UNIVERSITY OF CALIFORNIA

Los Angeles

Development and Evaluation of a Hybrid Dynamical-Statistical Downscaling Method

A dissertation submitted in partial satisfaction
of the requirements for the degree Doctor of Philosophy
in Atmospheric and Oceanic Sciences

by

Daniel Burton Walton

2014

ABSTRACT OF THE DISSERTATION

Development and Evaluation of a
Hybrid Dynamical-Statistical Downscaling Method

by

Daniel Burton Walton

Doctor of Philosphy in Atmospheric and Oceanic Sciences
University of California, Los Angeles, 2014
Professor Alexander Dean Hall, Chair

Abstract

Regional climate change studies usually rely on downscaling of global climate model (GCM) output in order to resolve important fine-scale features and processes that govern local climate. Previous efforts have used one of two techniques: (1) dynamical downscaling, in which a regional climate model is forced at the boundaries by GCM output, or (2) statistical downscaling, which employs historical empirical relationships to go from coarse to fine resolution. Studies using these methods have been criticized because they either dynamical downscaled only a few GCMs, or used statistical downscaling on an ensemble of GCMs, but missed important dynamical effects in the climate change signal. This study describes the development and evaluation of a hybrid dynamical-statstical downscaling method that utilizes aspects of both dynamical and statistical downscaling to address these concerns. The first step of the hybrid method is to use dynamical downscaling to understand the most important physical processes that contribute to the climate change signal in the region of interest. Then a statistical model is built based on the patterns and relationships identified from dynamical downscaling. This statistical model can be used to downscale an entire ensemble of GCMs quickly and efficiently. The hybrid method is first applied to a domain covering Los Angeles Region to generate projections of temperature change between the 2041-2060 and

1981-2000 periods for 32 CMIP5 GCMs. The hybrid method is also applied to a larger region covering all of California and the adjacent ocean. The hybrid method works well in both areas, primarily because a single feature, the land-sea contrast in the warming, controls the overwhelming majority of the spatial detail. Finally, the dynamically downscaled temperature change patterns are compared to those produced by two commonly-used statistical methods, BCSD and BCCA. Results show that dynamical downscaling recovers important spatial features that the statistical methods miss. This confirms that the dynamical downscaling provides a more credible fine-scale signal for use in the hybrid method.

The dissertation of Daniel Walton is approved.

Robert G. Fovell

David J. Neelin

Glen Michael MacDonald

Alexander Dean Hall, Committee Chair

University of California, Los Angeles

2014

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

Chapter 2 is a version of Walton D. B., F. Sun, A. Hall, S. Capps: A Hybrid Dynamical-Statistical Downscaling Technique, Part I: Development and Validation of the Technique. J. Climate, in press. Fengpeng Sun ran the regional climate simulations with WRF, with help from Scott Capps. Alex Hall was the project manager.

Support for this work was provided by the City of Los Angeles and the US Department of Energy as part of the American Recovery and Reinvestment Act of 2009. Additional funding was provided by the National Science Foundation (Grant #1065864, "Collaborative Research: Do Microenvironments Govern Macroecology?") and the Southwest Climate Science Center.

Vita

EDUCATION

University of California, Los Angeles

M.S., Atmospheric and Oceanic Sciences, June 2012

M.A., Mathematics, June 2009

Harvey Mudd College

B.S., Mathematics, May 2007

FELLOWSHIPS AND AWARDS

Selected to represent UCLA at UC Graduate Research Day. One of four UCLA graduate students to share research with California State Senators and Assembly Members. Sacramento, CA, March 2014.

National Science Foundation Graduate Research Fellowship. University of California, Los Angeles, September 2007 – August 2010.

Stavros Busenberg Prize in Applied Mathematics. Harvey Mudd College, 2006-2007.

RIF Scholarship for Exceptional Mathematical Ability. Harvey Mudd College, 2003-2004.

ACADEMIC SERVICE

Chi Epsilon Pi – Student Faculty Representative 2012-2013, 2013-2014

INVITED PUBLIC LECTURES

Daniel B. Walton (March 2014). "Climate Change in the Los Angeles Region." Presentation given at Third Annual Climate Change Forum, Pasadena, CA.

Daniel B. Walton, Marla Schwartz (September 2013). "Global climate change and impacts." Presentation given at UCLA Law School.

Daniel B. Walton (March 2013). "Mid-Century warming in the Los Angeles region." Presentation given at Climate Change in The Southland: Local Effects and Opportunities, Los Angeles, CA.

Daniel B. Walton (December 2012). "Mid-Century warming in the Los Angeles region." Presentation given at Faith and Climate Change event, Los Angeles, CA.

CONFERENCE PRESENTATIONS AND POSTERS

Daniel B. Walton, F. Sun, A. Hall (December 2014). Statistical Downscaling: Lessons Learned. Poster resented at the American Geophysical Union Annual Fall Meeting, San Francisco, CA.

Daniel B. Walton, F. Sun, A. Hall, and X. Qu (January 2013). A Combined Dynamical/Statistical Technique for Regionalizing an Ensemble of Climate Change Signals. Poster presented at the American Meteorological Society Annual Meeting, Austin, TX.

PUBLICATIONS

Berg, Neil, Alex Hall, Fengpeng Sun, Scott Capps, Daniel B. Walton, David Neelin, Baird Langenbrunner. Mid-21st Century precipitation changes over the Los Angeles region. J. Climate, in press.

Walton, D. B., The impact of air-sea coupling on simulating SST variability in the California Current System. (2012) Master's Thesis, University of California, Los Angeles. Available online: http://www.escholarship.org/uc/item/7pg1t8zr

Benjamin, A. T., D. B. Walton (2010). Combinatorially composing Chebyshev polynomials. Journal of Statistical Planning and Inference, 140(8): 2161-2167. DOI: 10.1016/j.jspi.2010.01.012

Benjamin, A. T., D. B. Walton (2009). Counting on Chebyshev Polynomials. Mathematics Magazine, 82(2): 117-126.

Burger, E. B., J. Gell-Redman, R. Kravitz, D. Walton, and N. Yates (2008). Shrinking the period lengths of continued fractions while still capturing convergents. Journal of Number Theory, 128(1), 144-153.

1 Research Overview

The overarching goal of this work is to improve projections of regional climate change. Typically, global climate models (GCMs) are the tools of choice for simulating climate change on a global scale. Unfortunately, these models have horizontal resolutions on the order of 1°- 2°($\sim$ 100 - 200 km), which prevent them from being able to resolve important physical features and processes that take place on smaller spatial scales (Giorgi and Mearns 1991, Leung et al. 2003). Therefore downscaling techniques have been developed to determine how the large-scale climate change signals in these GCMs are expressed on finer scales. This work looks to evaluate and improve these downscaling techniques in order to make more accurate projections of climate change at the regional scale.

Chapter 2 concerns the development of a new downscaling technique applied to the Los Angeles Region. Previous studies making regional climate change projections have either employed dynamical or statistical downscaling. Under dynamical downscaling, a regional climate model is forced at the boundaries by GCM data and run at high resolution in order to generate baseline and future data sets. Dynamical downscaling is often used because it produces a physically-consistent, complete set of variables. However, it takes significant computational resources, limiting studies to only a handful of GCMs (e.g. Duffy et al. 2006, Pierec et al. 2013). This prevents full sampling of the range of possible outcomes represented by the entire ensemble of GCMs, and the multiple emission scenarios or time periods that can be used. In order to downscale multiple GCMs, scenarios, and time periods, statistical downscaling is often used, since it exploits mathematical relationships between large-scale predictors and small-scale predictands, that are relatively computationally inexpensive to apply. However, statistical methods may miss important climate change features (Salathé et al. 2010). (This is a topic explored further in Chapter 4.) The hybrid downscaling method overcomes these obstacles by combining the ability of dynamical downscaling to capture fine-scale dynamics with the computational savings of a statistical model. Previously, dynamical and statistical downscaling techniques were only applied separately. Here, dynamical down-

scaling is used to inform a simple, region-specific statistical model that incorporates only the most important processes creating the fine-scale climate change pattern. The hybrid technique accomplishes two important goals: (1) it allows for downscaling of multiple GCMs for multiple time periods and scenarios, which are required to make probabilistic estimates of most-likely outcomes and uncertainty, (2) it captures the dominant physical processes as identified by dynamical downscaling, leading to enhanced credibility.

In Chapter 3, the new hybrid downscaling technique is applied to a different domain, one encompassing the entire state of California and the adjacent ocean. The goal here is to evaluate how well this technique works in a an alternate location and identify any modifications the could be made. In the Los Angeles Region, the fine-scale climate change pattern for temperature was characterized to a high degree of accuracy by just two factors: the regional mean temperature change and the difference between the land and ocean temperature change (the land-sea contrast in the warming). The full California domain also contains both ocean and land areas, leading to the hypothesis that the regional mean warming and the land-sea warming contrast would again be the two most important factors. Our results supported this hypothesis, with principal component analysis (PCA) showing that once the regional mean had been removed the first principal component (accounting for 89% of the spatial variance) had a clear land-sea warming dipole. This lead us to wonder if the first principal component would change if the ocean grid cells were removed before PCA were performed. Perhaps another feature besides land-sea contrast would be revealed as more important if only California's land areas were considered. Interestingly, our results showed that the spatial pattern of the first principal component was a nearly identical over land ($r = 0.99$), regardless of whether the ocean was included in PCA. Not only did the technique work successfully over this domain, it confirmed that the land-sea contrast in the warming was the dominant spatial pattern.

When the hybrid method was applied to the full California domain, some modifications were to the method itself were made. Most notably, a Monte-Carlo cross-validation

procedure was implemented to (a) diagnose the sensitivity of the hybrid statistical model parameters to the GCMs included in its training, and (b) evaluate the ability of the statistical model to predict the temperature change patterns for months or GCMs that were not included in its training data. When compared with the more naïve method of linear interpolation, the statistical model shows gains over land areas. The statistical model and linear interpolation both do very well over the ocean, where there is much less complexity to the warming pattern. The results of this chapter suggest that the hybrid downscaling method could be widely applicable, and is likely to particularly effective in regions where a single dominant factor provides spatial heterogeneity in the climate change pattern.

The hybrid method relies on dynamical downscaling (with WRF) to obtain the appropriate climate change patterns. Chapter 4 investigates how WRF's regional climate change projections differ from other downscaling methods. Temperature change patterns produced by dynamical downscaling are compared to those produced by two commonly used statistical downscaling techniques, Bias Correction with Spatial Disaggregation (BCSD) and Bias Correction with Constructed Analogs (BCCA). Surprisingly, even though BCSD produces realistic fine-scale spatial detail in both historical and future climate simulations, it has no spatial detail in the climate change signal (calculated by differing future and historical simulations). This is an important finding, since the high-resolution of the future and historical projections may mislead users into believing that the climate change signal contained in the data is also high-resolution. In fact, the climate change signal that BCSD produces is identical to the bias-corrected, regridded GCM. Therefore BCSD cannot capture any of the important spatial details. On the other hand, BCCA has the capacity to produce meaningful spatial detail. BCCA warming projections have pronounced land-sea contrast, although the exact locations and magnitudes of the features differ from WRF. While both methods capture the land-sea contrast to some extent, they produce very different results in the mountains, where the BCCA warming shows no signs of snow-albedo feedback. Snow-albedo feedback is understood to be an important part of the local climate change response,

in which reductions in snow cover (and, to a lesser extent, snow albedo) lead to increases in absorbed downward solar radiation, further enhancing the warming (Giorgi et al. 1997, Qu and Hall 2007). The magnitude of this enhancement is shown to vary from 2 °C to 7 °C per 100% snow loss (fully snow-covered to snow-free), depending on the month. Because BCCA uses *daily* historical analogs to simulate future temperatures, it inherently underestimates the snow-cover loss associated with climate change by up to a factor of five. Based on this investigation, we conclude that WRF captures important, physical details that may not be captured by common statistical downscaling methods. We also suggest potential ways to ensure that the warming enhancement due to snow-albedo feedback is properly included in statistically downscaled warming patterns.

2 Development of a Hybrid Dynamical-Statistical Downscaling Method and its Application to the Los Angeles Region

2.1 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. 2001, 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 of 1° to 2° ($\sim 100 - 200$ 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 multi-model ensemble from the World Climate Research Programme's Fifth Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012; Table 2.1).

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 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 (inter-model spread). Instead, information from a larger ensemble is preferred (Georgi 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 NCAR 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 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 multi-model downscaling. Due to its much lower computational cost, statistical downscaling is almost always used for multi-model downscaling (e.g., Giorgi et al. 2001, Tebaldi et al. 2005, Pierce et al. 2013). Unfortunately, statistical methods may not be able to capture important fine-scale changes in climate shaped by topography and mesoscale dynamics. Dynamical downscaling can capture such effects, provided the regional model resolution is high enough (e.g. Caldwell et al. 2009, Salathé et al. 2010, Arritt and Rummukainen 2011, Pierce et al. 2013). Pierce et al. (2013) found that when a pair of global model 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. 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 most important fine-scale warming features 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 fine-scale warming 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 within 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 that have already been 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.

The construction of a statistical model that mimics dynamical model behavior forces us to understand the physical mechanisms underpinning the regional patterns of change, adding an additional layer of credibility to the results. This 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.

The hybrid method is used to generate the warming patterns for 32 GCMs in the CMIP5 ensemble. These patterns represent our best estimate of what the warming would be if dynamical downscaling had been performed on these remaining GCMs.

2.2 Dynamical Downscaling

2.2.1 Model Configuration

Dynamical downscaling was performed using the Advanced Research Weather Research and Forecasting Model version 3.2 (WRF; Skamarock et al. 2008). WRF has been successfully applied to the California region in previous work (e.g. Caldwell 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: Kain-Fritsch (new Eta) cumulus scheme (Kain 2004); Yonsei University boundary layer scheme (Hong et al. 2006); Purdue Lin microphysics scheme (Lin et al. 1983); Rapid Radiative Transfer Model longwave radiation (Mlawer et al. 1997); 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.

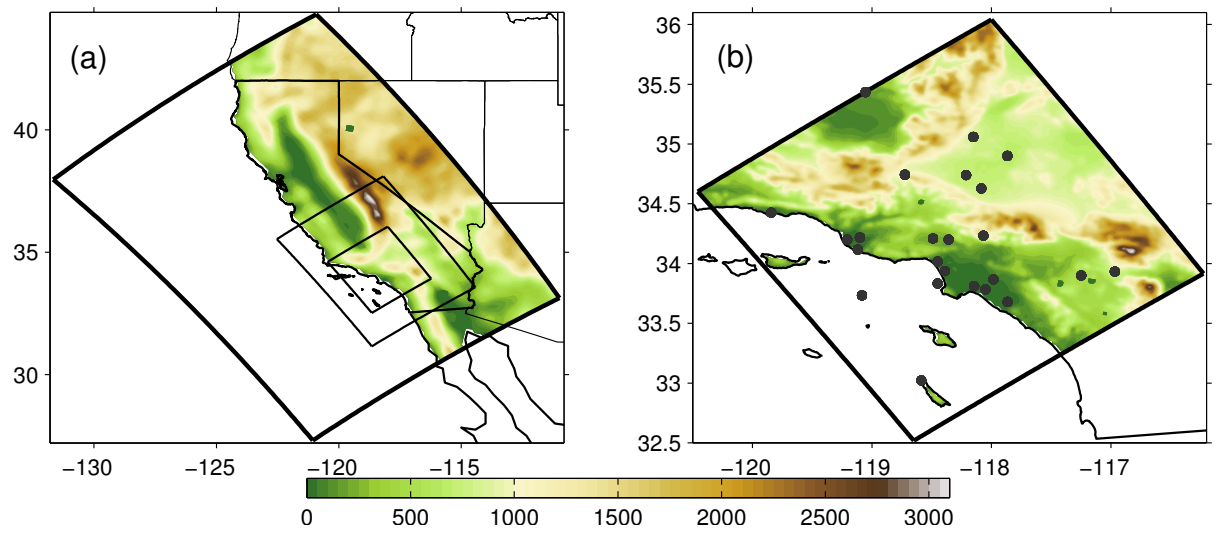

Figure 2.1: (a) Model setup, with three nested WRF domains at resolutions of 18, 6, and 2 km. Topography is shown in color (units: meters). (b) The innermost domain of the regional simulation, with 2 km resolution. Topography is shown in meters. Black dots indicate the locations of the 24 stations used for surface air temperature validation.

The three nested domains for the simulations are shown in Fig. 2.1. 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.

2.2.2 Baseline Simulation and Validation

Using this model configuration, we performed a baseline simulation whose purpose is two-fold: (1) to validate the model's ability to simulate regional climate, and (2) to provide

a baseline climate state against which a future climate simulation could be compared, to quantify the change in climate. This simulation is a dynamical downscaling of the National Centers for Environmental Prediction North America Regional Reanalysis (NARR; Mesinger et al. 2006) over the period 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. 2.1). It also provides surface boundary 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 during this time period, at 2-km resolution. The model was reinitialized each year in August, and run from September to August. Because each year was initialized separately, the time period could be divided into one-year 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 the available observational measurements from a network of 24 weather stations and buoys. These quality-controlled, hourly, near-surface meteorological observations were obtained from the National Climatic Data Center (NCDC; http://www.ncdc.noaa.gov/). The point measurements 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. 2.1). 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. 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. 2.2a). 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 ($r > 0.9$ in both seasons), although the model exhibits

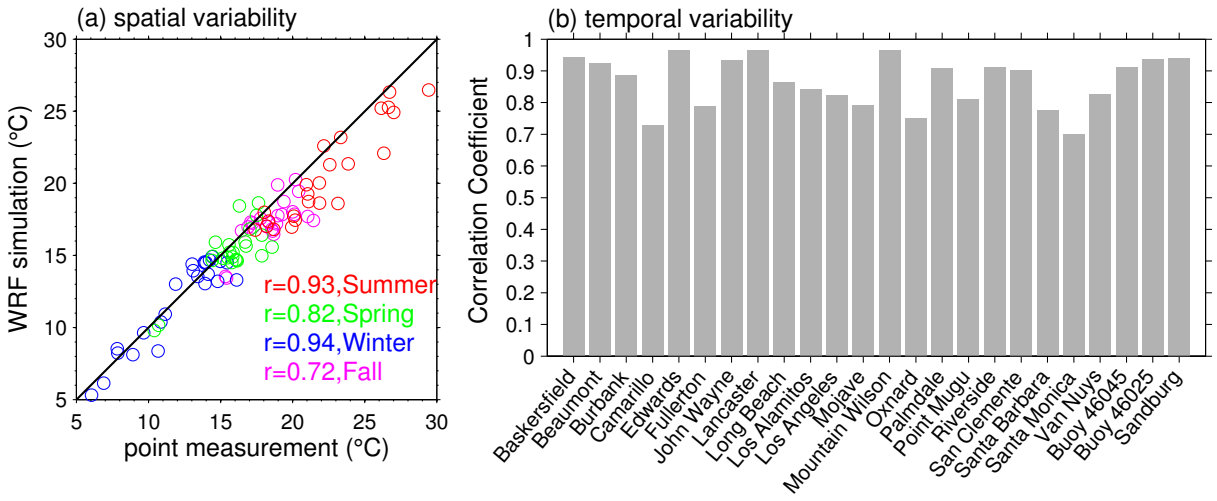

Figure 2.2: Validation of WRF dynamical downscaling against a network of 24 stations, for the period 1995-2001. (a) Simulated versus observed seasonal average climatological temperatures, for each of station. (b) Correlations at each station between simulated and observed monthly temperature anomalies. (Anomalies are relative to the composite monthly climatology.)

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. The model's ability to simulate temporal variability on monthly timescales and longer is also assessed. 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. 2.2b). Temporal variability is very well-simulated by the model, with high correlations at all locations.

Fig. 2.2 demonstrates that the model gives approximately the right spatial and temporal variations in surface air temperature at specific point locations where trustworthy observational data are available. This gives a high degree of confidence that the model is also producing the correct temperature variations in the rest of the region, where observations are absent. And most importantly for this study, it gives confidence that when it comes to surface air temperature, the model provides a realistic downscaling of the regional pattern implicit in the coarser resolution forcing data set. Thus, the dynamically downscaled climate change

patterns presented here 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.

2.2.3 Future Simulations

With the same model configuration as in the baseline simulation, we performed a second set of dynamical downscaling experiments designed to simulate the regional climate state corresponding to the mid–21st century. We applied the pseudo-global warming method (PGW; see Rasmussen et al. 2011 and references therein; also Sato et al. 2007, Kawase et al. 2009) to five GCMs in the CMIP5 ensemble corresponding to this time period and the RCP8.5 emissions scenario (see Table 2.1). To simulate the future period, we started by calculating the difference between future and baseline monthly climatologies (2041–2060 minus 1981–2000) for each GCM. These differences are the GCM climate change signals of interest. All model variables are included in the calculation of the climate change signal (i.e., 3-dimensional atmospheric variables such as temperature, relative humidity, zonal and meridional winds, and geopotential height and 2-dimensional surface variables such as temperature, relative humidity, winds and pressure). To produce the boundary conditions for the future period, we perturbed NARR data corresponding to the baseline period (September 1981–August 2001) by adding the change in monthly climatology. The resulting simulation can then be compared directly with the baseline regional simulation to assess the effect of the GCM climate change signals when they are included in the downscaling. 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 shift in mean climate. In addition to imposing a mean climate change perturbation at the lateral boundaries, CO_2 concentrations were also increased in WRF to match the equivalent radiative forcing in the RCP8.5 scenario.

Table 2.1: Details of the WCRP CMIP5 global climate models used in this study.

