# Detection and Attribution of Temperature Changes in the Mountainous Western United States

CÉLINE BONFILS,* BENJAMIN D. SANTER,* DAVID W. PIERCE,[+] HUGO G. HIDALGO,[+] GOVINDASAMY BALA,* TAPASH DAS,[+] TIM P. BARNETT,[+] DANIEL R. CAYAN,[+,#] CHARLES DOUTRIAUX,* ANDREW W. WOOD,[@] ART MIRIN,* AND TORU NOZAWA[&]

*Lawrence Livermore National Laboratory, Livermore, California
[+]Scripps Institution of Oceanography, La Jolla, California
[#]U.S. Geological Survey, La Jolla, California
[@]University of Washington, Seattle, Washington
[&]National Institute for Environmental Studies, Tsukuba, Japan

(Manuscript received 3 January 2008, in final form 22 April 2008)

## ABSTRACT

Large changes in the hydrology of the western United States have been observed since the mid-twentieth century. These include a reduction in the amount of precipitation arriving as snow, a decline in snowpack at low and midelevations, and a shift toward earlier arrival of both snowmelt and the centroid (center of mass) of streamflows. To project future water supply reliability, it is crucial to obtain a better understanding of the underlying cause or causes for these changes. A regional warming is often posited as the cause of these changes without formal testing of different competitive explanations for the warming. In this study, a rigorous detection and attribution analysis is performed to determine the causes of the late winter/early spring changes in hydrologically relevant temperature variables over mountain ranges of the western United States. Natural internal climate variability, as estimated from two long control climate model simulations, is insufficient to explain the rapid increase in daily minimum and maximum temperatures, the sharp decline in frost days, and the rise in degree-days above 0°C (a simple proxy for temperature-driven snowmelt). These observed changes are also inconsistent with the model-predicted responses to variability in solar irradiance and volcanic activity. The observations are consistent with climate simulations that include the combined effects of anthropogenic greenhouse gases and aerosols. It is found that, for each temperature variable considered, an anthropogenic signal is identifiable in observational fields. The results are robust to uncertainties in model-estimated fingerprints and natural variability noise, to the choice of statistical downscaling method, and to various processing options in the detection and attribution method.

## 1. Introduction

Winter and spring temperatures over the western United States have warmed significantly during the past 50 yr, as witnessed by earlier flower blooms (Cayan et al. 2001). This regional warming has been associated with a change of atmospheric circulation over the North Pacific (Dettinger and Cayan 1995), the origins of which are still under investigation (e.g.: Shindell et al. 2001; Gillett et al. 2005; Bonfils et al. 2008). This warm-

ing has been linked to a rise in the number of wildfires in western U.S. forests (Westerling et al. 2006) and an increase in the forested area burned over Canada (Gillett and Weaver 2004). Continued warming is likely to impact crop growth and development (Lobell et al. 2006), to exacerbate air pollution and heat waves (Hayhoe et al. 2004), and to affect the availability of water resources.

The water resources of the western United States depend on snowpack (e.g., Andreadis and Lettenmaier 2006), which stores precipitation during cold months and supplies meltwater to river basins during warm months. Water managers must balance the competing goals of meeting water demands while minimizing flood risk. How climate change will affect this delicate balance is a critical issue for the western United States

*Corresponding author address:* Céline Bonfils, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, P.O. Box 808, Mail Stop L-103, Livermore, CA 94550.
E-mail: bonfils2@llnl.gov

(Maurer et al. 2007). In the past, most decisions on water supply planning relied on the assumption of a stationary climate. This assumption is now being challenged (Gleick et al. 2000; Milly et al. 2008) by the emerging evidence of human-induced changes in the earth's hydrological cycle (Lambert et al. 2005; Gedney et al. 2006; Zhang et al. 2007; Santer et al. 2007; Willett et al. 2007) and by various temperature-driven regional hydrological changes. The snow/rain partitioning of precipitation, for instance, has changed with more precipitation falling as rain instead of snow, particularly in regions where mean winter minimum temperature rises above $-5°C$ and a warming trend bring these temperatures near freezing (Knowles et al. 2006). A pervasive decrease in snow water content and earlier snow melting occur in lower and midelevation mountain areas, particularly at the proximity of the snow line (where winter temperature is close to $0°C$; Mote et al. 2005). Associated with an increase in January–March temperatures, the March fraction of total annual streamflow rises, while the April–July flows drop, shifting streamflow peaks to an earlier date in the year (Dettinger and Cayan 1995; Stewart et al. 2005).

All these changes may have significant socioeconomic impacts on the population of the western United States, and underscore the urgent need to ensure that the best available scientific information on climate change is well integrated into long-term water management choices. Informed decision making on water management choices therefore requires a better understanding of the primary causes of the above-described observed changes. While all the cited studies suggest that these changes may be related to large-scale human-induced warming, this has not yet been demonstrated in a formal detection and attribution (D&A) study.

The current investigation is one of a series of studies focusing on the detection and attribution of changes in the hydrology of the western United States (Barnett et al. 2008; Pierce et al. 2008; Hidalgo et al. 2008, manuscript submitted to *J. Climate*). We perform a rigorous model-based detection and attribution analysis to determine the causes of the recent late winter/early spring changes in hydrologically relevant temperature variables. The D&A of changes in snowpack and the timing of streamflow are investigated in Pierce et al. (2008) and Hidalgo et al. (2008, manuscript submitted to *J. Climate*), respectively. The multivariate D&A analysis that combines the changes in temperature, snowpack, and timing of the peak streamflow into a single detection variable is discussed in Barnett et al. (2008).

Our study addresses the following questions: What are the salient characteristics of the temperature increase recently observed over the mountainous regions of the western United States? Why are temperatures changing? Are the temperature increases primarily naturally driven or human induced? Since variations in snow content can arise from temperature changes, precipitation changes, or a combination of the two (e.g., Groisman et al. 1994; Mote 2006), can we determine whether earlier snowmelt is primarily caused by temperature or precipitation changes?

There are a number of reasons why these questions are difficult to answer. First, the climate of the western United States is strongly influenced by such natural climate variations as the El Niño–Southern Oscillation (ENSO) and the Pacific decadal oscillation (PDO; Mantua et al. 1997). These modes of variability can strongly affect the behavior of temperature and hydrological variables, and hence complicate the identification of slow-evolving climate responses to external forcings. For example, Mote et al. (2005), Stewart et al. (2005), and Knowles et al. (2006) found that some portion of the changes in snowpack, streamflow timing, and snow/rain partitioning could be explained by fluctuations in the PDO. ENSO events influence predominantly the interannual variability of western U.S. temperatures, extreme precipitation (Cayan et al. 1999), snowfall (Smith and O'Brien 2001), snowpack (Cayan 1996), and streamflow (Cayan et al. 1999; Dettinger and Cayan 1995).

Second, observed climate changes represent the net response of the climate system to multiple forcing factors, plus additional noise from natural internal variability. Use of observational data alone, even in conjunction with sophisticated statistical tools, does not permit us to unambiguously separate the climate change contributions from different forcings. Such separation can be performed with numerical models, which are frequently used for the systematic experimentation that we cannot conduct in the real world. However, all models have errors in both the forcings they are run with and the climate responses to those forcings. Over the topographically complex western United States, for example, there are small-scale climate features that cannot be resolved with coarse-resolution global climate models. Furthermore, model experiments often neglect forcing mechanisms that are known to be important in the real world, such as the widespread land-use changes associated with changes in agriculture, urbanization, and irrigation, which can have a marked influence on the climate of the western United States (Bonfils and Lobell 2007).

There have been a number of attempts to detect human effects on North American climate (Stott 2003; Zwiers and Zhang 2003; Karoly et al. 2003; Karoly and Wu 2005). Karoly et al. (2003) found that neither cli-

mate noise nor natural forcings could explain the large observed increase in annual-mean North American surface air temperature (30° and 65°N) from 1950 to 1999. Christidis et al. (2007) reported a significant anthropogenic contribution to North American growing season length, largely because of an earlier date of spring onset. Finally, Bonfils et al. (2008) showed that the recent wintertime warming over California was inconsistent with purely natural climate fluctuations and required contributions from one or more external forcings to be explained. The same study concluded that global climate model simulations fail to reproduce the strong seasonality of Californian temperature trends, probably because of their coarse resolution and lack of time-varying land-use forcings.

To our knowledge, no formal D&A study to date has focused on the temperature changes occurring over mountainous areas of the western United States, a critical area for the regional hydrology. In the present work, we conduct a formal D&A analysis over nine mountainous regions of the west (Fig. 1) using four different hydrologically related surface air temperature variables. In the detection phase, we investigate whether the observed changes in these variables can be fully explained by the background "noise" of natural internal climate variability, as estimated from statistically downscaled control simulations performed with two different global models. In the attribution phase, we examine whether the observed changes are consistent with the twentieth-century climate simulations that include anthropogenic greenhouse gases (GHGs), ozone, and aerosol effects and inconsistent with simulations that incorporate solar and volcanic forcing only.

Our D&A analysis relies on climate models that have been selected for their ability to capture important features of the climate of the North Pacific and western United States, such as the mean state and the variability associated with the PDO and ENSO (see sections 2b and 3e). Since the effect of finescale orography on temperature cannot be adequately represented in coarse-resolution global climate models, we use two different statistical "downscaling" techniques to transform data from global models to the small spatial scales of interest here. One underlying assumption in our study is that the neglect of changes in irrigation or urbanization in the model simulations is of less concern in mountainous regions that are of interest here. The detection variables we consider are all directly relevant for understanding changes in surface hydrology and snowmelt, and include the seasonal averages [January through March (JFM)] of daily minimum and maximum temperatures ($T_{min}$, $T_{max}$), the number of frost days (JFM FD), and the number of degree-days above 0°C (JFM

DD $> 0$). The last variable is a simple proxy for temperature-driven snowmelt (see section 2a).

In section 2, we define the temperature indices and introduce the observational and model data used in our study. A brief description of the two downscaling techniques is also provided. In section 3, observed and simulated temperature trends are compared for each of the nine mountain regions. Our D&A technique is described and applied in section 4. Our focus here is on the estimated detection times and signal-to-noise ($S/N$) ratios for an anthropogenic fingerprint together with the sensitivity of our results to various datasets and processing choices. Discussions and conclusions are presented in section 5.

## 2. Observational and model data

### a. Observational data

Spatial and temporal variations in maximum and minimum temperature for the period 1950–99 were obtained from the University of Washington (UW) Land Surface Hydrology Research group in the form of a gridded dataset (Hamlet and Lettenmaier 2005). The UW dataset primarily includes daily-mean data from the National Climatic Data Service's Cooperative Observer network (Coop) and monthly-mean data from the U.S. Historical Climatology Network (USHCN; Karl et al. 1990). The USHCN data are long records corrected for changes in time of observation, station location, instrumentation, and land use. They are used to include adjustments for temporal inhomogeneities in the gridded Coop data. The data were interpolated to a regular grid with $1/8° \times 1/8°$ latitude–longitude resolution (about 140 km$^2$ per grid cell).

