From Climate-Change Spaghetti to Climate-Change Distribution

DISCUSSION PAPER

Prepared For:
California Energy Commission
Public Interest Energy Research Program

Prepared By:
Michael Dettinger
U.S. Geological Survey,
Scripps Institution of Oceanography

February 2004

500-04-028

Prepared For:
California Energy Commission

Guido Franco,
Project Manager

Kelly Birkinshaw,
Program Area Lead

Terry Surles,
PIER Program Manager

Robert L. Therkelsen,
Executive Director

Prepared By:
Michael Dettinger
U.S. Geological Survey
Scripps Institution of Oceanography
La Jolla, California

Acknowledgments

This study was funded by the California Energy Commission's California Climate Change Center at Scripps Institution of Oceanography, and by the Geological Survey's Hydroclimatology Program.

Contents

List of Figures

Abstract

The uncertainties associated with climate-change projections for California are unlikely to disappear any time soon, and yet important long-term decisions will be needed to accommodate those potential changes. Projection uncertainties have typically been addressed by analysis of a few scenarios, chosen based on availability or to capture the extreme cases among available projections. However, by focusing on more common projections rather than the most extreme projections (using a new resampling method), new insights into current projections emerge: (1) uncertainties associated with future emissions are comparable with the differences among models, so that neither source of uncertainties should be neglected or underrepresented; (2) twenty-first century temperature projections spread more, overall, than do precipitation scenarios; (3) projections of extremely wet futures for California are true outliers among current projections; and (4) current projections that are warmest tend, overall, to yield a moderately drier California, while the cooler projections yield a somewhat wetter future. The resampling approach applied in this paper also provides a natural opportunity to objectively incorporate measures of model skill and the likelihoods of various emission scenarios into future assessments.

1.0 Introduction

Projections of climate change due to increasing greenhouse-gas concentrations in the twenty-first century are inevitably uncertain because of the chaotic nature of the global climate system, because of models imperfections, and because of uncertainties regarding how society's emissions of greenhouse gases and other atmospheric contaminants will proceed in the future. In the midst of our uncertainties, however, current (climate model) projections exhibit some key commonalities that demand near-term attention from California's resource-management communities: (1) Even the most benign of the projected climate changes are sufficient to significantly alter the California's landscape, hydrology, and land and water resources, and (2) those alterations are likely to become significant within roughly the next 25 years (Barnett et al. 2004; Dettinger et al. 2004; van Rheenen et al. 2004). Thus, California—like the rest of society—is faced with the prospect of an uncertain array of climate changes that may be expected to develop within time frames that are comparable to the planning and implementation horizons of any major resource-management decisions that might respond to those changes.

To date, technical responses to this dilemma primarily have involved development and preliminary applications of tools for assessing the potential climate-change impacts and the efficacy of various possible adaptation or accommodation strategies. In part, this response has been motivated by the assumption that projection uncertainties will be reduced sufficiently in the near term to justify putting off more intensive and detailed assessments until later. However, the projected changes include important near-term impacts, and the slim likelihood that projection uncertainties will decline precipitously in the near term may not justify prevarication. For example, two highly respected climate modelers, David Randall and Akio Arakawa, recently opined that "a sober assessment suggests that with current approaches the cloud parameterization problem [the most vexing aspect of climate and climate-change modeling at present] will not be 'solved' in any of our lifetimes" (Randall et al. 2003). Thus, we should not depend on large reductions of projection uncertainties in time to make needed initial responses to the changing climate, and new strategies for more completely accommodating projection uncertainties are needed.

The development of the required uncertainty-based strategies will be challenging, but will offer the opportunity to focus more on the likelihoods, rather than just the uncertainties, of climate change. That is, as Myles Allen (2003) has recently commented, "Climate modelers need to start saying what changes can be ruled out as unlikely, rather than simply ruled in as possible." Indeed, it is perhaps time for California analysts to focus on what is more likely rather than on what is just possible. If this distinction can be determined, accommodation strategies and impact assessments will become more focused and practical.

This paper is an attempt to illustrate some of the insights that become possible as we progress from the recent relatively narrow emphasis on a few outlying projections to approaches that characterize the overall distributions of available climate-change projections. Our view of the future climate is clouded by uncertainties from model imperfections and uncertainties about how rapidly greenhouse gases will accumulate in the atmosphere, together with the naturally unpredictable variations of the global climate system. Preliminary depictions of how these uncertainties cloud projections of California's future climate are already possible, as demonstrated here using an existing multiple-model, multiple-emissions collection of climate-change projections. The resulting analysis suggests that, among currently available projections, California would most likely experience about +5ºC warming, with relatively little precipitation change. The sign of projected precipitation changes appears to depend on whether one considers one of the warmer or cooler projections, which generally yield modestly drier or modestly wetter outcomes, respectively. Thus, even this preliminary step toward characterizing the likelihoods of various climate changes provides new insights into the available projections. The approach used here also offers new opportunities for objectively weighting the various projections to improve our understanding still further.