MODEL	COUNTRY	INSTITUTE
ACCESS1.0	Australia	Commonwealth Scientific and Industrial Research Organization
ACCESS1.3	Australia	Commonwealth Scientific and Industrial Research Organization
BCC-CSM1.1	China	Beijing Climate Center, China Meteorological Administration
BNU-ESM	China	College of Global Change and Earth System Science, Beijing Normal University
Can-ESM2	Canada	Canadian Centre for Climate Modelling and Analysis
CCSM4	USA	National Center for Atmospheric Research
CESM1(BGC)	USA	National Science Foundation, Department of Energy, National Center for Atmospheric Research
CESM1(CAM5)	USA	National Science Foundation, Department of Energy, National Center for Atmospheric Research
CESM1(WACCM)	USA	National Science Foundation, Department of Energy, National Center for Atmospheric Research
CMCC-CM	Italy	Centro Euro-Mediterraneo per I Cambiamenti Climatici
CNRM-CM5	France	Centre National de Recherches Meteorologiques
CSIRO-Mk3.6.0	Australia	Commonwealth Scientific and Industrial Research Organization
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	USA	NOAA Geophysical Fluid Dynamics Laboratory
GFDL-ESM2M	USA	NOAA Geophysical Fluid Dynamics Laboratory
GFDL-ESM2G	USA	NOAA Geophysical Fluid Dynamics Laboratory
GISS-E2-H	USA	NASA Goddard Institute for Space Studies
GISS-E2-R	USA	NASA Goddard Institute for Space Studies
HadGEM2-AO	UK	Met Office Hadley Centre
HadGEM2-CC	UK	Met Office Hadley Centre
HadGEM2-ES	UK	Met Office Hadley Centre
INMCM4	Russia	Institute for Numerical Mathematics
IPSL-CM5A-LR	France	Institut Pierre Simon Laplace
IPSL-CM5A-MR	France	Institut Pierre Simon Laplace
MIROC-ESM	Japan	AORI (U. Tokyo), NIES, JAMESTEC
MIROC-ESM-CHEM	Japan	AORI (U. Tokyo), NIES, JAMESTEC
MIROC5	Japan	AORI (U. Tokyo), NIES, JAMESTEC
MPI-ESM-LR	Germany	Max Planck Institute for Meteorology
MRI-CGCM3	Japan	Meteorological Research Institute
NorESM1-M	Norway	Norwegian Climate Center

We first downscaled CCSM4 for a 20-year 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-year 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 (2041–2060 minus 1981–2000) from the GCM to the three-year period of NARR corresponding to September 1998–August 2001.

2.2.4 Warming Patterns

In this section, we examine monthly-mean warming patterns (future minus baseline) simulated from dynamical downscaling. Fig. 2.3 shows these warming patterns averaged over the five dynamically downscaled GCMs. There are two prominent features 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 et al. 2010) and the observational record (Sutton et al. 2007, Lambert and Chiang 2007, Drost et al. 2011). Greater warming over land is evident on the continental scale in both transient and equilibrium climate change experiments due to greater heat capacity and availability of moisture for evaporative heat loss over the ocean compared to land (Manabe et al. 1991, Sutton et al. 2007, Joshi et al. 2008). Moisture availability is particularly low in arid and semi-arid regions, including a large swath of the southwestern United States adjacent to the Los Angeles Region.

Land-sea contrast in the warming is present on large scales in each global model's climate change signal, but how is this contrast expressed on the regional scale? Local

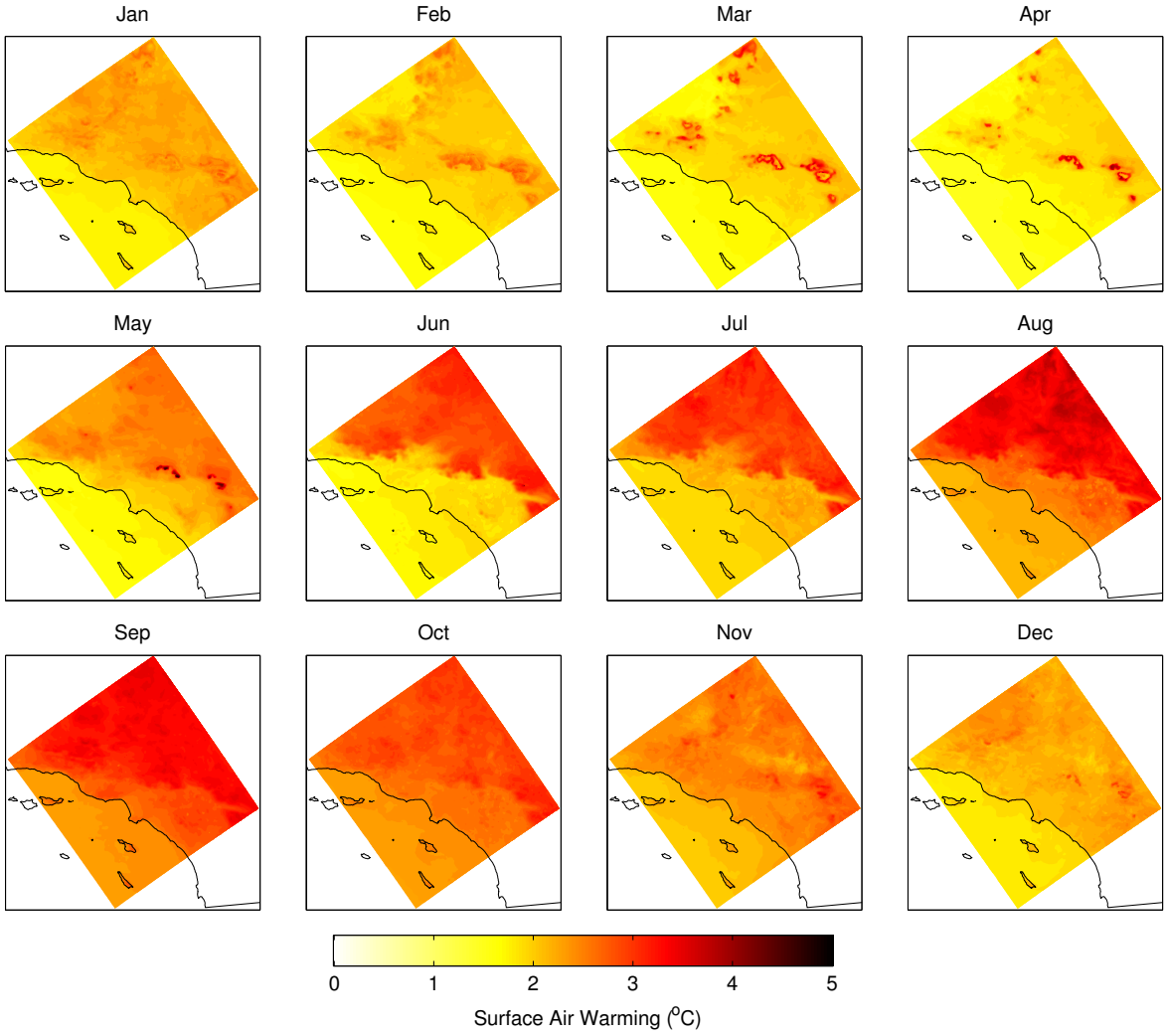

Figure 2.3: Monthly-mean surface air temperature change (units:°C) for the mid-century period (2041-2060) relative to the baseline period (1981-2000) averaged over the five dynamically downscaled GCMs.

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 (land 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 global models.

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. 2.1. During winter and spring months, snow-albedo feedback occurs in mountainous areas, a feature also observed previously in 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). Early in the year, snow cover at elevations near the snow line is more sensitive to temperature changes than at higher elevations. The decreased snow cover near the approximate snow line results in rings of enhanced warming in March and April. In May and June, snow cover at all elevations may be sensitive to temperature change, leading to larger warming extending all the way up to the mountain peaks.

2.3 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.

2.3.1 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. 2.4). Although PCA is typically applied to temporal anomalies to identify common modes of variability relative to temporal mean, instead we perform PCA on the spatial anomalies to find how warming at in the region differs from the regional average. The

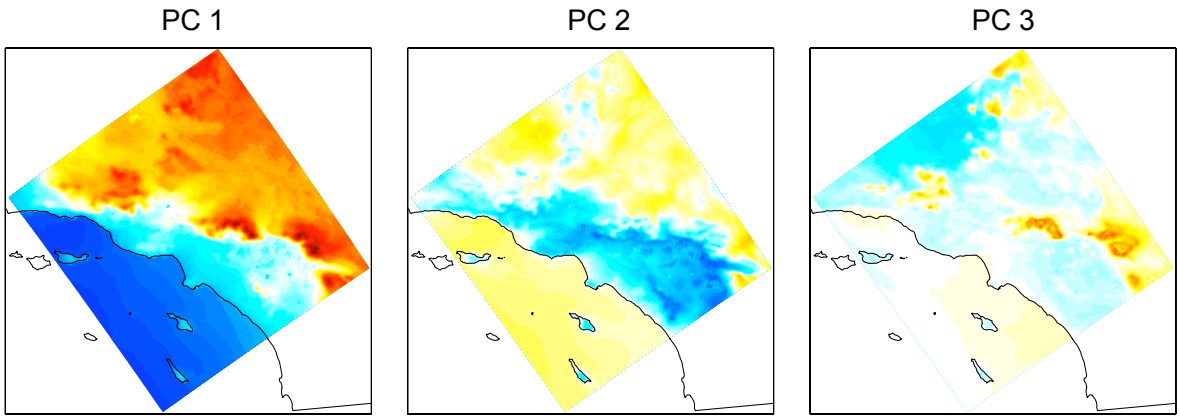

Figure 2.4: The spatial patterns associated with the three largest principal components, in descending order of size. PCA was performed on the monthly warming patterns from the five dynamically downscaled models, with the monthly domain averages removed. PC1 accounts for 74% of the variance. This component is referred to as the Coastal-Inland Pattern (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.

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%, 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. 2.5), 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 specific 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 2.2.4, 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.

The CIP and the regional mean can be linearly combined to closely approximate the dynamically downscaled warming patterns for each month and for each GCM. When linear regression is used to calculate the combination of the regional mean and the CIP that is closest to the dynamical warming pattern, the resulting approximate warming patterns have RMSE within 0.19 °C of their dynamical counterparts, on average. (When we repeated this calculation omitting the contribution of the CIP, the error more than doubled to 0.39°C, indicating the importance of including spatial variations.) Furthermore, at each point in the domain, we calculated the correlation between the dynamically downscaled warming and the linearly approximated warming. The domain average of these correlations is 0.98. This confirms that we can capture nearly all variations in warming just by combining appropriate scalings of the regional mean and the CIP. Therefore, we use these two factors as the basis for our statistical model, as discussed in the next section.

2.3.2 Finding Optimal Sampling Locations

In order to statistically downscale each of the remaining GCMs in the CMIP5 ensemble, we need to obtain approximate values of the regional mean and land-sea contrast, which we do by sampling the large-scale warming. To find the optimal sample locations, we examined the five GCMs we have dynamically downscaled and identified the points in the large-scale

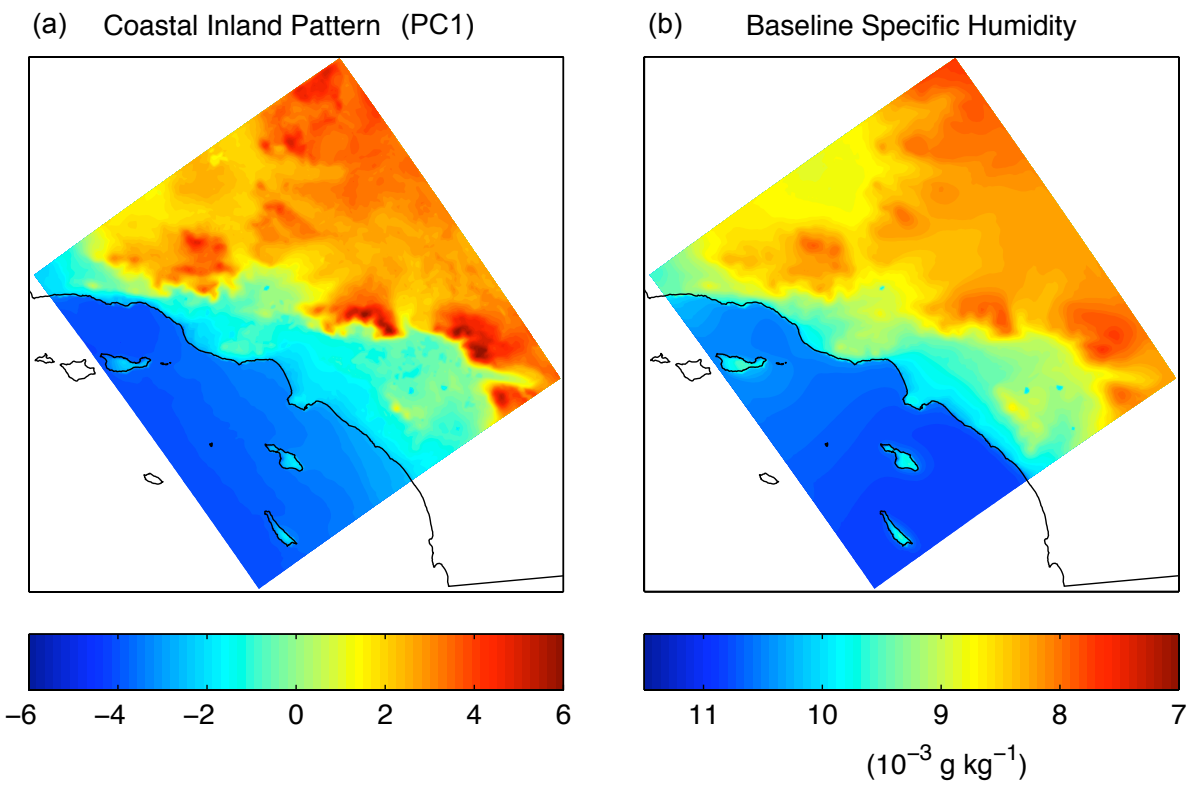

Figure 2.5: Coastal-Inland Pattern (left) and surface specific humidity climatology of the baseline period (right; units: 10^{-3} g $\cdot$ kg^{-1}). The two spatial patterns are highly correlated (r = -0.97).

domain that are best correlated with the dynamically downscaled regional mean and land-sea contrast. Since the GCMs have different resolutions, we first interpolated the GCM monthly warming patterns to a common grid (our outermost WRF grid, with 18 km resolution, Fig. 2.1). The highest correlations between the large-scale warming and the regional mean are found over the adjacent ocean and along the coast (Fig. 2.6a). Since these correlations were calculated using the monthly averages from each of the five GCMs, they indicate the degree to which sampling at that location would capture both inter-monthly and inter-model variations in the 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, due to 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

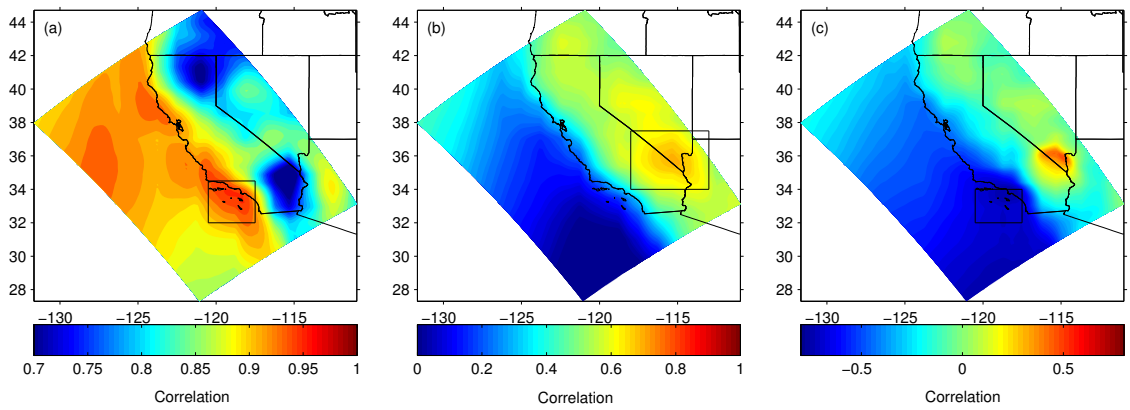

Figure 2.6: Correlations between GCM warming (interpolated to an 18-km grid) and the dynamically downscaled (a) regional mean warming and (b) land-sea contrast in the warming (i.e. loadings associated with PC1). The sampled regional mean warming and inland warming are calculated as averages over the warming in the black boxes in panels (a) and (b), respectively. Panel (c) shows 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 box in (c).

just the best-correlated point. The GCM regional mean warming, $\Delta T^{GCM}_{RegMean}$, is calculated as the average over all the points a rectangular region with longitude bounds [120.5°W, 117.5°W] and latitude bounds [32°N, 34.5°N] shown in Fig. 2.6a (black 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. The PC1 loadings (a proxy for land-sea contrast) were correlated with the GCM warming interpolated to the common 18-km grid (Fig. 2.6b). The correlations are highest over the high desert of Southern California and Southern Nevada, northeast of our 2-km domain. The GCM inland warming is calculated as the average warming over the rectangular area with longitude bounds [118°W, 113°W] and latitude bounds [34°N, 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 removed (Fig. 2.6c). 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 bounds [120.5°W, 117.5°W] and latitude bounds [32°N, 34°N]. The GCM land-sea contrast, is calculated as the GCM inland warming, ΔT^{GCM}_{Inland}, minus the GCM ocean warming ΔT^{GCM}_{Ocean}. 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.

2.3.3 Hybrid Statistical Model Formula

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

$$\Delta T^{stat}(i,j) = \alpha + \beta \Delta T^{GCM}_{RegMean} + \gamma(\Delta T^{GCM}_{Inland} - \Delta T^{GCM}_{Ocean}) \cdot CIP(i,j) \tag{1}$$

where (i,j) are coordinates in the 2km grid and α, β, and γ are coefficients determined by linear regression (Fig. 2.7). The values of these coefficients are $\alpha = 0.14$°C, $\beta = 1.10, \gamma = 1.03$. α is greater than 0, and β is slightly larger than one, which indicates that the dynamically downscaled regional mean warming is shifted and scaled to a slightly larger value than the sampled regional mean warming in the GCM. This reflects the fact that the GCM is sampled over the ocean and coastal zone, while the dynamically downscaled region also encompasses inland areas, which tend to warm more. The dynamically downscaled and GCM-sampled land-sea contrasts are nearly the same, as their ratio is approximately one ($\gamma = 1.03$).

2.3.4 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 retrained on only four of the five GCMs, and then used to predict the warming of the remaining GCM. This involved first redoing the PCA to find the CIP (PC1). (These alternative

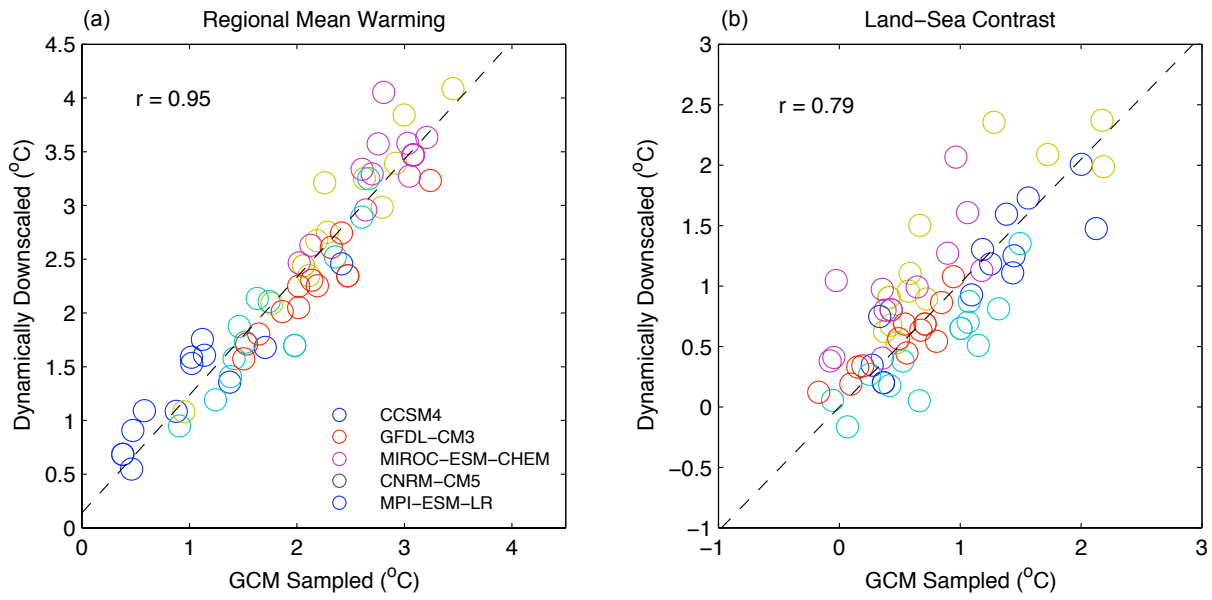

Figure 2.7: Dynamically downscaled regional-mean warming versus GCM-sampled regional-mean warming (left), and the dynamically downscaled land-sea contrast (i.e. the PC1 loadings) versus the GCM-sampled land-sea contrast (right). For each GCM (colors), the twelve monthly-mean warming values are shown. Linear approximations used by the statistical model are shown as black dashed lines.

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 α, β, and γ. Once the model was retrained, 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.

The statistical model consistently reproduces the dynamically downscaled warming pattern for the omitted GCM with a reasonable degree of accuracy (Fig. 2.8, rightmost columns). The average spatial is correlation between the dynamically and statistically gen-

erated annual-mean patterns is 0.95. The average root mean squared error (RMSE) in the annual-mean warming patterns is 0.28 °C over 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 whole domain 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 multi-model climate change study like this one. The statistical model is slightly less accurate at reproducing the monthly warming patterns (average RMSE is 0.39°C) due to greater variety of spatial details in the monthly patterns. Still, the error is an order of magnitude smaller than the range of the monthly-mean regional-mean warming (3.5°C). This gives additional confidence that the statistical model can capture even the monthly warming patterns to a reasonable level of accuracy.

2.3.5 Value of Incorporating Dynamical Information

The goal of this study is to provide an ensemble of projections, as if all 32 GCMs had been dynamically downscaled. Due to computational limitations, only five GCMs were dynamically downscaled and the remaining 27 GCMs were statistically downscaled using our statistical model that incorporates the dynamically downscaled output. A reasonable question is whether incorporating the dynamically downscaled output into the statistical model was helpful or if a simple statistical method would have sufficed. We answered this question by comparing the statistically generated warming patterns—generated via cross-validation—and the linearly interpolated GCM warming patterns, to see which method produced closer results to the five dynamically downscaled GCMs. Fig. 2.8 gives a comparison of warming patterns produced by dynamical downscaling, our statistical downscaling technique, and linear interpolation. A comparison to the warming at the nearest GCM grid point, is also included to give an idea of the result if raw GCM data are used, with no downscaling whatsoever.