The JFM $T_{min}$ and $T_{max}$ are simply computed by averaging daily $T_{min}$ and $T_{max}$ data from 1 January to 31 March of each year. The JFM FD index represents the total number of days over this period with a daily average temperature below the freezing point. The fourth and final index, DD $> 0$, measures the extent to which the daily average temperature exceeds the melting point (with daily values below 0°C set to zero), and JFM DD $> 0$ is the sum of each individual day's DD $> 0$. Quantifying snowmelt from DD $> 0$ is the simplest approach used in snowmelt-runoff models when data on surface energy balances are not available (see, e.g., Linsley 1943).[1] Our use of JFM DD $> 0$ as a proxy for snowmelt behavior (instead of the date of the onset of snowmelt, which is often difficult to define) is moti-

---

[1] In this technique, the daily snowmelt depth is computed by multiplying DD $> 0$ by a melt factor (in mm °C$^{-1}$ day$^{-1}$) that depends on the physical characteristics of the snow.
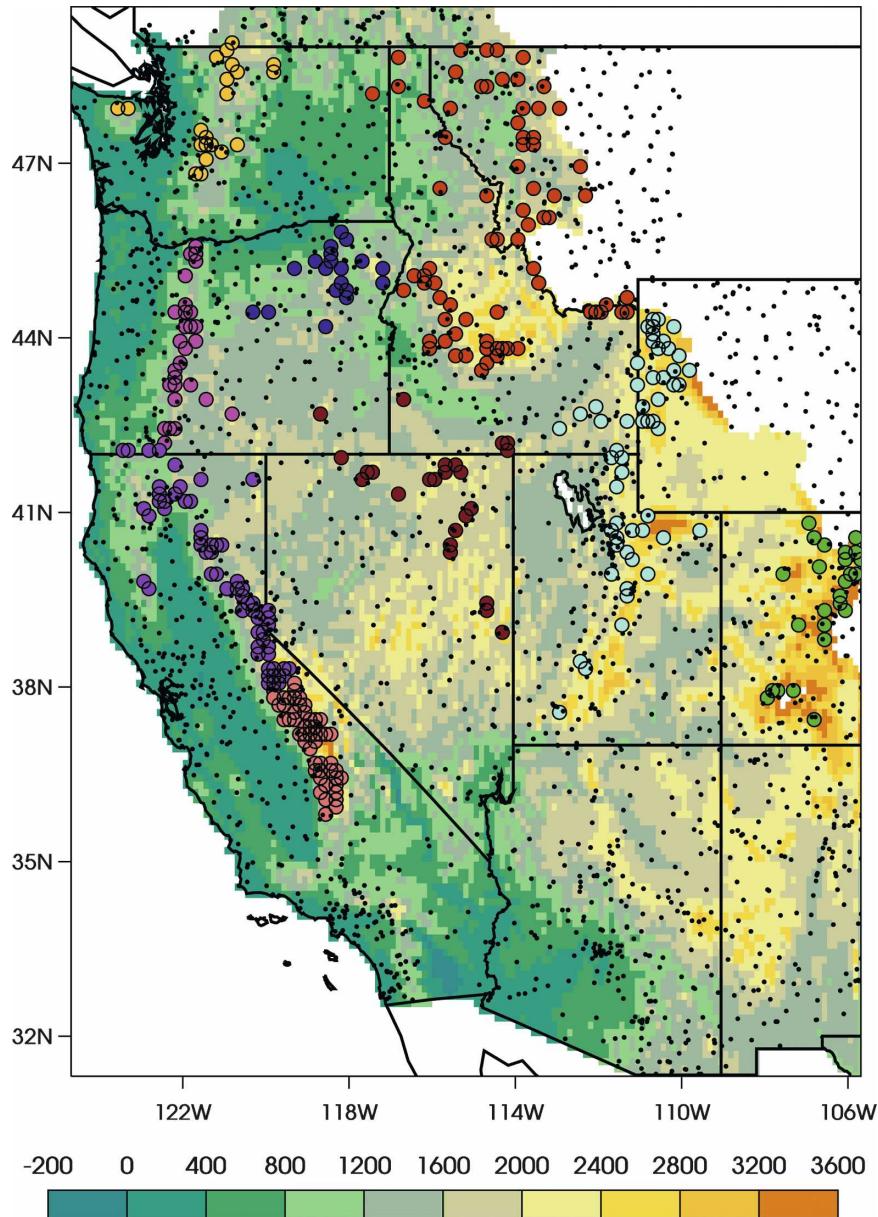
FIG. 1. Elevation (in meters) and location of the nine mountainous regions over which temperature indices are spatially averaged: the Washington Cascades (yellow), the northern Rockies (red), the Oregon Cascades (pink), the Blue Mountains (dark blue), the northern Sierras (purple), the southern Sierras (brown), the Great Basin (maroon), the Wasatch (light blue), and the Colorado Rockies (green). Each mountainous region is defined on the $1/8° \times 1/8°$ lat–lon grid of the UW observational temperature dataset. Colored circles denote the snow course locations with a climatological mean snow depth value of at least 1 cm on 1 Apr (Pierce et al. 2008; Barnett et al. 2008). The station network used to create the UW dataset is represented by black dots.

vated by the physics of snow: several days with temperatures above freezing may be required before an entire snowpack reaches 0°C and begins to melt. During the first few days with above-freezing conditions, meltwater produced at the surface of the snowpack may percolate through the snowpack and refreeze. Refreez-

ing can also occur at night, and it is only after several days of sustained above-zero temperatures that the snowpack structure changes and the melting is efficient.

We focus on mountainous regions of the western United States (which include 10 western states) in the vicinity of snow course stations with a climatological

mean snow depth value of at least 1 cm on 1 April (Fig. 1; see Pierce et al. 2008). Because these stations differ in terms of their latitude and elevation, we have defined the following nine subregions: the Washington Cascades, the northern Rockies, the Oregon Cascades, the Blue Mountains, the northern Sierras, the southern Sierras, the Great Basin, the Wasatch area, and the Colorado Rockies (Fig. 1). Observed time series and trends of the four temperatures indices averaged over these nine regions are displayed in Fig. 2. The network of stations used in the UW dataset has adequate coverage to estimate temperature changes in those nine regions (Fig. 1), although most temperature-observing sites are located at lower elevations than the snow course stations.

In the majority of stations in the United States, volunteer weather observers take temperature measurements once a day. At the time of observation (TO), the observers read both $T_{min}$ and $T_{max}$ recorded for the previous 24 h. At most U.S. stations, TO has shifted from afternoon to morning (Vose et al. 2003). Because morning observations produce more cold extremes, this has introduced a nonclimatic cold bias in extreme temperatures (DeGaetano and Allen 2002). Likewise, this bias may artificially weaken the positive trends in $T_{min}$ and $T_{max}$, attenuate the true increase in accumulated $DD > 0$, and lead to an underestimate of the true reduction in FD over time. As mentioned above, USHCN monthly temperature data were adjusted for TO bias (Karl et al. 1986) and used to correct the Coop data. Presumably, any residual TO bias would lead to an underestimate of a coherent warming signal caused by greenhouse gas increases. This would make the detection of such a signal more difficult.

### b. Global model simulations

This study uses data from three climate models, two downscaling techniques, and three sets of climate experiments (Table 1). The selection of climate models was primarily based on the availability of daily $T_{min}$, $T_{max}$ and precipitation data for the period 1950–99 and on the existence of multiple realizations of an experiment with anthropogenic forcing over the twentieth century (but no changes in solar or volcanic forcing).[2] A further selection criterion was model performance in representing the observed mean climate and its decadal



FIG. 2. Observed time series of $T_{min}$, $T_{max}$, frost days, and degree-days. Results are in the form of spatially averaged anomalies over the nine mountainous regions in our western U.S. study area, and are for the JFM season. For each variable, the least square best-fit linear trend is shown (solid line), as well as the minimum and maximum values among the 9 regional time series (gray envelop).

---

[2] Among the 18 CMIP3 models providing daily $T_{min}$ and $T_{max}$ data for simulations of twentieth-century climate change, only three have more than three ensemble members for the period 1961–99.

TABLE 1. Summary of the main characteristics of the simulations: $R$ represents the number of realizations. $Y_1$, $Y_2$, $L$ indicate the model-specific choices for the starting year, the ending year, and the length (in years) of the control runs, while $N_a$ and $N_b$ represent the number of nonoverlapping and overlapping 50-yr linear trends obtained from each of them. Forcings applied to the simulations are coded as below. A: greenhouse gases, B: ozone, C: direct effect of sulfate aerosols, D: indirect effect of sulfate aerosols, E: black carbon, F: organic carbon, G: mineral dust, H: sea salt, I: land-use change, J: solar irradiance, K: volcanic aerosols. Center for Climate System Research is CCSR, National Institute for Environmental Studies is NIES, Frontier Research Center for Global Change is FRCGC.

| Model designation | Downscaling technique | Initial resolution | Final resolution | Originating group(s) | $R$ | Forcings | $Y_1$ | $Y_2$ | $L$ | $N_a$ | $N_b$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| PCM-ANTH1 | BCSD | T42 | 1/8 × 1/8° | NCAR | 4 | ABC | 1871 | 1999 | 129 | — | — |
| PCM-ANTH2 | CA | T42 | 1/8 × 1/8° | NCAR | 4 | ABC | 1949 | 1999 | 51 | — | — |
| MIROC-ANTH | CA | T42 | 1/8 × 1/8° | CCSR/NIES/FRCGC | 10 | ABCDEFGHI | 1900 | 1999 | 100 | — | — |
| PCM-NAT | CA | T42 | 1/8 × 1/8° | NCAR | 2 | JK | 1949 | 1999 | 51 | — | — |
| CCSM3-CTL | CA | 1.0 × 1.25° | 1/8 × 1/8° | NCAR | 1 | None | 240 | 1089 | 850 | 17 | 81 |
| PCM-CTL | BCSD | T42 | 1/8 × 1/8° | NCAR | 1 | None | 451 | 1202 | 752 | 15 | 71 |
| Total | — | — | — | | 22 | — | — | — | 3424 | 32 | 152 |

variability over the western United States. Particular attention was focused on the fidelity with which models reproduced observed PDO and ENSO characteristics and on the impacts of these modes of variability on precipitation patterns.

Natural climate internal variability is estimated from two multicentury preindustrial control simulations performed with the finite-volume version of the National Center for Atmospheric Research/Department of Energy (NCAR/DOE) Community Climate System Model, version 3 (CCSM3-CTL; Bala et al. 2008) and the DOE/NCAR Parallel Climate Model (PCM-CTL; Washington et al. 2000; Collins et al. 2006). The atmospheric components of these models were run at 1.25° longitude × 1° latitude resolution and T42 spectral truncation, respectively. Both control runs have fixed preindustrial values of $CO_2$, sulfate aerosols, and tropospheric and stratospheric ozone, with no changes in solar irradiance or atmospheric burdens of volcanic aerosols. Our D&A analyses were performed with 850 yr of CCSM3-CTL data and 750 yr of PCM-CTL data. The control runs are relatively stationary over the selected analysis periods (Bala et al. 2008).

As noted above, the PDO[3] and ENSO are important sources of variability in western U.S. climate. Both CCSM3 and PCM capture key features of the spatial and temporal structure of these natural modes of variability, and are therefore suitable for estimating internal climate noise (see Fig. 2 of Pierce et al. 2008; Alexander et al. 2006; Meehl and Hu 2006). The leading mode of ENSO variability has spatial structure similar to that of observations, although it extends too far into the west Pacific (Pierce et al. 2008). The spatial pattern

of the PDO in the CCSM3-CTL and PCM-CTL captures the observed "horseshoe" shape over the North Pacific Ocean. The peak loading is correctly positioned over the central Pacific in the PCM-CTL but is displaced toward Japan in CCSM3-CTL (Pierce et al. 2008).

