SPECIAL

**Regional Climate Change**

COLLECTION

# A Hybrid Dynamical–Statistical Downscaling Technique. Part I: Development and Validation of the Technique

Daniel B. Walton, Fengpeng Sun, Alex Hall, and Scott Capps

*Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California*

(Manuscript received 11 March 2014, in final form 2 March 2015)

## ABSTRACT

In this study (Part I), the mid-twenty-first-century surface air temperature increase in the entire CMIP5 ensemble is downscaled to very high resolution (2 km) over the Los Angeles region, using a new hybrid dynamical–statistical technique. This technique combines the ability of dynamical downscaling to capture finescale dynamics with the computational savings of a statistical model to downscale multiple GCMs. First, dynamical downscaling is applied to five GCMs. Guided by an understanding of the underlying local dynamics, a simple statistical model is built relating the GCM input and the dynamically downscaled output. This statistical model is used to approximate the warming patterns of the remaining GCMs, as if they had been dynamically downscaled. The full 32-member ensemble allows for robust estimates of the most likely warming and uncertainty resulting from intermodel differences. The warming averaged over the region has an ensemble mean of 2.3°C, with a 95% confidence interval ranging from 1.0° to 3.6°C. Inland and high elevation areas warm more than coastal areas year round, and by as much as 60% in the summer months. A comparison to other common statistical downscaling techniques shows that the hybrid method produces similar regional-mean warming outcomes but demonstrates considerable improvement in capturing the spatial details. Additionally, this hybrid technique incorporates an understanding of the physical mechanisms shaping the region's warming patterns, enhancing the credibility of the final results.

## 1. Introduction

To make informed adaptation and mitigation decisions, policymakers and other stakeholders need future climate projections at the regional scale that provide robust information about most likely outcomes and uncertainty estimates (Mearns et al. 1999; Leung et al. 2003; Schiermeier 2010; Kerr 2011). The main tools available for such projections are ensembles of global climate models (GCMs). However, GCMs have grid box scales from 1° to 2.5° (~100–200 km), often too coarse to resolve important topographical features and mesoscale processes that govern local climate (Giorgi and Mearns 1991; Leung et al. 2003; Caldwell et al. 2009; Arritt and Rummukainen 2011). The inability of GCMs to provide robust predictions at scales small enough for stakeholder purposes has motivated numerous efforts to regionalize GCM climate change signals through a variety of downscaling methods (e.g., Giorgi et al. 1994; Snyder et al. 2002; Timbal et al. 2003; Hayhoe et al. 2004; Leung et al. 2004; Tebaldi et al. 2005; Duffy et al. 2006; Cabré et al. 2010; Salathé et al. 2010; Pierce et al. 2013). The aim of this study is to develop downscaling techniques to recover the full complement of warming signals in the greater Los Angeles region associated with the multimodel ensemble from phase 5 of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012; Table 1) of the World Climate Research Programme.

Regional downscaling attempts have been met with significant criticism (e.g., Schiermeier 2010; Kerr 2011, 2013). One major critique is that the downscaled output is constrained by the limitations of the GCM input. By itself, any single GCM may give a misleading picture of the true state of knowledge about climate change, including in the region of interest. Results from downscaling this single GCM will likewise be misleading. Furthermore, the high resolution and realistic appearance of the downscaled results may give a false impression of accuracy. This perception of accuracy at the regional scale is especially problematic if a very small

*Corresponding author address:* Daniel B. Walton, 7229 Math Sciences Building, 405 Hilgard Ave., Los Angeles, CA 90095.
E-mail: waltond@atmos.ucla.edu

TABLE 1. Details of the WCRP CMIP5 global climate models used in this study. Check marks indicate which scenarios are used. Five models were dynamically downscaled (shown in boldface). All available models are statistically downscaled using the hybrid method. Expansions of model names are available online at http://www.ametsoc.org/PubsAcronymList.

| Model | Country | Institute | RCP2.6 | RCP8.5 |
|---|---|---|---|---|
| ACCESS1.0 | Australia | Commonwealth Scientific and Industrial Research Organisation | | ✔ |
| ACCESS1.3 | Australia | Commonwealth Scientific and Industrial Research Organisation | | ✔ |
| BCC_CSM1.1 | China | Beijing Climate Center, China Meteorological Administration | ✔ | ✔ |
| BNU-ESM | China | College of Global Change and Earth System Science, Beijing Normal University | ✔ | ✔ |
| CanESM2 | Canada | Canadian Centre for Climate Modelling and Analysis | ✔ | ✔ |
| **CCSM4** | **United States** | **National Center for Atmospheric Research** | ✔ | ✔ |
| CESM1(BGC) | United States | National Science Foundation, U.S. Department of Energy, and National Center for Atmospheric Research | | ✔ |
| CESM1(CAM5) | United States | National Science Foundation, U.S. Department of Energy, and National Center for Atmospheric Research | ✔ | ✔ |
| CESM1(WACCM) | United States | National Science Foundation, U.S. Department of Energy, and National Center for Atmospheric Research | | ✔ |
| CMCC-CM | Italy | Centro Euro-Mediterraneo per I Cambiamenti Climatici | | ✔ |
| **CNRM-CM5** | **France** | **Centre National de Recherches Météorologiques** | ✔ | ✔ |
| CSIRO Mk3.6.0 | Australia | Commonwealth Scientific and Industrial Research Organisation | ✔ | ✔ |
| EC-EARTH | Europe | EC-Earth Consortium | ✔ | ✔ |
| FGOALS-s2 | China | LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences | ✔ | ✔ |
| FIO-ESM | China | The First Institute of Oceanography | | ✔ |
| **GFDL CM3** | **United States** | **NOAA/Geophysical Fluid Dynamics Laboratory** | ✔ | ✔ |
| GFDL-ESM2M | United States | NOAA/Geophysical Fluid Dynamics Laboratory | ✔ | ✔ |
| GFDL-ESM2G | United States | NOAA/Geophysical Fluid Dynamics Laboratory | ✔ | ✔ |
| GISS-E2-H | United States | NASA Goddard Institute for Space Studies | ✔ | ✔ |
| GISS-E2-R | United States | NASA Goddard Institute for Space Studies | ✔ | ✔ |
| HadGEM2-AO | United Kingdom | Met Office Hadley Centre | ✔ | ✔ |
| HadGEM2-CC | United Kingdom | Met Office Hadley Centre | | ✔ |
| HadGEM2-ES | United Kingdom | Met Office Hadley Centre | ✔ | ✔ |
| INM CM4 | Russia | Institute of Numerical Mathematics | | ✔ |
| IPSL-CM5A-LR | France | L'Institut Pierre-Simon Laplace | ✔ | ✔ |
| IPSL-CM5A-MR | France | L'Institut Pierre-Simon Laplace | ✔ | ✔ |
| MIROC-ESM | Japan | AORI (University of Tokyo), NIES, and JAMESTEC | ✔ | ✔ |
| **MIROC-ESM-CHEM** | **Japan** | **Atmosphere and Ocean Research Institute (AORI, University of Tokyo), National Institute for Environmental Studies (NIES), and JAMESTEC** | ✔ | ✔ |
| MIROC5 | Japan | AORI (University of Tokyo), NIES, and JAMESTEC | ✔ | ✔ |
| **MPI-ESM-LR** | **Germany** | **Max Planck Institute for Meteorology** | ✔ | ✔ |
| MRI-CGCM3 | Japan | Meteorological Research Institute | ✔ | ✔ |
| NorESM1-M | Norway | Norwegian Climate Centre | ✔ | ✔ |

number of GCMs are downscaled, since the uncertainty is dramatically undersampled. In this case, the downscaled output may not reflect the most likely climate outcomes in the region, and it certainly does not provide information about how the uncertainty associated with the GCM ensemble manifests itself at the regional scale. Typically, previous studies have downscaled only two global models (e.g., Hayhoe et al. 2004; Duffy et al. 2006; Cayan et al. 2008; Salathé et al. 2010). This is too small an ensemble to obtain meaningful statistics about the most likely (ensemble mean) warming and uncertainty (intermodel spread). Instead, information from a larger ensemble is preferred (Giorgi and Mearns 2002; Kharin and Zwiers 2002). The CMIP3 and CMIP5 ensembles

(Meehl et al. 2007; Taylor et al. 2012), with a few dozen ensemble members, are usually seen as large enough to compute a meaningful ensemble mean and span the climate change uncertainty space.

While downscaling of a large ensemble is desirable to compute most likely outcomes and fully characterize uncertainty, this can be impractical because of the high computational cost. Dynamical downscaling, in particular, is an expensive technique, and most studies that perform it have only applied it to a few global models. For example, Duffy et al. (2006) downscaled PCM and HadCM2 over the western United States, and Pierce et al. (2013) downscaled GFDL CM2.1 and CCSM3 over California. There are other examples of dynamical downscaling of

multiple GCMs, such as the Coordinated Regional Downscaling Experiment (CORDEX; Giorgi et al. 2009), but these are very large undertakings that require coordination of multiple research groups. Furthermore, they tend to span large geographic areas at lower resolutions (roughly 50 km) than needed for the region of interest here. Areas of intense topography and complex coastlines typically need a model resolution finer than 10–15 km (Mass et al. 2002). The greater Los Angeles region contains minor mountain complexes, such as the Santa Monica mountains, that have a significant role in shaping local climate gradients. These mountain complexes have spatial scales of just a few kilometers, so even higher resolution, with correspondingly higher computational costs, would be needed here. Thus, for the purposes of this study, dynamical downscaling alone is an impractical answer to the need for multimodel downscaling.

Because of its much lower computational cost, statistical downscaling is almost always used for multimodel downscaling (e.g., Giorgi et al. 2001; Tebaldi et al. 2005; Pierce et al. 2013). Statistical downscaling relies on empirical mathematical relationships to go from large-scale predictors to finescale predictands. These relationships are often much faster to apply than dynamical downscaling, which makes them ideal for downscaling large ensembles of GCMs for multiple time periods or scenarios. However, they are subject to the stationarity assumption that the relationship between the predictors and predictands continues to hold, even in a changed climate (Wilby and Wigley 1997). Although statistical models are valuable tools for downscaling multimodel ensembles, they do not produce a full complement of variables like dynamical downscaling. Dynamical downscaling allows for an investigation of the physics that underlie the local climate response, which leads to enhanced credibility. Furthermore, the physical realism and the ability to explicitly simulate complex local processes potentially allows dynamical downscaling to capture important finescale changes in climate that otherwise might not be included (e.g., Caldwell et al. 2009; Salathé et al. 2008, 2010; Arritt and Rummukainen 2011; Pierce et al. 2013). For example, Salathé et al. (2008) found that December–February warming in the mountains of the Pacific Northwest could be approximately twice the GCM value at locations experiencing snow–albedo feedback. Snow–albedo feedback is also important in the Southern California domain considered, with the presence of a number of mountain ranges with wintertime snow coverage, including the San Gabriel and San Bernardino Mountain ranges. Pierce et al. (2013) found that when a pair of GCMs was dynamically downscaled, the average difference in the annual warming between the Southern

California mountains and coast was twice that of two common statistical downscaling techniques, bias correction with spatial disaggregation (BCSD) and bias correction with constructed analogs (BCCA). This suggests that statistical downscaling alone may be insufficient in order to capture sharp gradients in temperature change in our region of interest.

Here we provide a hybrid downscaling technique that allows us to fully sample the GCM ensemble with the physical credibility of dynamical downscaling but without the heavy computational burden of dynamically downscaling every GCM. In this technique, dynamical downscaling is first performed on five GCMs. Then, the results from dynamical downscaling are used to identify the common finescale warming patterns and how they relate to the major GCM-scale warming features. Based on these relationships, a simple statistical model is built to mimic the warming patterns produced by the dynamical model. In the statistical model, the common finescale patterns are dialed up or down to reflect the regional-scale warming found in the particular GCM being downscaled. While scaling of regional climate change patterns has been around since Mitchell et al. (1990) and Santer et al. (1990), the scaling has primarily been relative to the global-mean warming and only for a single GCM (e.g., Cabré et al. 2010). The statistical model described here is more versatile because 1) it works for any GCM, not just those dynamically downscaled; 2) the downscaled warming is dependent on the GCM's regional-mean warming characteristics, not the global-mean warming; and 3) this dependence is allowed multiple degrees of freedom, based on the physical processes at play in this particular region.

To build a statistical model that mimics dynamical downscaling, the physical mechanisms underpinning the regional climate change pattern must be understood. This process-oriented approach addresses another concern about regional downscaling, namely that it is difficult to determine if the regional climate change patterns are credible even if they appear realistic and visually appealing, because the dynamics underpinning them are unclear, undiagnosed, or unknown.