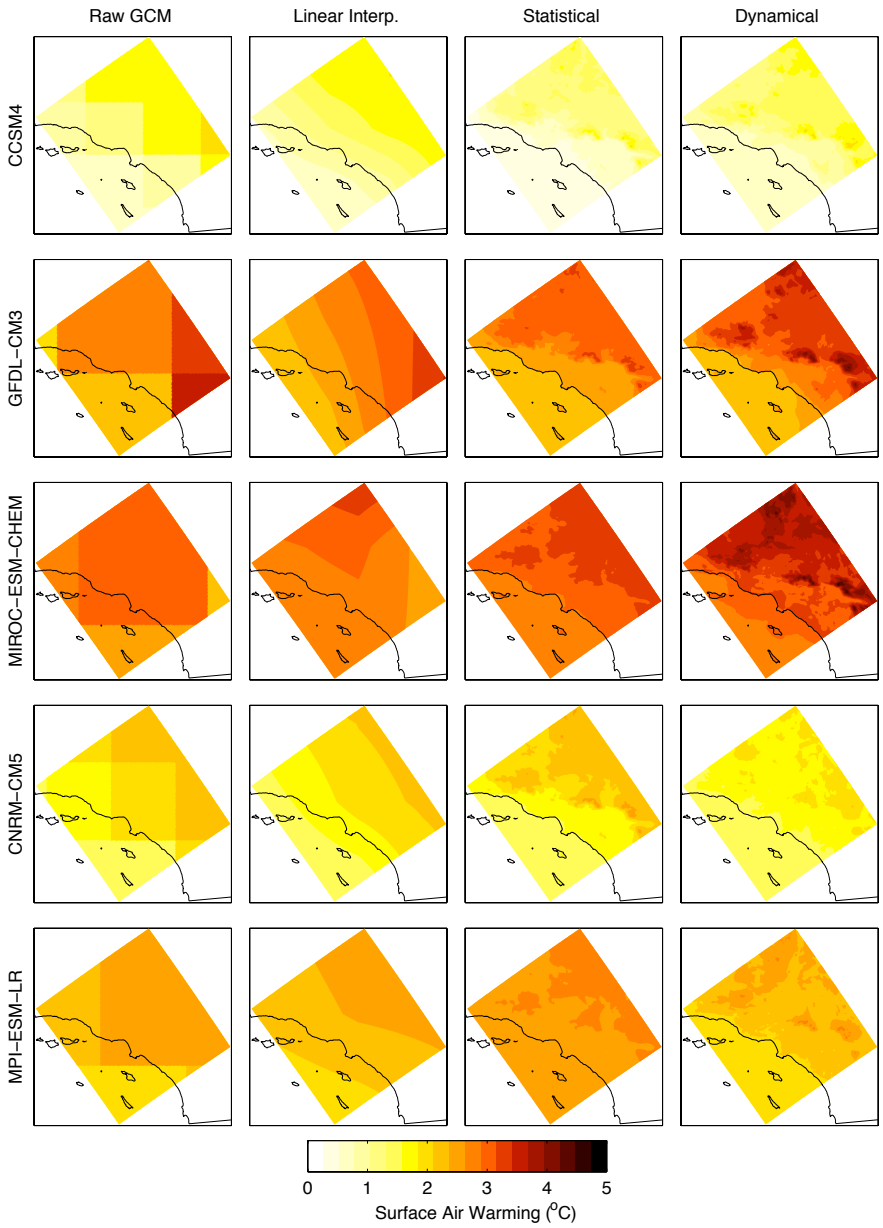

Figure 2.8: Annual-mean warming projections (units: °C) for five GCMs produced by four different methods: nearest GCM grid box (first column), linear interpolation of GCM (second column), statistical downscaling with the hybrid technique (third column), and dynamical downscaling (fourth column). Projections are for mid-century (2041-2060) relative to the baseline period(1981-2000) under the RCP8.5 scenario. The statistically generated pattern for a particular GCM was created using the hybrid technique trained on the other four GCMs.

Table 2.2: Comparison of the spatial correlation and root mean squared error (RMSE) for the raw GCM, linear interpolated and the statistically downscaled warming patterns, relative to the dynamically downscaled warming. By virtue of their orthogonality, errors in regional mean and spatial pattern are shown separately.

	Annual Error			Monthly Error		
Method	Spatial Correlation	Spatial RMSE (°C)	Regional-Mean RMSE (°C)	Spatial Correlation	Spatial RMSE (°C)	Regional-Mean RMSE (°C)
Raw GCM	0.64	0.26	0.21	0.49	0.35	0.26
Linear Interpolation	0.79	0.20	0.22	0.60	0.29	0.26
Statistical Model	**0.95**	**0.14**	0.23	**0.67**	**0.24**	0.27

The statistically downscaled warming patterns are clearly the 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 RMSE (divided into errors in the regional mean and errors in the spatial pattern), shown in Table 2.2. The average spatial correlation between the statistically downscaled annual-mean warming patterns and the dynamically downscaled patterns is 0.95, compared to 0.79 and 0.64 for the linear interpolated and raw GCM warming patterns, respectively. This demonstrates that the statistical method is superior to linear interpolation at predicting the spatial variations and sharp gradients in warming. For the monthly average patterns, the statistical model also provides added value over linear interpolation. The value added is somewhat smaller because the statistical model dials up or down only one spatially varying pattern (the CIP), while each month has a slightly different characteristic spatial pattern (Fig. 2.3).

The second metric is root mean squared error (RMSE). For the annual warming, the statistical model adds value by capturing the spatial variations. The statistical model's spatial error is 0.14°C, which is a substantial improvement over linear interpolation and the raw GCM values of 0.20 °C and 0.26°C, respectively. The statistical model is also an improvement over interpolation of the monthly patterns, though the improvement is

somewhat smaller. Again, this is likely due to the simplicity of using a single spatial pattern for all calendar months. We experimented with using different spatial patterns for each month. However, the gains in accuracy due to 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 statistical model provides no added value in predicting the regional mean warming because the GCM warming averaged over the innermost domain is already a good predictor of the dynamical downscaled regional mean.

The biggest advantage of the statistical model comes when we consider the ensemble-mean annual-mean warming. As we have seen, the statistically downscaled, linearly interpolated, and raw GCM warming patterns all have biases relative to dynamical downscaling. However, when we aggregate the approximate warming patterns into a five-model ensemble, the statistical model's errors cancel out, while those from the other methods do not (Fig. 2.9). In fact, the statistically downscaled ensemble-mean is nearly an unbiased estimator of the dynamically downscaled ensemble mean. The only bias is a slight one at the highest elevations. In contrast, the other two methods have systematic biases as large as 1 °C in magnitude. These 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. Thus, there are large swathes of the region where the statistical model is necessary to provide an accurate characterization of the most likely warming outcome.

We note that the error estimates in Table 2.2 and the patterns in Figs. 2.8 and 2.9 are based on the statistical model built on only four GCMs and their associated regional warming patterns. Since each GCM has a unique combination of regional mean and land-sea contrast (Fig. 10), 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 2.3.3, produces results of even higher quality. In fact, using linear regression guarantees that the statistically and

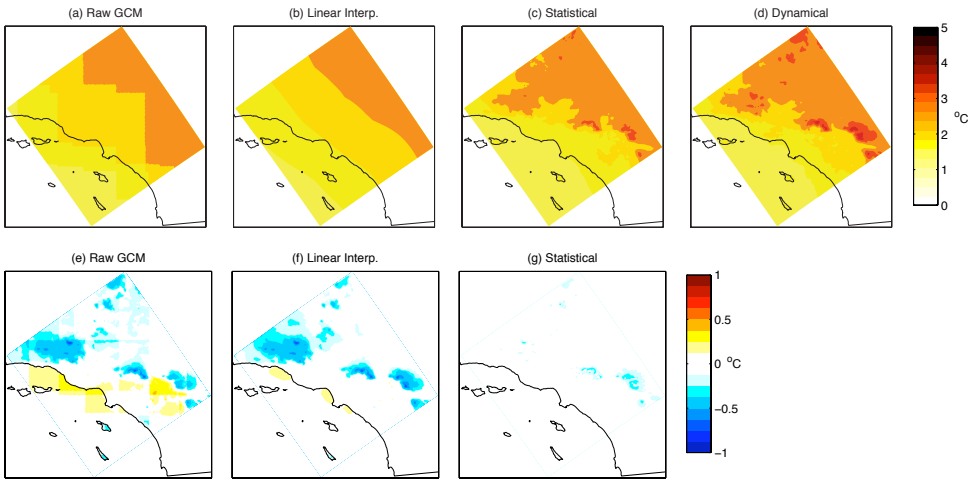

Figure 2.9: Annual-mean warming (°C) averaged over five GCMs downscaled using four different methods: (a) nearest GCM grid box, (b) linear interpolation of GCM, (c) statistical downscaling with hybrid technique, (d) dynamical downscaling. Bias of first three methods relative to dynamical downscaling (°C) shown in (e)-(g).

dynamically downscaled mean warming match, causing the biases in the ensemble-mean annual-mean warming to be negligible. The final statistical model is used to generate the results discussed from Section 2.4 onward.

2.4 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. 2.10). 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. 2.10, 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. 2.11. There is considerable variation among these warming patterns, underscoring the

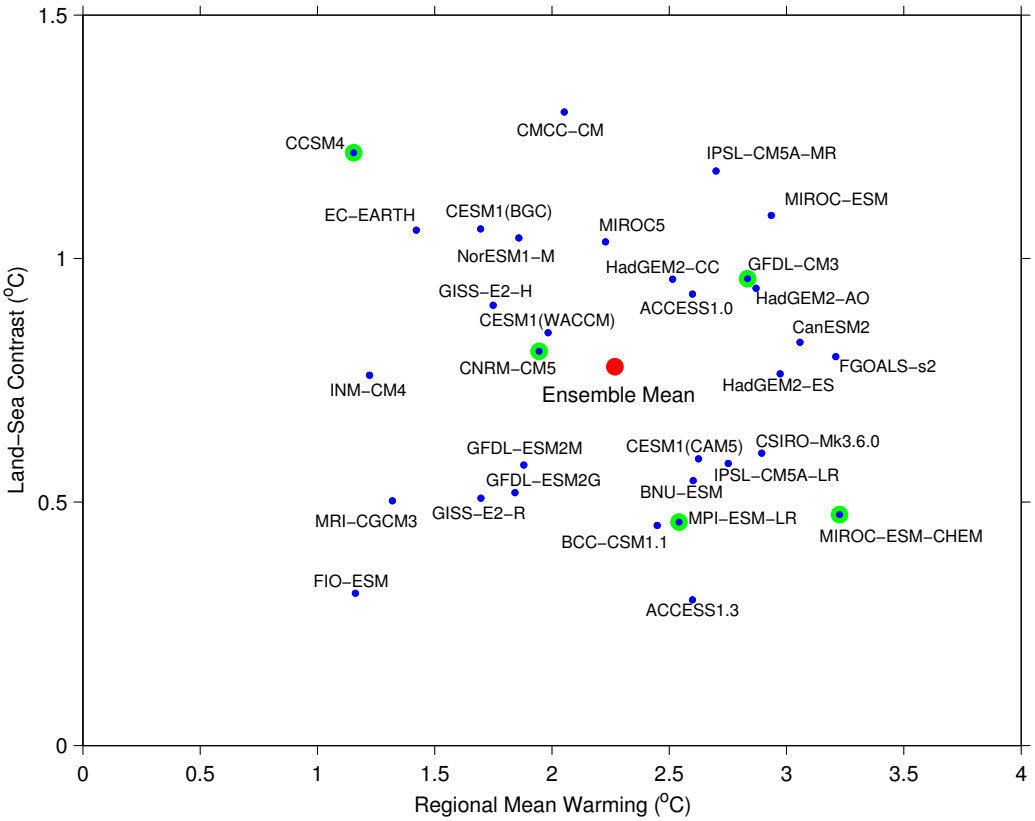

Figure 2.10: Annual-mean values of regional-mean warming and land-sea contrast (°C) for each GCM (blue dots) with the ensemble mean (red dot). The five GCMs that are also dynamically downscaled are highlighted in green.

importance of considering multiple global models 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. 2.12. 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 inter-model 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. 2.13). The average August difference between the inland and coastal areas is 0.6 °C , with certain locations

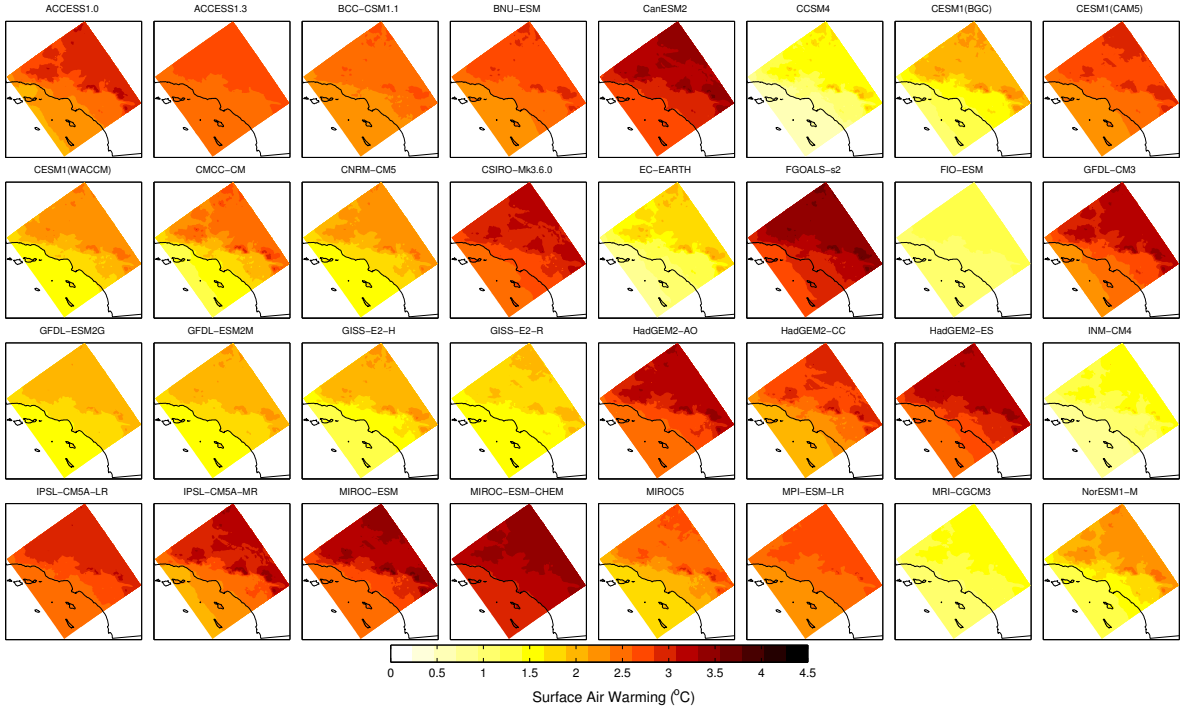

Figure 2.11: Annual-mean warming patterns (°C) generated by applying the statistical model to all 32 GCMs. Warming patterns are shown for the mid-century period (2041-2060) relative to the baseline period(1981-2000), under the RCP8.5 scenario.

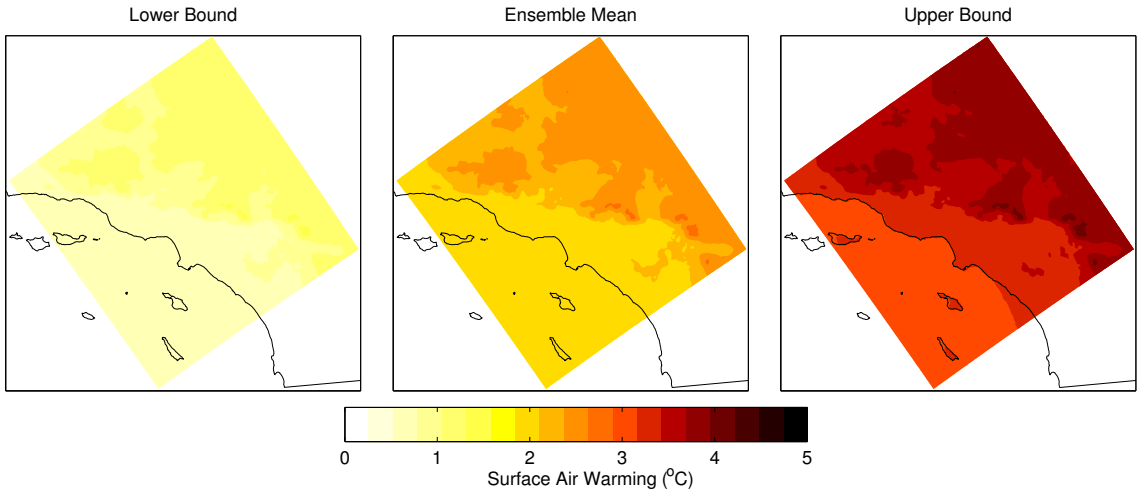

Figure 2.12: . Ensemble-mean annual-mean warming and upper and lower bounds (°C), based on a 95% confidence interval, for 32 statistically downscaled GCMs run with the RCP8.5 scenario.

showing warming elevated above the coastal values by as much as 1.2 °C (+62%).

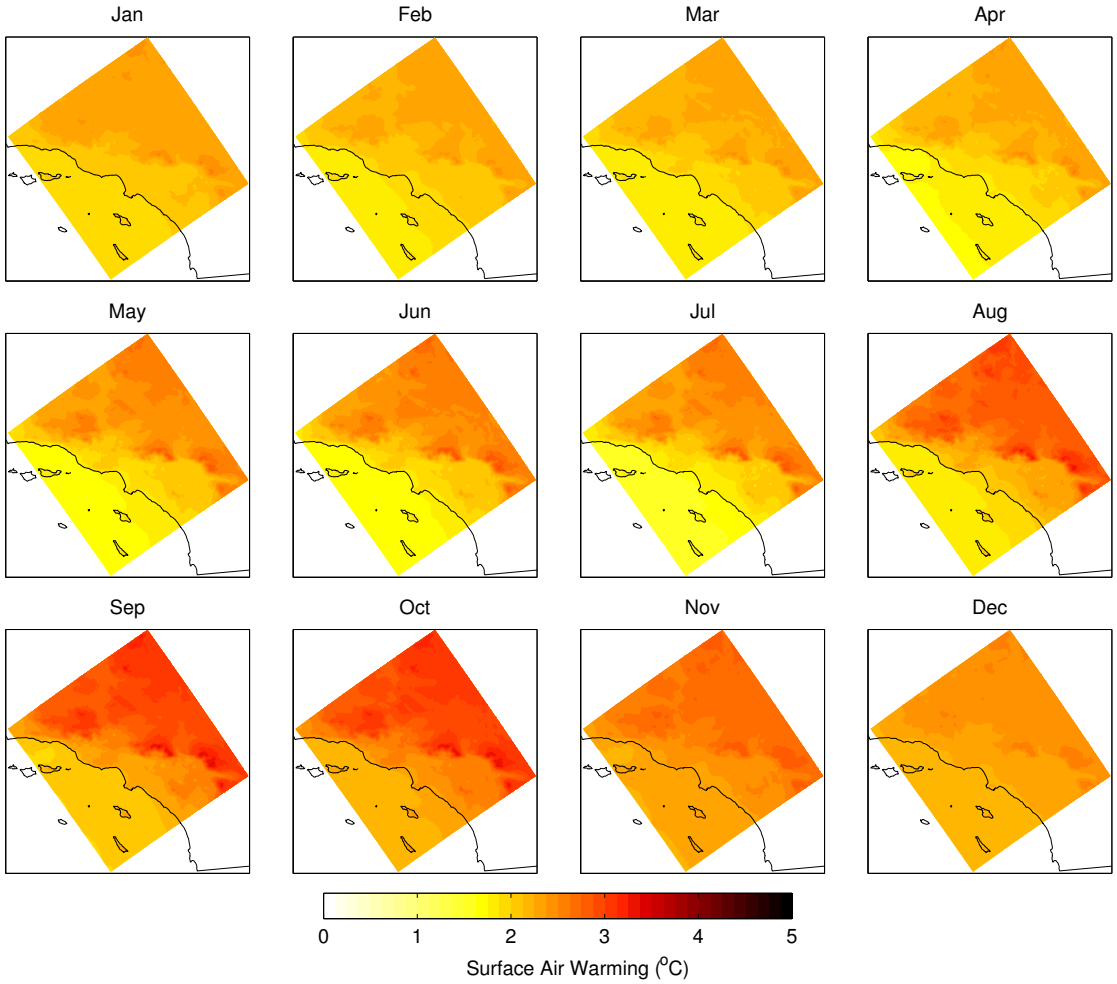

Figure 2.13: Ensemble-mean monthly-mean warming (°C) computed by averaging the monthly statistically downscaled warming patterns over 32 CMIP5 GCMs.

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 (compare Figs. 2.3 and 2.13), because the statistical model underestimates some warming due to snow-albedo feedback. Based on comparisons between the dynamically and statistically downscaled warming patterns for spring (MAM), the springtime ensemble-mean warming would be as much as 0.5 °C or larger in the San Bernardino and San Gabriel Ranges. This is consistent with the bias in the statistical model seen at the highest elevations (Fig. 2.9g).

2.5 Discussion and Conclusions

In this paper, we present a hybrid dynamical-statistical method to downscale the mid-century warming signal in 32 CMIP5 GCMs. First, we used dynamical downscaling to produce warming patterns associated with five GCMs. Then, to save computational resources, a statistical model was built that scales the characteristic dynamically derived patterns according to the regional warming and land-sea contrast sampled from the global model. This statistical model was then used to approximate the warming that would result if the remaining global models were dynamically downscaled. The ensemble-mean regional-mean warming was projected to be approximately 2.3 °C , with 95% confidence that the warming is between 1.0 °C and 3.6 °C . Thus, the inter-model differences in the GCM outcomes create significant uncertainty in projections of warming over Southern California.