These features of the PDO are also apparent in all Coupled Model Intercomparison Project phase 3 (CMIP3) twentieth-century runs performed with PCM and the T85 version of CCSM3. Spatial correlations between the observed PDO pattern and the patterns simulated in the CCSM3 and PCM twentieth-century runs typically range from 0.86 to 0.91, respectively. On decadal time scales, the amplitude of SST variability in the PDO region is roughly 25%–30% higher than observed in PCM and 15% lower than observed in CCSM3. There is no evidence, therefore, that the two control simulations significantly underestimate either ENSO or PDO variability (see further discussion in section 3e). PCM also captures features of the observed secular changes in sea surface temperatures, such as the prominent "regime change" in the mid-1970s. In the model, this shift is primarily due to a combination of internally generated variability and anthropogenic forcing (Meehl et al. 2009). Finally, we note that both PCM and CCSM3 successfully replicate the pattern of climatological mean December–February (DJF) precipitation over the western United States, as estimated from the National Centers for Environmental Prediction (NCEP) Global Reanalysis (Kalnay et al. 1996) over the 1949–98 period ($r > 0.9$).

The anthropogenic signal was estimated from two ensembles of historical simulations: a 4-member ensemble[4] performed with PCM (PCM-ANTH) and a 10-member ensemble generated with the T42-resolution

---

[3] The PDO is defined here as the leading EOF of wintertime northern Pacific sea surface temperatures. The regional mean sea surface temperatures is subtracted prior to calculation of EOFs.

[4] The runs analyzed were B06.22, B06.23, B06.27, and B06.28.

("Medres") version of the Model for Interdisciplinary Research on Climate 3.2 (MIROC3.2) model (MIROC-ANTH). PCM-ANTH runs include changes in well-mixed greenhouse gases, tropospheric and stratospheric ozone, and the direct scattering effects of sulfate aerosols. MIROC-ANTH (K-1 model developers 2004; Nozawa et al. 2007) additionally includes the direct effects of carbonaceous aerosols, and some indirect effects of both sulfate and carbonaceous aerosols on clouds (see Santer et al. 2007 and Table 1 for a complete list of forcings). All CMIP3 twentieth-century runs performed with the MIROC T42 model replicate the observed structure of the PDO, although peak loadings of the leading EOF are misplaced and slightly underestimated (not shown). The spatial correlation with observations ($r = 0.72$) is slightly lower than for CCSM3 or PCM, but the MIROC T42 models still ranks among the top CMIP3 models in terms of its representation of the spatial structure of the PDO. The MIROC model depicts a PDO frequency spectrum comparable to observations (Shiogama et al. 2005).

Finally, to characterize the climate response to natural external forcings, we analyze two PCM simulations which are forced solely by historical changes in solar irradiance and volcanic aerosols (PCM-NAT; cases B06.68 and B06.69).

### c. Downscaling techniques

The western United States is climatologically complex; its topography is not well represented by the coarse resolution of most climate models: even the best models display climate biases at regional scales (Maurer and Hidalgo 2007). To better capture the nuances of the climate changes over mountainous regions and at the spatial scales of large river basins as required by our study, the daily precipitation, $T_{min}$, and $T_{max}$ data from all climate simulations were first statistically downscaled, and then used to force the variable infiltration capacity (VIC) hydrological model (Liang et al. 1994; Cherkauer and Lettenmaier 2003). Two downscaling approaches were employed here: the constructed analogs technique (CA; H. G. Hidalgo et al. 2008) and the bias-correction and spatial downscaling procedure (BCSD; Wood et al. 2004). The use of two different methods provides useful information on the sensitivity of D&A results to the choice of statistical downscaling technique.

The CA technique estimates a daily "target pattern" of $T_{min}$, $T_{max}$, or precipitation from a climate model. This target is a linear combination of observed daily patterns that have been aggregated to the climate model resolution (the analog). The downscaled data are generated by applying the estimated regression co-efficients for the "target pattern" to the corresponding 1/8°-resolution daily observed patterns.

The BCSD method generates climatological cumulative distribution functions of the monthly-mean climate model data (over the period 1950–99) and maps their quantiles onto those of gridded observations aggregated to the climate model resolution. Anomalies of the bias-corrected model variables are then formed relative to the climatological reference period, interpolated to 1/8° resolution, and added to the 1/8° gridded observational means. The final step of this procedure is to generate daily forcing fields by a resampling and rescaling technique. The BCSD method yields bias-corrected daily temperature and precipitation outputs that preserve the mean, variability, and temporal evolution of the monthly-mean data (Maurer et al. 2007). Details of the techniques are given in H. G. Hidalgo et al. (2008) and Wood et al. (2004), and their performance is compared in Maurer and Hidalgo (2007).

To investigate whether the anthropogenic signal is sensitive to the choice of downscaling technique, the four PCM anthropogenic runs were downscaled using both the BCSD approach (PCM-ANTH1) as well as the CA method (PCM-ANTH2). All other simulations were downscaled with one of the two techniques only: BCSD was used for the PCM-CTL run, while the CA method was applied for the CCSM3-CTL run, the MIROC anthropogenic runs, and PCM-NAT experiment (see Table 1). The four temperature variables employed in the D&A analysis were calculated from the statistically downscaled $T_{min}$ and $T_{max}$ data and (in the case of the FD and DD > 0 indices) from the surface air temperature recalculated by the VIC model after estimating a diurnal cycle of temperature from $T_{min}$ and $T_{max}$.

### d. Correlation of temperature indices with SWE/P

Each of the selected temperature indices is expected to show some relationship with changes in aspects of the hydrological cycle. For each of the nine regions, we computed the correlation between 1 April snow water equivalent (SWE) divided by the accumulated October–March precipitation (SWE/$P$; see Pierce et al. 2008) and the four different JFM temperature indices over the period 1950–99 (Fig. 3). SWE is divided by $P$ to reduce the influence of precipitation fluctuations on the variability and trends in SWE, and hence highlight trends in the temperature-driven component of SWE.

In both observations (Fig. 3a) and the PCM-ANTH1 results (Fig. 3b), SWE/$P$ changes over 1950–99 are inversely correlated with temperature changes. In PCM-ANTH1 and the observations, the DD > 0, FD, and
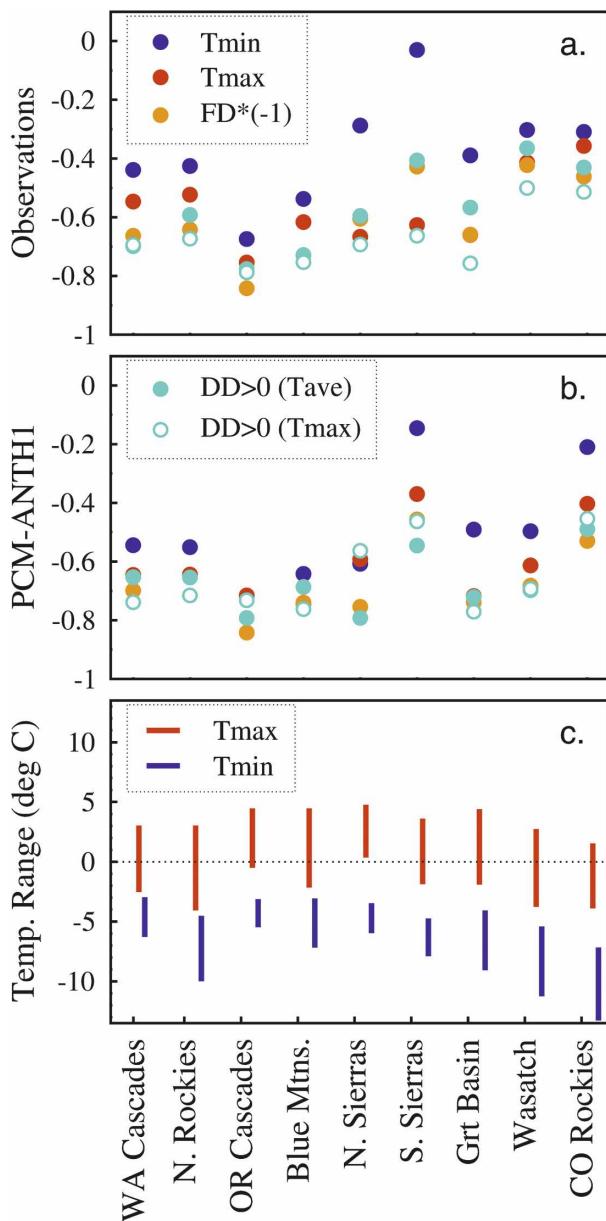
FIG. 3. Correlation between SWE/P and five different JFM temperature indices averaged over the period 1950–99 from (a) observations and (b) PCM-ANTH1 ensemble mean. (c) Range of daily observed climatological data for $T_{min}$ and $T_{max}$ computed over the period 1950–99.

$T_{max}$ temperature indices generally show the largest negative correlations with SWE/P changes. Correlations between $T_{min}$ and SWE/P are consistently lower. SWE/P changes are more sensitive to $T_{max}$ than $T_{min}$ fluctuations since $T_{max}$ reaches and exceeds the melting point threshold, while $T_{min}$ remains below this threshold during the entire JFM season (Fig. 3c). As a sensitivity test, we computed DD > 0 from daily maximum

temperatures instead of daily average temperatures (the former threshold is sometimes adopted in snowmelt-runoff models). The new index (DDx > 0) shows a very large inverse correlation with SWE/P in both the observations and the PCM-ANTH1 results (Figs. 3a,b).

Temperature changes in April may also influence the rate of snowmelt, especially in northern and colder states. However, correlations between time series of observed SWE/P changes and changes in temperature indices averaged over January to April are consistently lower than those previously discussed for JFM (except for $T_{min}$). This provides some justification for our analysis of JFM temperature changes. It also allows a more direct comparison with the SWE/P results in Pierce et al. (2008), who analyzed SWE as measured on 1 April.

## 3. Observed and simulated trends

### a. Spatial patterns of trends

Between 1950 and 1999, observed JFM $T_{min}$ and $T_{max}$ have increased over the entire western United States by 1.83°C and 1.54°C (Table 2), with 95% confidence intervals of ±0.68°C and ±0.61°C, respectively. Over the same period, JFM FD decreased by 7.6 days, while JFM DD > 0 increased by 74.8°. All of these trends are significantly different from zero at the 5% level. While the $T_{min}$ and $T_{max}$ trends show spatially and elevationally coherent patterns of warming over the western United States (Figs. 4a,b), patterns of trends in FD and DD > 0 more closely reflect the underlying topography of the region (Figs. 4c,d), indicating a sensitivity of FD and DD > 0 to elevation. For example, trends in FD are generally large except over the Central Valley and desert areas, where the frost events are rare occurrences. Trends in DD > 0 are small or close to zero at very high elevations, where temperatures are too low to exceed the threshold of 0°C.

Removing PDO and ENSO variability[5] from the temperature indices reduces the amplitude of the overall trends, but does not noticeably modify the coherence of the warming pattern (for $T_{min}$ and $T_{max}$) or the coherence of the decrease in FD and the increase in DD > 0 (see Fig. 4, 2nd column). Even after removal of PDO and ENSO effects, all trends remain significantly

---

[5] We linearly regressed the temperature time series against the contemporaneous December–February Niño-3.4 and PDO indices and then removed the temperature variability associated with each of these indices. Niño-3.4 anomalies were calculated with respect to the climatology over the base period 1971–2000.