After the statistical model undergoes a rigorous cross-validation procedure and assessment of value added, it is applied to generate the warming patterns for the remaining GCMs in the CMIP5 ensemble. These statistically generated warming patterns represent our best estimate of what the warming would be if dynamical downscaling had been performed on these remaining GCMs. The efficiency of the hybrid technique allows us to downscale multiple emission scenarios and multiple time periods. In this study (Part I), we downscale 32 GCMs for the midcentury period (2041–60) under representative concentration pathway 8.5 (RCP8.5). In Sun
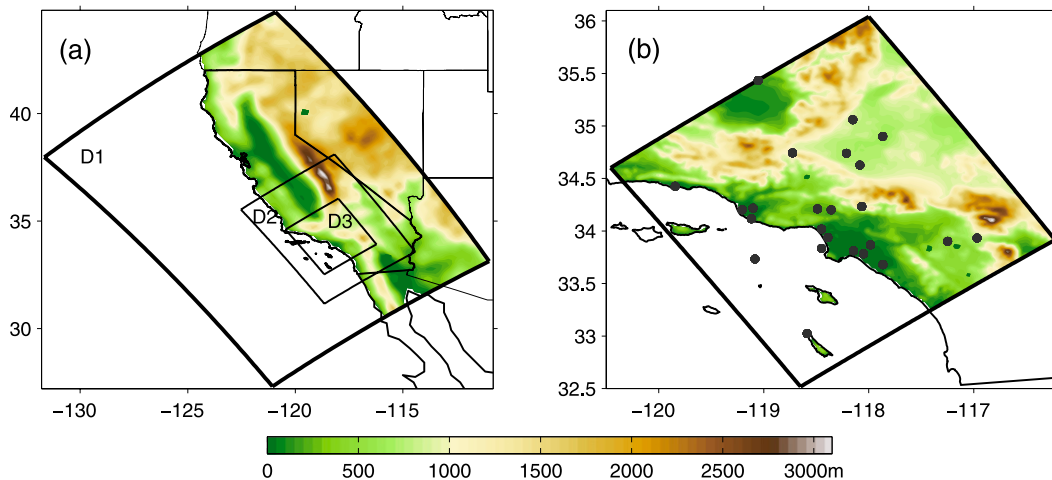
FIG. 1. (a) Model setup with three nested WRF domains (D1, D2, and D3) at resolutions of 18, 6, and 2 km (area sizes $1188 \times 1566$, $486 \times 612$, and $276 \times 288\,km^2$, respectively). Topography (m) is shown in color. (b) The innermost domain (D3) of the regional simulation, with 2-km resolution. Black dots indicate the locations of the 24 stations used for surface air temperature validation.

et al. (2015, hereinafter Part II), we expand this analysis by downscaling the full ensemble for RCP8.5 end of century (2081–2100) and representative concentration pathway 2.6 (RCP2.6) midcentury and end of century.

## 2. Dynamical downscaling

### a. Model configuration

Dynamical downscaling was performed using the Advanced Research Weather Research and Forecasting Model version 3.2 (hereinafter WRF; Skamarock et al. 2008). WRF has been successfully applied to the California region in previous work (e.g., Caldwell et al. 2009; Pierce et al. 2013). For this study, we optimized it for the California region with sensitivity experiments using various parameterizations, paying particular attention to the model's ability to simulate low cloud in the marine boundary layer off the California coast. The following parameterization choices were made: the Kain–Fritsch (new Eta) cumulus scheme (Kain 2004), the Yonsei University boundary layer scheme (Hong and Lim 2006), the Purdue–Lin microphysics scheme (Lin et al. 1983), the Rapid Radiative Transfer Model longwave radiation (Mlawer et al. 1997), and the Dudhia shortwave radiation schemes (Dudhia 1989). The Noah land surface model (Chen and Dudhia 2001) was used to simulate land surface processes including vegetation, soil, snowpack, and exchange of energy, momentum, and moisture between the land and atmosphere.

The three one-way nested domains for the simulations are shown in Fig. 1. The outermost domain covers the entire state of California and the adjacent ocean at a horizontal resolution of 18 km, the middle domain covers roughly the southern half of the state at a horizontal resolution of 6 km, and the innermost domain encompasses Los Angeles County and surrounding regions at a horizontal resolution of 2 km. In each domain, all variables in grid cells closer than five cells from the lateral boundary in the horizontal were relaxed toward the corresponding values at the lateral boundaries. This procedure ensures smooth transitions from one domain to another. Each domain has 43 sigma levels in the vertical. To provide a better representation of surface and boundary layer processes, the model's vertical resolution is enhanced near the surface, with 30 sigma levels below 3 km.

### b. Baseline simulation and validation

We followed a previously established dynamical downscaling method [see Rasmussen et al. (2011) and references therein; see also Sato et al. (2007) and Kawase et al. (2009)] in designing our baseline and future simulations. Under this approach, only a single baseline simulation is performed. The purpose of this baseline simulation is twofold: 1) to validate the model's ability to simulate regional climate and 2) to provide a baseline climate state against which the future climate simulations can be compared, to quantify climate change. This simulation is a dynamical downscaling of the National Centers for Environmental Prediction North American Regional Reanalysis (NARR; Mesinger et al. 2006) over the period from September 1981 to August 2001. This dataset has 32-km resolution and provides lateral boundary conditions at the outer boundaries of the outermost domain (Fig. 1). It also provides surface boundary
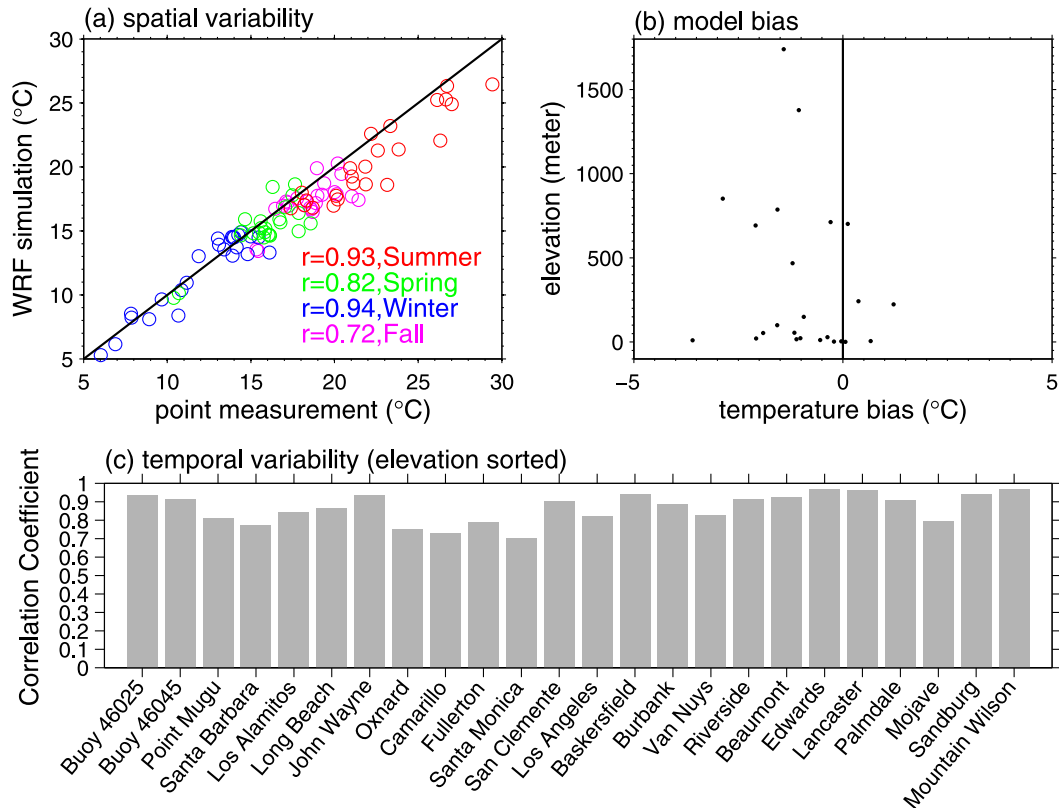
FIG. 2. Comparison of the WRF baseline simulation against a network of 24 stations, for the period 1995–2001. (a) Simulated vs observed seasonal average climatological temperatures, for each station. Spatial correlations reported for each season (colors). (b) WRF minus observed annual-mean temperature at each station, scattered against elevation. (c) Correlations at each station between the time series of simulated and observed monthly temperature anomalies. Anomalies are relative to the monthly climatology. Stations are listed in increasing order of elevation.

conditions over the ocean (i.e., sea surface temperature) in each of the three domains. The simulation is designed to reconstruct the regional weather and climate variations that occurred in the innermost domain at a 2-km resolution during this time period. The model was re-initialized each year in August and run from September to August. Because each year was initialized separately, the time period could be divided into 1-yr runs performed in parallel.

The regional model's ability to reproduce climate variations during the baseline period was assessed by comparing the output from the baseline climate simulation to two types of data: 1) point measurements from a network of 24 weather stations and buoys and 2) a spatially complete, observationally based gridded product. The point measurements are quality-controlled, hourly, near-surface meteorological observations obtained from the National Climatic Data Center (NCDC; http://www. ncdc.noaa.gov/). The stations are located in a variety of elevations and distances from the coast, and are numerous enough to provide a sampling of the range of

temperatures seen across the region (Fig. 1). However, both the length and completeness of observational temperature records vary by location. Most locations have reasonably complete records after 1995, so validation is performed over the 1995–2001 period.

First, we check the realism of the spatial patterns seen in surface air temperature climatology against the station measurements. Spatial patterns simulated by the model are highly consistent with observations, as indicated by high correlations between observed and simulated temperatures within each season (Fig. 2a). This confirms that for each season, the model simulates spatial variations in climatological temperature reasonably well. The spatial pattern is particularly well represented in summer and winter (correlation of $r >$ 0.9 in both seasons), although the model exhibits a slight cold bias in the summer. During the transition seasons, the model and observed spatial patterns are still in broad agreement, with correlations greater than 0.7. In Fig. 2b, the model's annual-mean bias relative to observations is scattered versus elevation. Overall,
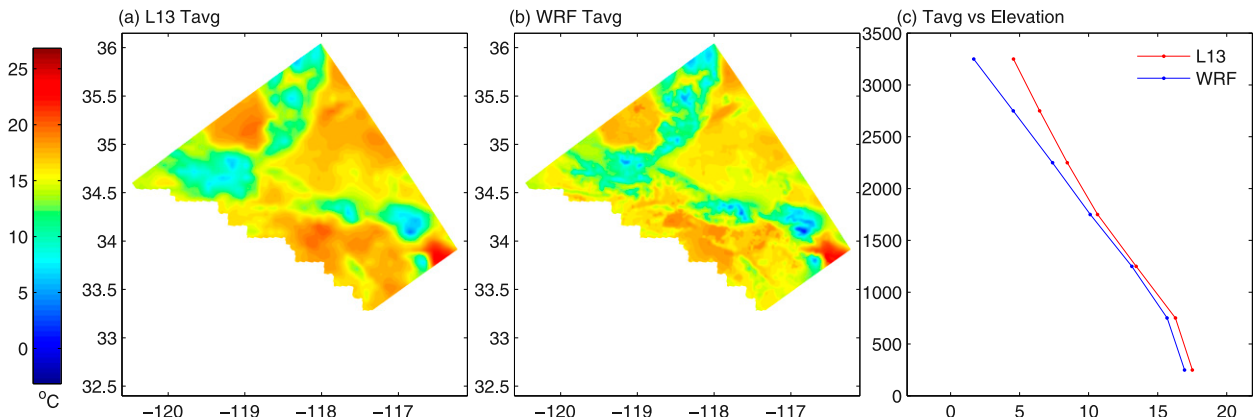
FIG. 3. Annual-mean temperature climatology (°C) for the 1981–2001 period for (a) the L13 observationally based gridded product, linearly interpolated from its native 1/16° resolution to the 2-km resolution innermost WRF grid, and (b) the WRF baseline simulation. (c) Temperature averaged over each 500-m elevation band for L13 (red) and WRF (blue).

WRF tends to have a cold bias, averaging roughly 1°C over the 24 stations, and there is no clear trend in the bias with height. We also assessed the model's ability to simulate temporal variability on monthly time scales and longer. At each of the 24 locations, the correlation was computed between the observed and modeled time series of monthly mean temperature anomalies, after first removing a composite seasonal cycle (Fig. 2c). Temporal variability is very well simulated by the model, with high correlations at all locations.