In this hybrid method, statistical downscaling is employed a unique way. First, while statistical models typically relate GCM output to observations, ours relates GCM output to dynamically downscaled output. This is because our statistical model is designed to be an approximate dynamical model. The differences between the dynamically and statistically downscaled patterns are an order of magnitude smaller than inter-model variations in the warming. This means that our statistical downscaling projections for the ensemble-mean warming and spread are reasonable approximations to the projections that would have resulted from dynamically downscaling all 32 models.

The second difference is that our 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 stationarity concerns that the relationship between predictor and predictand may not hold in the future period (Wilby and Wigley 1997). 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—that the remaining GCMs have values of mid-century 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.

The statistical model adds value by capturing the fine-scale spatial variations in the warming. Inland and mountain locations are expected to warm up considerably more than coastal areas, especially during the summer months. When we compared the statistically downscaled patterns to the raw and linearly interpolated GCM patterns, the statistical model captured these spatial variations much more accurately. Furthermore, when we take an ensemble average of the warming patterns, the errors in the statistically downscaled patterns nearly cancel out, revealing only minor biases. In contrast, the raw and linear interpolated GCM warming patterns have large systematic biases, especially along the coastline and in topographically complex regions that are not resolved well in the GCMs. The statistical model does not improve upon the GCM estimates of regional mean warming estimates, because the GCM warming averaged over our innermost domain is already a good predictor of the dynamically downscaled regional mean.

Another 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, our method first employs dynamical downscaling, which allows us to identify two important physical mechanisms controlling the warming. The first is that the local atmospheric circulation leads to warming over the coastal ocean similar to that seen over the ocean in GCMs, warming over the coastal zone that is slightly elevated above the ocean values, and much higher warming over inland areas separated from the coast by mountain complexes. The second mechanism, smaller in spatial scope, is snow-albedo feedback, which leads to enhanced warming in the

mountains. With this knowledge, we built a statistical model that scales the characteristic spatial pattern (which contains 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 northeast Pacific Ocean over the coming decades, as is likely if GCM projections are correct. Given the realistic behavior of the WRF model in distributing humidity and temperature across the landscape, it seems very likely that the associated warming pattern in the 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.

3 Application and Modification of the Hybrid Method to a Larger Domain Covering All of California

3.1 Introduction

The previous chapter focuses on the development of a new hybrid downscaling method. This method employs a region-specific statistical model that is constructed from the spatial patterns found in dynamical downscaling, and the relationships between those patterns and the GCM climate change signal. Broad questions remain about the portability of this method. Can it be effectively applied to other domains? If so, will the statistical model rely on only a single spatial mode?

Here, the hybrid method is applied to a much larger domain encompassing the entire state of California and the adjacent ocean. Although their sizes differ, the full California domain also contains both land and ocean areas. The work of Di Luca et al. (2012) suggests that in these coastal regions that are roughly half-ocean, half-land, the contrast in warming between the two portions is the dominant contributor to spatial variations. Thus the expectation is that the first principal component will be associated with the land-sea contrast. However, the Los Angeles domain spanned a width of approximately 280 km in either direction, which is covered by only a handful of grid cells for many GCMs. Because the California domain covers a larger area, the corresponding warming patterns in the GCMs likely show considerably more spatial variations, which could require multiple large-scale modes, in addition to the land-sea dipole. Furthermore, a larger domain could include a wider variety of local effects, like snow-albedo feedback in the Sierra Nevada mountains and low clouds effects along the coast, that might not both be present in a more localized domain. This too could require the use of multiple spatial modes, especially for local effects that are not tied with the land-sea contrast. Also explored here is the effect of excluding the ocean in the principal component analysis. Since the majority of the spatial details of interest are over the land areas, principal component analysis is repeated for the land areas only.

This chapter also attempts to improve the hybrid method. One aspect missing from the application of the method to the Los Angeles Region was a thorough investigation of the sensitivity of the statistical model to the exact choice of warming patterns used to train it. Here Monte-Carlo simulations are performed to as part of a sensitivity analysis of the optimal sample locations for the large-scale predictors and the statistical model coefficients. Other small modifications were made in order to make the method more objective.

Layout of this chapter is as follows. In Section 3.2, the dynamically downscaled data that forms the basis for the statistical model is explored. The baseline simulation is evaluated against a gridded observational product, and PCA is performed on the warming patterns. Section 3.3 covers the construction and evaluation of the statistical model. The statistical model is then applied to the 32 GCMs listed in Table 2.1. Finally, in Section 3.4, we discuss the applicability of the method, characterize the regions for which it would best-suited, and make recommendations regarding modifications that might be necessary.

3.2 Dynamical Downscaling

The hybrid statistical model is based on spatial patterns and relationships determined from dynamical downscaling. In this application, our dynamically downscaled data comes from simulations using the Weather Research and Forecasting (WRF) regional model applied to a domain encompassing the state of California.

3.2.1 WRF Model Setup

WRF version 3.5 (Skamarock et al. 2008) was run over two nested domains of 27 km and 9 km horizontal resolution (Fig 3.1). WRF was coupled to the Noah-MP land surface model version 1.1 (Niu et al. 2011) to simulate land surface processes, including vegetation, soil, snowpack, and exchange of energy, momentum, and moisture between the land and atmosphere. Based on sensitivity experiments the following parameterizations were chosen: Thompson's MP microphysics scheme (Thompson et al., 2008); Kain-Fritsch cumulus scheme

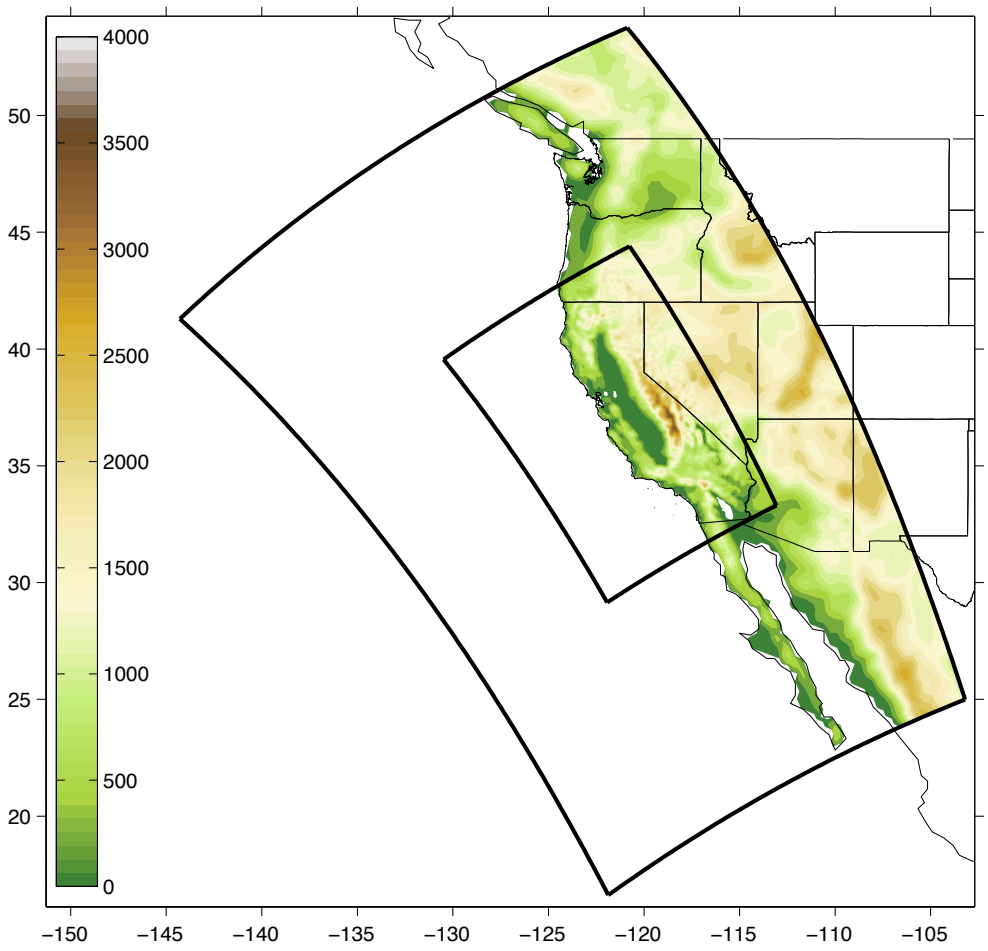

Figure 3.1: Boundaries (thick black lines) and topography (color) of WRF domains 1 and 2. The resolutions of domains 1 and 2 are 18 km and 6 km, respectively. Coastline and political boundaries are also included (thin black lines).

(Kain, 2004); MYNN planetary boundary layer scheme (Nakanishi and Niino 2004).

This WRF setup was used to perform multiple simulations, following the pseudo-global warming approach (PGW; see Rasmussen et al. 2011). The first simulation (the "baseline" simulation) is a reconstruction of historical climate over the 1981-2000 period. The National Center for Environmental Prediction's North American Regional Reanalysis (NARR; Mesinger et al. 2006) of the 1981-2000 period was used for the lateral and surface boundary conditions for the baseline simulation. This simulation can be directly compared to observations of the historical climate to evaluate model skill in reproducing temperature climatology and variability (Section 3.2.2). The second set of simulations (the "future"

simulations), represent predictions of the future climate over the 2081-2100 period, based on the climate change signals present in four GCMs: CNRM-CM5, GFDL-CM3, INMCM4, and IPSL-CM5A-LR (see Table 2.1 for details). Boundary forcing for each future simulation is generated by taking the same NARR data corresponding to the 1981-2000 period and adding on the change in the GCM monthly climatology between the 2081-2100 and 1981-2000 periods. The added perturbation includes changes in 3D temperature, winds, relative humidity, and geopotential height. The details of the calculation of the perturbation are the same as in the Los Angeles case (see Section 2.2.3).

Spectral nudging was applied above the boundary layer (roughly above 800 hPa) in both the baseline and future runs. Zonal and meridional winds, temperature, and geopotential height were nudged back to NARR values in the baseline simulation, and to NARR plus the change in GCM monthly climatology in the future simulations. Nudging was only used in the outer domain and only to waves with wavelength greater than 200 km. Nudging was performed every 360 seconds, with a nudging coefficient of .0003 s^{-1}.

3.2.2 Baseline Temperature Validation

To evaluate the skill of our dynamical model in reproducing baseline temperature climatology and variations, we compare the baseline simulation to the Livneh et al. (2013; henceforth L13) observationally-based gridded product. The L13 1/16-degree dataset is an update to the Maurer et al. (2002) gridded product based on data from NOAA Cooperative Observer (COOP) stations from around the country. L13 data include daily maximum (Tmax) and minimum (Tmin) temperatures, but not daily average temperatures (Tavg). For comparison purposes, Tmax and Tmin are averaged over each month to create monthly-average Tavg values. L13 Tavg was resampled (using nearest-neighbor) to the WRF 9km resolution grid. (Note: L13 data is only available over land areas.) Snapshots of WRF temperatures were recorded at 3-hourly intervals for the beginning portion of the simulation. This temporal resolution was determined to be too low to properly resolve Tmax and Tmin values. So,

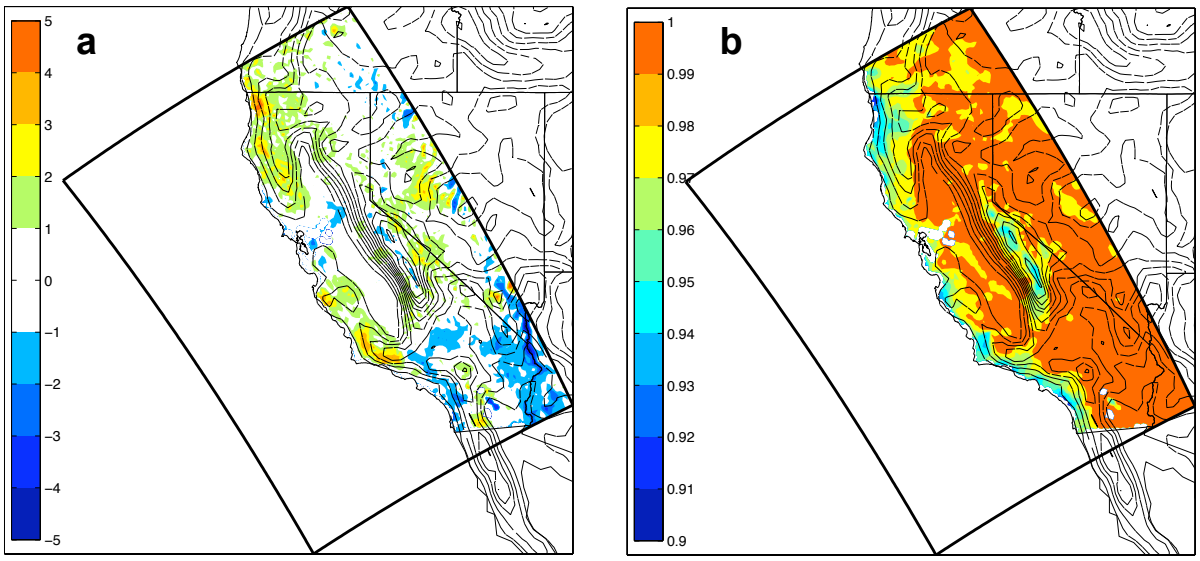

Figure 3.2: Evaluation of model mean climatology and temporal variability versus L13 gridded product over the period 1995-1999. (a) Difference in annual-mean surface air temperature climatology (WRF minus L13, units: °C). (b) Correlation between monthly-average temperature anomalies. (Anomalies are relative to the composite monthly-mean climatology). Low-resolution (27 km) topography also shown (black contours).

beginning in year 1995, snapshots were recorded at 30 minute intervals to allow for more accurate calculations of Tmax and Tmin. Thus our comparison is limited to the years 1995-2000. WRF daily Tmax and Tmin values were averaged to create monthly Tavg values.

First, we examine the skill of WRF in reproducing the 1995-2000 annual-mean Tavg climatology. The majority of land areas in our domain have WRF climatological Tavg values within 1 °C of L13 values (Fig 3.2a). The largest discrepancies are in (1) the coastal mountains, where WRF produces baseline Tavg values 1 to 3 °C larger than L13; and (2) the Colorado River Basin (along the border with Arizona), where WRF produces Tavg values as much 4 °C colder than L13. These differences are generally in areas with complex topography. We suspect that differences are due either to the difficulty resolving complex features in WRF or problems associated with creating a gridded temperature field in complex terrain from sparse observations.

Monthly Tavg anomalies were computed relative to the 1995-2000 monthly climatology in order to examine temporal variability. Correlations of WRF and L13 Tavg anomalies are above 0.9 for all locations, with many locations inland of the major coastal mountain ranges having correlations over 0.99 (Fig 3.2b). These universally high correlations suggest that WRF successfully captures temporal variations in the region. Correlations are slightly lower along the coast and in the Sierra Nevada mountains. In the coastal zone, WRF may have difficulty simulating variability in low clouds or the marine layer. In the Sierra Nevada, temperature variations may depend on snowpack size, area, and persistence, which are complicated and challenging to model.

Overall, WRF agrees well with L13 in both climatology and variability during the baseline period. Agreement is excellent over interior land areas. Agreement is still good in the coastal zone and mountain areas, where complex topography and challenging dynamical processes are present. This comparison with gridded data gives us confidence that when a realistic signal is imposed at the boundaries, WRF produces the appropriate response throughout the domain.

3.2.3 Temperature Change Patterns

The future WRF simulations representing the 2081-2100 period are forced at the boundaries by 1981-2000 NARR data plus the change in monthly GCM climatology between the two time periods. Therefore it is likely that temperature change signal computed by differencing the future and baseline WRF climatologies will be similar to that of the GCM climatological difference. Figure 3.3 shows the linearly interpolated monthly surface air temperature change in each GCM. Clear patterns exist across the GCMs, including the land-sea contrast in the warming discussed at length in the previous chapter. Land-sea warming contrast is also present in the WRF dynamically downscaled monthly-mean warming (Fig 3.4). While the broad warming patterns are often similar, WRF is not completely constrained to follow the warming pattern of the GCM. WRF provides its own treatment of local effects. For example,

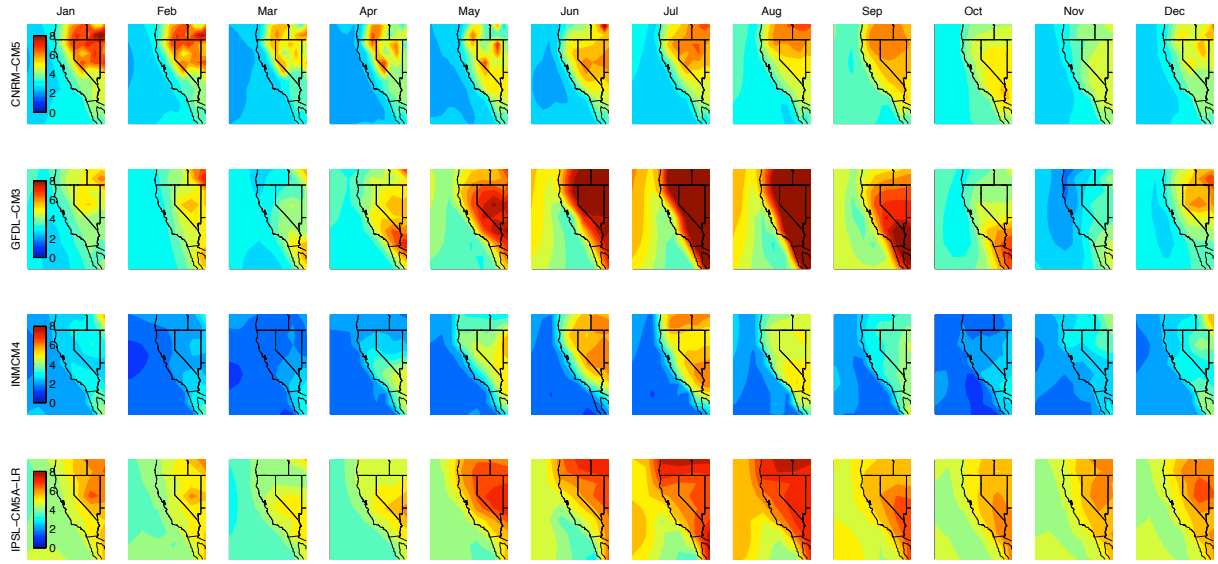

Figure 3.3: Monthly surface air temperature climatology difference (°C) between the future (2081-2100) and baseline (1981-2000) periods for five GCMs.

the large patches of snow-albedo feedback warming enhancement found in CNRM-CM5 in March, April, and May are not present at the same location and strength in the dynamically downscaled simulation. WRF does predict warming enhancement due to reduced snow cover in May and June, but the magnitude is much less than suggested by the GCM.

Principal component analysis (PCA) was performed on the 48 monthly spatial anomalies (relative to regional mean) to determine the preferred spatial modes resulting from dynamical downscaling. Typically PCA is performed on temporal anomalies, resulting in principal components that are time series and associated loadings that form a spatial pattern. Here PCA is performed on the spatial anomalies, where the principal components are spatial patterns, and the loadings are functions of month and GCM. (Note that grid cells nearest to the boundary of the domain have been removed in order to prevent un-physical boundary effects from unduly affecting the results of PCA.) Figure 3.5 shows the three largest principal components (PCs) and their associated loadings, as well as the regional-mean warming for each GCM and month. On average, the regional mean warming is highest in late summer and early fall. The first principal component (PC1) explains 89% of the spatial variance and

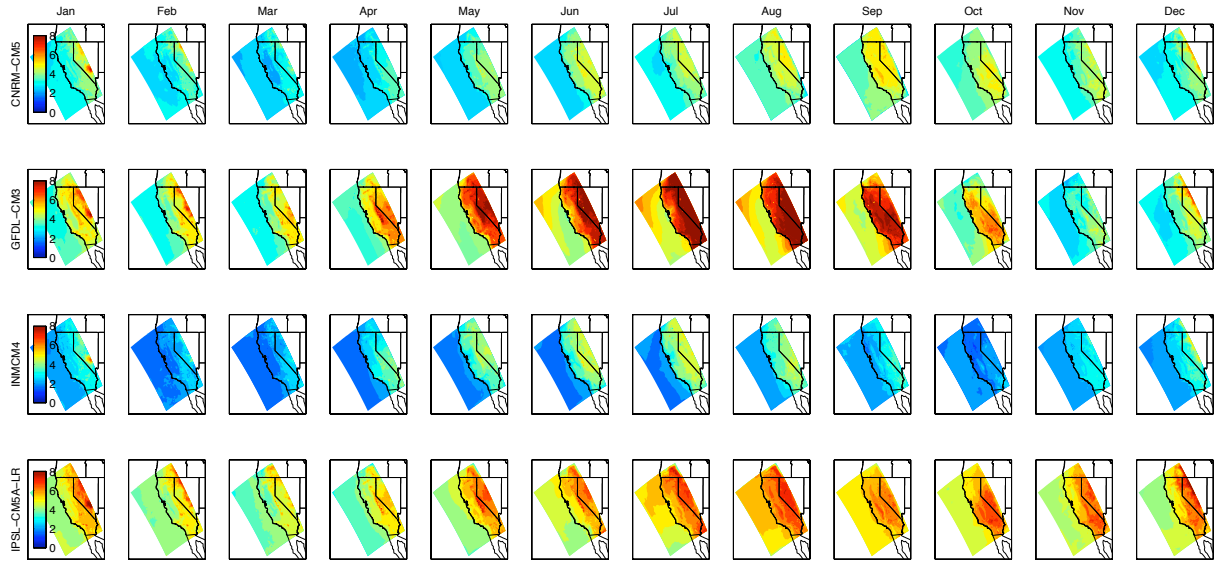

Figure 3.4: Monthly surface air temperature climatology changes (°C) between the future (2081-2100) and baseline (1981-2000) periods for four dynamically downscaled GCMs. Results shown for domain 2.

is clearly the most important spatial pattern. Land-sea contrast is the primary feature of PC1, which is consistent with the results for the Los Angeles Region, which used a domain comprised of roughly equal parts land and ocean. The remaining PCs account for only a small fraction of the variance (3.5% and 2.8%, respectively), but could very well represent physical modes of spatial variability in the climate change patterns. PC2 and PC3 likely reflect how large-scale gradients in the GCM warming patterns oriented over land areas are expressed in the presence of the fine-scale topography.