2.0 Problem

The most common approach for analyzing climate-prediction uncertainties is analysis of ensembles of predictions, wherein each prediction differs from the others due to some prescribed model condition. Ensembles of climate projections often are used to describe prediction uncertainties associated with model constructions, initial conditions, and future emissions of greenhouse gases into the global atmosphere. In weather- and climate-prediction applications, several studies have argued that ensemble means are better predictors than are any individual members of the contributing ensembles (e.g., Krishnamurti et al. 2000; Richardson 2001; Zhu et al. 2002), and this finding may eventually be found to extend also to climate-change projections. Thus, climate-change ensemble means (or approximations thereof) might reasonably be analyzed, although we have heard of few uses of even this strategy. Rather, the more common strategy, to date, has been to analyze one or two example scenarios (often determined more by logistics and availability than by their representativeness).

More determined efforts are responding to today's limited ensembles of available projections, and to the burdens imposed on some impact studies by each additional ensemble member, by trying to, at least, "bookend" the climate-change possibilities by analyzing only the upper and lower bounds of the available projections (e.g., recent reports such as *Global Climate Change and California: Potential Implications for Ecosystems, Health, and the Economy,* at www.energy.ca.gov/pier/ reports/500-03-058cf.html and *Climate Warming & California's Water Future,* at http://cee.engr.ucdavis.edu/faculty/lund/CALVIN/ ReportCEC/CECReport2003.pdf, as well as the ongoing coordinated efforts by the Joint Department of Water Resources/U.S. Bureau of Reclamation Climate Change Work Team in California). Such approaches ultimately say little about the true uncertainties facing scientists and decision makers. Neither the ensemble means, nor the most extreme predictions (often true outliers), describe the real scatter among current projections.

Ensembles of predictions also presumably contain information about the overall likelihoods of various scenarios and about higher-order statistics of the projection scatter. In reality, the ensemble scatter is more descriptive of models and emissions than of real-world climate changes under multiple stresses and processes, but the ensemble statistics remain our best (and essentially, only available) avenue for quantitatively representing overall climate-change uncertainties in the immediate future.

Ideally, then, strategies for interpreting and using climate-change projections would be informed by a more complete synthesis of available ensembles. Although “majority rules” does not apply to climate-change projections, it is nonetheless questionable whether the common focus of impact studies on outliers (the least commonly projected outcomes) is at all a more useful strategy. Instead, it would be more useful and intelligible to plan and work from a more complete depiction of the scatter in current projection ensembles.

We will continue to be uncertain as to the true probability distribution of future climate, viewing the future as we do mostly through the lenses of imperfect and continually evolving climate models and emissions scenarios, but approaches that really quantify the scatter among current models and emissions scenarios will provide more complete understanding of the commonalities and contrasts among present-day projections than is being brought to bear in most studies and decision making at present.

3.0 Providing Ensemble Projection Distribution Functions (pdfs)

The typical ensemble, whether numbering tens of members or a very few, offers the analyst and decisionmaker a “spaghetti” of simulated futures (e.g., as shown later in Fig. 1). This representation of an ensemble is useful and simple, giving a qualitative sense of scatter, commonalities, and trends. Done correctly, the spaghetti provides a sense of how trends compare with shorter term natural variations in the systems considered.

However, our eyes are naturally drawn to outliers out of proportion to their significance, and clusters in the morass may receive less consideration than is their due. A more even-handed and quantitative view of the spaghetti of a typical ensemble requires estimation of the probability distribution from which the ensemble was sampled. In reality, since we are working with imperfect models and forcings, the distribution that our ensembles are sampled from are not the same as the distribution of future real-world climates. Thus, we are only able to estimate the distribution of projections of future climates and cannot directly estimate the actual probabilities of various future climates. We can only estimate (what we will refer to as) *projection distribution functions* (pdfs) as the best available approximations of the true climate-change probability distributions.

If the ensemble includes many members, then characterizing the ensemble trends and scatter can be as simple as ranking the predictions for each time and binning the results to directly form histograms or crude model-scenario pdfs. Even if the pdfs so estimated are crude, they can provide useful measures for comparing projections to observations and can provide a useful basis for comparing different ensembles (e.g., in the weather-prediction sense) (Toth et al. 2003).