The WRF baseline simulation was compared to an observationally based gridded product in order to ascertain its skill in reproducing temperatures in the rest of the domain, where no station observations are available. The gridded product used here is from Livneh et al. (2013, hereinafter L13). The L13 dataset includes daily maximum temperature (Tmax) and minimum temperature (Tmin) that are averaged here to produce climatological values (Tavg). L13 has a native resolution of 1/16° (about 6 km at this latitude), and it has been linearly interpolated down to the 2-km resolution WRF domain for comparison purposes. This gridded product also has the advantage of being available for the whole baseline simulation. The 1981–2001 L13 and WRF baseline temperature climatologies (Figs. 3a,b) both show large variations of up to 20°C over the complex topography in the domain. WRF simulates colder temperatures in the mountain peaks, presumably because it has finer resolution (leading to higher and colder peaks) and because it explicitly simulates snow at the high elevations. The L13 dataset only incorporates a few stations located in snow-covered areas in this domain. Instead, it relies on temperatures at lower elevation locations and uses an assumed lapse rate to calculate the temperatures at higher elevations. This, combined with the elevation mismatch, could explain why WRF is

colder than L13 at elevations above about 1500 m (Fig. 3c). Overall, WRF captures the complex spatial variations in the temperature climatology found in the L13 data, which leads us to believe that it could also capture the spatial variations found in the climate change patterns.

Figure 2 demonstrates that the model gives approximately the right spatial and temporal variations in surface air temperature at specific point locations where station observations are available. Figure 3 shows that the model is also producing the correct temperatures in the rest of the region, where direct observations are absent. When taken in combination, these figures give high confidence that when it comes to surface air temperature, the model can provide realistic downscaling of the regional pattern implicit in the coarser-resolution forcing dataset. Thus, the dynamically downscaled climate change patterns presented below are very likely a true reflection of how the atmosphere's dynamics would distribute the warming across the region if climate change signals seen in the global models occurred in the real world.

c. Future simulations

Following the approach of Rasmussen et al. (2011), we used the same model configuration as in the baseline simulation to perform a second set of dynamical downscaling experiments designed to simulate the regional climate state corresponding to the mid-twenty-first century. Five global climate models in the CMIP5 ensemble corresponding to this time period and the RCP8.5 emissions scenario were downscaled (see Table 1). The approach we are using provides important advantages including 1) a single baseline run that can be directly compared with observations to assess model skill, 2) computational savings by requiring only a single

baseline simulation, and 3) computational savings by requiring only short future simulations, since the same monthly perturbation is added to each year. Each future simulation is a dynamical downscaling of the NARR baseline reanalysis plus a perturbation: the difference in GCM climatology between the baseline and future periods. The perturbation contains the GCM climate change signals of interest, and is calculated by taking the differences between future and baseline monthly climatologies (2041–60 minus 1981–2000) for each GCM. All model variables are included in the calculation of the climate change signal (i.e., three-dimensional atmospheric variables such as temperature, relative humidity, zonal and meridional winds, and geopotential height and two-dimensional surface variables such as temperature, relative humidity, winds, and pressure). The monthly varying perturbations are linearly interpolated in time to match the NARR temporal resolution and then added to the NARR data corresponding to the baseline period (September 1981–August 2001). Because we downscaled the mean climate change signal in each GCM rather than the raw GCM data, we did not downscale changes in GCM variability. Thus, any future changes in variability in the regional simulations are solely the result of WRF's dynamical response to a mean change at the boundaries. In addition to imposing a mean climate change perturbation at the boundaries, $CO_2$ concentrations were also increased in WRF to match $CO_2$-equivalent radiative forcing in the RCP8.5 scenario.

We first downscaled CCSM4 for a 20-yr period and then performed sensitivity testing to see if it was necessary to downscale such a long period to recover the regional temperature change signal. (Using a shorter period when downscaling the other GCMs conserves scarce computational resources.) Because we perturbed each year in the future period with the same monthly varying change signal from CCSM4, we expected the warming patterns for each year to be relatively similar. In fact, the warming patterns were nearly identical each year: We could have dynamically downscaled only three years and recovered an average warming signal within 0.1°C of the 20-yr value. Therefore, the remaining four GCMs were only downscaled for three years. For each of these GCMs, the boundary conditions for the future run were created by adding the mean climate change signal (2041–60 minus 1981–2000) from the GCM to the 3-yr period of NARR corresponding to September 1998–August 2001.

### d. Warming patterns

The five future simulations were differenced with the baseline simulation to determine the high-resolution

monthly mean temperature changes. Figure 4 shows the warming averaged over the five dynamically downscaled GCMs. There are two prominent features in the warming patterns that can be understood through underlying physical processes. First, the warming is greater over land than ocean. This is true for all months, but the effect is particularly evident in the late spring, summer, and early fall. Differences between warming over the ocean and land surfaces have been well documented in GCMs (Manabe et al. 1991; Cubasch et al. 2001; Braganza et al. 2003, 2004; Sutton et al. 2007; Lambert and Chiang 2007; Joshi et al. 2008; Dong et al. 2009; Fasullo 2010) and the observational record (Sutton et al. 2007; Lambert and Chiang 2007; Drost et al. 2012). Greater warming over land is evident on the continental scale in both transient and equilibrium climate change experiments (Sutton et al. 2007). In transient experiments, the greater heat capacity of the ocean results in a slower temperature increase, leading to a land–sea contrast in the warming. A number of explanations—primarily based on the difference in moisture availability—have been proposed for the existence of land–sea contrast in equilibrium experiments (e.g., Sutton et al. 2007; Joshi et al. 2008). Moisture availability is particularly low in arid and semiarid regions, including a large swath of western North America adjacent to the greater Los Angeles region, which explains the strong land–sea contrast present in the warming signal.

Land–sea contrast in the warming is present on large scales in each GCM's climate change signal, but how is this contrast expressed on the regional scale? Local topography and the circulation simulated by WRF govern which areas have warming that is more ocean-like or land-like. The land–sea breeze brings marine air and its characteristics to the coastal zone on a daily basis (Hughes et al. 2007) which suppresses warming there, keeping it at or near ocean levels. This suppression is limited to the coastal zone because marine air masses cannot easily penetrate the surrounding mountain complexes. Meanwhile, the inland areas separated from the coast by a mountain complex are not exposed to marine air and have similar warming as interior land areas in the GCMs.

The second prominent feature is the enhanced warming at high elevations, which can be seen by comparing the warming to the domain topography shown in Fig. 1. During winter and spring months, snow–albedo feedback occurs in mountainous areas, a feature also observed previously in regional simulations of California's mountainous areas by Kim (2001). In a warmer climate, reductions in snow cover result in an increase in absorbed solar radiation, which are balanced, in part, by increased surface temperatures (Giorgi et al. 1997).
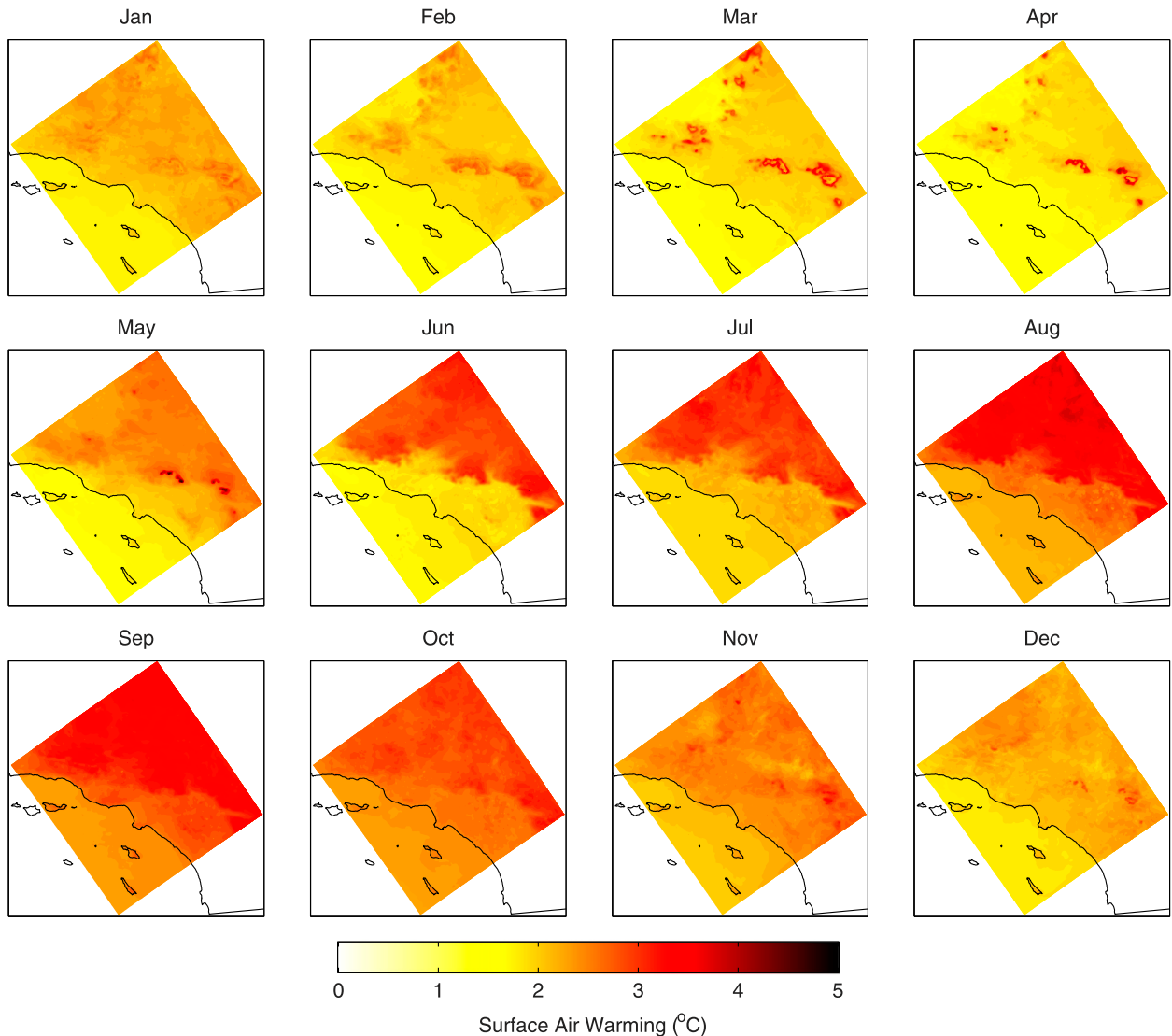
FIG. 4. Monthly mean surface air temperature change (°C) for the midcentury period (2041–60) relative to the baseline period (1981–2000) averaged over the five dynamically downscaled GCMs.

Early in the year, only snow near the snow line is warm enough to be sensitive to an increase in temperature. The reduction in snow cover results in a band of enhanced warming between the baseline and future snow lines visible in the months of March and April. In May and June, any small amount of remaining snow is sensitive to temperature change, leading to warming even at the mountain peaks.

## 3. Statistical downscaling

We constructed a statistical model to accurately and efficiently approximate the warming patterns that would have been produced had dynamical downscaling been performed on the remaining GCMs. The statistical model scales the dominant spatial pattern (identified through principal component analysis of the dynamical warming patterns) and the regional mean so they are consistent with the regionally averaged warming over the Los Angeles region as well as the land–sea contrast in the warming.