3.3 Hybrid Statistical Model

Statistical downscaling relies on mathematical relationships between large-scale predictors and fine-scale predictands. In this domain, the fine-scale warming can be distilled down to two predictands: the regional mean warming and the PC1 loading (a proxy for the land-sea warming contrast). The statistical model approximates the dynamical warming patterns as a linear combination of the regional mean warming and the land-sea contrast multiplied

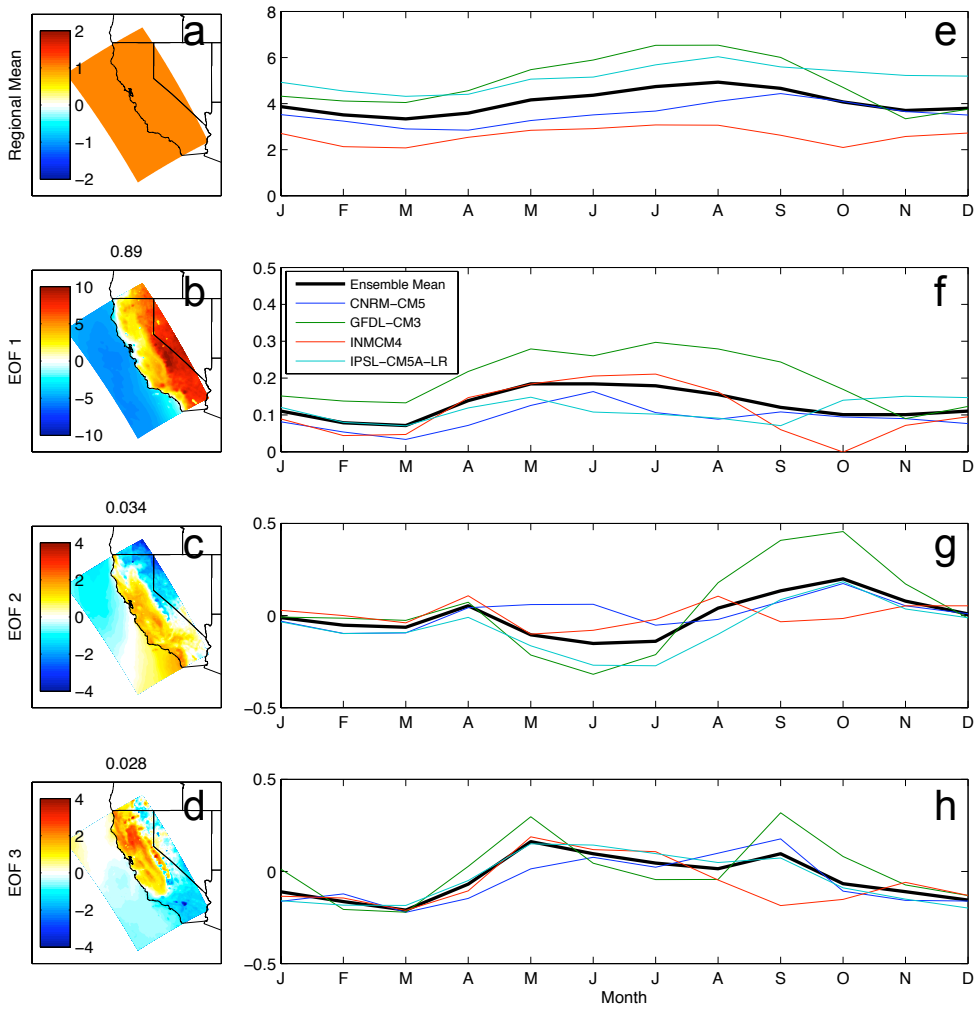

Figure 3.5: Regional-mean warming (a) spatial pattern and (e) values for each month and dynamically downscaled GCM (colored lines – individual GCMs; solid black line – four-model mean). (b-d) First three principal components, and (f-h) associated loadings, based on PCA of 48 monthly warming patterns.

by PC1. The values of the dynamically downscaled regional-mean warming and land-sea warming contrast generally appear to be inherited from the large-scale GCM warming pattern (compare Fig 3.3 and Fig 3.4). Thus it makes sense to sample the linearly interpolated GCM warming in order to estimate the dynamically downscaled regional mean warming and land-sea warming contrast. Once these sample locations have been identified, the statistical model can be configured.

3.3.1 Finding Optimal Sample Locations

The monthly warming patterns for each GCM were linearly interpolated to a 1° x 1° grid and then correlated with values of the dynamically downscaled regional mean warming (Fig 3.6a,d). Correlations are highest over the ocean west of our domain, though a number of remote areas also show high correlations. The highest correlated point, [129° W, 41° N], is used as the sample location for the regional mean warming. Next, the linearly interpolated warming was correlated with the PC1 loadings. We expect that PC1 loadings can be predicted by the land-sea contrast in the GCM, which means eventually selecting two points, one over land and one over the ocean and using their difference as a the predictor. The first location was selected just based on the highest correlations of the linearly interpolated GCM warming and PC1 loadings, which were found over the Mojave Desert (Fig 3.6b,e), at [116° W, 35° N]. The highest correlated point is used as the inland warming sample point. To select the location to sample the ocean warming, partial correlations were computed between the linearly interpolated GCM warming and the PC1 loadings, with the effect of the warming sampled inland location removed (Fig 3.6c,f). The ocean warming is sampled at the point of strongest negative correlation, found at [124° W, 37° N], off the coast of California.

3.3.2 Model Formula

The statistical model approximates the dynamically downscaled warming as the sum of the regional mean warming term having no spatial variation and a land-sea contrast term where PC1 serves as the spatial pattern. Mathematically, the approximate warming at a location (i,j), for a given month and GCM, can be expressed as

$$\Delta T^{stat}(i,j) \;=\; a\Delta T^{GCM}_{RegMean} + b + PC1(i,j)\cdot(c\Delta T^{GCM}_{Inland} + d\Delta T^{GCM}_{Ocean} + e) \qquad (2)$$

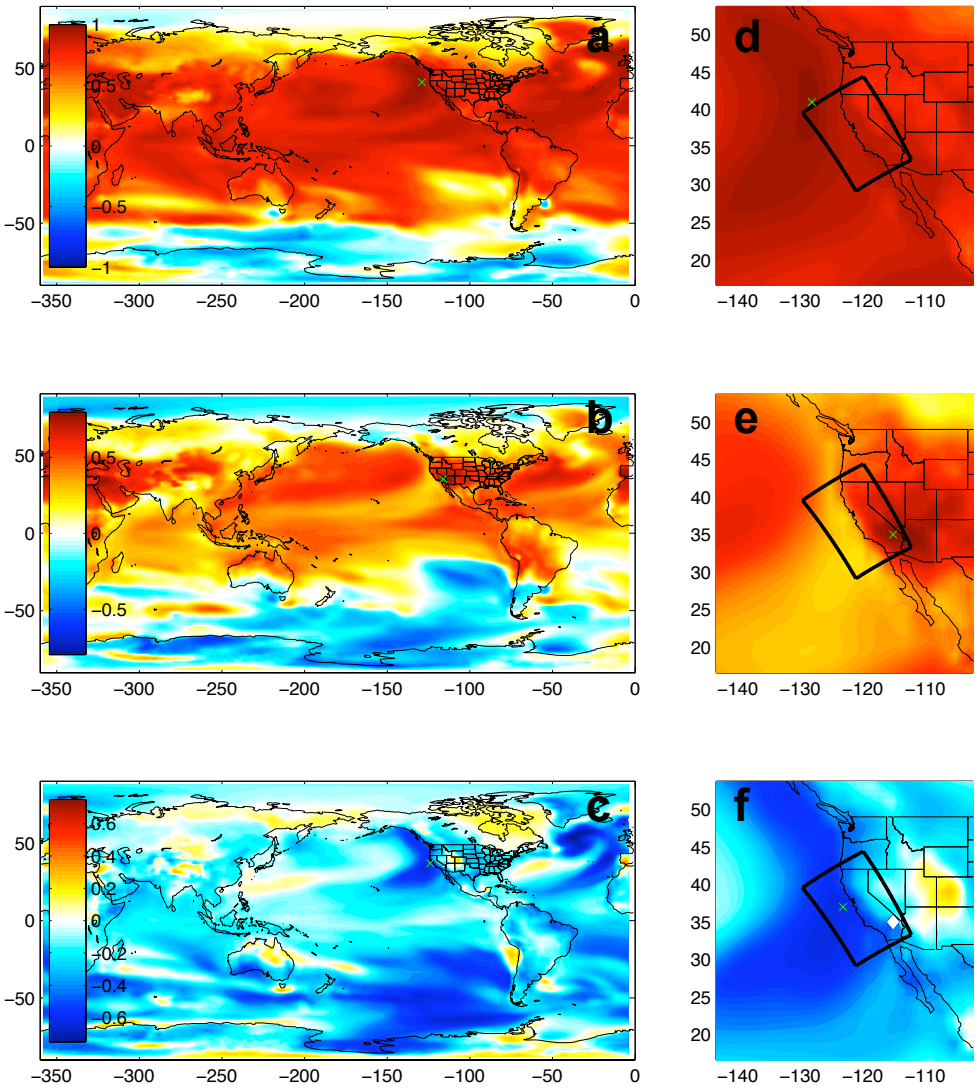

Figure 3.6: (a) Correlation between interpolated GCM monthly warming and dynamically downscaled monthly regional-mean warming (average taken over Domain 2). Regional-mean sampling location shown in green. (b) Correlation between interpolated GCM monthly warming and loadings associated with the first principal component. The highest correlated point (green) is used as the first sampling locaiton. (c) Partial correlations between interpolated GCM monthly warming and loadings associated with the first principal component, with the effect of the first sampling location removed. (d-f) Same as (a-c) but zoomed in to show detail.

where coefficients a-e are calculated from linear regression: $a = 1.02$, $b = 0.63\ ^{\circ}\mathrm{C}$, $c = 0.046$, $d = -0.031$, and $e = -0.014$.

3.3.3 Cross-Validation

Ideally, testing of the hybrid statistical model would involve comparing the dynamical downscaling warming of a new GCM to the warming predicted by applying the hybrid statistical model to that GCM's large-scale pattern. However, the main purpose of developing the hybrid technique is to limit the number of GCMs that need to be dynamically downscaled, since each one takes considerable computational resources. An alternative is leave-one-out cross-validation (LOOCV) where the hybrid statistical model is retrained using the downscaled data from all but one GCM, and is then used to statistically downscale the remaining one. If the hybrid statistical model has good predictive power, then the statistically downscaled warming pattern should be relatively close to the dynamically downscaled pattern. Unfortunately, there are only four total dynamically downscaled GCMs, and three GCMs may be too small of a sample size to properly train a functional statistical model. To avoid this possibility, we implement an alternative technique. A total of 500 Monte Carlo simulations were performed in which 36 of the 48 monthly dynamically downscaled warming patterns were selected at random (without replacement) and used to train the hybrid statistical model. Then the new statistical model was tested on the remaining 12 warming patterns. These Monte Carlo simulations serve two separate interests: (1) evaluating the sensitivity of PC1, the sample locations, and regression coefficients to the choice of months/GCMs included in statistical model calibration, and (2) assessing the predictive power of the statistical model.

3.3.4 Sensitivity of Model Components

The first principal component generated by PCA is remarkably insensitive to the choice of month/GCM used (not shown). This is because land-sea contrast accounts for the dominant spatial pattern in this region for nearly every month and GCM. Sample locations are also fairly consistently located (Fig 3.7). The point of highest correlation with the dynamically downscaled regional mean warming is the same for over 300 of 500 Monte Carlo simulations. The remaining 200 simulations have their highest correlations at adjacent grid points. Simi-

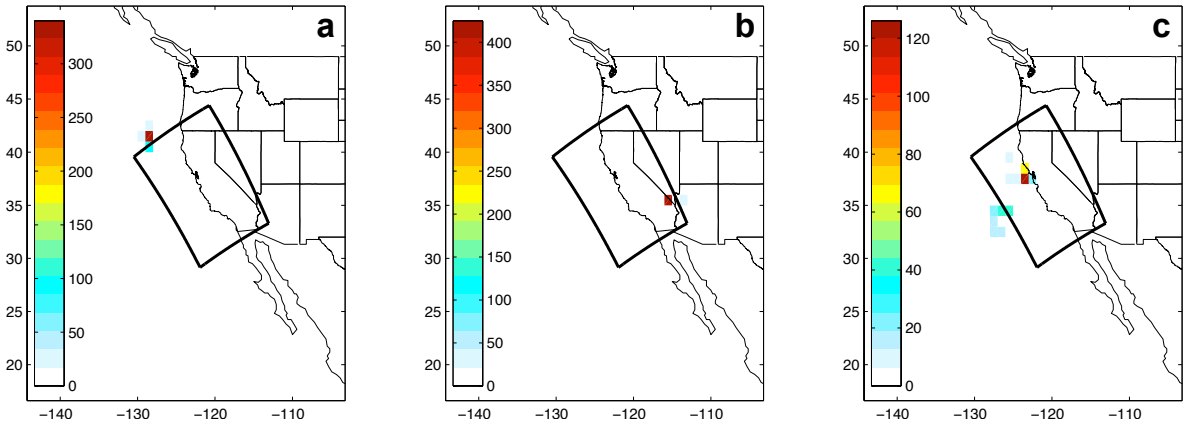

Figure 3.7: Frequency that the interpolated GCM warming at a given grid cell had (out of 500 total Monte Carlo simulations) (a) the highest correlation with the dynamically downscaled regional mean warming, (b) the highest correlation with the loadings associated with PC1, (c) the highest partial correlations with the loadings for PC1, with the effect of the warming at the point described in (b) removed.

larly, the point having highest correlation with the PC1 loadings (where the inland warming is sampled) is consistently located in the Mojave Desert, with over 400 of the simulations sharing the same highest correlated point. The point having highest partial correlations with PC1(where the ocean warming is sampled) shows considerably more variability, with many simulations showing values adjacent to the California coast, while others are farther out over the ocean, at the edge of the domain.

The coefficients generated via regression show different amounts of variation (Fig 3.8). The spreads are smallest in a and c (the scaling factors applied to the regional mean and inland warming), which is consistent with the fact that their associated sample locations are less variable. In contrast, d and e are dependent on sample locations of the inland warming, ocean warming, which widens their distributions.

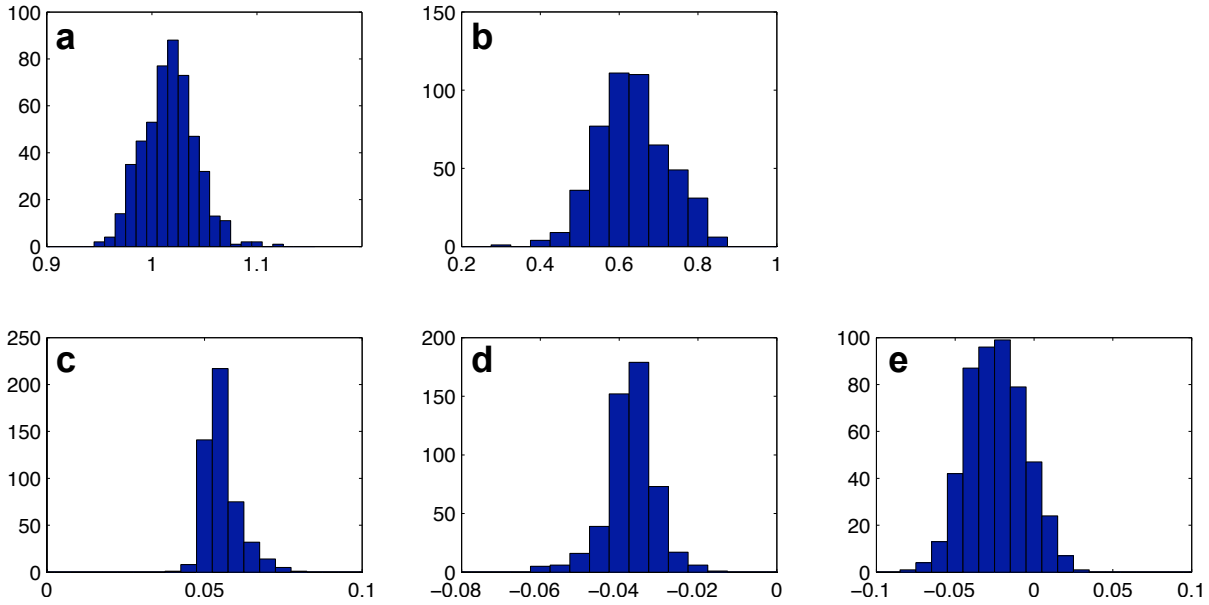

Figure 3.8: (a-e) Distribution of hybrid statistical model regression coefficients a-e in Equation 2 for 500 Monte Carlo simulations.

3.3.5 Predictive Power

Predictive power is assessed based on the mean absolute error (MAE) between the twelve dynamically and statistically generated warming patterns. Distributions of MAE are based on the 500 Monte Carlo simulations. Errors were subdivided into errors over land areas, over the ocean, over the entire domain. For comparison purposes, we include the errors for a naïve statistical downscaling technique: linear interpolation. Linear interpolation takes into account no high resolution spatial information and thus serves as a clear baseline upon which the hybrid statistical model should improve.

The hybrid statistical model improves over linear interpolation for land areas, with median values of MAE at 0.43 °C and 0.57 °C (Figs 3.9, 3.10). On the other hand, both the hybrid model and linear interpolation have lower, roughly equal errors over the ocean. The sea surface temperature (SST) at the ocean surface for WRF is set to NARR values in the baseline simulation and NARR plus the linearly interpolated GCM SST perturbation in future simulations. Since the surface air temperatures are closely tied to SSTs in both WRF and GCM, the WRF surface air warming pattern over the ocean is very similar to

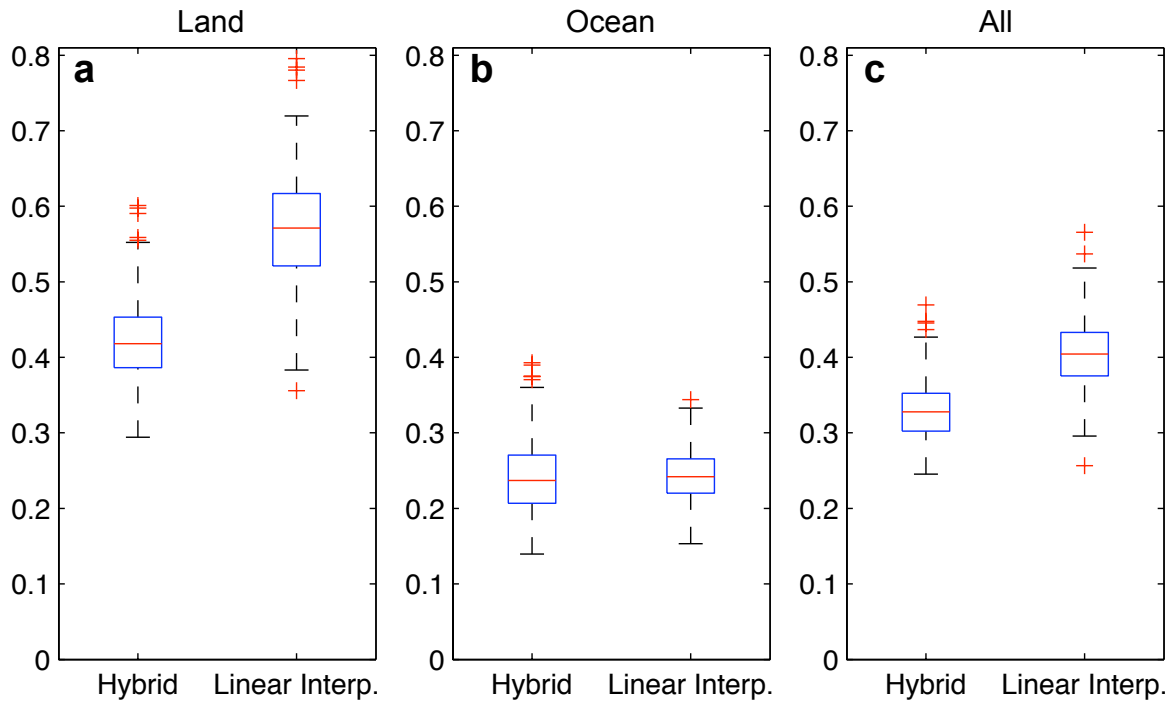

Figure 3.9: Box plots showing the distribution of mean absolute error for the hybrid statistical model and linear interpolating in capturing the dynamically downscaled warming over (a) land areas, (b) ocean, (c) all areas, based on 500 Monte Carlo simulations. Red line shows median, blue box contains the 25th to 75th percentile. All data is contained with the black bars except for the most extreme values, which are plotted individually.

the linearly interpolated GCM surface air warming pattern. Thus linear interpolation of the GCM pattern is already a good predictor over the ocean. (Although this is true for our particular WRF experimental design, regional climate change experiments with coupled ocean and atmosphere models would probably find more complex, realistic structure in the ocean warming fields, and it is less likely that linear interpolation would still be as accurate.)