TABLE 2. Observed trends in five JFM temperature indices over the period 1950–99. All trends are total linear changes over 50 yr, together with their 95% confidence intervals (adjusted for temporal autocorrelation). Results are spatially averaged over the entire western U.S. study area.

| | $T_{min}$ (°C) | $T_{max}$ (°C) | FD (days) | DD > 0 (°C day) | DDx > 0 (°C day) |
|---|---|---|---|---|---|
| Raw data | 1.83 ± 0.68 | 1.54 ± 0.61 | −7.59 ± 2.84 | 74.80 ± 22.29 | 108.48 ± 40.05 |

different from zero at the 5% level. For comparison, spatial trends for one of the four downscaled PCM-ANTH1 realization (B06.28) are also presented (Fig. 4, 3rd column). The general structure and amplitude of the simulated trends is consistent with observations. The most noticeable difference is that the simulated patterns of $T_{min}$ and $T_{max}$ changes show less spatial heterogeneity than the corresponding observed results. The observed sensitivity of the amplitude of FD and DD > 0 trends amplitudes to elevation is well captured by the model. Regression-based removal of PDO and ENSO signals from these temperature indices has a larger impact in the selected PCM simulation than in observations (Fig. 4; compare 2nd and 4th columns).

### b. Comparison of observed and unforced trends in mountainous regions

Observed trends over 1950–99 in the nine mountain regions were computed for all four temperature indices[6] (Fig. 2). A standard statistical test of trend significance (Santer et al. 2000a) reveals that in all cases except one (the $T_{max}$ trend in the Colorado Rockies) observed trends are significantly different from zero at the 5% level.[7] A related question is whether natural internal climate variability, as simulated by the CCSM3 and PCM models, could explain the observed 50-yr trends. To address this question, we calculated (separately for each index and each region) 50-yr trends from the CCSM3-CTL and PCM-CTL runs. This was done for both nonoverlapping trends and for trends that overlapped by all but 10 yr. We then pooled results from both control runs to form 36 sampling distributions of unforced trends (9 regions × 4 indices). For the case of overlapping 50-yr trends, this procedure yields a total of 152 samples (see Table 1), from which we can estimate the probability (i.e., the $p$ value) that the observed trend could be due to climate noise alone. Note that the use of overlapping trends yields smoother estimates of the underlying sampling distributions but has relatively little influence on the estimated $p$ values.

Assuming that the model-based estimates of internal climate noise are reliable (see sections 2b and 3e), we conclude that, in 27 (31) out of 36 cases, observed trends are significantly different from internal climate variability at the 5% level using a two-tailed test (using a one-tailed test).[8] Intriguingly, the southern Sierras yields nonsignificant trends for three of the four temperature indices. This is the only region that experiences a slight (although nonsignificant) increase in snowpack (Pierce et al. 2008). In summary, our simple trend comparison analysis suggests that, for most variables and in most regions, internal climate variability alone cannot explain the observed long-term temperature changes over the last 50 yr.

### c. Comparison of unforced and anthropogenically forced trends

We consider next whether the observed changes are consistent with model results from anthropogenically forced simulations. Before addressing this question, we would like to assess (for each region and temperature index) whether the distribution of ANTHRO trends differs significantly from the pooled distribution of unforced trends obtained from the CCSM3-CTL and PCM-CTL runs.

As expected, trends from individual ANTH realizations show considerable between-realization variability (see Fig. 5). For example, while trends from three PCM ANTH realizations (B06.22, B06.27, and B06.28) have the same sign as the observed trends for all four temperature indices, one realization (B06.23) has changes of opposite sign. Trends obtained with CA downscaling are weaker than those obtained with BCSD, in agreement with the results of Maurer and Hidalgo (2007). Of the 10 MIROC-ANTH simulations, 5 yield changes that are consistent in sign with the observations. The

---

[6] Observed time series for $T_{min}$ in the Washington Cascades, northern Rockies, and northern Sierras are shown in Barnett et al. (2008).

[7] This test of trend significance accounts for temporal autocorrelation of the regression residuals in estimating both the standard error of the trend and the degrees of freedom for the $t$ test of trend significance.

[8] This is a reasonable assumption since we a priori expect a warming in North America (Karoly et al. 2003) and in California (Bonfils et al. 2008), with a subsequent reduction in FD and increase in DD > 0.
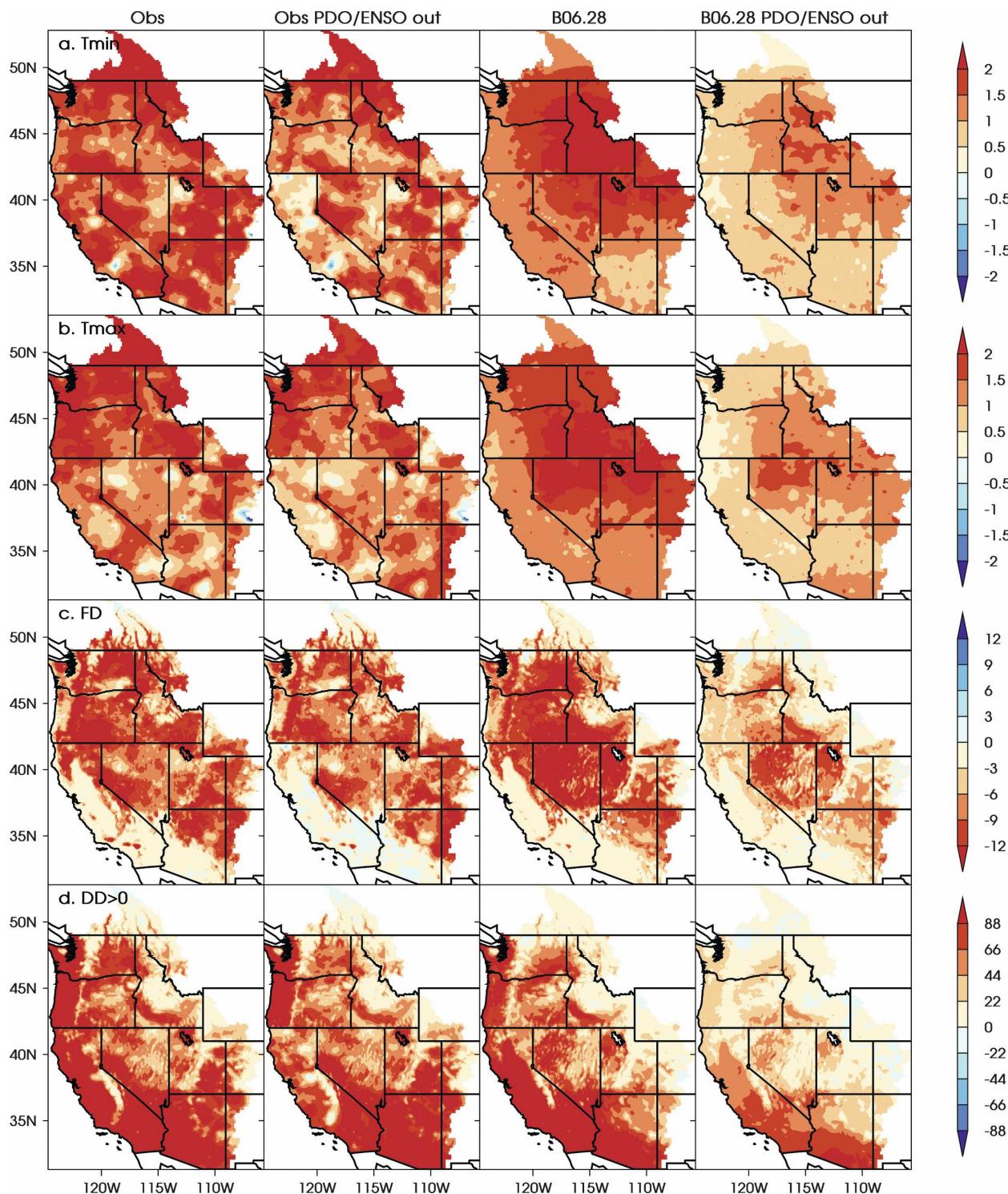
FIG. 4. Observed and one example of simulated 1950–99 changes in (a) JFM $T_{min}$, (b) $T_{max}$, (c) FD, and (d) DD > 0 over the western United States. Trends have been computed from the UW and one PCM-ANTH1 realization (B06.28), using raw data (columns 1 and 3) and data in which observed or simulated PDO and ENSO are regressed out (columns 2 and 4). The color scale is chosen such that red indicates an increase in $T_{min}$ and $T_{max}$, an increase of DD > 0, and a decrease in FD.
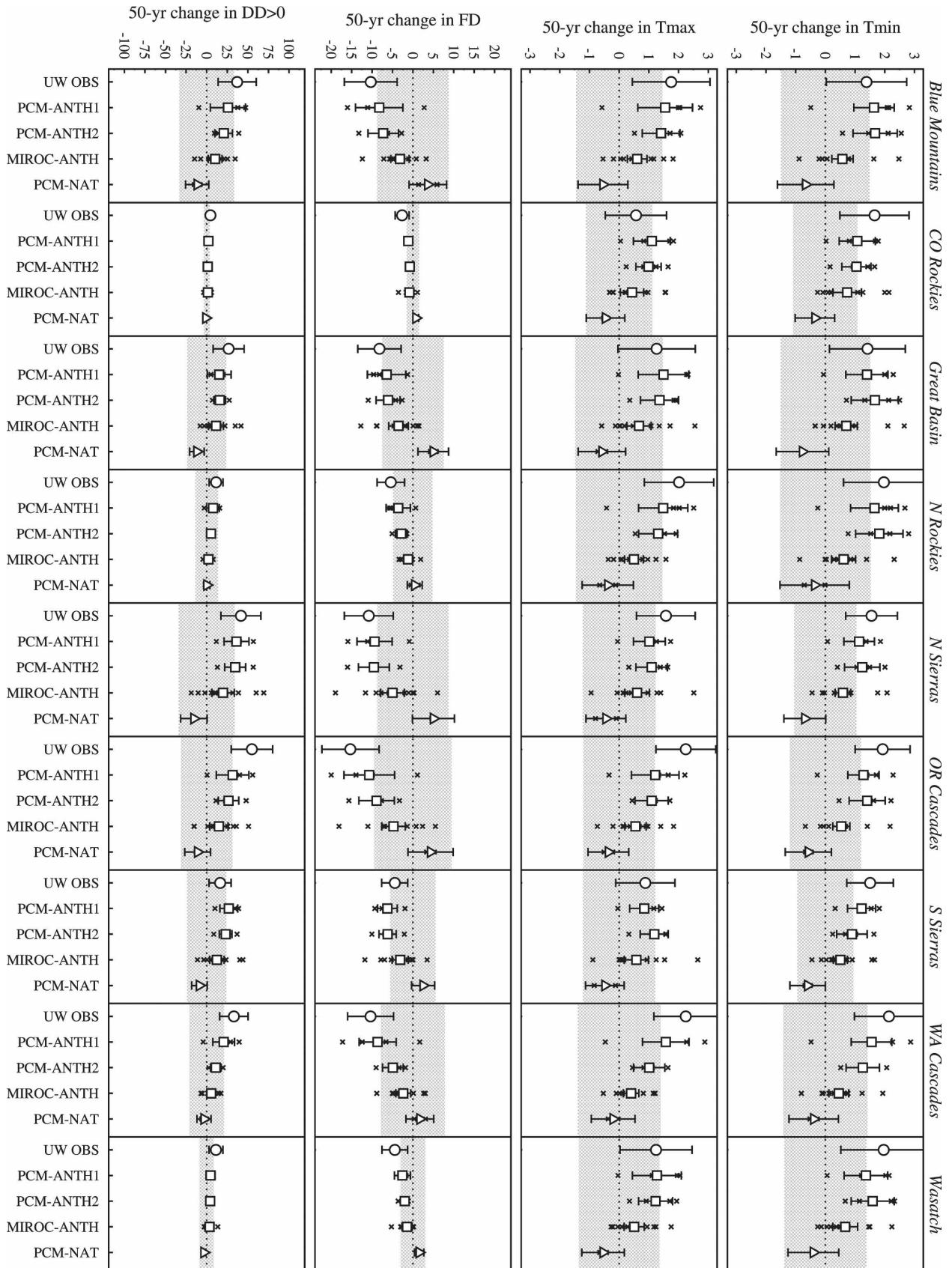
other five realizations show cooling or virtually no overall change in temperatures. Although significantly different from zero (Fig. 5), the trends estimated from the MIROC-ANTH ensemble mean are always smaller than those obtained from PCM-ANTH1 and PCM-ANTH2. This is likely due to the large negative forcing associated with the indirect effect of aerosols on clouds, which is included in the MIROC-ANTH forcings but ignored in the PCM-ANTH forcings (see, e.g., Fig. 5 in Santer et al. 2007). The large between-realization variability in both the PCM and MIROC ANTH results illustrates the difficulty of reliably estimating the response to external forcings in the presence of substantial low-frequency chaotic variability.