### a. Principal component analysis of spatial patterns

Principal component analysis (PCA) was performed on the 60 monthly warming patterns (five models, each with 12 monthly warming patterns) with their regional means removed (Fig. 5). Here PCA involves eigenvalue decomposition of the covariance matrix $\mathbf{X}^T\mathbf{X}$, where
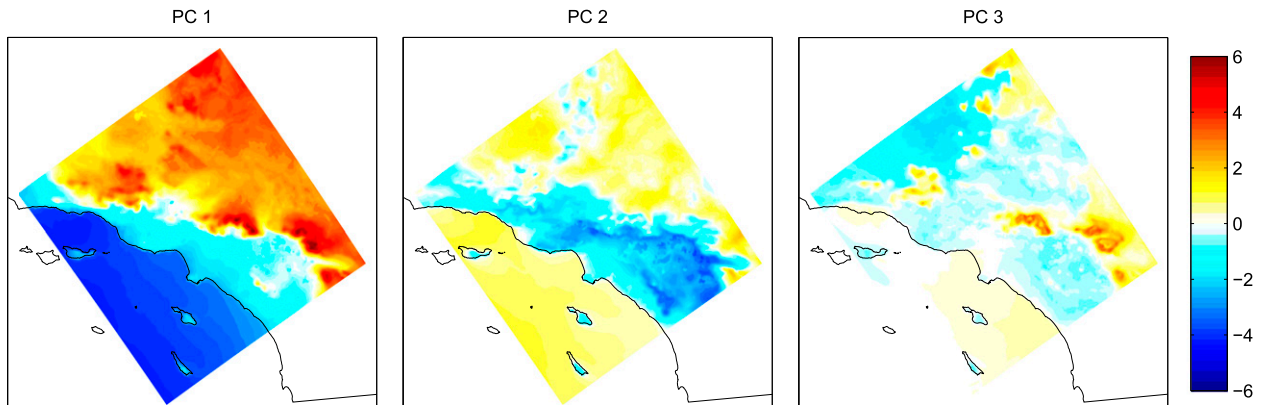
FIG. 5. The three largest principal components, in descending order of size. PCA was performed on the monthly warming patterns from the five dynamically downscaled models, with the monthly domain averages removed. PC1 accounts for 74% of the variance. This PC is referred to as the CIP because of its negative loadings over the coastal land areas and positive loadings inland. PCs 2 and 3 account for 13% and 5% of the variance, respectively.

each column of **X** is a warming pattern for a month of a dynamically downscaled GCM. The matrix **X** has size $19\,872 \times 60$, where $19\,872$ is the number of grid points in the domain and 60 is the number of monthly dynamically downscaled warming patterns. The columns of **X** are centered to mean zero, corresponding to subtracting the regional mean from each warming pattern. The resulting principal components (PCs) are spatial patterns, with associated loadings that are functions of month and GCM. [Note that this differs from the way that PCA is often performed on spatiotemporal data, in which the matrix **X** has dimensions $n \times p$, corresponding to $p$ locations in space observed at $n$ times. Usually **X** is centered so that the mean of each column is zero, which is equivalent to subtracting the time mean (climatology) at each location. The resulting principal components are time series with associated loadings in space. The spatial loadings are usually referred to as empirical orthogonal functions.] The reason that we apply PCA in this fashion is that it produces the spatial modes that explain the largest fractions of the spatial variance (as opposed to time series that explain the most temporal variance).

The leading principal component (PC1) explains 74% of the spatial variance. It is referred to as the coastal–inland pattern (CIP) henceforth because of its strong positive loadings inland and negative loadings over the coastal zone and ocean. The second and third PCs (13% and 5% variance explained) may also represent important physical phenomena, but their roles in shaping the warming patterns are much smaller, and we ignore them for the remainder of this paper.

The CIP arises from local dynamics modulating the basic contrast in climate between the land and ocean. These dynamics are apparent in other basic variables shaping the region's climate. For example, there is a very

strong negative correlation ($r = -0.97$) between the CIP and the baseline period annual-mean specific humidity (Fig. 6), a climate variable that also exhibits a significant land–sea contrast in this region. This relationship arises because the ocean is by far the most consistent source of water for evaporation in this region. Air masses over the ocean are rapidly and continuously resupplied with water vapor as necessary to maintain high relative humidity levels. Meanwhile, dry air masses over the desert interior remain cut off from moisture sources. In the coastal zone, land–sea breezes and synoptically driven alternations of the onshore and offshore flow pattern (Conil and Hall 2006) generally lead to intermediate moisture levels. Very similar dynamics mediate the warming distribution, as described in section 2d, with relatively small warming over the ocean, intermediate warming over the coastal zone, and larger warming inland. Thus the CIP is an expression of local atmospheric circulation patterns endemic to the region. Because the mechanisms that create the CIP are independent of the particular GCMs we have chosen, we are confident that the CIP can be used to downscale other GCMs.

For each month and GCM, the dynamically downscaled warming patterns can be closely approximated as the sum of the dynamically downscaled regional-mean warming and the CIP scaled by the loadings derived from principal component analysis. The average of the root-mean-square error (RMSE) between the resulting approximate warming patterns and their dynamically downscaled counterparts is 0.19°C. (When this calculation was repeated omitting the contribution of the CIP, the average RMSE more than doubled to 0.39°C, indicating the importance of including spatial variations.) Furthermore, at each location in the domain, the approximate warming was correlated with the dynamically
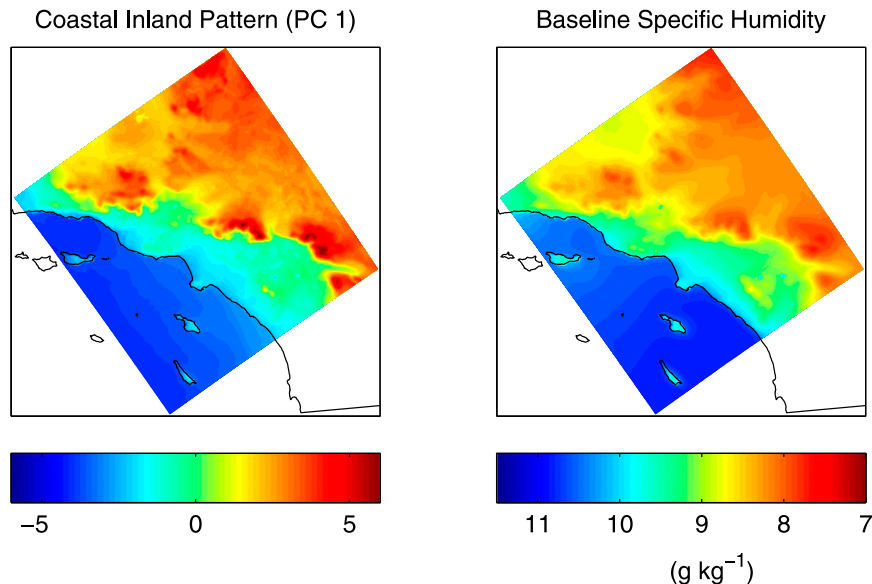
Coastal Inland Pattern (PC 1)

Baseline Specific Humidity



FIG. 6. (left) Coastal–inland pattern and (right) surface specific humidity climatology (g kg$^{-1}$) of the baseline period. The two spatial patterns are highly correlated ($r = -0.97$).

downscaled warming. The domain average of these correlations is 0.98. This confirms that nearly all variations in the dynamically downscaled warming can be captured by combining the regional mean and the scaled CIP. Therefore, these two factors are used as the basis for the statistical model. In section 3b, we determine the optimal locations in the GCM warming patterns to use as predictors of the dynamically downscaled regional-mean warming and of the appropriate scaling for the CIP.

### b. Finding large-scale predictors

To statistically downscale each of the remaining GCMs in the CMIP5 ensemble, we need to identify large-scale predictors of the dynamically downscaled regional mean and land–sea contrast. To do this, we identified the locations where the GCM warming is best correlated with the dynamically downscaled regional mean and land–sea contrast. Since the GCMs have different resolutions, we first linearly interpolated the GCM monthly warming patterns to a common grid (our outermost WRF grid, with 18-km resolution; Fig. 1). The highest correlations between the interpolated GCM warming and the dynamically downscaled regional mean are found over the adjacent ocean and along the coast (Fig. 7a). Since these correlations were calculated using the monthly warming patterns from each of the five GCMs, they indicate the degree to which sampling at that location would capture both intermonthly and intermodel variations in the dynamically downscaled regional mean. If this exercise could be undertaken for all 32 GCMs in the ensemble, the location of the optimal sampling point might be slightly

different, because of variations in resolution and grid placement between the GCMs. To build in a tolerance for such ensemble-size effects, we sampled over a region encompassing the highest correlated points, rather than just the best-correlated point. The predictor of the dynamically downscaled regional mean, RgMean(gcm, month), is the average warming over all the points a rectangular region with longitude bounds 120.5°–117.5°W and latitude bounds 32°–34.5°N shown in Fig. 7a (black and white dashed box).

A similar procedure was used to find the optimal locations to sample the land and the ocean warming for calculation of the GCM land–sea contrast. First, the exact values of the land–sea contrast from the dynamically downscaled warming patterns were calculated by taking the dot product of the monthly mean warming patterns with the CIP. These values were then correlated with the GCM warming interpolated to the common 18-km grid (Fig. 7b). The correlations are highest over the high desert of Southern California and southern Nevada, northeast of our 2-km domain. The predictor of the dynamically downscaled inland warming is calculated as the average warming over the rectangular area with longitude bounds 118°–113°W and latitude bounds 34°–37.5°N. To find the location to sample the ocean warming, we repeated this procedure, but using partial correlations with the effect of inland warming predictor removed (Fig. 7c). These partial correlations identify the optimal ocean sampling location to use in conjunction with our previously selected inland location. The GCM ocean warming is calculated as the warming averaged over a rectangular area with longitude
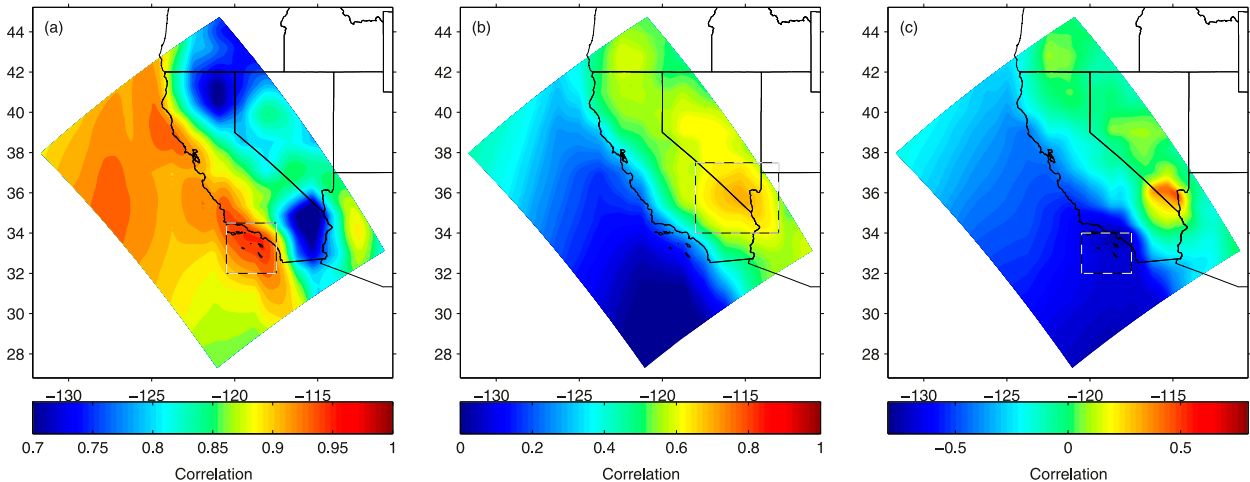
FIG. 7. Correlations between GCM warming (interpolated to an 18-km grid) and the dynamically downscaled (a) regional-mean warming and (b) land–sea contrast in the warming. The sampled regional-mean warming and inland warming are calculated as averages over the warming in the black and white dashed boxes in (a) and (b), respectively. (c) Partial correlations between the interpolated GCM warming and the dynamically downscaled land–sea contrast with the effect of the sampled inland warming removed. The ocean warming is calculated as the average over the black and white dashed box in (c).

bounds 120.5°–117.5°W and latitude bounds 32°–34°N. The predictor of the dynamically downscaled land–sea contrast, LandSeaContrast(gcm, month), is calculated as the inland warming predictor minus the GCM ocean warming predictor. If the procedure is reversed, and the optimal ocean location is selected before the optimal inland location, they still end up in nearly identical spots.

## c. The prediction equation

The statistical model approximates the dynamically downscaled warming as a linear combination of the scaled regional-mean warming in the GCM and the product of the GCM's land–sea contrast with the coastal–inland pattern. The prediction equation for the statistically downscaled warming is

$$\Delta T(gcm, month, i, j) = \alpha + \beta \times \text{RgMean}(gcm, month)$$
$$+ \gamma \times \text{LandSeaContrast}(gcm, month) \times \text{CIP}(i, j),$$

where $(i, j)$ are coordinates in the 2-km grid and $\alpha$, $\beta$, and $\gamma$ are coefficients determined by linear regression of the dynamically downscaled values of the regional-mean warming and land–sea contrast onto their large-scale predictors, RgMean and LandSeaContrast (Fig. 8). The values of these coefficients are $\alpha = 0.14°C$, $\beta = 1.10$, and $\gamma = 1.03$. Since $\alpha > 0$, and $\beta$ is slightly larger than one, the sampled GCM warming must be shifted up and inflated to match the dynamically downscaled regional-mean warming. This reflects the fact that the predictor

(the warming over the coast and adjacent ocean in the GCMs) must be shifted to a slightly greater value to match the dynamically downscaled regional mean, which encompasses inland areas as well. The dynamically downscaled and GCM-sampled land–sea contrasts are nearly the same, as their ratio is approximately one ($\gamma = 1.03$).