3.3.6 Ensemble Warming Projections

The final hybrid statistical model – calibrated using all 48 dynamically downscaled monthly warming patterns – was applied to 32 CMIP5 GCMs run under the RCP8.5 scenario (full details on the GCM included in this ensemble can be found in Table 1). The resulting annual-

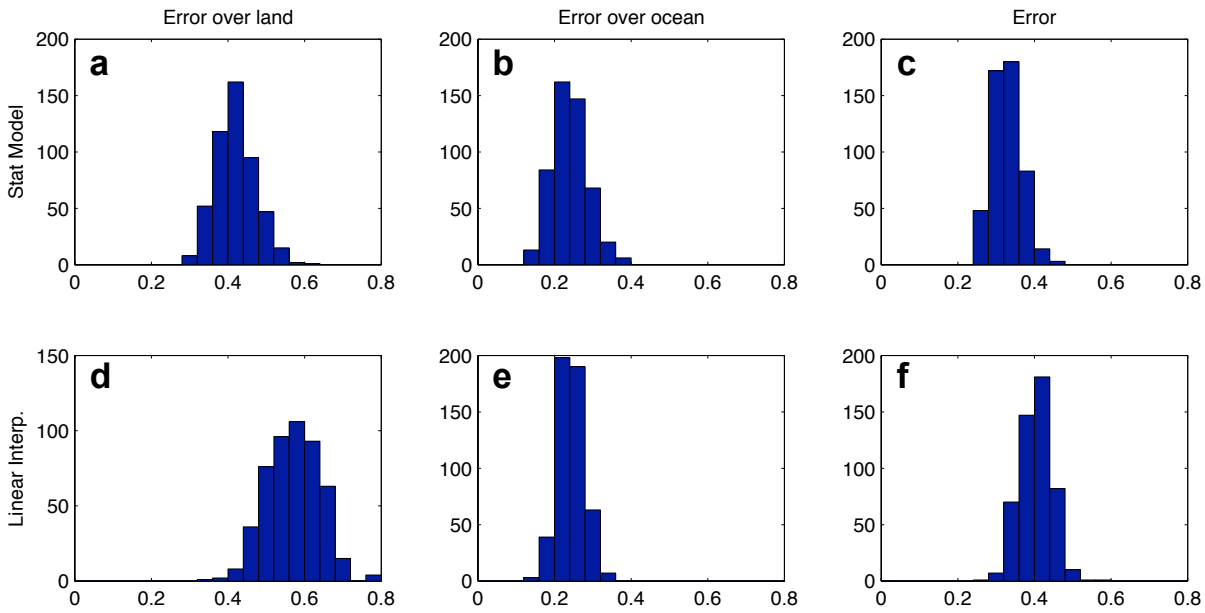

Figure 3.10: Histograms showing the distribution of mean absolute errors (units: °C)relative to dynamical downscaling from 500 Monte Carlo simulations of hybrid statistical model output (a) over land areas, (b) over the ocean, (c) over the entire domain. (d-f) same as (a-c) but for linear interpolation.

mean warming patterns (2081-2100 minus 1981-2000) show a variety of regional mean and land-sea contrast values (Fig 3.11). We expect that the true annual-mean patterns would have more spatial heterogeneity than those presented, because we only allow for one spatial degree of freedom in the statistical model. However, PCA shows that a single degree of freedom could account for 89% of the spatial variance in the monthly warming patterns of dynamically downscaled GCMs. This gives us confidence that the statistical model captures the most important features.

The purpose of performing ensemble mean projections is to produce probabilistic estimates of most likely outcomes and their uncertainty. The statistical model is nearly an unbiased estimator of the ensemble-mean annual-mean warming for the four dynamically downscaled models (Fig 3.12). In contrast, linear interpolation has systematic biases where the coastline and complex terrain create fine-scale warming variations that cannot be captured by the GCM. This leads us to believe that the statistical model's prediction of

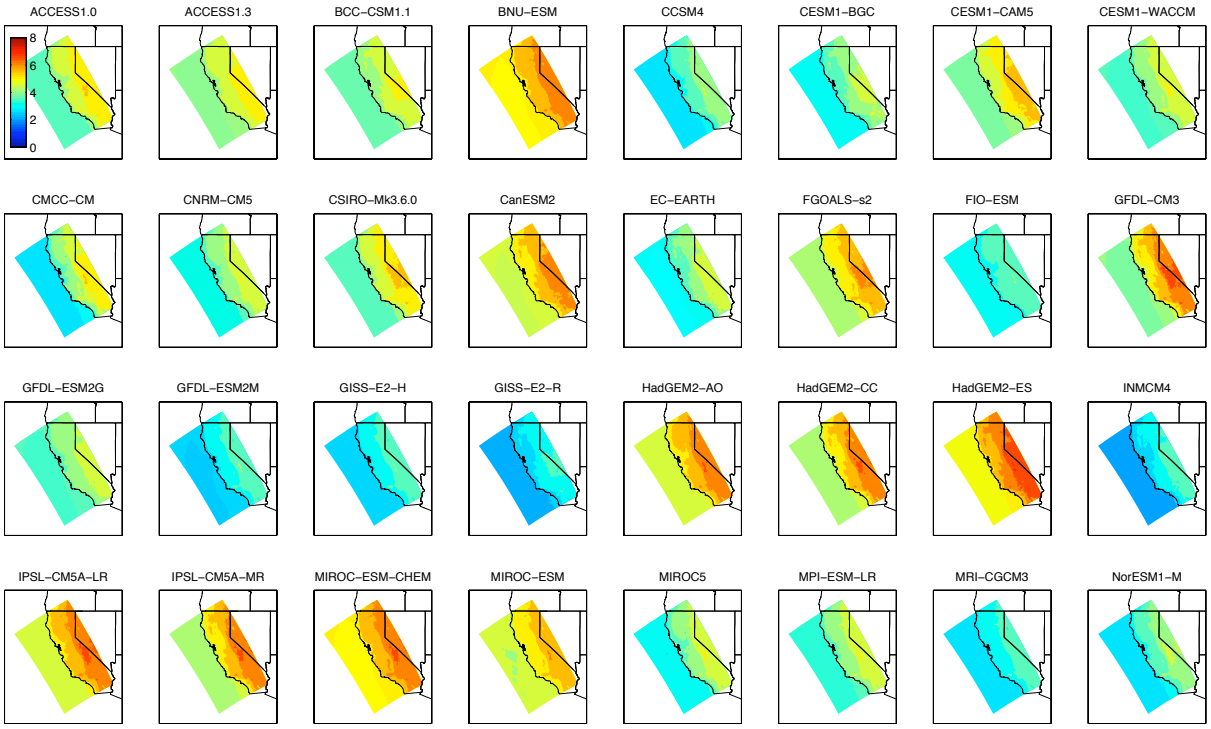

Figure 3.11: Annual-mean warming patterns (2081-2100 minus 1981-2000) produced by the hybrid statistical model for 32 CMIP5 GCMs.

the annual-mean warming averaged over the entire 32-GCM ensemble would also be more accurate.

The statistical model also improves upon linear interpolation in capturing the spread in warming predictions. Inter-model range in the warming was calculated for the four dynamically downscaled GCMs, and it is clear that linear interpolation leads to underestimation of the spread, especially over land areas (Fig 3.13). Dynamical downscaling gives a model spread of more than 2.6 C over all land areas, with high elevation areas in the Sierras having spreads of as much as 3.6 C, while linear interpolation gives values in the 2.0-2.4 C range over California. Hybrid statistical downscaling does not recover all of the spread (especially since only one degree of spatial freedom is used), however some improvement is made. Average values of the spread over California are more in line with dynamical downscaling, with values as high as 3.0 C in the Sierras. Assuming that the four dynamically downscaled GCMs are representative subset of the 32 GCMs, then the hybrid statistical model's pro-

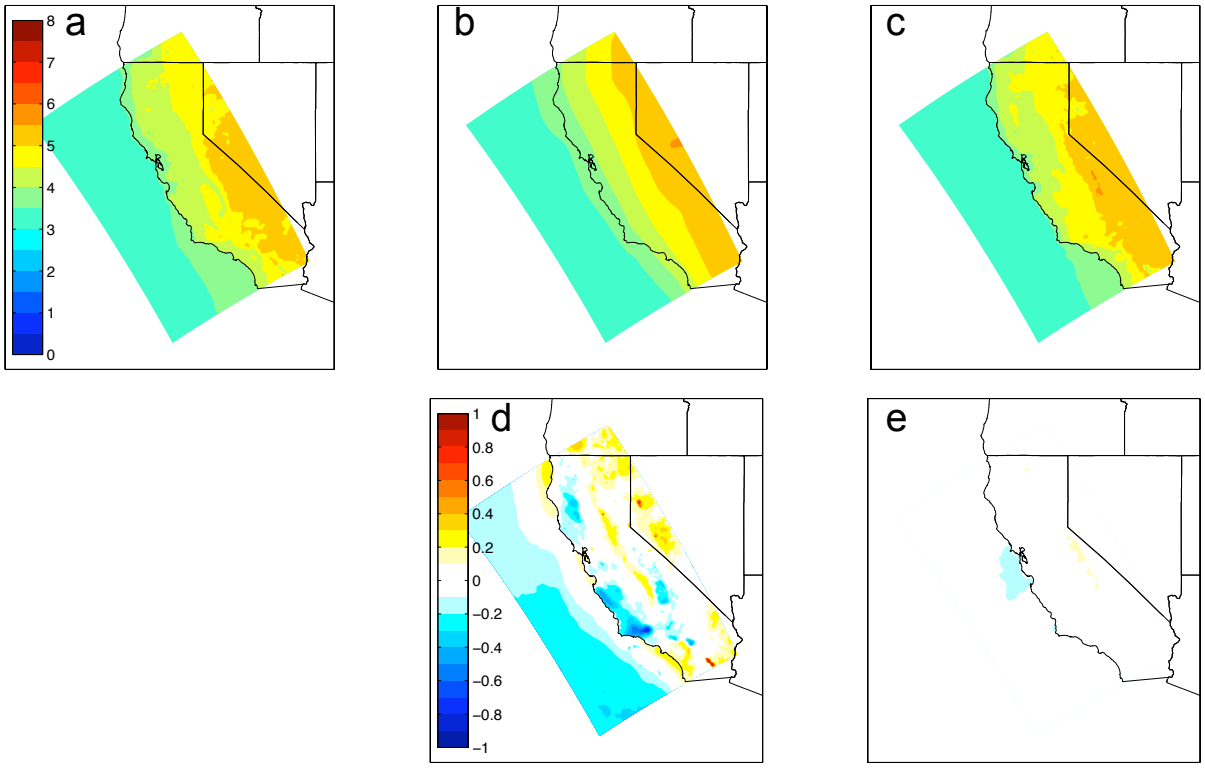

Figure 3.12: (a) Annual-mean warming (units: °C) averaged over the four dynamically downscaled GCMs. (b) Annual-mean warming produced by linearly interpolating the four corresponding GCM warming patterns. (c) Corresponding annual-mean warming produced by hybrid statistical downscaling. (d) Mean bias of (b) relative to (a). (e) Mean bias of (c) relative to (a).

jected inter-model spread of the ensemble is also likely to be an improvement over linear interpolation.

3.4 Discussion

The main goal of this chapter is to evaluate how well the hybrid statistical method works in an alternate domain. In this larger domain, covering all of California, the method appears to work well based on a couple of different metrics. The method shows improved predictive power versus linear interpolation over land areas (Sec 3.3.5), and elimination of biases in the ensemble-mean warming, and improved estimation of the inter-model spread (Sec 3.3.6). Over the ocean, there is relatively little spatial detail in the warming patterns, so the hybrid

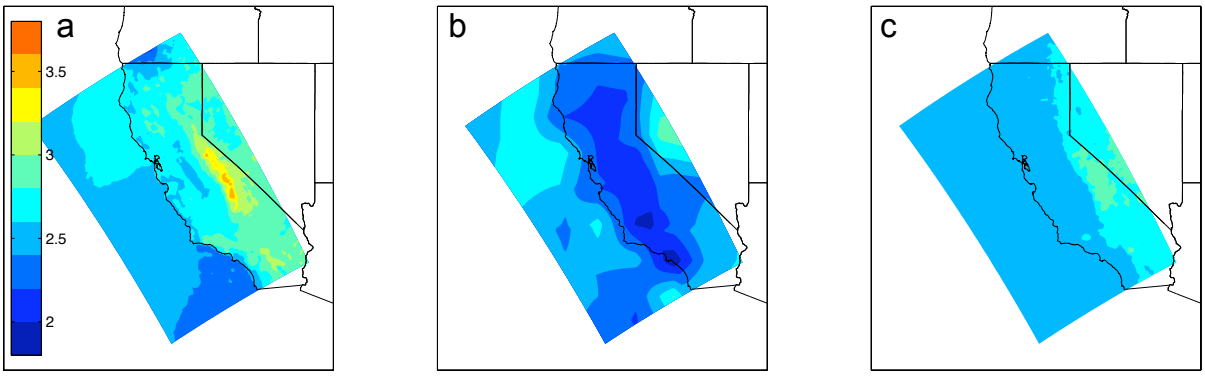

Figure 3.13: Inter-model range in the annual-mean warming (units: °C) for (a) the four dynamically downscaled GCMs, (b) the corresponding linearly interpolated GCMs, (c) the corresponding warming patterns produced by hybrid statistical downscaling.

method and linear interpolation both do equally well (though this might be different if WRF were coupled to a high resolution ocean model instead of specifying NARR SST at the ocean surface).

The first principal component for this domain exhibits a land-sea dipole in the warming, similar to that found in the Los Angeles Region. Our initial suspicion was that this was because each domain consists of roughly half land and half ocean points (in the Los Angeles domain there was slightly more land, around 2/3 of the domain). Because the ocean warms considerably less than land in nearly all of our simulations, the difference in warming between these accounts for the majority of the spatial variance. This makes us question how domain dependent the results of the statistical model are. If the ocean were to be excluded in PCA, would the first PC still be a land-sea dipole? In fact, when the ocean is excluded, the first five principal components produced by PCA are remarkably similar to those produced by PCA using the entire domain (Fig 3.14). When the entire-domain PCs were re-centered so that the mean of the scores over the land areas was zero (to be consistent with the land-only PCs), the sign of land-only PC2 was flipped, and the order of land-only PCs 4 and 5 were switched, the correlations between the two sets of patterns over the land areas were generally very high ($r = 0.996, 0.78, 0.61, 0.72, 0.90$ for the five pairs of PCs). In

particular PC1 was nearly identical over the land areas ($r = 0.996$). The fact that these two patterns are so similar indicates that whether that warming is more ocean-like or land-like provides the main spatial heterogeneity not just between land and ocean, but also between land points. In other words, even if they ocean is excluded, the contrast between coastal and inland areas is still the dominant mode of spatial variability. However, the fraction of variance explained by PC1 in the land-only case, is less (75% versus 89% for the entire domain). The difference in warming between coastal and inland land areas is smaller than between ocean and land areas. So, when the ocean is removed, PC1 explains less variance than before. Meanwhile, the other PCs are relatively unaffected when the ocean points are not included, since their patterns are primarily centered over the land areas.

These results suggest that the hybrid downscaling method is likely to be useful in coastal areas, since land-sea contrast is the dominant spatial pattern, regardless of whether the domain explicitly includes ocean. Land-sea warming contrast is strongest in the subtropics (Sutton et al. 2007), which explains why the method works so well over the Los Angeles domain and the larger California domain. We hypothesize more generally that the method will be useful in any domain in which there is a dominant dipole in the warming that is also present in the GCM. (However, if the dipole is not also present in the GCM warming pattern, then it could be difficult to create a mathematical relationship between the large-scale predictor and the fine-scale patterns, which would reduce the skill of the hybrid statistical model.)

Some adjustments to the method were made when it was applied to the larger California domain. Some of these were practical considerations. For example, single grid points were used to sample the predictors for the regional mean and PC1 loadings, instead of averages over larger regions. This adjustment was made to make the method more objective, since selecting the appropriate averaging box required subjective human input, and could not be easily performed 500 times as needed for the Monte-Carlo simulations. Despite selecting a single location, instead of using a broader averaging area designed to make the

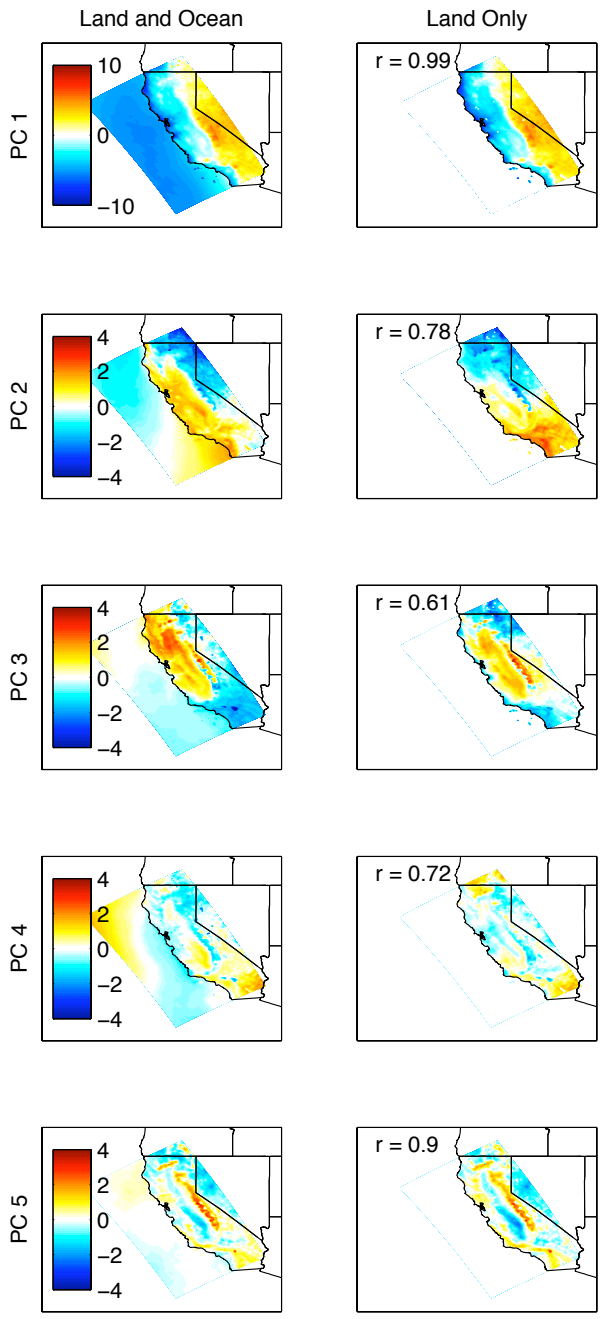

Figure 3.14: Comparison of first five PCs produced by PCA over entire domain (left column) and over just the land areas (right column). To facilitate a better visual comparison, (i) the PCs on the left have been re-centered so that the mean over the land areas is zero, (ii) the sign of land-only PC2 has been flipped, and (iii) the order that land-only PCs 4 and 5 appear have been switched. Correlations are shown between the pairs.

method more robust, the statistical model was still successful, based on the assessment of predictive power.

The use of Monte-Carlo simulations was another adjustment to the method. Using Monte-Carlo simulations added an aspect that had been missing from the previous application of this method, an assessment of the sensitivity of the parameters to the months and GCMs included in the training of the statistical model. This addition allowed us to see that the optimal sampling location stayed in nearly the same locations for all but a handful of simulations, giving us confidence in the robustness of the statistical model. The drawback of using Monte-Carlo simulations is in the evaluation of predictive power. Although it allows for more samples, using a randomly selected subset of months and GCMs is a less appropriate way of evaluating the statistical model's purpose. The statistical model is intended to predict the warming of all months of new GCMs, not new months for GCMs are already included. For instance, when the July and September dynamically downscaled warming patterns for GFDL-CM3 are included in the training of the statistical model, then the statistical model probably does a very good job predicting the outcome for August under GFDL-CM3, because the model inputs and outputs are similar. A truer test would be to use the leave-one-out cross-validation procedure (LOOCV) from the Los Angeles application, in which one GCM is completely left out and the statistical model is trained on the remaining GCMs. Ultimately, with only four GCMs to work with here, leaving only three for training seemed too small a sample to train on, and the Monte-Carlo technique was used. Our recommended best practice is to dynamically downscale as many GCMs as is feasible, 5 or more, with 6 or 7 preferable. This will allow for LOOCV to be used without radically changing the statistical model. Regardless, Monte-Carlo simulations are still recommended to ascertain the sensitivity of model parameters.

4 Comparison of Dynamically and Statistically Downscaled Warming Patterns

4.1 Introduction

Recently, archives have been made available of CMIP3 and CMIP5 GCMs temperature and precipitation data downscaled using two prominent statistical downscaling methods: Bias Correction with Constructed Analogs (BCCA) and Bias Correction with Spatial Disaggregation (BCSD). GCM surface air temperature and precipitation fields have been downscaled to 0.125° resolution to recover significant sub-grid scale features that cannot be captured by the GCMs themselves. Data covering the entire continental United States are available online at `http://gdo-dcp.ucllnl.org/downscaled_cmip_projections` (Reclamation 2013; henceforth R13) and are likely to be widely used as inputs a range of regional climate change impacts models (Maurer et al. 2007). It is important for users of these archives to understand the potential drawbacks and limitations of the data and the effect of those limitations on subsequent work. This study is limited to an investigation of the downscaled temperature data.

Previous studies have noted in passing that statistical downscaling temperature patterns may miss effect of the snow-albedo feedback, leading to an underestimation of the future warming compared to dynamical downscaling. Snow-albedo feedback happens when warmer temperatures lower the effective surface albedo, either by reducing snow cover or changing the reflective properties of the snow, and result in an increase in absorbed downward short-wave radiation, leading to further warming. Qu and Hall (2007) show that, of the two pathways to reducing albedo, reduction in snow cover is dominant. Salathe et al. (2008) found that dynamical downscaling of the ECHAM5 global model over the Pacific Northwest produced temperature changes in mountain areas that were significantly larger than with statistical downscaling. The dynamically downscaled results had warming enhancements due to snow cover reductions by as much 1.5 °C above what was not found in

the GCM or through statistical downscaling. Kim (2001) and Kim et al. (2002) also noted the clear presence of snow-albedo feedback in dynamically downscaled simulations. Pierce et al. (2013) generated projections of temperature and precipitation changes over California using statistical methods BCSD and BCCA, and dynamical downscaling with multiple RCMs, although because the results were aggregated over climate regions, the differences in warming due to presence of snow-albedo feedback could not be judged.