We use a Kolmogorov–Smirnov (KS) test to determine whether the trend distributions inferred from the ANTH and CTL experiments differ significantly (Press et al. 1992).[9] For each variable and region, we first compare the pooled distribution of 14 PCM-ANTH1 and MIROC-ANTH trends to the pooled distribution of 32 nonoverlapping PCM-CTL and CCSM3-CTL unforced trends (Table 3). In all cases except one, the distributions of anthropogenic and unforced trends are significantly different at the 5% level. When this test is repeated with the individual PCM and MIROC ANTH ensembles, the distribution of PCM-ANTH trends still differs significantly from the distribution of unforced trends, irrespective of the downscaling method, the region, and the variable under consideration. The MIROC-ANTH trend distributions also differ significantly from unforced trends distributions (except for DD > 0). KS tests conducted with the individual CCSM3-CTL and PCM-CTL distributions lead to simi-

lar conclusions.[10] Nonsignificant results primarily involve the combination of MIROC-ANTH and CCSM3-CTL distributions.

### d. Comparison of observed trends with anthropogenically forced and naturally forced trends

To assess whether the observed trends are consistent with the distribution of anthropogenically forced signals, we conducted a simple test of the significance of trend differences (Santer et al. 2000b).[11] In 158 out of 180 cases, the null hypothesis that the observed trend is drawn from the same distribution as the ANTH trends cannot be rejected (see Table 3, "UW observation" column); in other words, the observed trends are consistent with those expected because of anthropogenic forcing.

We also considered whether the observed trends for all regions and temperature indices are consistent with the simulated response to natural external forcing alone. The observed and PCM-NAT trends are of opposite sign (see Fig. 5), and the PCM-NAT trends are not significantly different from zero. Our results suggest that historical changes in volcanic aerosols and solar irradiance cannot explain observed changes in $T_{min}$, $T_{max}$, DD > 0, and FD.

Because downscaled data were available from only two PCM-NAT runs, we were not able to obtain a reliable estimate of the true response of western U.S. temperature indices to natural external forcings. To improve this estimate, we also analyzed changes in monthly-mean $T_{min}$ and $T_{max}$ data (spatially averaged

---

[9] The KS test compares the cumulative distribution A(x) of the ANTH trends with the distribution C(x) of the CTL unforced trends. Since we expect the $T_{min}$, $T_{max}$, and DD > 0 ANTH trends to exceed the unforced trends, we test the null hypothesis that $A(x) \geq C(x)$. For FD, the null hypothesis is that $A(x) \leq C(x)$. We compute the maximum distance between C(x) and A(x) and the associated one-tailed p value.

[10] Using all possible model configurations for estimating the anthropogenic fingerprint and the noise ensures that fingerprint identification is robust to uncertainties in the applied forcings and to structural uncertainties in the models (related to differences in physics, parameterizations, resolution, etc.).

[11] The standard errors are computed from observed and ensemble mean time series, accounting for temporal autocorrelation of the regression residuals (Santer et al. 2000b).

---

←

Fig. 5. Changes in temperature indices over the period 1950–99 computed from observations (circle), from PCM-ANTH1, PCM-ANTH2, MIROC-ANTH ensemble means (squares), and PCM-NAT ensemble mean (triangle). Individual realizations are indicated by "x." The model-derived estimate of the two-tailed 95% confidence interval natural internal variability is represented by the shaded area, the limit of which is calculated as 1.96 × sE, where sE is the standard error of the sampling distribution of combined CCSM3-CTL and PCM-CTL 50-yr unforced trends. Vertical bars represent the 95% confidence interval for the 50-yr observed and ensemble mean trends (i.e., their standard error × 1.641 for a one-tailed t test). For most geophysical variables, there is a positive lag-1 correlation in the regression residuals, showing that samples are not totally independent. We account for the effect by reducing the sample size, which inflates the standard error estimate (Santer et al. 2000a). Sample sizes are, however, not corrected when autocorrelation in the regression residuals is negative.

TABLE 3. Results of KS test to determine whether the anthropogenic signal strengths are statistically different from the population of unforced trends. Tests are conducted using individual or combined runs (the "combined" case in the second column combines PCM-ANTH1 and MIROC-ANTH ensemble means). A cross indicates the anthropogenic and unforced trends are statistically different at the 5% significance level. A circle indicates the observed trend is consistent with the distribution of trends obtained with anthropogenic forcing (i.e., the null hypothesis that the observed trend can be drawn from the same distribution as the ANTH trends is not rejected at the 5% significance level, using a test of significance of the difference between two populations' means; see text for details).

| | | Combined CTL | | | | | | | | | FV-CTL | | | | | | | | | PCM-CTL | | | | | | | | | UW OBSERVATION | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | BM | CR | GB | NR | NS | OR | SS | WA | WH | BM | CR | GB | NR | NS | OR | SS | WA | WH | BM | CR | GB | NR | NS | OR | SS | WA | WH | BM | CR | GB | NR | NS | OR | SS | WA | WH |
| $T_{min}$ | Combined | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH1 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH2 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | MIROC-ANTH | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| $T_{max}$ | Combined | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH1 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH2 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | MIROC-ANTH | × | × | × | × | × | × | × | × | × | × | × | × | | | × | × | × | × | × | × | × | | | × | × | × | × | ○ | ○ | ○ | | | ○ | ○ | ○ | ○ |
| FD | Combined | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH1 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH2 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | | | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | MIROC-ANTH | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| DD > 0 | Combined | × | × | × | × | × | × | × | × | × | × | × | × | | × | × | × | × | × | × | × | × | | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH1 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | | | ○ | ○ | ○ | ○ |
| | PCM-ANTH2 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | | | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | MIROC-ANTH | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | | | ○ | ○ | ○ | ○ |
| DDx > 0 | Combined | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH1 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | PCM-ANTH2 | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| | MIROC-ANTH | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | × | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

over the entire western United States) in 4 PCM-NAT runs and in 32 additional PCM-based estimates of "total" natural variability.

This procedure yields a total of 36 estimates[12] of the changes in $T_{min}$ and $T_{max}$ in response to natural external forcing. We then tested the null hypothesis that the observed trend is consistent with the model NAT results. We find that the null hypothesis is rejected at the 5%. These results show that natural external forcing alone is an inadequate explanation for the observed changes in $T_{min}$ and $T_{max}$ over the western United States. A similar comparison of simulated and observed trends (based on monthly-mean, spatially averaged temperature data) reveals that the observations *are* consistent with the PCM-ALL trends (Fig. 6).

### e. Reliability of the model-based noise estimates

The strength of the conclusions of the detection analysis depends on the ability of the models to simulate the observed climatological mean and internal climate variability. An evaluation of these statistics is conducted using JFM $T_{min}$ spatially averaged for the nine mountainous regions. The 50-yr climatological mean, the interannual variability, and the low-frequency variance are computed for UW, ANTH, NAT, and CTL (Fig. 7). CCSM3-CTL and PCM-CTL simulations are analyzed using their 17 and 15 nonoverlapping segments. UW, ANTH, and NAT time series are first detrended to compute the variability in an attempt to remove the slow-evolving signal from the records (retaining this trend would otherwise amplify the variability in UW, ANTH, and NAT and bias the comparison with CTL variability). While the models show biases in their simulation of the means in some regions, we find no evidence that the models systematically underestimate the observed high-frequency (not shown) and low-frequency variability (Fig. 7) for this temperature variable. Similar results are found with $T_{max}$, FD and DD > 0 (not shown). We conclude the models used here provide an adequate representation of natural internal climate variability for our detection and attribution work.

## 4. Detection and attribution analysis

### a. Methodology

The trend comparisons presented in section 3 strongly suggest that, in most of the regions examined here, observed changes in hydrologically related temperature variables are inconsistent with natural internal

---

[12] 4 PCM-NAT + 16 (S + V) + 16 (ALL-ANTH).
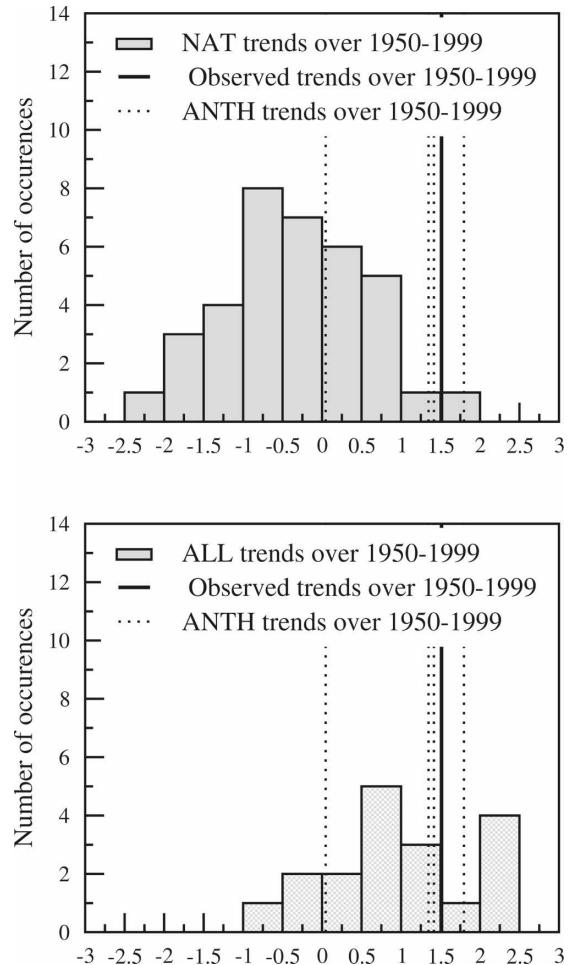