## d. Validation of the statistical model

Cross-validation was performed to assess how accurately the statistical model replicates the warming patterns produced by the dynamical model. The entire statistical model was rebuilt using only four of the five GCMs, and then used to predict the warming of the remaining GCM. This involved first redoing the principal component analysis to find the CIP. (These alternative patterns are nearly identical no matter which model is left out: the correlation between any two is greater than 0.98. This is additional evidence for the robustness of this pattern in regional warming.) Next, the optimal sampling locations were recalculated. They were similarly located in each case. Finally, linear regression was performed to recalculate the parameters $\alpha$, $\beta$, and $\gamma$. Once the model was rebuilt, it was applied to the remaining GCM. This procedure was performed five times in all, with each GCM taking a turn being omitted from calibration and used for testing. This cross-validation technique gives us five sets of predicted warming patterns that are compared to their dynamical counterparts. These warming patterns are also used later to assess value added (section 3e).
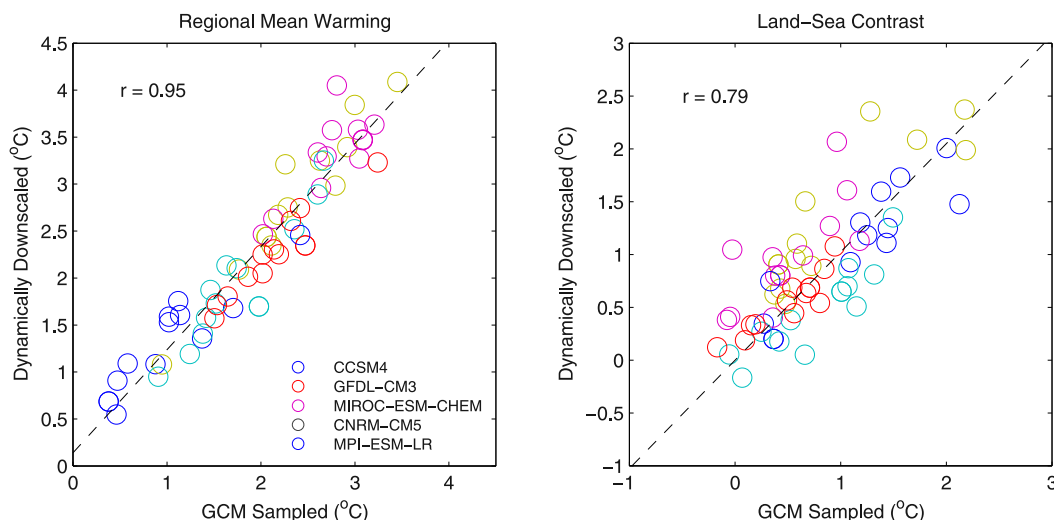
FIG. 8. Scatterplots of (left) dynamically downscaled regional-mean warming vs GCM-sampled regional-mean warming and (right) dynamically downscaled land–sea contrast vs the GCM-sampled land–sea contrast. For each GCM (colors), the 12 monthly mean warming values are shown. Approximations used by the statistical model are shown as black dashed lines.

The statistical model consistently reproduces the dynamically downscaled warming pattern for the omitted GCM with a reasonable degree of accuracy (Fig. 9, leftmost columns). When the ocean is excluded, the average spatial correlation between the dynamically and statistically generated annual-mean patterns is 0.88. The average mean absolute error in the annual-mean warming patterns is 0.27°C over the land areas in the five models. This error has to be viewed in the context of the variations the statistical model is intended to capture. The range of the five annual means averaged over the land areas is 2.1°C, about an order of magnitude larger than the error. This error is small enough that substituting the statistical model output for that of the dynamical model does not significantly affect the mean or spread of the ensemble, two of the most important outcomes of a multimodel climate change study like this one. The statistical model is slightly less accurate at reproducing the monthly warming patterns (average RMSE is 0.35°C) resulting from greater variety in the monthly patterns. Still, the error is an order of magnitude smaller than the range of the monthly mean land-mean warming (3.9°C). This gives additional confidence that the statistical model can capture even the monthly warming patterns to a reasonable level of accuracy.

*e. Comparison with other statistical downscaling techniques*

A reasonable question is whether other statistical downscaling methods would produce warming patterns that are equally close to the dynamically downscaled

patterns as the hybrid method. We compare the hybrid method to four other methods: BCCA (Hidalgo et al. 2008; Maurer and Hidalgo 2008; Maurer et al. 2010), BCSD (Wood et al. 2002, 2004; Maurer 2007), linear interpolation of the GCM warming pattern, and the warming at the nearest GCM grid point, which gives an idea of the result if raw GCM data are used, with no downscaling whatsoever. The BCCA and BCSD data can be found online (http://gdo-dcp.ucllnl.org/downscaled_cmip_projections; Reclamation 2013) and have been linearly interpolated from their native ⅛° resolution to the 2-km resolution innermost WRF grid.

The hybrid warming patterns are most visually similar to the dynamically generated warming patterns. However, it is important to verify this observation using objective measures of model skill. We used two metrics: spatial correlation and mean absolute error (further divided into errors in the regional mean and errors in the spatial pattern), shown in Table 2. Comparisons were made over the land areas only, because BCCA and BCSD do not generate values over the ocean. Interestingly, BCSD produces nearly the same warming pattern as linear interpolation, so the results of all of these comparisons are nearly identical for the two methods. The average spatial correlation between the hybrid statistically downscaled annual-mean warming patterns and their dynamically downscaled counterparts is 0.88, compared to 0.61, 0.61, 0.61, and 0.35 for BCCA, BCSD, the linear interpolated, and the raw GCM warming patterns, respectively. This demonstrates that the hybrid method is superior to any other method at
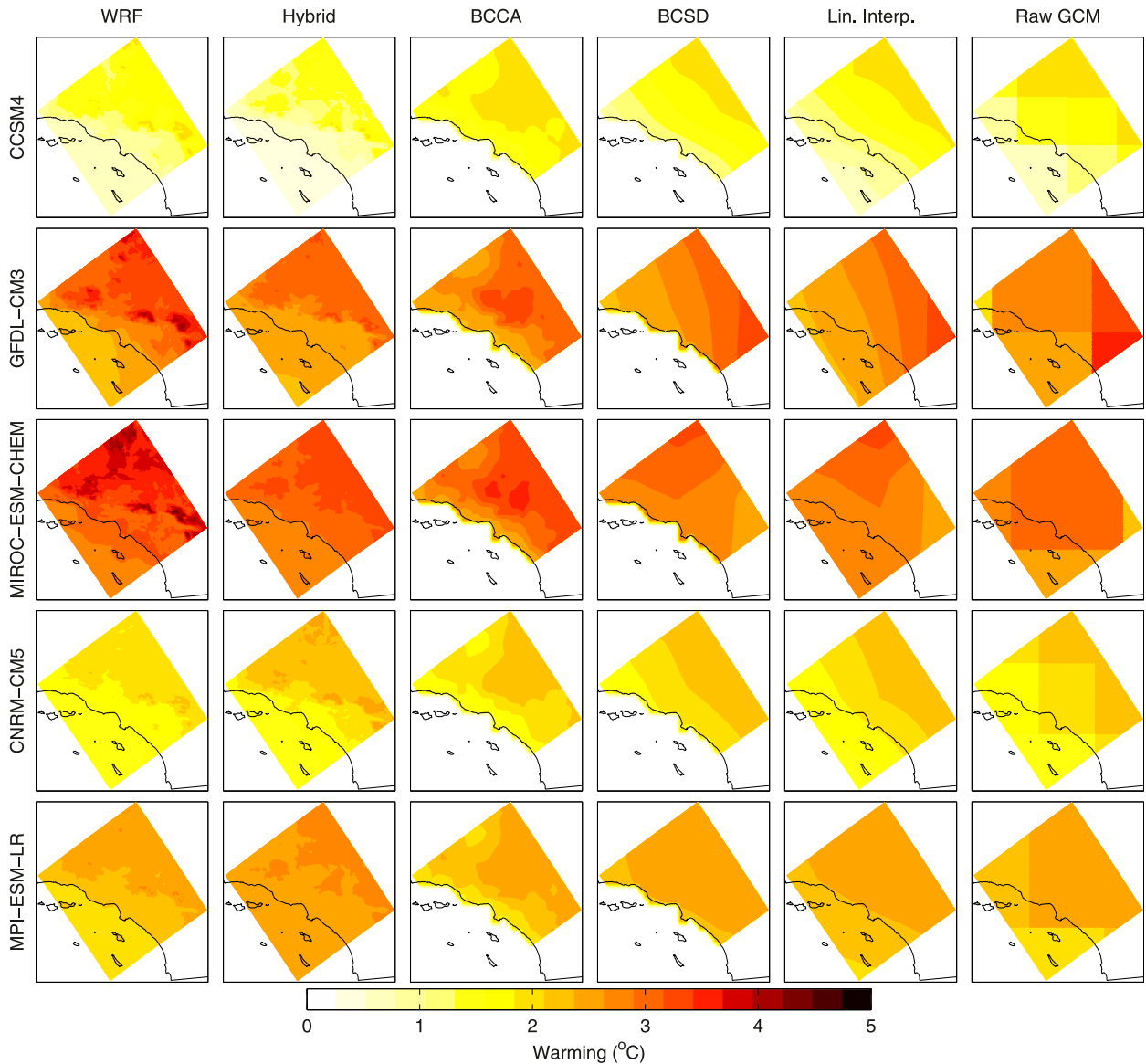
FIG. 9. Annual-mean warming projections (°C) for five GCMs produced by six different methods: (left)–(right) dynamical downscaling, statistical downscaling with the hybrid technique, statistical downscaling with BCCA, statistical downscaling with BCSD, linear interpolation of GCM, and nearest GCM grid box. Projections are for midcentury (2041–60) relative to the baseline period (1981–2000) under the RCP8.5 scenario. Note that the hybrid downscaled patterns presented in the column labeled "Hybrid" are the results of the cross-validation exercise.

accurately predicting the locations of the spatial features in the warming. For the monthly average patterns, the hybrid model also has higher correlations. The correlations are somewhat lower because the statistical model dials up or down only one spatially varying pattern (the CIP), whereas each month has a slightly different characteristic spatial pattern (Fig. 4).

The second metric is mean absolute error (MAE), taken over the land areas. The hybrid method and BCCA have the same MAE (0.27°C), while the other

methods have slightly higher errors in the 0.29°–0.32°C range (Table 2). So the statistical downscaling methods are all somewhat close to the dynamically downscaled results according to this metric. The fact that the purely statistical techniques approximate these warming patterns fairly accurately suggests that the stationarity assumption holds in this situation. When we split each annual pattern into its regional mean and spatial pattern, we found that all methods did comparably in capturing the regional mean (MAE range 0.24°–0.27°C), but

TABLE 2. Average spatial correlation and MAE between the dynamically downscaled annual and monthly warming patterns and those generated by the raw GCM, linear interpolation, BCSD, BCCA, and the hybrid method.

| | Annual | | | | Monthly | | | |
|---|---|---|---|---|---|---|---|---|
| | Spatial correlation | Spatial MAE (°C) | Regional MAE (°C) | Total MAE (°C) | Spatial correlation | Spatial MAE (°C) | Regional MAE (°C) | Total MAE (°C) |
| Raw GCM | 0.35 | 0.29 | 0.26 | 0.32 | 0.27 | 0.27 | 0.31 | 0.42 |
| Linear interpolation | 0.61 | 0.21 | 0.27 | 0.29 | 0.42 | 0.22 | 0.31 | 0.38 |
| BCSD | 0.61 | 0.21 | 0.27 | 0.29 | 0.41 | 0.23 | 0.33 | 0.40 |
| BCCA | 0.61 | 0.24 | 0.24 | 0.27 | 0.36 | 0.26 | 0.36 | 0.44 |
| Hybrid | 0.88 | 0.14 | 0.27 | 0.27 | 0.53 | 0.19 | 0.29 | 0.35 |

the hybrid method does considerably better at capturing the spatial pattern MAE with 0.14°C versus 0.21°–0.29°C for the other methods. Thus the hybrid method's strength is its ability to capture the spatial details, not the regional mean. The hybrid method also outperforms the other methods for the monthly patterns. The monthly mean errors are larger than in the annual-mean case, which is likely due to the simplicity of using a single spatial pattern for all calendar months. Although we experimented with using different spatial patterns for each month, the gains in accuracy were offset by problems arising from small sample sizes. (We had only five dynamically downscaled warming patterns each month to calibrate each monthly varying model, rather than the 60 patterns used for the original model.)