Capturing the warming enhancement due to snow-albedo feedback is important for hydrology impact studies in mountain areas. For example, in mountain ranges like California's Sierra Nevada, snow accumulates over the wet season and melts in the late spring and summer, playing an important role in supplying water for the region. Underestimating the warming here could lead to underestimates of the loss in future snowpack and shift in timing of the snowmelt. Furthermore, snowmelt impacts other hydrological variables like stream flow and runoff, which are used for flood-risk assessments and agricultural studies.

In this study, we determine if the statistically downscaled warming patterns by R13 underestimate the warming enhancement due to snow-albedo feedback and if so, by how much. To do this, the warming patterns produced by BCSD and BCCA are compared to those produced by dynamical downscaling with WRF performed in Chapter 2. Details about the data presented in Section 4.2. In Section 4.3, the warming patterns are compared for four CMIP5 GCMs over a Southern California domain that includes a number of mountain complexes including the San Gabriel, San Bernardino, and Southern Sierra Nevada mountains where snow-albedo effects occur. We demonstrate that the most important discrepancy between the statistically and dynamically downscaled warming projections in this region comes from snow-albedo feedback. We also calculate the magnitude of the warming sensitivity in WRF and suggest a new metric for quantifying the impact of the warming enhancement. In Section 4.4, we propose explanations for why neither BCSD nor BCCA effectively capture the warming enhancement and suggest potential solutions. Finally, Section 4.5 contains a summary of our findings.

4.2 Downscaled Climate Change Data

The data used in this study reflect multiple ways to downscale the climate change signals from GCMs in Fifth Coupled Model Intercomparion Project (CMIP5; cite). Details about each of the four GCMs used in this study are provided in Table 2.1. The RCP8.5 scenario was selected for all future runs.

4.2.1 Statistical Downscaling with BCSD and BCCA

Statistically downscaled data were downloaded from the Downscaled CMIP5 Climate and Hydrology Projections, available online at `http://gdo-dcp.ucllnl.org/downscaled_cmip_projections`. Projections at 0.125° resolution are derived using two different statistical techniques.

The first technique, Bias Correction and Spatial Disaggregation (BCSD), is described by Wood et al. (2002), Wood et al. (2004), and Maurer (2007). BCSD involves two steps. In the bias correction step, monthly GCM data is first regridded to 1° resolution and then bias-corrected relative to the 1950-1999 monthly climatology from an observationally-based gridded product. The observationally-based gridded product used in this case is the Maurer et al. (2002) 0.125° gridded data set, coarsened to 1°. Bias correction is performed using a quantile-mapping procedure that corrects for biases between the cumulative distribution functions of the GCM and gridded product. In the spatial disaggregation step, the bias-corrected and regridded GCM data is differenced with the 1950-1999 climatology to produce monthly anomalies. These anomalies are linearly interpolated to 0.125° resolution and added to the 0.125° observed climatology to create monthly downscaled data.

Hidalgo et al. (2008), Maurer and Hidalgo (2008), and Maurer et al. (2010) describe the second technique: Bias Correction with Constructed Analogs (BCCA). This technique produces daily values of maximum and minimum temperature. The first step of BCCA is to regrid the GCM data to 1° and bias correct relative to the daily Maurer et al. (2002) gridded product. Daily bias corrections are performed based on cumulative distribution functions

created from the historical daily Tmax (Tmin) data within a window of ±15 Julian days. The second step, constructed analogs, involves comparing these 1° regridded, bias-corrected GCM patterns for a given day to a library of coarsened daily patterns (analogs) from the 1950-1999 Maurer data. This library is seasonally confined, consisting of the daily Tmax (Tmin) patterns within ±45 days of the target day, for every year of the 1950-1999 period (∼2300 total patterns). For a given day, the 30 closest analogs to the GCM pattern are chosen from this library of historical analogs based on the root mean squared error. Linear regression is used to calculate the linear combination of these 30 analogs that best approximates the target pattern. Finally, to generate the high-resolution, 0.125° downscaled result, the same linear combination is applied to the corresponding 0.125° observed patterns for those same 30 historical days. Since this technique produces daily maximum and minimum temperatures, these were averaged over the month to produce monthly-mean temperatures.

For both BCSD and BCCA, data from two time slices were used in this study: (1) "baseline" 1981–2000, and (2) "future" 2041–2060. Future and baseline monthly surface air temperature climatologies were differenced to produce monthly varying climate change patterns. While data is available for the entire continental U.S., here we narrow our scope to a region that includes Southern California and the Southern Sierra Nevada mountains. This subset was selected to correspond with the domains of the available dynamically downscaled data. Likewise, GCMs were selected to correspond with available dynamically downscaled GCMs. Data is only available for land areas.

4.2.2 Dynamical Downscaling with WRF

Dynamically downscaled data comes from simulations using the Advanced Research Weather Research and Forecasting Model version 3.2 (WRF; Skamarock et al. 2008) over California performed in Chapter 2. The simulations were done with three nested domains. 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 covers Los Angeles County. WRF is coupled to the Noah land surface model (Chen and Dudhia 2001). These runs consist of a single baseline run and multiple future runs. The baseline run is intended to be a simulation of the historical period September 1, 1981 through August 31, 2001, whose realism can be evaluated through comparison to observations. (Chapter 2 and Berg et al. (in press) assess the skill of this run in simulating observed temperature and precipitation, respectively.) The lateral and ocean surface boundaries for the baseline run come from the North American Regional Reanalysis dataset (NARR, Mesinger et al. 2006) for the specified 1981-2001 period. The four future runs are intended to simulate the 2041-2061 period based on the climate change signals found in four CMIP5 GCMs: CCSM4, CNRM-CM5, GFDL-CM3, MIROC-ESM-LR. Table 2.1 gives the details of the four GCMs used. Boundary conditions for these runs were created by adding the differences between future (2041-2061) and baseline (1981-2001) monthly-mean climatology from each of the GCMs onto the 1981-2001 NARR-derived boundary conditions. All relevant meteorological variables were included in these calculations, as well as sea surface temperature. (See Chapter 2 for full details of the experimental design.)

For comparison with the statistically downscaled data, the WRF output in the outer (18-km resolution), and middle (6-km resolution) domain were used. Data from these domains was regridded to 0.125° resolution. The regridded middle domain data was used where possible, since its native resolution is higher. The middle domain covers a majority of the comparison region, including the mountain areas central to our analysis. However, for completeness, the regridded outer domain data is used in the northwest and northeast corners of the comparison region, which are not included in the middle domain. We limited our analysis to land areas, as the statistically downscaled data is only available there.

Precipitation and temperature validation of the baseline simulation (performed in Chapter 1 and by Berg et al. (in press)), showed that WRF is able to capture both the mean climatology and temporal variations reasonably well. Here we extend the validation to

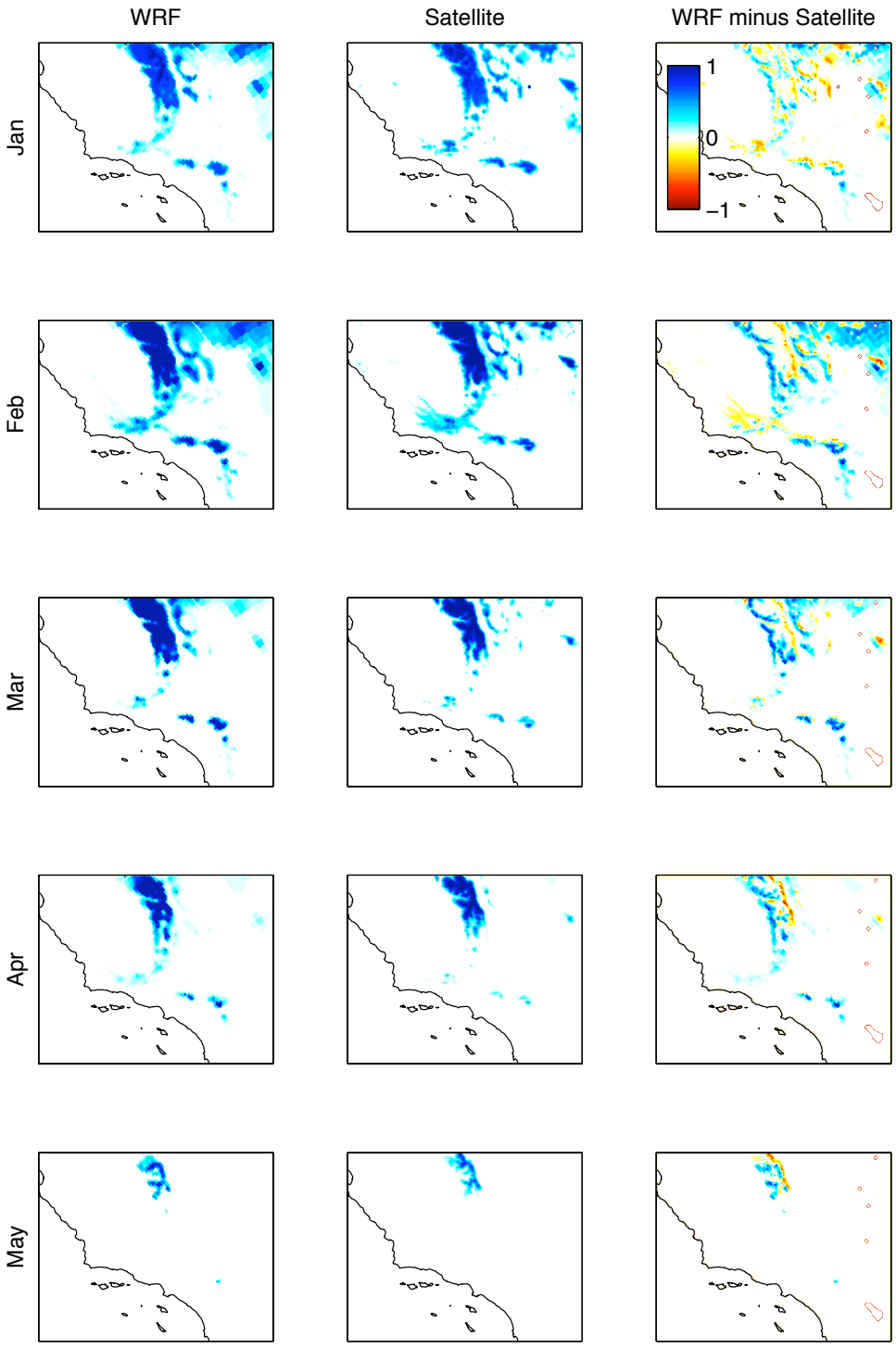

Figure 4.1: Monthly average snow-covered fraction for January-May 2001 in WRF (left), MODIS/Terra (middle), and the difference between them (right).

snow-covered fraction by comparing against MODIS/Terra Snow Cover Monthly L3 Global 0.05° CMG v.5 satellite data (Hall et al. 2006). WRF snow cover for the months January-May of was linearly interpolated to the 0.05° common grid and compared to MODIS/Terra snow cover (Fig 4.1). In general, WRF captures the seasonal progression of the snow extent very well. The main areas of disagreement occur right near the snow line, where snow cover would be most sensitive temperature. WRF snow cover extends to lower elevations on the windward slopes of the Sierra Nevada, San Gabriel, and San Bernardino mountains, while there is too little snow cover on the leeward side. These narrow bands of disagreement are similar to those found by Wrzesien et al. (2014), who comparing WRF V3.4/Noah-MP and MODSCAG over the Sierra Nevada. These could be due to WRF tendency to overestimate windward precipitation and underestimate leeward precipitation (Caldwell et al. 2009; Rögnvaldsson et al. 2011). A second, unrelated problem occurs in the northeast corner of the domain, where the outer, 18-km resolution domain is used. Here WRF clearly overestimates the snow cover and fails to resolve the complicated snow cover features present in the MODIS/Terra data. Based on this comparison, we trust WRF to simulate the timing and magnitude and snow cover over the majority of the domain, except for slight biases near the snow line and large overestimation of snow cover extent in the northeast corner of the domain.

4.3 Warming Disagreement in Snow-Sensitive Areas

Analysis begins with a comparison of the ensemble-mean monthly warming patterns for each method. The main systematic difference between them occurs during the late winter and spring (Fig 4.2). In areas of snow cover loss, WRF projects the warming to be amplified by as much 50% relative to the other methods. Areas most strongly affected include the slopes of Sierra Nevada, San Gabriel, and San Bernardino mountain ranges. Bands of enhanced warming between the historical and future snow lines are visible, as noted in Chapter 2. The enhancement is especially prominent in the March. While all methods produce an increase

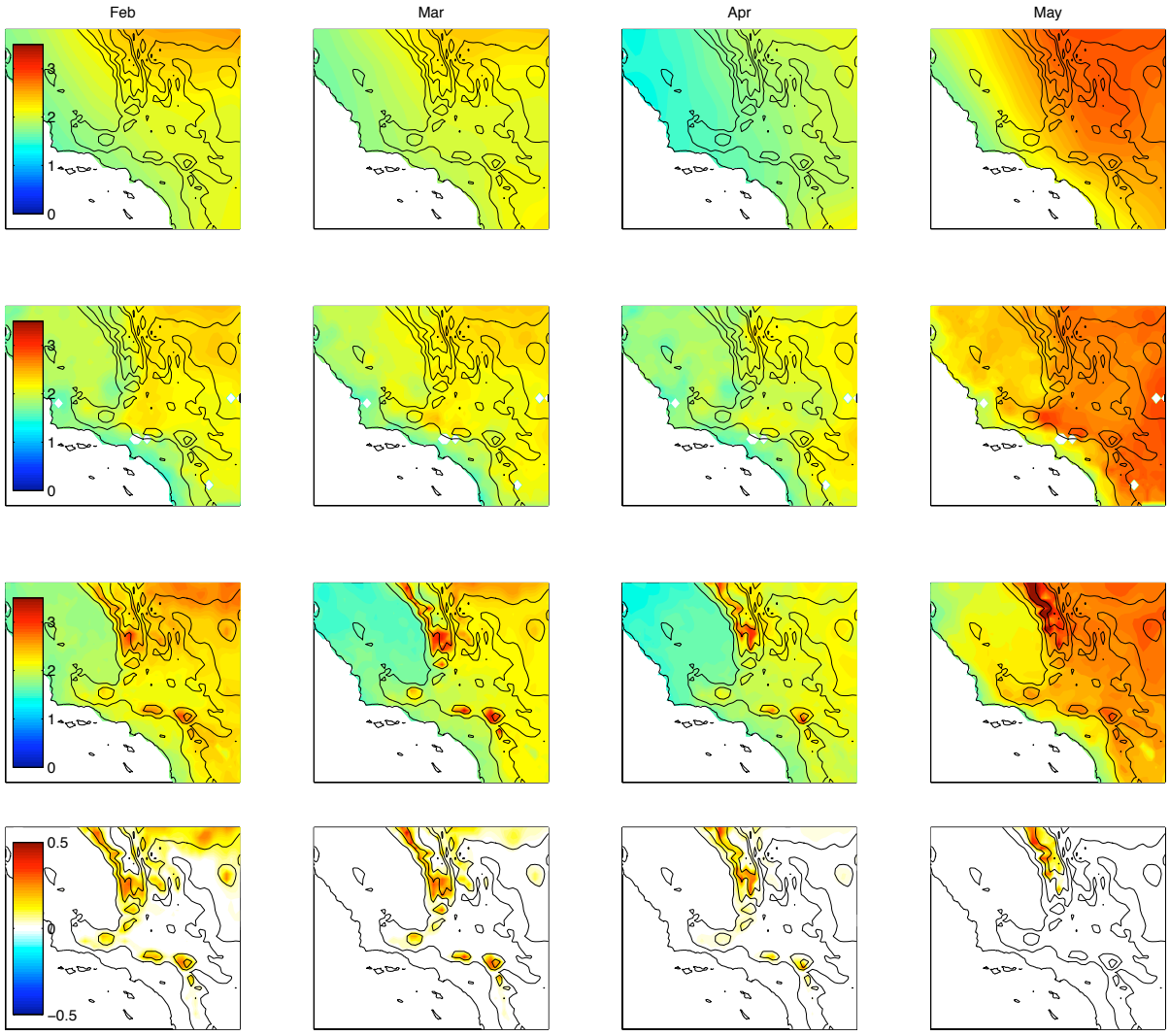

Figure 4.2: Ensemble-mean temperature increase for the months of February through May from BCSD (row 1), BCCA (row 2), and WRF (row 3). WRF ensemble-mean decrease in snow-covered fraction (row 4).

in March domain-average temperature of around 2.3 °C, WRF shows warming as high as 3.5 °C at locations with snow cover reductions. While the statistically generated patterns show some gradients in the warming, neither of them matches the location or intensity of the warming enhancement found in WRF.

When we further examine the WRF data by comparing the warming to the change in snow-covered covered fraction for each of the four dynamically downscaled GCMs, we

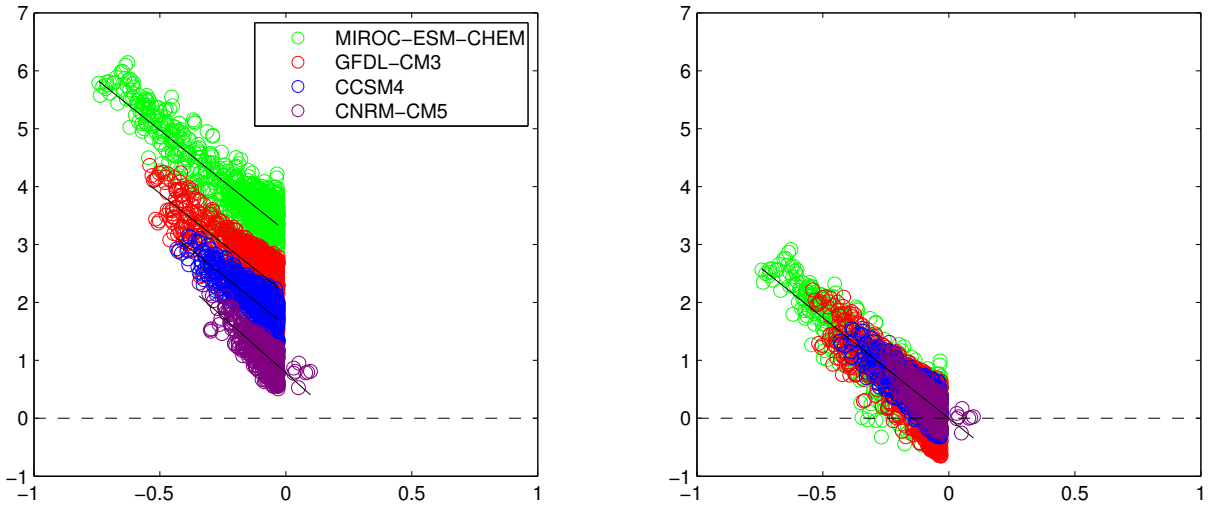

Figure 4.3: March warming versus change in snow covered fraction over the inner-most WRF domain. Note: only points where $\Delta\text{SCF} > 0.07$ are shown. Black lines are generated by performing linear regression for each downscaled GCM. (b) Same as (a) but the warming has been shifted to intercept zero. The shifted values reflect the warming enhancement. The black line is generated from linear regression of the combined data from all four downscaled GCMs..

see a linear relationship between the two variables (Fig 4.3a). The linear relationship has a different intercept but nearly the same slope in each case. When the data from each GCM is shifted to intercept zero (Fig 4.3b), the values represent the warming enhancement due to snow cover feedbacks. Note that the linear relationship of each set of points is now very similar, justifying their treatment as a single population. The slope of the line in Fig 4.3b represents the sensitivity of the warming enhancement to the snow cover change, found to be -3.5 ± 0.1°C. Thus WRF predicts that a location that goes from full snow coverage to bare ground would warm an additional 3.5 °C. Indeed, there are locations that lose over 50% snow cover in both the GFDL-CM3 and MIROC-ESM-CHEM runs, resulting in warming enhancements of more than 2 °C.

The warming enhancement due to snow cover changes is a phenomenon present during the entire October-June period, but with varying geographic scope and intensity. Based on statistics of the four-model mean, snow-covered area in our inner domain decreases the most in February (Fig 4.4a). However the sensitivity to snow-cover change appears to follow the

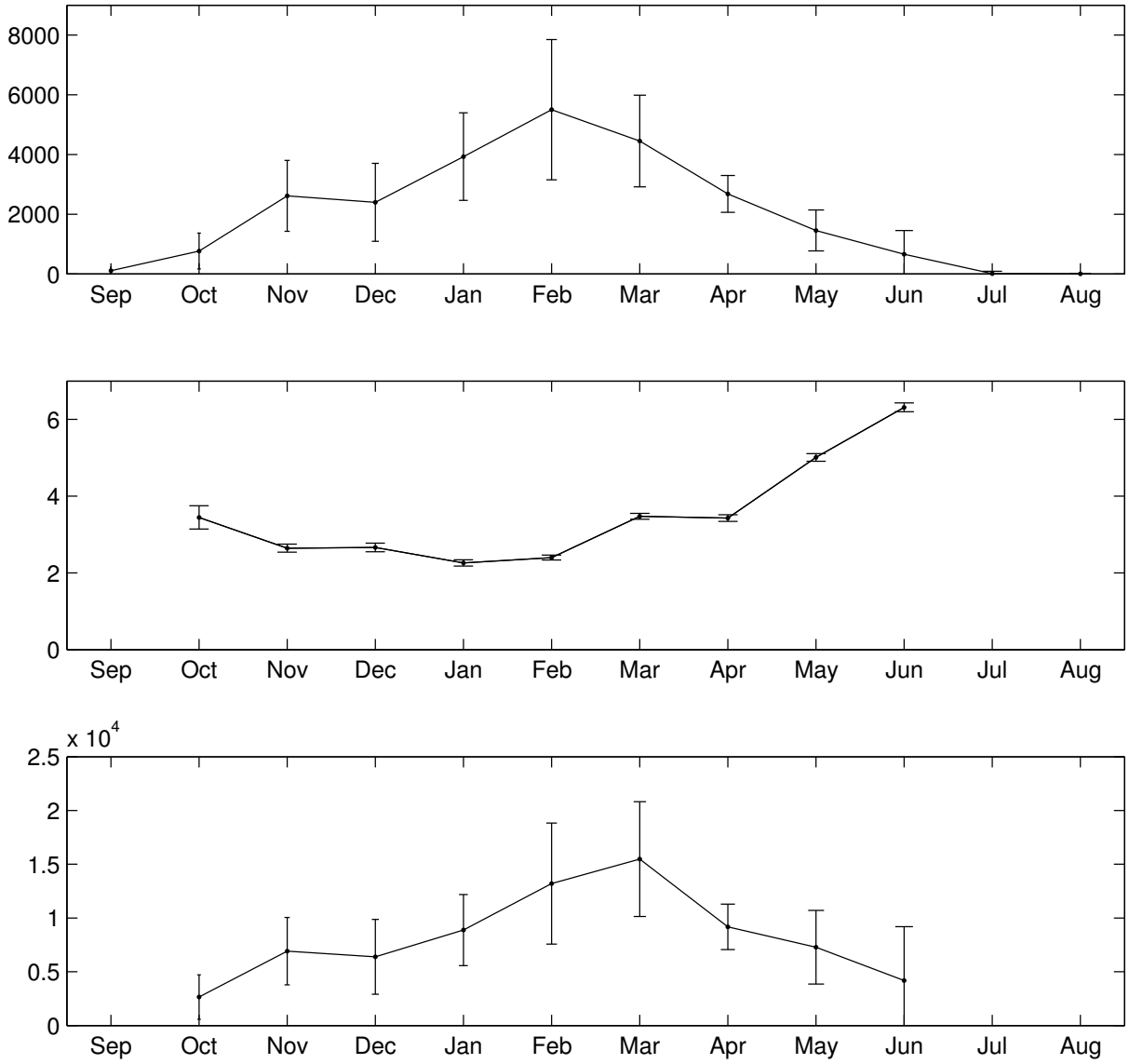

Figure 4.4: WRF projected ensemble mean (a) decrease in snow-covered area (units: km^2), (b) warming enhancment per change in snow-covered fraction (units: K), (c) warming enhancment multiplied by change in snow-covered area (units: $K \cdot km^2$)

seasonality of insolation, with a peak in June (Fig 4.4b). Thus snow coverage decreases are larger earlier in the season, but the sensitivity of the warming to those changes is relatively low. Meanwhile, later in the season, fewer areas are affected because there is less snow, but those that are affected have a larger warming enhancement. Multiplying the snow-covered area decrease by the warming sensitivity (Fig 4.4c) is one way to measure which months are

most heavily impacted by the warming enhancement. The peak of this metric is in March, when snow-covered losses and sensitivity are simultaneously relatively high.