FIG. 6. Comparison of observed (solid line) and ANTH (dotted lines) trends with the sampling distributions of "NAT" and "ALL" trends. All trends are linearly fitted to JFM $T_{max}$ changes spatially averaged over the western United States for the period 1950–99. (a) The "NAT" distribution is obtained from 36 monthly-mean PCM-based estimates of "total" natural variability (where "total" includes the internally generated variability and the variability caused by natural external forcing). This ensemble is composed of 4 PCM realizations including both solar and volcanic forcings, of 16 additional estimates generated by adding temperature changes from each of 4 "solar only" and each of 4 "volcano only" realizations, and of 16 estimates obtained by subtracting each of 4 "ANTH" realizations from each of 4 "ALL" realizations (Santer and Wigley 2008). The resulting variability estimates in these 16 "S + V" and 16 "ALL − ANTH cases" are noisier than in the 4 PCM-NAT realizations. This is because the noise is amplified by the addition of S and V realizations (or the subtraction of ALL and ANTH realizations) with uncorrelated internal variability. (b) The "ALL" distribution is obtained from 20 monthly-mean PCM-based estimates of changes in temperature in response to all forcings (anthropogenic and natural). This ensemble is composed of 4 PCM realizations including all forcings and of 16 linear combinations of SV and ANTH realizations. Test of the significance of trend differences indicates that the observed temperature changes are consistent with the distributions of ALL trends at the 5% level but inconsistent with natural external forcing alone. Similar results are obtained with JFM $T_{min}$ changes.
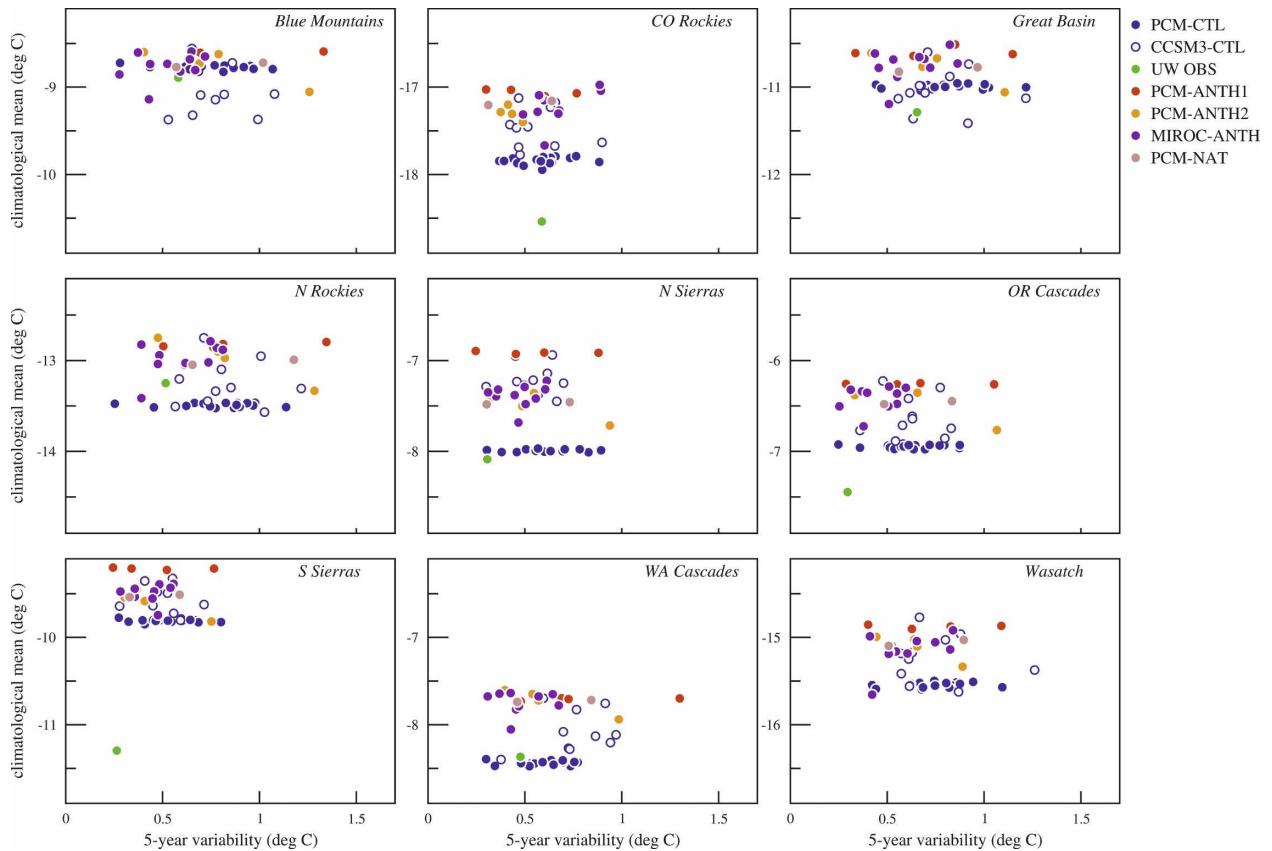
FIG. 7. The 50-yr climatological mean vs the 5-yr variability in the time series of JFM $T_{min}$ from UW, ANTH, NAT, and CTL. For CCSM3-CTL and PCM-CTL, their 17 and 15 nonoverlapping segments are used. The time series from UW, ANTH, and NAT are first detrended to compute the variability to remove the slow-evolving signal from the records.

variability alone but are consistent with the response to anthropogenic forcings. These trend comparisons were performed for each of nine individual regions. In contrast, the formal D&A analysis presented in the following section (Santer et al. 1995, 2007) relies on the combined information from all nine regions. The D&A method applied follows Hasselmann (1979), in which the expected pattern of climate response to anthropogenic forcings (the "fingerprint") is searched for in the observations.[13] We determine the time at which this fingerprint is statistically identifiable in the UW temperature data.

Our D&A analysis requires three components: an ANTH ensemble (either PCM-ANTH1, PCM-ANTH2, or MIROC-ANTH) to estimate the fingerprint, the

UW observational dataset which is projected onto the fingerprint, and a long control simulation (either CCSM3-CTL or PCM-CTL) for assessing the statistical significance of results. The input UW and ANTH data consist of 9 space points (one for each of the 9 regions) and 50 time samples (for the 1950–99 period). The control run data have either 750 or 850 time samples (for PCM and CCSM3, respectively). All data for individual regions are expressed in anomaly form with respect to climatological means calculated over the entire period. The fingerprint is defined as the leading EOF of the ANTH ensemble-mean anomalies, based on the $(9 \times 9)$ temporal covariance matrix.[14] Ensemble averaging and use of the leading EOF are both noise reduction techniques. The projections of the observational and CTL

---

[13] In Hasselmann's 1979 paper, the model-derived fingerprint is "optimized"—that is, it is rotated away from high noise directions in order to enhance S/N ratios and detection times. Here, no attempt is made to optimize the searched-for fingerprint.

[14] Note that our analysis focuses on the spatial structure of the fingerprint: we do not explicitly include information on its temporal evolution. An implicit assumption in our approach is that the *amplitude* of the fingerprint pattern changes over time, but its gross structure does not.

anomaly data onto the fingerprint yield the signal and noise time series $Z(t)$ and $N(t)$, respectively.

### b. Example of detection of changes in JFM $T_{min}$

In the following, we analyze the JFM $T_{min}$ case, using the mean of the PCM-ANTH1 ensemble to estimate the anthropogenic fingerprint, the CCSM3-CTL to assess the statistical significance, and the UW observations. In the baseline "no weighting" case, we assign each region unit weight despite differences in their spatial coverage.

The leading EOF of PCM-ANTH1 explains 89.5% of the overall space–time variance and has uniform sign in all regions (Fig. 8a). The projection of the observations onto this fingerprint, $Z(t)$, has a pronounced trend (not shown), suggesting that the fingerprint primarily captures slow secular warming common to all regions. The leading modes of the PCM and CCSM3-CTL simulations are structurally very similar to the PCM-ANTH1 fingerprint (Fig. 8a). This is because, in both the control and ANTH runs, the spatial coherence of temperature variability at interannual and longer time scales is large relative to the size of the study area.

We then project the CCSM3-CTL run data onto the PCM-ANTH1 fingerprint to obtain the noise time series $N(t)$. Least squares linear trends of increasing length $L$ ($L = 10, 11, 12, \ldots 50$ yr) are fitted to $Z(t)$ and to nonoverlapping $L$-length segments of the $N(t)$ time series. The start date for the $Z(t)$ trends is fixed at 1950—the start date of the observations and ANTH experiments used here. This allows us to compare the signal trend over increasingly longer time periods (1950–59, 1950–60, 1950–61, ... 1950–99) against the standard deviation of the sampling distribution of noise trends on the appropriate time scales (i.e., 10, 11, 12, ... 50 yr). The resulting signal-to-noise ($S/N$) ratio is a function of the trend length $L$. In our formalism, we define the "detection time" as the time at which the $S/N$ ratio first exceeds (and then remains above) some stipulated significance threshold (typically 5% or 1%) and $S/N_{50}$ as the $S/N$ ratio for the 50-yr trend length. One underlying assumption is that the $L$-length distributions of trends in $N(t)$ are Gaussian.

In our specific example, the PCM-ANTH1 fingerprint is identifiable in the observed JFM $T_{min}$ data as early as 1986—that is, 36 yr after the beginning of the observational records used here (Fig. 9). The $S/N$ ratio shows an almost monotonic increase, with superimposed high-frequency noise. For the $T_{min}$ case, the detection time is not sensitive to the choice of model used to estimate the fingerprint (PCM or MIROC) or to the choice of statistical downscaling method for the PCM fingerprint (CA or BCSD).
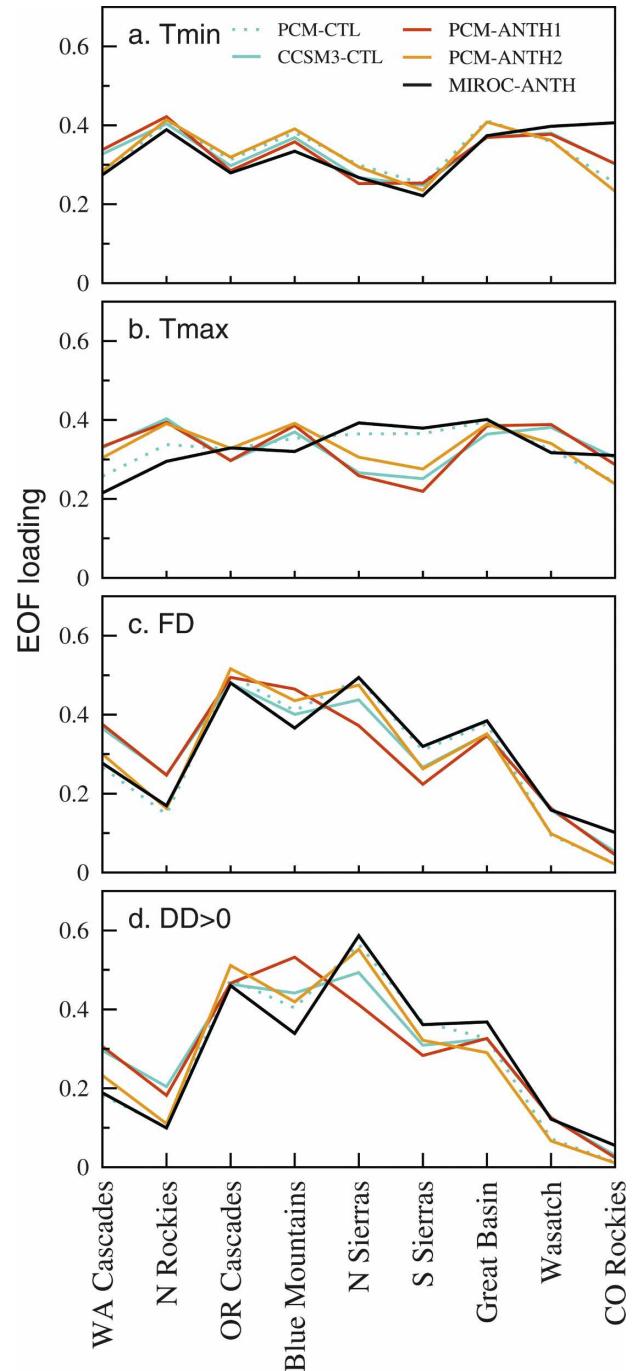