When the number of ensemble members is smaller, however, developing even a rough estimate of the pdfs involves assumptions about the character of the projection uncertainties sampled by the ensemble. One approach is to sort and rank the ensemble predictions, use them as mileposts of the pdf (e.g., the median projection value at a given lead time marks the median in the pdf), and then smooth algebraically to fill in interpolated values. Alternatively, one can attribute error bars of some weight and shape to each ensemble member and then essentially sum the error bars from all the ensemble members to arrive at the overall ensemble pdf (but then important assumptions need to be made regarding the growth rate of the error bars for the individual ensemble members). Both of these approaches have the advantage that they are simple, but have the disadvantage that they require subjective choices or assumptions by the analyst.

In this paper, a third alternative is applied that, in its simplest form, has no subjectively tunable parameters. However, because the particular method used to estimate the pdfs is probably less important than the effect of viewing pdfs (rather than spaghetti), details of this particular alternative are left to an Appendix. The method requires no tunable parameters, because it characterizes the ensemble spread by a data-adaptive principal components analysis and then resamples the independent components obtained from that analysis as often as necessary to provide a smooth pdf. Using the orthogonality properties that are designed into principal components analysis (PCA), the resampling method provides an almost unlimited number of other realizations that are statistically independent of each other but that retain the essential characteristics of the ensemble members, including evolving ensemble means and standard deviations and all the lag and intervariable correlations. The method readily handles ensembles that bifurcate along two or more trajectories and handles heavy-tailed distributions as a matter of course.

4.0 Climate-Change Distributions for Northern California

As an illustration of the difference between describing climate-change ensembles as pdfs and as spaghetti, the component-resampling procedure described in the Appendix is applied here to an ensemble of climate-change projections of twenty-first century (2001–2099) climate. The ensemble considered here was compiled from six climate models, each simulating responses to each of three specified greenhouse-gas-plus-sulfate-aerosols emissions scenarios (Fig. 1). The ensemble includes three projections each by the U.S. PCM, Canadian CCCM, German ECHAM4, British HadCM3, Japanese NIES, and Australian CSIRO coupled ocean-atmosphere global climate models; the emissions scenarios are the A2, B2, and IS92a SRES scenarios (Houghton et al. 2001), which represent projections of relatively rapid, intermediate, and moderate rates of twenty-first century emissions increases, respectively.

By considering this even balancing of models and scenarios, no model or scenario is emphasized over the others. This even-handed treatment of the models and scenarios was valued here sufficiently so that other models that did not have all the scenarios available, and scenarios that have not been run in all the models, were excluded from the present analysis. Ideally, in the absence of known deficiencies in one or another of the ensemble members, the climate-change pdfs should reflect, in an even-handed way, the combination of uncertainties associated with models and the uncertainties associated with future emissions. However, not all models (or scenarios) are equally skillful at reproducing or projecting climate variations. A simple extension of the resampling procedure to allow an uneven treatment of the models (i.e., to weight the most skillful models the most and the least skillful models the least) is outlined at the end of this section.

The 18 ninety-nine-year-long (future) climate projections of Northern California climate change compiled in Figure 1 all share rapid warming tendencies after about 1970 and, by about 2020, temperatures have all warmed beyond most of the background of historical temperature variability. The general spread of temperatures by 2001 is from +2.5°C to +9°C. Notably, the scatter among scenarios is not substantially larger or different than the scatter among models considered here. Emission scenarios (e.g., the A1 and B1 scenarios of Houghton et al. 2001) that diverge even more than the scenarios analyzed here might be different enough to spread the projections considerably more. Projections of precipitation in the twenty-first century are less unanimous, with some projections becoming much wetter (the wettest projections are both from the Canadian model) and some drier. Plotted in this way, the eye naturally focuses on the outliers in the ensemble, and many studies have been constructed to address the bounds of such projection ensembles, rather than exploring the more common results.

To improve visualization, interpretation, and—for some applications—the usefulness of this ensemble, the 18 projections of temperature and precipitation were resampled according to the procedure described in the Appendix. In this application of the component-resampling procedure, mixing of the ensemble loading patterns was restricted to only allow projections by a single model to be intermixed. This restriction prevents the possibly inappropriate mixing of incompatible components from the projections by very different climate models. The restriction is easily accomplished by beginning each resampling cycle with the choice of one of the models at random, followed by random sampling among only the several amplitude series for that model, to obtain the new realization. With this restriction, the number of independent combinations is about 6×3^{20} or about 10^{10}, certainly a sufficient number for applications.