4.4 Challenges in Capturing Snow-Albedo Feedback

Dynamical downscaling with WRF shows large warming enhancement due to snow-albedo feedback. Why don't the statistical methods generate similar results? BCSD produces a smoothed warming pattern for all months that is nearly identical that generated by linearly interpolating the GCMs (Fig 4.2). This is a result of how BCSD generates the fine-scale patterns by linearly interpolating the coarse-scale anomalies and adding them on to the fine-scale climatology. Thus when the baseline is subtracted from the future, the fine-scale climatology cancels out, essentially leaving the linearly interpolated coarse-scale GCM pattern. Small differences may arise due to non-linearities in the bias-correction process. However these differences are small and do not add any high-resolution information, since the bias correction happens at the coarse scale. Therefore, we can't expect BCSD to pick up snow-albedo feedback any better than the GCM itself. Snow-albedo feedback occurs in GCMs (e.g. Hall and Qu, 2006), but only on spatial scales corresponding to their horizontal resolution (100 – 200 km). The width of many of the warming enhancement features found in the WRF patterns are much smaller, sometimes as little as a single WRF grid box, requiring resolutions O(10 km) or less. Since GCMs cannot resolve these features, they are not present in the BCSD patterns.

In contrast, BCCA shows some fine-scale structure to the warming pattern. Its warming patterns have some similar features to the WRF patterns, with strong spatial gradients related to the mountains and coastline (Fig 4.2). One might expect, a priori, that by utilizing historical analogs, BCCA should have implicit sensitivity to snow-covered fraction. Warm days during the historical period should have less snow cover on average, and hence enhanced warming. However, BCCA shows no evidence of snow-albedo feedback similar WRF in the Sierra Nevada, San Gabriel, and San Bernardino mountain ranges.

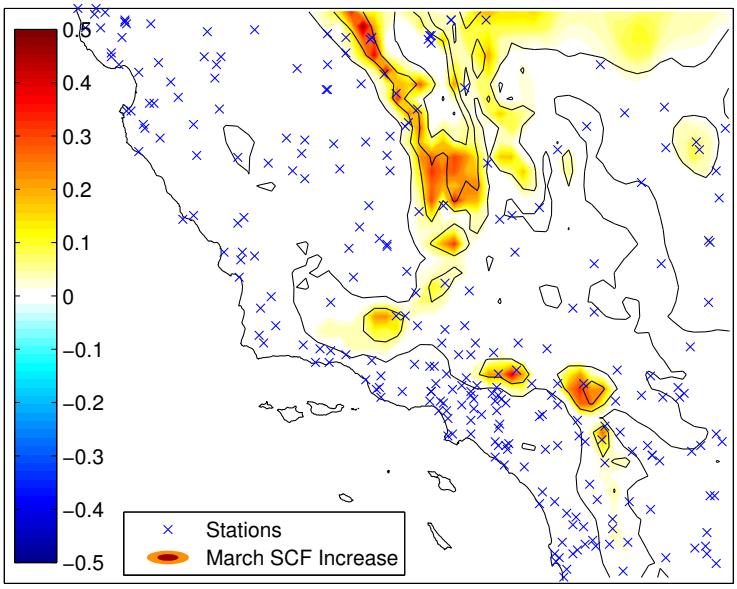

Figure 4.5: Stations locations used to the create the temperature field for the Maurer et al. (2002) data (blue crosses) and March ensemble-mean snow-cover decrease. Topography and coastline shown (black contours).

One explanation for this may be that the relatively narrow bands of localized snow-albedo feedback warming enhancement are at locations that are not well resolved in the observations. BCCA relies on a library of daily temperature patterns from the Maurer et al. (2002) 0.125° gridded product. This data set is based on data from stations whose locations are at locations shown in Fig 4.5 (blue crosses). The areas where WRF shows snow cover changes between baseline and future runs he month of March (Fig 4.5, red shading) are less densely sampled than other areas of the domain. Locations in the Southern Sierras are especially under-sampled. If the historical patterns are insufficiently sampled to capture the fine-scale features in these areas, then surely the climate change pattern produced by BCCA cannot either.

The above theory could explain why many areas where WRF predicts snow cover change do not experience enhanced warming under BCCA. However, there are clearly some stations located in areas of snow loss, and even these locations do not show the warming enhancement found in WRF (Fig 4.2). Even if BCCA were based on data sets with ideal

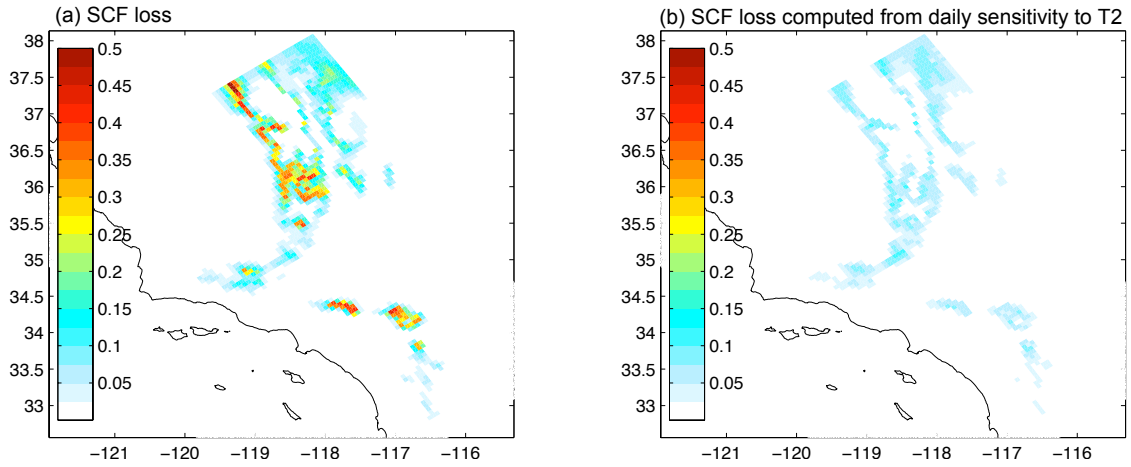

Figure 4.6: (a) Ensemble-mean March snow-covered fraction loss for the inner domain. (Ensemble-mean warming for March was 2.0 °C averaged over the land areas in this domain.) (b) Estimated ensemble-mean change in March snow-covered fraction assuming an increase in land-average temperature of a 2.0°C, using the daily sensitivity of March snow-cover to land-average temperature from the 1985-2001 period.

spatial sampling it would likely still suffer from a snow feedback that is too weak. This is because BCCA matches future days with warm days in the historical period, which means that a future day is treated as an anomaly relative to the baseline climatology instead of occurring as part of a new climatology where a majority of the days of the year are warmer than normal. This is an important difference for snow cover, which depends at any given moment on the integrated history of snowfall and snowmelt over the previous portion of the snow season. For example, we would expect the snow cover on April 1st to be much lower if we knew the previous months of the snow season were on average 2 °C warmer than normal, rather than if we only knew that the day of April 1st were 2 °C warmer than normal. Essentially, BCCA relies on the daily sensitivity of snow cover to temperature which is likely to produce much less snow cover loss than is projected by the WRF simulations. Therefore the future temperatures predicted by BCCA are probably based on temperature patterns associated with only minimal snow cover loss, resulting in little warming enhancement.

To test the plausibility of this theory, we compared the snow cover change found

in the WRF to the snow cover change that would be expected based on using the daily sensitivity of snow cover to temperature. The sensitivity daily snow-covered fraction to the corresponding daily domain-average temperature anomaly was calculated for the month of March from the WRF baseline run. Then this daily sensitivity was used to calculate the expected snow loss based on a domain-average warming of 2.0 °C. (This is the value is the ensemble-mean warming for the month of March in WRF.) The snow-covered fraction decrease predicted by the daily sensitivity is far less than the snow-covered fraction decrease seen in the WRF climate change runs, by as much as a factor of five (Fig 4.6). Thus, the snow loss implicit in daily historical patterns is indeed much smaller than the snow loss that would result if the same temperature increase were experienced over the entire snow season up to that point. This suggests that even with higher sampling of spatial patterns, BCCA would still largely miss the warming enhancement.

It is unclear if there is an easy fix to incorporate the snow-albedo feedback into the statistically-downscaled warming projections. On one hand, there is the possibility of directly adding in this missing warming enhancement to the statistically downscaled output. The warming enhancement could be parameterized explicitly as the sensitivity of temperature to snow cover (a constant factor depending on season, quantified above) multiplied by the snow cover change. However the snow cover change itself depends on the temperature change, complicating the picture. Thus an iterative method may be needed if the warming enhancement is included explicitly.

Alternatively, statistical methods like constructed analogs could be adapted to compare more than just daily patterns. Analog methods that match daily temperature patterns implicitly use only the daily feedback loop between temperature and snow cover. An ideal historical analog would match both the target day's temperature and the history of that temperature over the previous months of the snow season. This way the temperature pattern in the historical analog reflects the appropriate change in snow expected under climate change, where the majority of the snow season is warmer as well. Adapting the method

to account for these effects would be challenging, especially when projecting large changes in warming. For example, it is unlikely to find even a few years in the historical period where the entire snow season averages 5 °C above the baseline climatology, even though some GCMs predict warming levels that high by the end of the 21st Century (Sun et al. in press). Thus requiring temperature anomalies to persist for the entire snow season would dramatically reduce the number of suitable historical analogs. Indeed, this speaks to the larger issue of likening the effects of a large climate shift to short-lived historical climate anomalies: the larger the magnitude and longer the duration of anomaly required, the less chance there is of finding a relevant historical analog.

4.5 Summary of Findings

This study compares three techniques for downscaling GCM temperature changes over Southern California. The two statistical downscaling techniques are Bias Correction with Spatial Disaggregation (BCSD) and Bias Correction with Constructed Analogs (BCCA). The third technique is dynamical downscaling performed using the WRF regional model. For much of the domain, the three methods produce similar results. The major systematic difference between the warming patterns is a strong warming enhancement in the mountain areas in WRF due to snow-albedo feedback. The sensitivity of the warming enhancement to snow cover loss in WRF is found to be between 2 °C and 7 °C per 100% loss of snow cover (completely snow-covered to snow-free) depending on the month. Winter months exhibit lower sensitivities but more total snow cover loss, while later spring months are more sensitive, but show less snow cover loss.

Since snow-albedo feedback is an important, physically credible process, we would expect its effects to be visible in downscaled climate projections. However, both statistical downscaling methods largely miss the additional warming due to snow-albedo feedback, each for its own reasons. BCSD produces warming patterns that are nearly identical to the linearly interpolated GCM warming patterns. Since the GCMs are too coarse to resolve the snow

processes at the necessary spatial detail, the warming enhancement is not found in the BCSD patterns. While BCCA has the capability to produce warming patterns with much greater spatial detail, it sill largely misses the warming associated with snow-albedo feedback. This appears to be because BCCA captures only the effect of that day's temperature anomaly on snow cover, effectively underestimating the sensitivity of snow cover to a warmer climate. Snow cover depends on the past history of the snow season's accumulation and melts events, not just the temperature on that day. In fact, based on the WRF data, the snow cover decrease accompanying a 2 °C warmer climate is as much as five times greater than the snow cover decrease accompanying a daily temperature fluctuation of the same magnitude.

While the comparison region is limited to Southern California in this study, these same snow albedo effects occur throughout the mountainous western United States. For example, Salathe et al. (2008) showed that a large portion of the Pacific Northwest was affected by snow-albedo feedback. A proper assessment of how any of these snow-covered locations might change must account for the warming enhancement due to snow-albedo feedbacks.

Users of the BCSD and BCCA downscaled data sets should be aware that WRF projects 50% more mid-century surface air warming in mountain areas due to snow-albedo feedbacks. This could have an important effect on impacts studies that use downscaled data as inputs for surface hydrology models and other processes. Some adjustments to the statistical methods are explored, but each has foreseeable drawbacks. Ultimately, due to the complexity the most effective method may be to use dynamical model with a coupled land surface component for doing regional climate change studies in snow-covered areas.

5 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. Capps, D. Walton, D. Neelin, and B. Langenbrunner: Mid-21st Century Precipitation Changes over the Los Angeles Region. J. Climate, in press.

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:10.007/s00382-002-0286-0.

Braganza, K., D. J. Karoly, A. C. Hirst, 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:10.007/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, K. Hayhoe, 2008: Climate change scenarios for the California region. Climatic Change, 87, (Suppl 1):S21–S42, doi:10.1007/s10584-007-9377-6.

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.

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.

Di Luca, A., R. de Elia, R. Laprise, 2012: Potential for small scale added value of RCM's downscaled climate change signal. Clim. Dyn., 40(3-4), 601-618.

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.

Drost, F., D. Karoly, and K. Braganza, 2011: 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.

Duffy, P. B., R. W. Arritt, J. Coquard, W. Gutowski, 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.

Fasullo, J. T., 2010: Robust land–ocean contrasts in energy and water cycle feedbacks. J. Climate, 23, 4677–4693.

Fowler, H. J., S. Blenkinsop, C. Tebaldi, 2007: Linking climate change modeling to impacts studies: Recent advances in downscaling techniques for hydrological modeling. Int. J. Climatol., 27, 1547–1578.

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

Giorgi, F., and L. O. Mearns, 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:http://dx.doi.org/10.

1175/1520-0442(2002)015<1141:COAURA>2.0.CO;2

Giorgi, F., C. S. Brodeur, 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: http://dx.doi.org/10.1175/1520-0442(1994)007<0375:RCCSOT>2.0.CO;2

Giorgi, F., 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: http://dx.doi.org/10.1175/1520-0442(1997)010<0288:EDOTSC>2.0.CO;2

Giorgi, F., et al., 2001: Regional climate information—evaluation and projections. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (J.T. Houghton et al., Eds). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 583–638.

Giorgi, F., C. Jones, and G. R. Asrar, 2009: Addressing climate information needs at the regional level: the CORDEX framework. World Meteorological Organization (WMO) Bulletin, 58, 175–183.

Hall, D. K., V. V. Salomonson, and G. A. Riggs. 2006. MODIS/Terra Snow Cover Monthly L3 Global 0.05 Deg CMG. Version 5. Boulder, Colorado USA: National Snow and Ice Data Center.

Hayhoe, K., D. R. Cayan, and Coauthors, 2004: Emissions pathways, climate change, and impacts on California. Proceedings of the National Academy of Sciences of the United States of America, 101, 12422–12427.

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.

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.

Kain, J, S., 2004: The Kain–Fritsch convective parameterization: An update. J. Appl. Meteor. Soc., 43, 170–181. doi:http://dx.doi.org/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.

Kerr, R.A., 2013: Forecasting regional climate change flunks its first test. Science, 339, 638.

Kharin, V. V., and F.W. Zwiers, 2002: Climate predictions with multimodel ensembles. J. Climate, 15(7), 793.

Kim, J., 2001: A nested modeling study of elevation-dependent climate change signals in California induced by increased atmospheric CO2, Geophys. Res. Letters, 28, 15, 2951-2954.

Kim, J., T.K. Kim, R. W. Arritt, and N. L. Miller, 2002: Impacts of Increased Atmospheric CO2 on the Hydroclimate of theWestern United States. J. Climate, 15, 1926–1942. doi: http://dx.doi.org/10.1175/1520-0442(2002)015<1926:IOIACO>2.0.CO;2

Lambert, F. H., and Chiang, J. C., 2007: Control of land-ocean temperature contrast by ocean heat uptake. Geophysical Research Letters, 34(13), L13704.

Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby, 2003: Regional climate research: Needs and opportunities. Bull. Amer. Meteor. Soc. January 2003, 89–95.

Leung, L. R., 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(1-3), 75–113.

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:http://dx.doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2

Livneh, B., and Coauthors, 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.

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 CO2. Part I. Annual mean response. J. Climate, 4(8), 785–818.

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(3), 407–430.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A Long-Term Hydrologically Based Dataset of Land Surface Fluxes and States for the Conterminous United States. J. Climate, 15, 3237–3251.

Maurer, E. P., L. Brekke, T. Pruitt, and P. B. Duffy, 2007: Fine-resolution climate projections enhance regional climate change impact studies, Eos Trans. AGU, 88(47), 504.

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, and J.F. Mitchell, 2007: The WCRP CMIP3 multimodel dataset: A new era in climate change research. Bull. Amer. Meteor. Soc. 88(9), 1383–1394.

Mesinger, F., G. DiMego, E. Kalnay, K. Mitchell, P.C. Shafran, W. Ebisuzaki, and W. Shi, 2006: North American regional reanalysis. Bull. Amer. Meteor. Soc., 87(3), 343–360.

Mitchell, J. F. B., S. Manabe, T. Tokioka, and V. Meleshko, and Coauthors 1990: "Equilibrium Climate Change," in Houghton, J. T., Jenkins, G. J., and Ephraums, J.

J. (Eds.), Climate Change: The IPCC Scientific Assessment, Cambridge University Press.

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. Journal of Geophysical Research: Atmospheres (1984–2012), 102(D14), 16663–16682.

Nakanishi M., and Niino H., 2004: An Improved Mellor–Yamada Level-3 Model with Condensation Physics: Its Design and Verification. Boundary-Layer Meteorology, 112, 1–31.

Niu, G.-Y., and Coauthors 2011: The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. J. Geophys. Res., 116, D12109, doi:10.1029/2010JD015139. Thompson 2008 MP microphysics

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. Clim. Dyn., 40(3-4), 839–856.

Qu X., and A. Hall, 2007: What Controls the Strength of Snow-Albedo Feedback?. J. Climate, 20, 3971–3981.

Rasmussen, R., C. Liu, K. Ikeda, D. Gochis, D. Yates, F. Chen, and E. Gutmann, 2011: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. J. Climate, 24(12), 3015–3048.

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, Colorado. 47pp.

Reclamation, 2014: Downscaled CMIP3 and CMIP5 Climate and Hydrology Projections: Release of Hydrology Projections, Comparison with preceding Information, and Sum-

mary of User Needs. Prepared by the U.S. Department of the Interior, Bureau of Reclamation, Technical Services Center, Denver, Colorado. 110 pp.

Rognvaldsson O., J. W. Bao, H. Agustsson, H. Olafsson, 2011: Downslope windstorm in Iceland – WRF/MM5 model comparison. Atmos. Chem. Phys. 11: 103–120, doi: 10.5194/acp-11-103-2011.

Salathé Jr., E. P., R. Steed, C. F. Mass, 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: http://dx.doi.org/10.1175/2008JCLI2090.1

Salathé Jr, E. P., L. R. Leung, Y. Qian, and Y. Zhang, 2010: Regional climate model projections for the State of Washington. Climatic Change, 102(1-2), 51–75.

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

Sato T., F. Kimura, and A. Kitoh, 2007: Projection of global warming onto regional precipitation over Mongolia using a regional climate model. J. Hydrology, 333, 144–154.

Schiermeier, Q., 2010: The real holes in climate science. Nature, 463(7279), 284–287.

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-4751+STR.

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. Geophysical Research Letters, 29(11), 9-1-9-4.

Sun, F., D. B. Walton, A. Hall: 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, in press.

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. Geo-

physical Research Letters, 34(2), L02701.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. Bull. Amer. Meteor. Soc., 93(4), 485–498.

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(10), 1524–1540.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization. Mon. Wea. Rev., 136, 5095–5115.

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(7-8), 807–823.

Walton D. B., F. Sun, A. Hall, S. Capps: A Hybrid Dynamical-Statistical Downscaling Technique, Part I: Development and Validation of the Technique. J. Climate, in press.

Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations. Progress in Physical Geography, 21(4), 530–548.

Wrzesien, M. L., T. M. Pavelsky, S. B. Kapnick, M. T. Durand, T. H. and Painter, 2014: Evaluation of snow cover fraction for regional climate simulations in the Sierra Nevada. Int. J. Climatol.. doi: 10.1002/joc.4136