FIG. 8. Comparison of the spatial structure of EOFs for four different variables. Results are for the PCM and CCSM3 control runs and the MIROC and PCM anthropogenic fingerprints; the latter are statistically downscaled using two different methods. All results are for the case of no area weighting of control run and ANTH temperature. The spatial structure is more pronounced for DD > 0 and FD than for $T_{min}$. Note that, for the purposes of EOF calculations, the ordering of the nine mountainous regions is arbitrary.
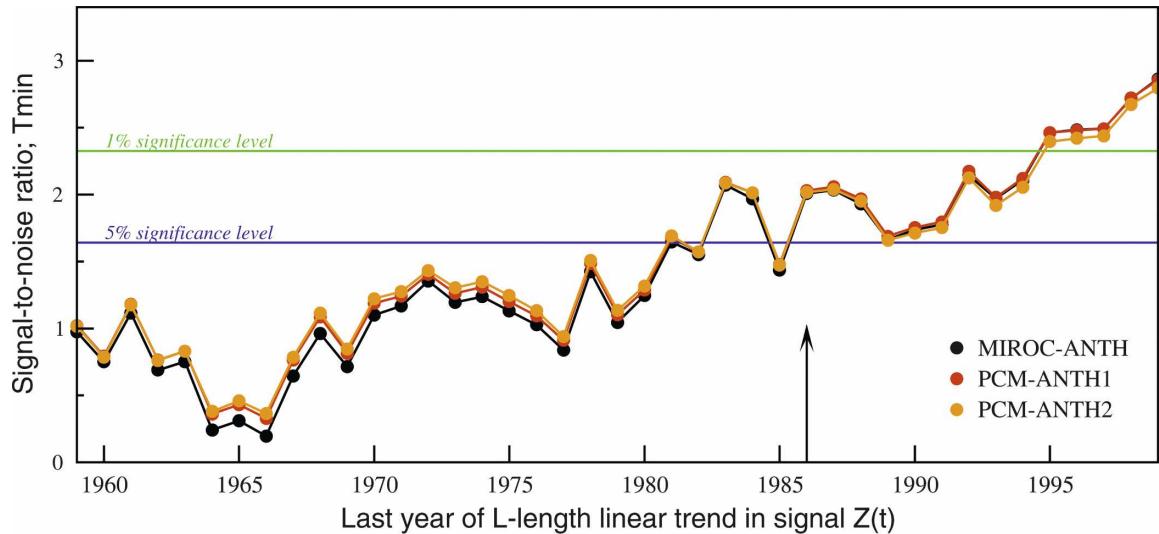
FIG. 9. Time-dependent $S/N$ ratios estimates for $T_{min}$, using CCSM3-CTL for statistical significance testing, and three different ensemble-means for estimating the ANTH fingerprint (unweighted case). Detection time (see arrow, estimated when $S/N$ exceeds and remains above the 5% significance level) is 1986, irrespective of the choice of model used to estimate the fingerprint. The $x$ axis represents the last year of L-length linear trend in the signal estimate (with 1950 as the first year).

## c. Sensitivity testing

In this section, we examine the sensitivity of our D&A results to a range of different processing choices. We find that the ANTH fingerprint can be positively identified in all four observed temperature indices, irrespective of the model used to estimate the fingerprint (two options) or the noise (two options), the method used to downscale the PCM fingerprint (two options), the end points values, and the applied areal weighting (two options; see Fig. 10). The weighting option, which has not been discussed previously, involves either unit weighting or appropriate areal weighting of the nine mountain regions.

In all 48 cases, the detection time for an anthropogenic fingerprint occurs as early as 1986 and no later than 1994. The $S/N_{50}$ results are sensitive to the noise configuration. In Fig. 10, all symbols are located to the right of the line of equal $S/N$ ratio, indicating that $S/N_{50}$ (as well as detection time) is always enhanced when the CCSM3-CTL run is used for assessing statistical significance. This systematic difference in $S/N_{50}$ ratios is due to the larger decadal variability in the PCM-CTL run (Pierce et al. 2008). The sensitivity of $S/N_{50}$ values to the choice of downscaling method and model used to estimate the fingerprint is much smaller than the sensitivity to the noise configuration (Fig. 10).

For the $T_{min}$ results with no areal weighting, the leading signal and noise modes show comparatively little spatial structure (Figs. 8a,b). Because of this, the $T_{min}$ $S/N_{50}$ results are relatively insensitive to the area-

weighting option. In the FD and DD > 0 cases, however, the dominant signal and noise modes have more pronounced spatial structures (Figs. 8c,d), and hence $S/N_{50}$ values are more sensitive to the applied weighting (Fig. 10). Interestingly, $S/N_{50}$ values obtained for $T_{min}$ are higher than those obtained for $T_{max}$, irrespective of the various processing options. The results are consistent with the detection results obtained by Bonfils et al. (2008) over the domain of California.

We also considered whether our results were sensitive to the choice of later start dates (1955, 1960, and 1965). For a 1955 start date, anthropogenic effects on temperature are consistently identifiable, independent of processing choices (see Fig. 10b). For later start dates, however (1960 and 1965), consistent detection is not achieved. This is primarily because of an increase in noise and degradation in $S/N$ ratios when trends are calculated over shorter periods of time (see, e.g., Santer et al. 2007).

## 5. Discussion and conclusions

In this work we have assessed whether observed warming in the western United States over the period 1950–99 is both outside the range expected because of natural internal climate variability and consistent with the expected effects of anthropogenic climate forcing. This was done with a model-estimated anthropogenic "fingerprint" of warming that includes the effects of well-mixed greenhouse gases, ozone, and direct and
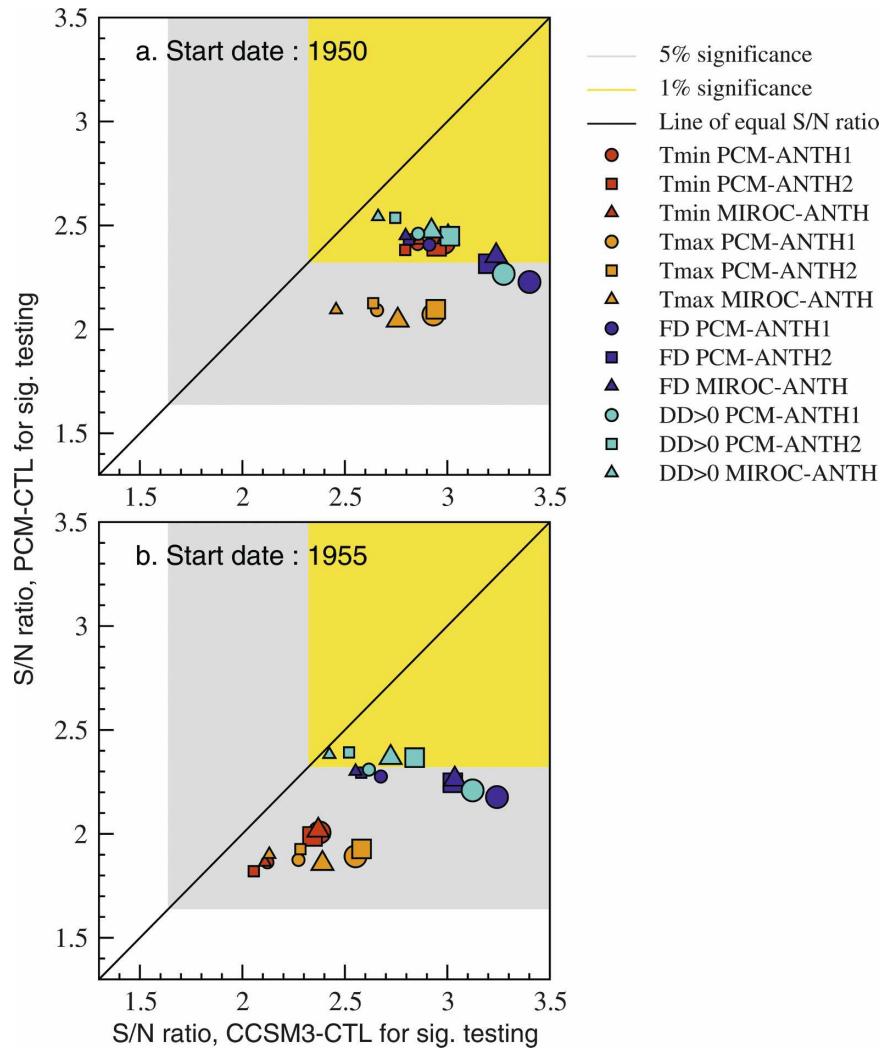
FIG. 10. Sensitivity of $S/N_{50}$ to the variable (color), the choice of the model used to estimate the fingerprint (symbol), the method to downscale the fingerprint (circle versus square), the areal-weighting option (size), and the start year: (a) 1950; (b) 1955. CCSM3-CTL (PCM-CTL) is used for assessing statistical significance in $x$ axis ($y$ axis).

some indirect aerosol effects. We examined variates of importance to the hydrological cycle in this generally arid region, including JFM $T_{min}$, the number of frost days, and the degree-days above freezing.

Overall, the anthropogenic fingerprint is robustly identifiable in all observed hydrologically related temperature records. Natural variability in the western climate system has been ruled out as the major cause of the changes. Solar variability and volcanic forcings are also ruled out by the analysis. Bonfils et al. (2008) found that late winter/early spring rise in $T_{min}$ and $T_{max}$ recently observed in California were too rapid to be explained by natural climate variability alone, and required one or more external forcing agents to be explained. We not only reach the same conclusion for the

mountainous regions of western United States in this study, we can also confidently attribute those changes to human activities.

The model-predicted anthropogenic fingerprints are statistically detectable in the DD > 0 observational field. Similar results are found using DDx > 0, another proxy of temperature-driven snowmelt (Fig. 3). These results are in agreement with the work of Pierce et al. (2008), which demonstrates that 1) observed reductions in SWE/$P$ (over the same nine regions) cannot be fully explained by climate noise or natural forcings, 2) half the SWE/$P$ reduction arises from human effects, and 3) the reduction in snowpack is driven by increasing temperature, not by decreasing precipitation. Therefore, historical changes in the atmospheric composition re-

sult in a human-induced warming over western U.S. mountainous regions, itself responsible for the observed reductions of snowpack (Pierce et al. 2008) and shifts in the timing of the streamflow (Hidalgo et al. 2008, manuscript submitted to *J. Climate*). Models of climate change unanimously project an acceleration of the warming in the western United States with temperature change projections of ~1°–3°C for 2050 and 2°–6°C by 2100 in California (Lobell et al. 2006). With such substantial warming, serious implications for water infrastructure and water supply sustainability can be expected over the western United States in the near future.