The biggest advantage of the hybrid method comes when we consider the ensemble-mean annual-mean warming. As we have seen, no method approximates each GCM's dynamically downscaled annual-mean warming pattern perfectly. However, when we aggregate each method's approximate warming patterns into a five-model ensemble mean, the hybrid method's errors cancel out, while those from the other methods do not

(Fig. 10). By construction, the hybrid statistically downscaled ensemble mean is nearly an unbiased estimator of the dynamically downscaled ensemble mean. The only reason it is not completely unbiased is that the CIP is not identical to the ensemble-mean annual-mean spatial pattern in the warming (although they are similar). In contrast, the other methods have systematic biases as large as 1°C in magnitude. BCCA shows amplified warming over the middle of the domain, but in a way that does not match the topography and circulation of this area, and completely misses amplified warming in the mountainous western part of the domain. In contrast, the other methods give overly smoothed land–sea contrasts that fail to resolve the sharp gradients in the warming over the mountains, along the coastline, and in the western part of the domain. Based on this comparison there are large swathes of the region where the hybrid statistical model is valuable in providing an accurate characterization of the most likely warming outcome.

We note that the error estimates in Table 2 and the patterns in Figs. 8 and 9 are based on the statistical model built on only four GCMs and their associated
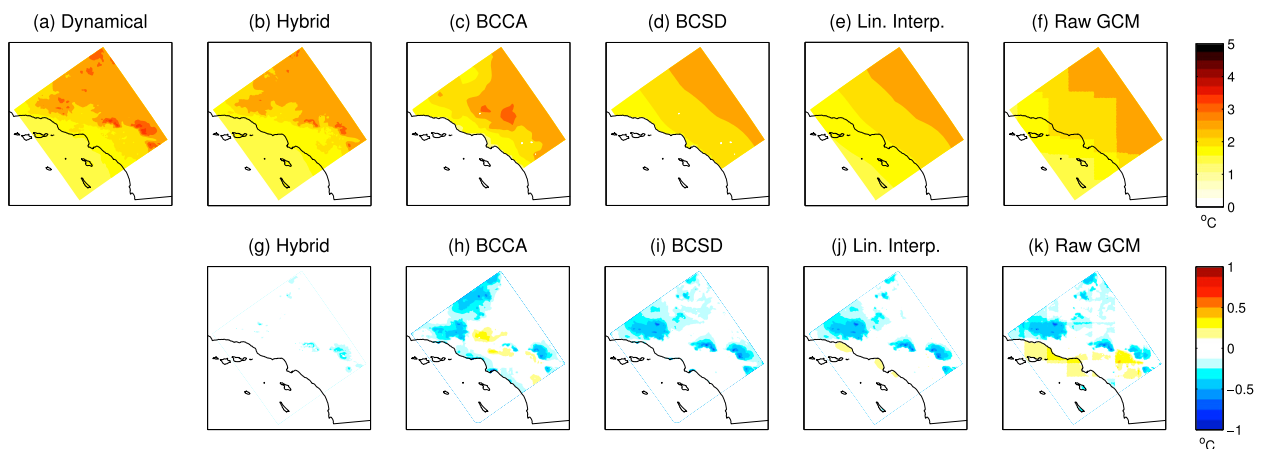


FIG. 10. Annual-mean warming (°C) averaged over five GCMs downscaled using four different methods: (a) dynamical downscaling, (b) statistical downscaling with the hybrid technique, (c) statistical downscaling with BCCA, (d) statistical downscaling with BCSD, (e) linear interpolation of GCM, and (f) nearest GCM grid box. (g)–(k) Bias of methods relative to dynamical downscaling (°C).
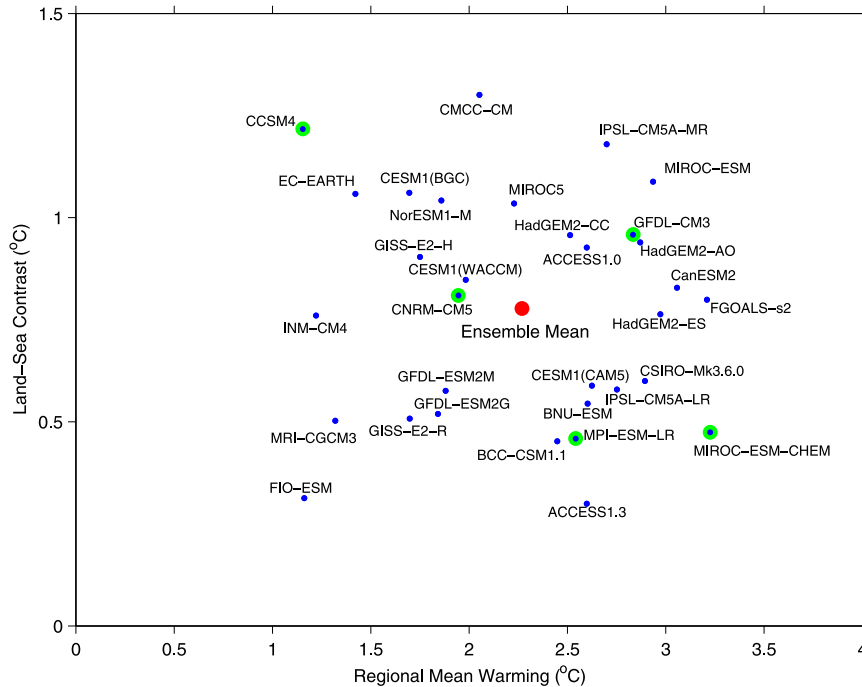
FIG. 11. Annual-mean values of regional-mean warming and land–sea contrast (°C) for each GCM (blue dots) with the ensemble mean (red dot). The five GCMs that are also dynamically downscaled are highlighted in green.

regional warming patterns. Since each GCM has a unique combination of regional mean and land–sea contrast (Fig. 11), when one is left out there is a large region of the parameter space that goes unrepresented in the calibration of the statistical model. Thus the final statistical model, calibrated using all five GCMs as described in section 3c, produces results of even higher quality. The final statistical model is used to generate the results discussed from section 4 onward.

## 4. Ensemble-mean warming and uncertainty

The final statistical model (calibrated using all five GCMs) was applied to all 32 CMIP5 GCMs with output available for the RCP8.5 scenario. The GCMs have widely varying values of the regional mean and land–sea contrast (Fig. 11). The regional mean values range from 1.4° to 3.3°C, and land–sea contrast ranges from 0.3° to 1.3°C. Notably, these two parameters are also uncorrelated, so pattern scaling using only a single of degree of freedom would be misleading here. The dynamically downscaled GCMs (Fig. 11, highlighted in green) approximately span the range of both parameters, confirming that the statistical model has been validated in the same parameter range in which it is applied. The annual-mean warming patterns that result

from plugging these parameters into the statistical model are shown in Fig. 12. There is considerable variation among these warming patterns, underscoring the importance of considering multiple GCMs when doing regional downscaling.

The ensemble-mean annual-mean warming pattern, as well as upper and lower bounds of the 95% confidence interval, are shown in Fig. 13. The regional-mean warming is 2.3°C, with a lower bound of 1.0°C and an upper bound of 3.6°C. This large intermodel spread indicates that the models disagree considerably on the magnitude of warming, even when using the same scenario. However, the global models share the characteristic of more warming inland than over the ocean. The difference in ensemble-mean warming between coastal and inland areas is especially dramatic in the summertime (Fig. 13). The average August difference between the inland and coastal areas is 0.6°C, with certain locations showing warming elevated above the coastal values by as much as 1.2°C (+62%).

The winter and spring warming that would occur in the mountains would likely be somewhat larger if we had done dynamical downscaling for all the global models (cf. Figs. 3 and 14), because the statistical model underestimates some warming resulting from snow–albedo feedback. Based on comparisons between the
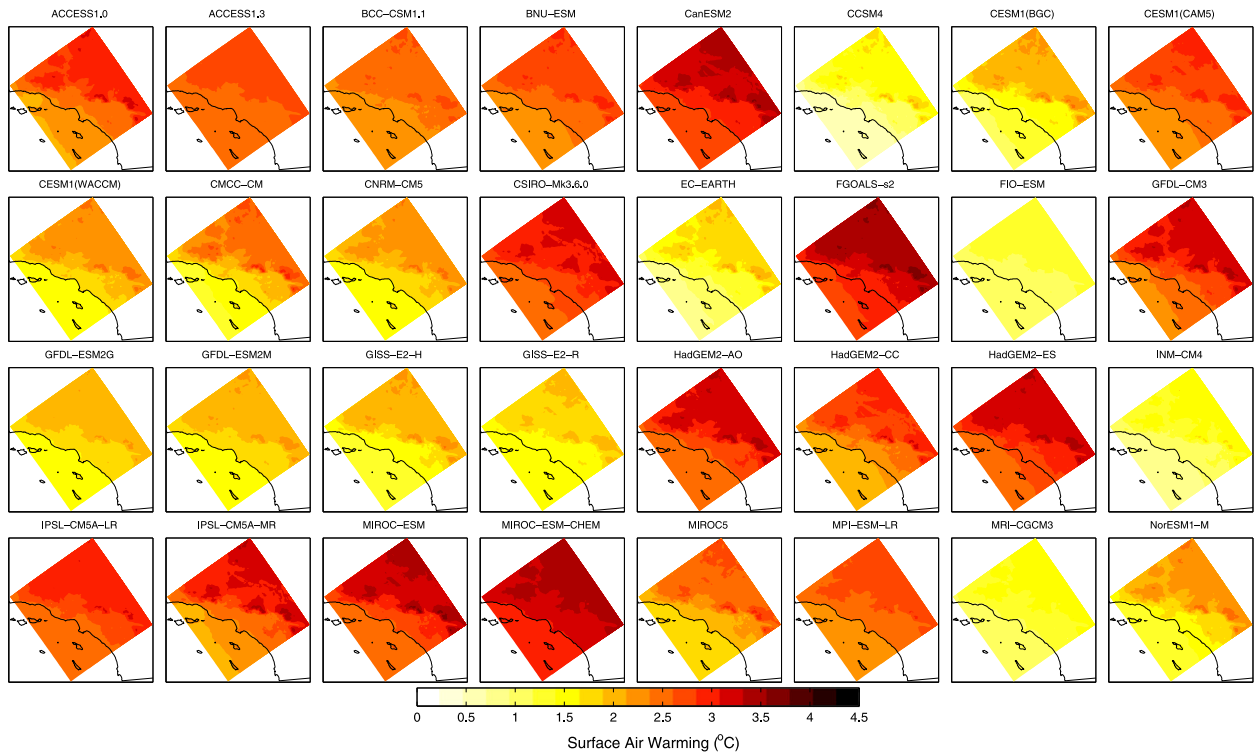
FIG. 12. Annual-mean warming patterns (°C) generated by applying the statistical model to all 32 GCMs. Warming patterns are shown for the midcentury period (2041–60) relative to the baseline period (1981–2000), under the RCP8.5 scenario.

dynamically and statistically downscaled warming patterns for spring (March–May), the springtime ensemble-mean warming would be as much as 0.5°C or larger in the San Bernardino and San Gabriel Mountain ranges. This is also consistent with the larger errors in the statistical model seen at the highest elevations in Fig. 10g.

## 5. Discussion

In this hybrid method, statistical downscaling is employed a unique way. First, while statistical models typically relate large-scale GCM output to observations, ours relates GCM output to dynamically downscaled output. This is because our hybrid statistical model is designed to be an *approximate* dynamical model. The second difference is that the hybrid statistical model was built to directly predict the temperature change, as opposed to predicting the future period temperatures and then differencing them with baseline temperatures, as is typically done. Normally, the empirical relationship employed by a statistical model is derived from the historical time period and then applied to a future time period. This leads to concerns about violating the stationarity assumption because the relationship between predictor and predictand may not hold in the future

period. In contrast, our statistical model uses a mathematical relationship between the temperature change in the GCM and the temperature change produced by dynamically downscaling. Therefore, we have a different stationarity assumption—one that is easier to satisfy—namely, that the remaining GCMs have values of midcentury regional-mean warming and land–sea contrast within the range of the five we dynamically downscaled. Since this condition is satisfied, we have confidence that the statistical relationships hold for all the GCMs that we downscale.