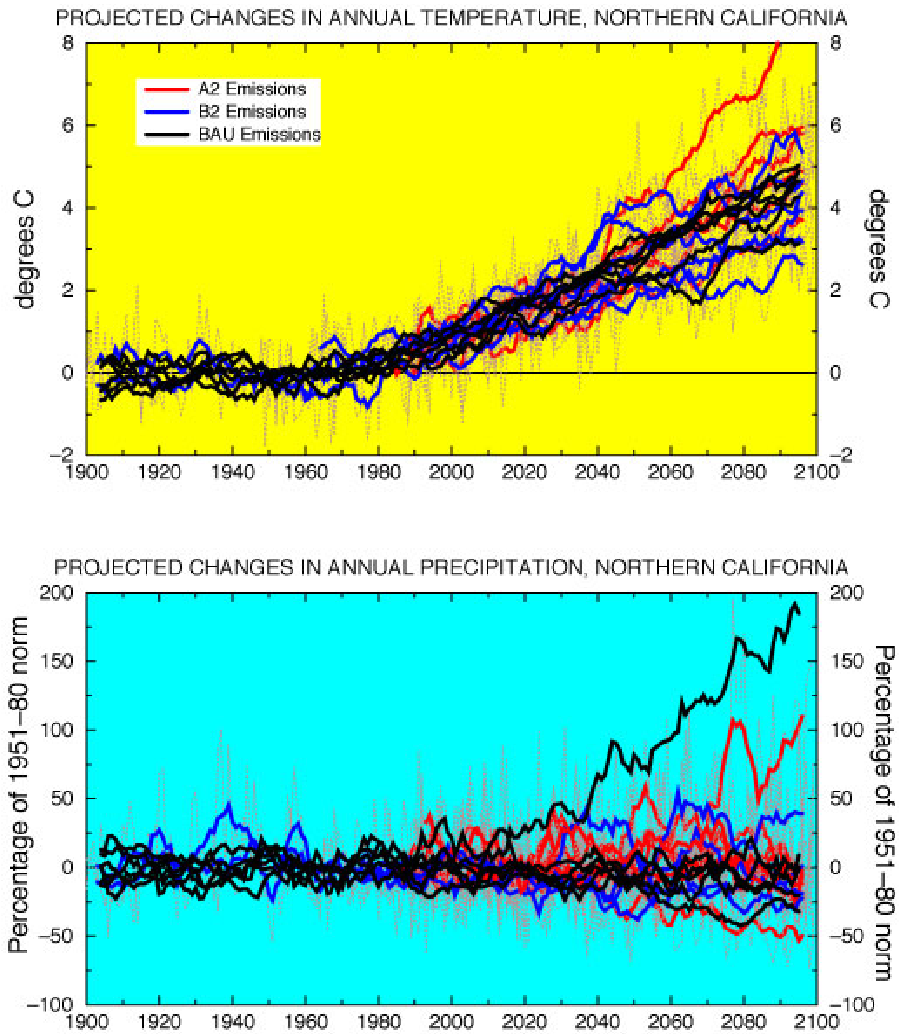


Figure 1. Ensembles of historical and future temperature and precipitation projections from six coupled ocean-atmosphere general-circulation models, each forced by historical scenarios, and then—in the twenty-first century—the A2, B2, and IS92a SRES emissions scenarios (Houghton et al. 2001). The dashed background of curves shows annual deviations from the 1951–1980 simulated means; whereas, heavy curves show 7-year moving averages. Projections are for a single model grid cell (ranging from 2.5°C to 5.5°C spatial resolution, depending on model) from each model centered over northern California.

Figure 2 shows the results of a 20,000-member resampling of the 18-member climate-change projection ensemble. The PCA applied in the first step of the procedure was extended so that temperature and precipitation changes were analyzed and resampled together. The pdfs shown are thus joint pdfs of temperature and precipitation. Consequently, for example, if a particular model has a tendency for excursions of temperature and precipitation to occur simultaneously, the component-resampled realizations will emulate those linkages.

Early in the twenty-first century, the projections are closely clustered, somewhat warmer, and somewhat drier on average than the 1951–1980 climatology (because, even by 2000, greenhouse forcings are larger than during that climatology period) (Dai et al. 2001). The ensembles spread over the course of the twenty-first century, until by 2099, temperature-change projections range (mostly) from about +2°C to +7°C, and precipitation-change projections range from about –30 to +25 cm/yr, with two outlying exceptions. The probability distributions shown are reflections of the joint variations of temperature and precipitation so that if, for example, the projections that were warmest overall tended also to be the wettest, and vice versa for cooler and drier models, the component-resampled realizations would maintain these tendencies faithfully.