## REFERENCES

Alexander, M., and Coauthors, 2006: Extratropical atmosphere–ocean variability in CCSM3. *J. Climate,* **19,** 2496–2525.

Andreadis, K. M., and D. P. Lettenmaier, 2006: Assimilating remotely sensed snow observations into a macroscale hydrology model. *Adv. Water Res.,* **29,** 872–886.

Bala, G., and Coauthors, 2008: Evaluation of a CCSM3 simulation with a finite volume dynamical core for the atmosphere at 1° latitude × 1.25° longitude resolution. *J. Climate,* **21,** 1467–1486.

Barnett, T. P., and Coauthors, 2008: Human-induced changes in the hydrology of the western United States. *Science,* **319,** 1080–1083.

Bonfils, C., and D. Lobell, 2007: Empirical evidence for a recent slowdown in irrigation-induced cooling. *Proc. Natl. Acad. Sci. USA,* **104,** 13 582–13 587.

——, P. Duffy, B. Santer, T. Wigley, D. B. Lobell, T. J. Phillips, and C. Doutriaux, 2008: Identification of external influences on temperatures in California. *Climatic Change,* **87,** 43–55.

Cayan, D. R., 1996: Interannual climate variability and snowpack in the western United States. *J. Climate,* **9,** 928–948.

—— K. T. Redmond, and L. G. Riddle, 1999: ENSO and hydrologic extremes in the western United States. *J. Climate,* **12,** 2881–2893.

——, S. A. Kammerdiener, M. D. Dettinger, J. M. Caprio, and D. H. Peterson, 2001: Changes in the onset of spring in the western United States. *Bull. Amer. Meteor. Soc.,* **82,** 399–415.

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.

Christidis, N., P. A. Stott, S. Brown, D. J. Karoly, and J. Caesar, 2007: Human contribution to the lengthening of the growing season during 1950–99. *J. Climate,* **20,** 5441–5454.

Collins, W. D., and Coauthors, 2006: The Community Climate System Model version 3 (CCSM3). *J. Climate,* **19,** 2122–2143.

DeGaetano, A. T., and R. J. Allen, 2002: Trends in 20th century temperature extremes across the United States. *J. Climate,* **15,** 3188–3205.

Dettinger, M. D., and D. R. Cayan, 1995: Large-scale atmospheric forcing of recent trends toward early snowmelt runoff in California. *J. Climate,* **8,** 606–623.

Gedney, N., P. M. Cox, R. A. Betts, O. Boucher, C. Huntingford, and P. A. Stott, 2006: Detection of a direct carbon dioxide effect in continental river runoff records. *Nature,* **439,** 835–838.

Gillett, N. P., and A. J. Weaver, 2004: Detecting the effect of climate change on Canadian forest fires. *Geophys. Res. Lett.,* **31,** L18211, doi:10.1029/2004GL020876.

——, R. J. Allan, and T. J. Ansell, 2005: Detection of external influence on sea level pressure with a multi-model ensemble. *Geophys. Res. Lett.,* **32,** L19714, doi:10.1029/2005GL023640.

Gleick, P. H., and Coauthors, 2000: Water: The potential consequences of climate variability and change for the water resources of the United States. The Report of the Water Sector Assessment Team of the National Assessment of the Potential Consequences of Climate Variability and Change, 160 pp.

Groisman, P. Ya., T. R. Karl, R. W. Knight, and G. L. Stenchikov, 1994: Changes of snow cover, temperature, and radiative heat balance over the Northern Hemisphere. *J. Climate,* **7,** 1633–1656.

Hamlet, A. F., and D. P. Lettenmaier, 2005: Production of temporally consistent gridded precipitation and temperature fields for the continental United States. *J. Hydrometeor.,* **6,** 330–336.

Hasselmann, K., 1979: On the signal-to-noise problem in atmospheric response studies. *Meteorology of Tropical Oceans,* D. B. Shaw, Ed., Royal Meteorological Society of London, 251–259.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. *Proc. Natl. Acad. Sci. USA,* **101,** 12 422–12 427.

Hidalgo, H. G., M. D. Dettinger, and D. R. Cayan, 2008: Downscaling with constructed analogues: Daily precipitation and

temperature fields over the United States. California Energy Commission Rep. CEC-500-2007-123, 62 pp.

——, and Coauthors, 2008: Detection and attribution of climate change in streamflow timing of the western United States. *J. Climate,* submitted.

K-1 model developers, 2004: K-1 Coupled GCM (MIROC) description. H. Hasumi and S. Emori, Eds., Center for Climate System Research, University of Tokyo, 38 pp.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Karl, T. R., C. N. Williams Jr., P. J. Young, and W. M. Wendland, 1986: A model to estimate the time of observation bias associated with monthly mean maximum, minimum, and mean temperature for the United States. *J. Climate Appl. Meteor.,* **25,** 145–160.

——, ——, F. T. Quinlan, and T. A. Boden, 1990: United States Historical Climatology Network (HCN) serial temperature and precipitation data. Environmental Science Division, Publication 3404, Carbon Dioxide Information and Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, 389 pp.

Karoly, D. J., and Q. Wu, 2005: Detection of regional surface temperature trends. *J. Climate,* **18,** 4337–4343.

——, K. Braganza, P. A. Stott, J. M. Arblaster, G. A. Meehl, A. J. Broccoli, and K. W. Dixon, 2003: Detection of a human influence on North American climate. *Science,* **302,** 1200–1203.

Knowles, N., M. D. Dettinger, and D. R. Cayan, 2006: Trends in snowfall versus rainfall in the western United States. *J. Climate,* **19,** 4545–4559.

Lambert, F. H., N. P. Gillett, D. A. Stone, and C. Huntingford, 2005: Attribution studies of observed land precipitation changes with nine coupled models. *Geophys. Res. Lett.,* **32,** L18704, doi:10.1029/2005GL023654.

Liang, X., D. P. Lettenmaier, E. F. Wood, and S. J. Burges, 1994: A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *J. Geophys. Res.,* **99,** 14 415–14 428.

Linsley, R. K., Jr., 1943: A simple procedure for the day-to-day forecasting of runoff from snowmelt. *Trans. Amer. Geophys. Union,* **24,** 719–736.

Lobell, D., C. Field, K. Nicholas-Cahill, and C. Bonfils, 2006: Impacts of future climate change on California perennial crop yields: Model projections with climate and crop uncertainties. *Agric. For. Meteor.,* **141,** 208–218.

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. *Bull. Amer. Meteor. Soc.,* **78,** 1069–1079.

Maurer, E. P., and H. G. Hidalgo, 2007: Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.,* **12,** 551–563.

——, I. T. Stewart, C. Bonfils, P. B. Duffy, and D. Cayan, 2007: Detection, attribution, and sensitivity of trends toward earlier streamflow in the Sierra Nevada. *J. Geophys. Res.,* **112,** D11118, doi:10.1029/2006JD008088.

Meehl, G. A., and A. Hu, 2006: Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multidecadal Pacific sea surface temperature anomalies. *J. Climate,* **19,** 1605–1623.

——, ——, and B. D. Santer, 2009: The mid-1970s climate shift in the Pacific and the relative roles of forced versus inherent decadal variability. *J. Climate,* in press.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, and R. J. Stouffer, 2008: Stationarity is dead: Whither water management? *Science,* **319,** 573–574.

Mote, P. W., 2006: Climate-driven variability and trends in mountain snowpack in western North America. *J. Climate,* **19,** 6209–6220.

——Mote, P. W., A. F. Hamlet, M. P. Clark, and D. P. Lettenmaier, 2005: Declining mountain snowpack in western North America. *Bull. Amer. Meteor. Soc.,* **86,** 39–49.

Nozawa, T., T. Nagashima, T. Ogura, T. Yokohara, N. Okada, and H. Shiogama, 2007: Climate change simulations with a coupled ocean–atmosphere GCM called the Model for Interdisciplinary Research on Climate: MIROC. CGER's Supercomputer Monograph Rep. Vol. 12, Center for Global Environmental Research, National Institute for Environmental Studies, Japan, 93 pp.

Pierce, D. W., and Coauthors, 2008: Attribution of declining western U.S. snowpack to human effects. *J. Climate,* **21,** 6425–6444.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, 1992: Kolmogorov-Smirnov test. *Numerical Recipes in FORTRAN: The Art of Scientific Computing.* 2nd ed. Cambridge University Press, 617–620.

Santer, B. D., and T. M. L. Wigley, 2008: Progress in detection and attribution research. *Climatic Change Science and Policy,* S. H. Schneider, A. Rosencranz, and M. D. Mastrandrea, Eds., Island Press, in press.

——, U. Mikolajewicz, W. Brüggemann, U. Cubasch, K. Hasselmann, H. Höck, E. Maier-Reimer, and T. M. L. Wigley, 1995: Ocean variability and its influence on the detectability of greenhouse warming signals. *J. Geophys. Res.,* **100,** 10 693–10 725.

——, T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor, 2000a: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res.,* **105,** 7337–7356.

——, and Coauthors, 2000b: Interpreting differential temperature trends at the surface and in the lower troposphere. *Science,* **287,** 1227–1232.

——, and Coauthors, 2007: Identification of human-induced changes in atmospheric moisture content. *Proc. Natl. Acad. Sci. USA,* **104,** 15 248–15 253.

Shindell, D. T., G. A. Schmidt, R. L. Miller, and D. Rind, 2001: Northern Hemisphere winter climate response to greenhouse gas, ozone, solar, and volcanic forcing. *J. Geophys. Res.,* **106,** 7193–7210.

Shiogama, H., M. Watanabe, M. Kimoto, and T. Nozawa, 2005: Anthropogenic and natural forcing impacts on ENSO-like decadal variability during the second half of the 20th century. *Geophys. Res. Lett.,* **32,** L21714, doi:10.1029/2005GL023871.

Smith, S. R., and J. J. O'Brien, 2001: Regional snowfall distributions associated with ENSO: Implications for seasonal forecasting. *Bull. Amer. Meteor. Soc.,* **82,** 1179–1191.

Stewart, I. T., D. R. Cayan, and M. D. Dettinger, 2005: Changes toward earlier streamflow timing across western North America. *J. Climate,* **18,** 1136–1155.

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.,* **30,** 1728, doi:10.1029/2003GL017324.

Vose, R. S., C. N. Williams, Jr., T. C. Peterson, T. R. Karl, and

D. R. Easterling, 2003: An evaluation of the time of observation bias adjustment in the U.S. Historical Climatology Network. *Geophys. Res. Lett.,* **30,** 2046, doi:10.1029/2003GL018111.

Washington, W. M., and Coauthors, 2000: Parallel Climate Model (PCM) control and transient simulations. *Climate Dyn.,* **16,** 755–774.

Westerling, A. L., H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, 2006: Warming and earlier spring increase western U.S. forest wildfire activity. *Science,* **313,** 940–943.

Willett, K. M., N. P. Gillett, P. D. Jones, and P. W. Thorne, 2007: Attribution of observed surface humidity changes to human influence. *Nature,* **449,** 710–712.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change,* **62,** 189–216.

Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature,* **448,** 461–465.

Zwiers, F. W., and X. Zhang, 2003: Towards regional climate change detection. *J. Climate,* **16,** 793–797.