One potential weakness of using the method of Rasmussen et al. (2011) to produce the baseline and future dynamically downscaled simulations is that only a mean perturbation was added to the future boundary conditions, so the weather patterns exhibited in the future simulation are very similar to those in the baseline simulation, and changes to variability are difficult to assess. However, mean changes are investigated here, and if the weather patterns change in the GCM in some mean fashion, that mean change is factored into the boundary conditions when the GCM climate change signal is added. Therefore, it is likely that changes in the mean climate presented here are not substantially affected by this particular choice of methodology.
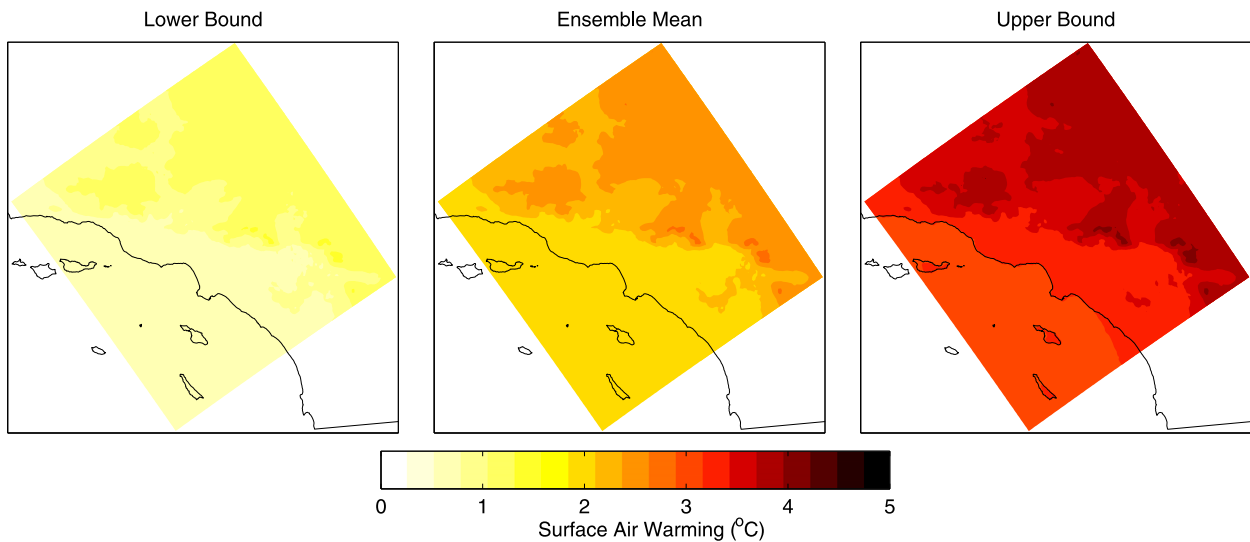
| Lower Bound | Ensemble Mean | Upper Bound |



FIG. 13. Ensemble-mean annual-mean warming and upper and lower bounds (°C), based on a 95% confidence interval, for 32 statistically downscaled GCMs run with the RCP8.5 scenario.

An important advantage of our hybrid method is that it reflects our understanding of regional climate dynamics. Some types of statistical models, like those based on artificial neural networks, have the effect of being "black boxes," where the mathematical relationships have no clear physical interpretation. Unlike those techniques, the hybrid method first performs dynamical downscaling, which allows us to identify that land–sea warming contrast and the snow–albedo feedback are the two important physical mechanisms controlling the warming. This knowledge is incorporated into the hybrid statistical model, which scales the characteristic spatial pattern (containing signatures of both mechanisms) to fit with the large-scale land–sea contrast and regional-mean warming. Because the warming patterns produced by the hybrid approach reflect physical understanding of the region's climate, they have an extra layer of credibility. Suppose, for instance, that the real climate does warm more over the interior of western North America than over the northeastern Pacific Ocean over the coming decades, as is likely if GCM projections are correct. Given the realistic behavior of WRF in distributing humidity and temperature across the landscape, it seems very likely that the associated warming pattern in the greater Los Angeles region would be characterized by sharp gradients separating the desert interior and coastal ocean, and that these gradients would be distributed across the landscape in a way very similar to the regional warming patterns we present here.

Since the strength of the hybrid method is capturing the spatial differences produced by dynamical downscaling, we expect that the hybrid method will be most applicable to other regions with strong spatial structure

in the mean warming. This includes most coastal domains, since land–sea contrast is a global phenomenon. Coastal domains in the subtropics may be especially appropriate, because the land–sea contrast is strongest at these latitudes as a result of the large difference in moisture between the land surface and the ocean. Other factors could cause important gradients in the warming, such as contrasts in elevation or surface properties. This method may also be particularly valuable in high elevation regions, since snow–albedo feedback can create important, sharp gradients in the warming. In the Los Angeles region considered here, snow–albedo feedback plays an important role, but only over a tiny fraction of the domain. On the other hand, the land–sea warming contrast affects the entire Los Angeles region, which is why it was explicitly factored into the hybrid statistical model. The effect of snow–albedo feedback was still included, but only as part of PC1, which is predominantly an expression of the land–sea contrast. If the domain considered were limited only to high-elevation areas, then it is likely that the first principal component would primarily feature the warming enhancement signature of snow–albedo feedback. Thus the processes included in the hybrid statistical model will depend on the region of interest. In fact, the method used to identify and characterize the process may also vary. Although principal component analysis was valuable here to capture the pattern associated with land–sea contrast, future users of the hybrid methodology need not be limited to using principal component analysis as performed here. Rotating the components, applying clustering, or even just using simple parameterizations may
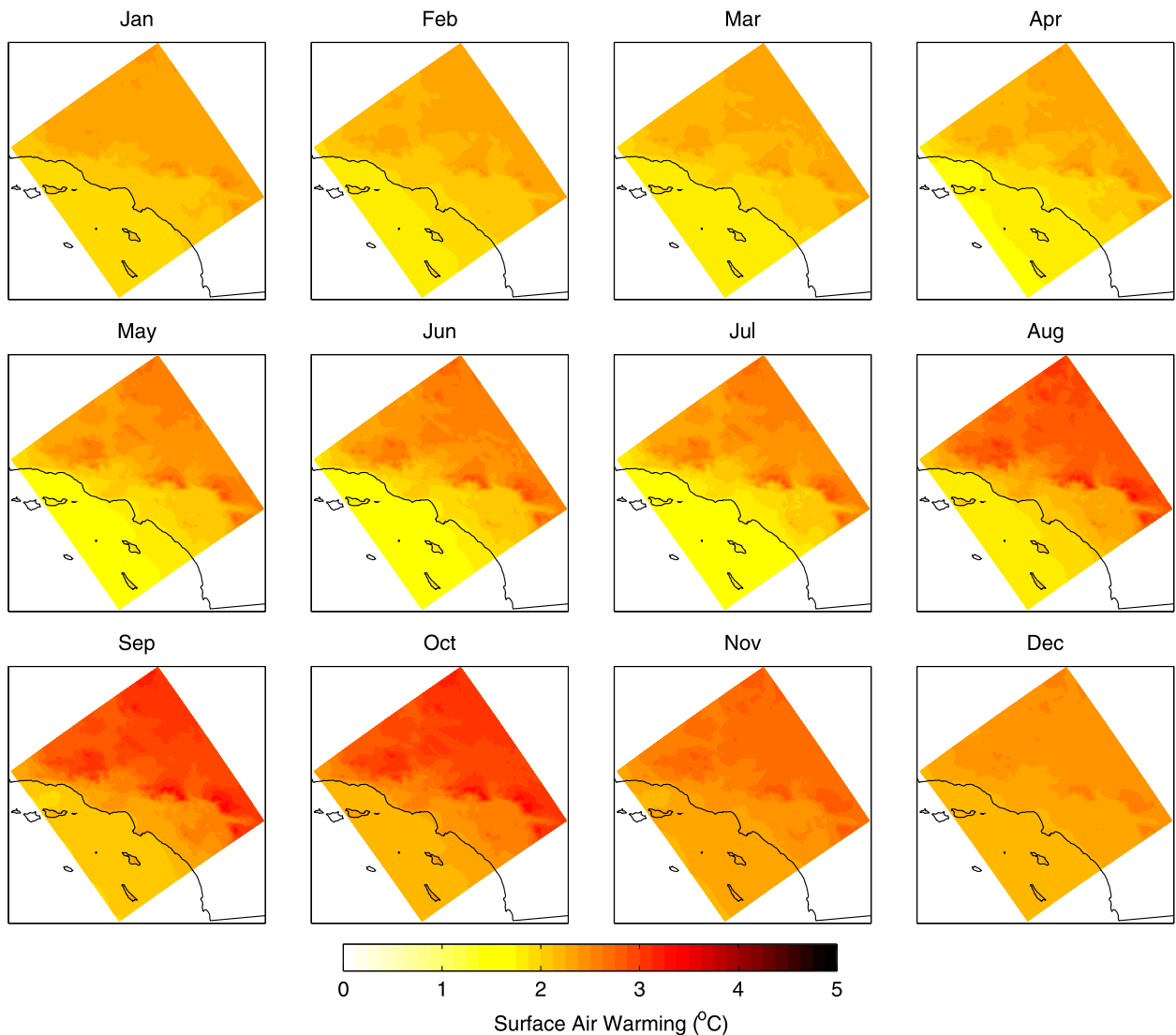
FIG. 14. Ensemble-mean monthly mean warming (°C) computed by averaging the monthly statistically downscaled warming patterns over 32 CMIP5 GCMs.

all be appropriate depending on the process considered. The hybrid methodology is defined not by the use of principal components, but by the identification of the most salient processes creating the regional climate change signal from dynamical downscaling, and the use of the dynamical output in formulating a process-oriented statistical model.

## 6. Conclusions

In this paper, we present a hybrid dynamical–statistical approach for downscaling the climate change signal from an entire ensemble of GCMs. The method is applied here to downscale the midcentury warming signal over the Los Angeles region in 32 CMIP5 GCMs. This approach starts

with the use of dynamical downscaling following a previously used method of Rasmussen et al. (2011). Under this method, a single baseline simulation is performed with WRF that represents the region's historical climate for the 1981–2000 period. The baseline simulation is compared both to station measurements and to the Livneh et al. (2013) spatially complete, observationally based gridded product. These comparisons reveal strong correlations between simulated and observed spatial patterns and time series, leading us to believe that WRF can do a reasonable job of approximating temporal and spatial variations in climate. Then, five future simulations are performed with WRF forced by the same boundary conditions from the baseline simulation, but with the monthly mean climate shifted according to the GCMs'

climate change signals. These five future simulations are differenced with the baseline simulation to produce the regionalized climate change signals corresponding to each GCM. Next, to save computational resources, a statistical model is built that scales the characteristic patterns derived from dynamical downscaling according to the regional warming sampled from the GCM. This statistical model is used to approximate the warming that would result if the remaining global models were dynamically downscaled. The ensemble-mean regional-mean warming is projected to be approximately 2.3°C, with 95% confidence that the warming is between 1.0° and 3.6°C. Thus, the intermodel differences in the GCM outcomes create significant uncertainty in projections of warming over the Los Angeles region.

One of the chief advantages of the hybrid method is its ability to capture the important finescale spatial variations in the warming. Based on the dynamically downscaled simulations with WRF, inland and mountain locations are expected to warm up considerably more than coastal areas, primarily during the summer months. The hybrid statistical model is the best able to capture the spatial variations found in WRF, based on a comparison with several other commonly used statistical downscaling techniques. Whether WRF's patterns are the most accurate representation of how the true future climate would change is difficult to assess, since there is no ground truth for the future. However, WRF's ability to accurately capture climate variations in the baseline period and produce realistic, physically consistent features in the climate change patterns gives us confidence in the veracity of its output. A further advantage of the hybrid method is that, when we average over the five simulated annual-mean warming patterns, the errors in the hybrid patterns nearly cancel out, revealing only minor biases. In contrast, the other statistical downscaling methods produce warming patterns with large systematic biases relative to WRF, especially along the coastline and in topographically complex regions that are not resolved well in the GCMs. The hybrid method has similar accuracy in approximating the regional-mean warming compared to the other techniques, likely because all methods have access to the GCM warming averaged over the region, which is already a good predictor of the dynamically downscaled regional mean. The fact that the hybrid method approximates WRF more accurately than the other statistical methods is likely due to the training of the hybrid statistical model on the climate change signal itself, not just on the baseline period, as is done typically with statistical methods, including the others used here. Were these other methods to be trained upon WRF climate change output, or even just the WRF baseline output instead of baseline gridded observation, they would likely improve their ability to approximate the WRF climate change patterns.

In Part II of this study, we apply the hybrid technique developed here to other scenarios and time periods. We examine the differences between midcentury (2041–60) and end-of-century (2081–2100) warming and demonstrate how emission scenario has a much larger effect at the end of the century. We also explore how warming effects the diurnal cycle and the number of extreme heat days. In a separate study, Berg et al. (2015) modify this hybrid dynamical–statistical approach to downscale the CMIP5 ensemble's midcentury precipitation projections to the greater Los Angeles region.

## REFERENCES

Arritt, R. W., and M. Rummukainen, 2011: Challenges in regional-scale climate modeling. *Bull. Amer. Meteor. Soc.,* **92,** 365–368, doi:10.1175/2010BAMS2971.1.

