout the state at the same time. The forecast capabilities described herein would help decision-makers plan for seasonal electricity demands and to maintain adequate supplies under different temperature scenarios.

Figure 4.1. Contingency analysis using empirical conditional probabilities between the observed MAM-PDO and (a), (b) JJA-Tmean, (c), (d) JJA-CDD. (a), (c) is for a MAM-PDO scenario Below Normal and (b), (d) is for a MAM-PDO scenario Above Normal. For this purpose, researchers used the stations that the California Energy Commission has already defined as representative of the different California Climate zones (see text). The conditional probabilities between persistence previous winter-spring seasons and summer PDO is also included in (a) and (b), in the right upper corner. Percentage values greater (lower) than 45, 51 and 57 (24, 18, and 13) are significant at the 0.10, 0.05 and 0.01 levels respectively for high (low) count bins in the contingency analysis.

6. References

- Barnett, T., and R. Preisendorfer, 1987: Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures determined by canonical correlation analysis. *Mon. Wea. Rev*., **115**, 1825-1850.
- Douville, H., 2003: Assessing the influence of soil moisture on seasonal climate variability with AGCMs. *J. Hydrometeor*., **4**, 1044-1066.
- Gershunov, A., and T. Barnett, 1998: Interdecadal modulation of ENSO teleconnections. *Bull. Amer. Meteor. Soc.*, **79**, 2715-2725.
- Gershunov, A., and D. Cayan, 2003: Heavy daily precipitation frequency over the contiguous United States: Sources of climate variability and seasonal predictability. *J. Climate*, **16**, 2752-2765.
- Groisman, P., R. Knight, T. Karl, D. Easterling, B. Sun and J. Lawrimore, 2004: Contemporary Changes of the Hydrological Cycle over the Contiguous United States: Trends Derived from *in situ* Observations. *Journal of Hydrometeorology*, **5**(1), 64–85.
- Higgins, R., A. Douglas, A. Hahmann, E. Berbery, D. Gutzler, J. Shuttleworth, D. Stensrud, J. Amador, R. Carbone, M. Cortez, M. Douglas, R. Lobato, J. Meitin, C. Ropelewski, J. Schemm, S. Schubert and C. Zhang, 2003. Progress in Pan American CLIVAR Research: The North American Monsoon System. *Atmosfera*, **16**(1), 29-65.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analysis of global sea surface temperatures 1856–1991. *J. Geophys. Res.*, **103**, 18567–18589.
- Mantua, N., S. Hare, Y. Zhang, J. Wallace and R. Francis, 1997: A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Amer. Meteor. Soc.*, **78**, 1069-1079.
- NCDC, 2003: Data documentation for data set 3200 (DSI-3200): Surface land daily cooperative summary of the day. National Climatic Data Center, Asheville, NC, 36 pp.
- Pierce, D., 2004: Effects of the North Pacific Oscillation and ENSO on seasonally averaged temperatures in California. *Draft Manuscript*, CRD-SIO-UCSD, La Jolla, California.
- Pierce, D., 2002: The role of sea surface temperatures in interactions between ENSO and the North Pacific Oscillation. *J. Climate*, **15**, 1295-1307.
- Westerling, A., A. Gershunov, D. Cayan and T. Barnett, 2002: Long lead statistical forecast of area burned in western U. S. wildfires by ecosystem province. *Int. J. Wild. Fire*, **11**, 257-266.