Berg, N., A. Hall, F. Sun, S. C. Capps, D. Walton, B. Langenbrunner, and D. Neelin, 2015: Twenty-first-century precipitation changes over the Los Angeles region. *J. Climate,* **28,** 401–421, doi:10.1175/JCLI-D-14-00316.1.

Braganza, K., D. J. Karoly, A. C. Hirst, M. E. Mann, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2003: Simple indices of global climate variability and change: Part I—Variability and correlation structure. *Climate Dyn.,* **20,** 491–502, doi:10.1007/s00382-002-0286-0.

——, ——, ——, P. Stott, R. J. Stouffer, and S. F. B. Tett, 2004: Simple indices of global climate variability and change. Part II: Attribution of climate change during the twentieth century. *Climate Dyn.,* **22,** 823–838, doi:10.1007/s00382-004-0413-1.

Cabré, M., S. A. Solman, and M. N. Nuñez, 2010: Creating regional climate change scenarios over southern South America for the 2020's and 2050's using the pattern scaling technique: Validity and limitations. *Climatic Change,* **98,** 449–469, doi:10.1007/s10584-009-9737-5.

Caldwell, P. M., H.-N. S. Chin, D. C. Bader, and G. Bala, 2009: Evaluation of a WRF based dynamical downscaling simulation over California. *Climatic Change,* **95,** 499–521, doi:10.1007/s10584-009-9583-5.

Cayan, D., E. P. Maurer, M. D. Dettinger, M. Tyree, and K. Hayhoe, 2008: Climate change scenarios for the California region. *Climatic Change,* **87** (Suppl. 1), 21–42, doi:10.1007/s10584-007-9377-6.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling

system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.,* **129,** 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Conil, S., and A. Hall, 2006: Local regimes of atmospheric variability: A case study of Southern California. *J. Climate,* **19,** 4308–4325, doi:10.1175/JCLI3837.1.

Cubasch, U., and Coauthors, 2001: Projections of future climate change. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 525–582.

Dong, B., J. M. Gregory, and R. T. Sutton, 2009: Understanding land–sea warming contrast in response to increasing greenhouse gases. Part I: Transient adjustment. *J. Climate,* **22,** 3079–3097, doi:10.1175/2009JCLI2652.1.

Drost, F., D. Karoly, and K. Braganza, 2012: Communicating global climate change using simple indices: An update. *Climate Dyn.,* **39,** 989–999, doi:10.1007/s00382-011-1227-6.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.,* **46,** 3077–3107, doi:10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Duffy, P. B., and Coauthors, 2006: Simulations of present and future climates in the western United States with four nested regional climate models. *J. Climate,* **19,** 873–895, doi:10.1175/JCLI3669.1.

Fasullo, J. T., 2010: Robust land–ocean contrasts in energy and water cycle feedbacks. *J. Climate,* **23,** 4677–4693, doi:10.1175/2010JCLI3451.1.

Giorgi, F., and L. O. Mearns, 1991: Approaches to regional climate change simulation: A review. *Rev. Geophys.,* **29,** 191–216, doi:10.1029/90RG02636.

——, and ——, 2002: Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method. *J. Climate,* **15,** 1141–1158, doi:10.1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2.

——, C. S. Brodeur, and G. T. Bates, 1994: Regional climate change scenarios over the United States produced with a nested regional climate model. *J. Climate,* **7,** 375–399, doi:10.1175/1520-0442(1994)007<0375:RCCSOT>2.0.CO;2.

——, J. W. Hurrell, M. R. Marinucci, and M. Beniston, 1997: Elevation dependency of the surface climate change signal: A model study. *J. Climate,* **10,** 288–296, doi:10.1175/1520-0442(1997)010<0288:EDOTSC>2.0.CO;2.

——, and Coauthors, 2001: Regional climate information—Evaluation and projections. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 583–638.

——, C. Jones, and G. R. Asrar, 2009: Addressing climate information needs at the regional level: The CORDEX framework. *WMO Bull.,* **58,** 175–183.

Hayhoe, K., and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. *Proc. Natl. Acad. Sci. USA,* **101,** 12 422–12 427, doi:10.1073/pnas.0404500101.

Hidalgo, H. G., M. D. Dettinger, and D. R. Cayan, 2008: Downscaling with constructed analogues: Daily precipitation and temperature fields over the Unites States. California Energy Commission Tech. Rep. CEC-500-2007-123, 48 pp.

Hong, S. Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.,* **42,** 129–151.

Hughes, M., A. Hall, and R. G. Fovell, 2007: Dynamical controls on the diurnal cycle of temperature in complex topography. *Climate Dyn.,* **29,** 277–292, doi:10.1007/s00382-007-0239-8.

Joshi, M. M., J. M. Gregory, M. J. Webb, D. M. Sexton, and T. C. Johns, 2008: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Climate Dyn.,* **30,** 455–465, doi:10.1007/s00382-007-0306-1.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor. Soc.,* **43,** 170–181, doi:10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Kawase, H., T. Yoshikane, M. Hara, F. Kimura, T. Yasunari, B. Ailikun, H. Ueda, and T. Inoue, 2009: Intermodel variability of future changes in the Baiu rainband estimated by the pseudo global warming downscaling method. *J. Geophys. Res.,* **114,** D24110, doi:10.1029/2009JD011803.

Kerr, R. A., 2011: Vital details of global warming are eluding forecasters. *Science,* **334,** 173–174, doi:10.1126/science.334.6053.173.

——, 2013: Forecasting regional climate change flunks its first test. *Science,* **339,** 638, doi:10.1126/science.339.6120.638.

Kharin, V. V., and F. W. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate,* **15,** 793–799, doi:10.1175/1520-0442(2002)015<0793:CPWME>2.0.CO;2.

Kim, J., 2001: A nested modeling study of elevation-dependent climate change signals in California induced by increased atmospheric $CO_2$. *Geophys. Res. Lett.,* **28,** 2951–2954, doi:10.1029/2001GL013198.

Lambert, F. H., and J. C. Chiang, 2007: Control of land–ocean temperature contrast by ocean heat uptake. *Geophys. Res. Lett.,* **34,** L13704, doi:10.1029/2007GL029755.

Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby, 2003: Regional climate research: Needs and opportunities. *Bull. Amer. Meteor. Soc.,* **84,** 89–95, doi:10.1175/BAMS-84-1-89.

——, Y. Qian, X. Bian, W. M. Washington, J. Han, and J. O. Roads, 2004: Mid-century ensemble regional climate change scenarios for the western United States. *Climatic Change,* **62,** 75–113, doi:10.1023/B:CLIM.0000013692.50640.55.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.,* **22,** 1065–1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Livneh, B., E. A. Rosenberg, C. Lin, B. Nijssen, V. Mishra, K. M. Andreadis, E. P. Maurer, and D. P. Lettenmaier, 2013: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States: Update and extensions. *J. Climate,* **26,** 9384–9392, doi:10.1175/JCLI-D-12-00508.1.

Manabe, S., R. J. Stouffer, M. J. Spelman, and K. Bryan, 1991: Transient responses of a coupled ocean–atmosphere model to gradual changes of atmospheric $CO_2$. Part I. Annual mean response. *J. Climate,* **4,** 785–818, doi:10.1175/1520-0442(1991)004<0785:TROACO>2.0.CO;2.

Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.,* **83,** 407–430, doi:10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2.

Maurer, E. P., 2007: Uncertainty in hydrologic impacts of climate change in the Sierra Nevada, California, under two emissions scenarios. *Climatic Change,* **82,** 309–325, doi:10.1007/s10584-006-9180-9.

——, and H. G. Hidalgo, 2008: Utility of daily vs. monthly large-scale climate data: An intercomparison of two statistical downscaling methods. *Hydrol. Earth Syst. Sci.,* **12,** 551–563, doi:10.5194/hess-12-551-2008.

——, ——, T. Das, M. D. Dettinger, and D. R. Cayan, 2010: The utility of daily large-scale climate data in the assessment of climate change impacts on daily streamflow in California. *Hydrol. Earth Syst. Sci.,* **14,** 1125–1138, doi:10.5194/hess-14-1125-2010.

Mearns, L. O., I. Bogardi, F. Giorgi, I. Matyasovszky, and M. Palecki, 1999: Comparison of climate change scenarios

generated from regional climate model experiments and statistical downscaling. *J. Geophys. Res.,* **104** (D6), 6603–6621, doi:10.1029/1998JD200042.

Meehl, G. A., C. Covey, K. E. Taylor, T. Delworth, R. J. Stouffer, M. Latif, B. McAvaney, and J. F. Mitchell, 2007: The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bull. Amer. Meteor. Soc.,* **88,** 1383–1394, doi:10.1175/BAMS-88-9-1383.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.,* **87,** 343–360, doi:10.1175/BAMS-87-3-343.

Mitchell, J. F. B., S. Manabe, T. Tokioka, and V. Meleshko, 1990: Equilibrium climate change. *Climate Change: The IPCC Scientific Assessment,* J. T. Houghton, G. J. Jenkins, and J. J. Ephraums, Eds., Cambridge University Press, 131–172.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-*k* model for the longwave. *J. Geophys. Res.,* **102** (D14), 16 663–16 682, doi:10.1029/97JD00237.

Pierce, D. W., T. Das, D. R. Cayan, E. P. Maurer, N. Miller, Y. Bao, and M. Tyree, 2013: Probabilistic estimates of future changes in California temperature and precipitation using statistical and dynamical downscaling. *Climate Dyn.,* **40,** 839–856, doi:10.1007/s00382-012-1337-9.

Rasmussen, R., and Coauthors, 2011: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *J. Climate,* **24,** 3015–3048, doi:10.1175/2010JCLI3985.1.

Reclamation, 2013: Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Downscaled CMIP5 Climate Projections, Comparison with preceding Information, and Summary of User Needs. Prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, CO, 47 pp. [Available online at http://gdo-dcp.ucllnl.org/downscaled_cmip_projections/dcpInterface.html.]

Salathé, E. P., Jr., R. Steed, C. F. Mass, and P. H. Zahn, 2008: A high-resolution climate model for the U.S. Pacific Northwest: Mesoscale feedbacks and local responses to climate change. *J. Climate,* **21,** 5708–5726, doi:10.1175/2008JCLI2090.1.

——, L. R. Leung, Y. Qian, and Y. Zhang, 2010: Regional climate model projections for the State of Washington. *Climatic Change,* **102,** 51–75, doi:10.1007/s10584-010-9849-y.

Santer, B. D., T. M. L. Wigley, M. E. Schlesinger, and J. F. B. Mitchell, 1990: Developing climate scenarios from equilibrium GCM results. MPI Rep. 47, 29 pp.

Sato, T., F. Kimura, and A. Kitoh, 2007: Projection of global warming onto regional precipitation over Mongolia using a regional climate model. *J. Hydrol.,* **333,** 144–154, doi:10.1016/j.jhydrol.2006.07.023.

Schiermeier, Q., 2010: The real holes in climate science. *Nature,* **463,** 284–287, doi:10.1038/463284a.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp.

Snyder, M. A., J. L. Bell, L. C. Sloan, P. B. Duffy, and B. Govindasamy, 2002: Climate responses to a doubling of atmospheric carbon dioxide for a climatically vulnerable region. *Geophys. Res. Lett.,* **29** (11), doi:10.1029/2001GL014431.

Sutton, R. T., B. Dong, and J. M. Gregory, 2007: Land/sea warming ratio in response to climate change: IPCC AR4 model results and comparison with observations. *Geophys. Res. Lett.,* **34,** L02701, doi:10.1029/2006GL028164.

Sun, F., D. B. Walton, and A. Hall, 2015: A hybrid dynamical–statistical downscaling technique. Part II: End-of-century warming projections predict a new climate state in the Los Angeles region. *J. Climate,* **28,** 4618–4636, doi:10.1175/JCLI-D-14-00197.1.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Amer. Meteor. Soc.,* **93,** 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tebaldi, C., R. L. Smith, D. Nychka, and L. O. Mearns, 2005: Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multimodel ensembles. *J. Climate,* **18,** 1524–1540, doi:10.1175/JCLI3363.1.

Timbal, B., A. Dufour, and B. McAvaney, 2003: An estimate of future climate change for western France using a statistical downscaling technique. *Climate Dyn.,* **20,** 807–823, doi:10.1007/s00382-002-0298-9.

Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: A review of methods and limitations. *Prog. Phys. Geogr.,* **21,** 530–548, doi:10.1177/030913339702100403.

Wood, A. W., E. P. Maurer, A. Kumar, and D. P. Lettenmaier, 2002: Long-range experimental hydrologic forecasting for the eastern United States. *J. Geophys. Res.,* **107** (D20), 4429, doi:10.1029/2001JD000659.

——, 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, doi:10.1023/B:CLIM.0000013685.99609.9e.