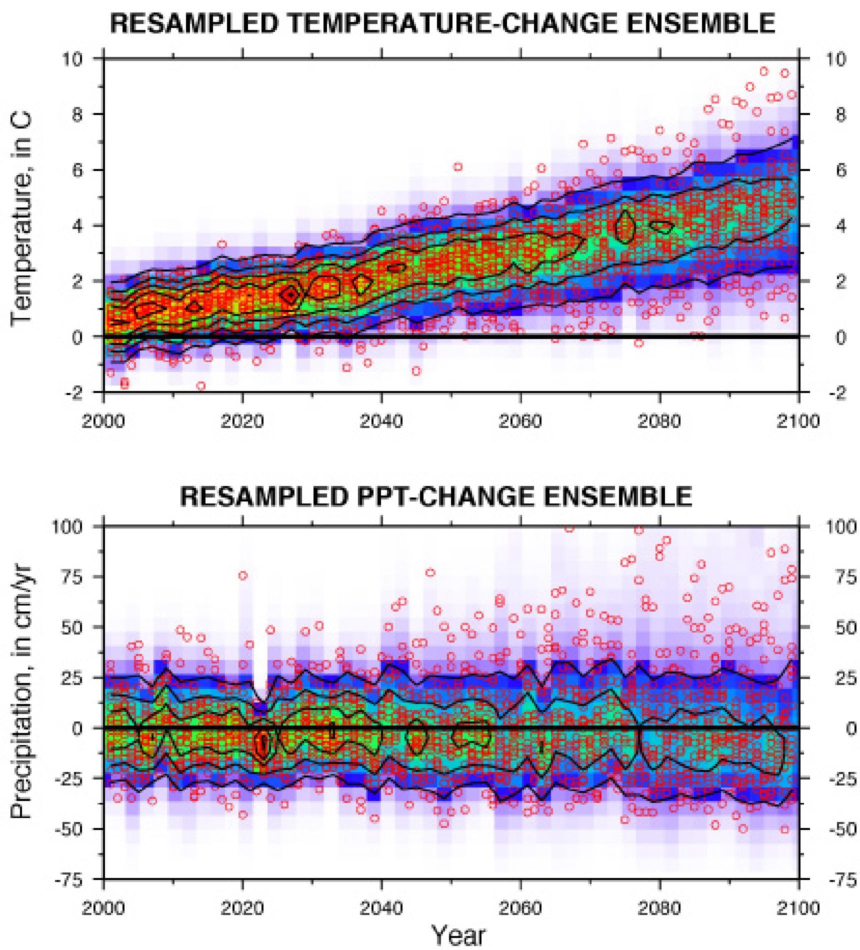
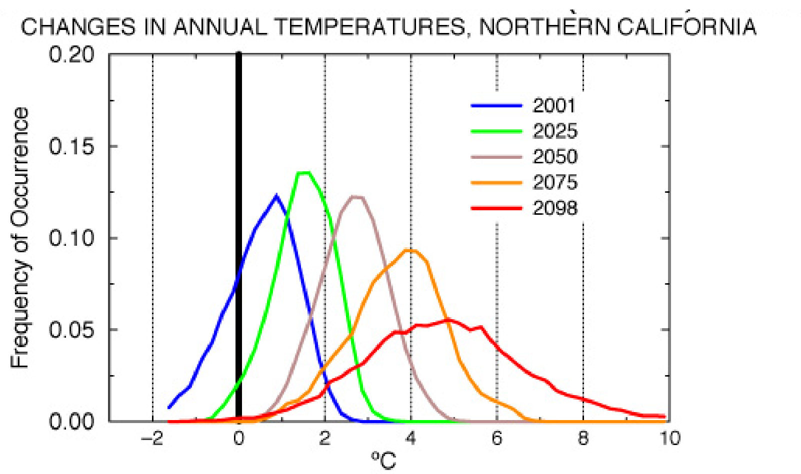


Figure 2. Distributions of original and component-resampled projections of annual twenty-first century surface-air temperatures and precipitation changes for a grid cell over Northern California (40°N 120°W), from the ensemble of projections shown in Fig. 1. Red circles show the raw ensemble projections; contours and shading show resampled joint temperature-precipitation probabilities, with a contour interval of 0.025.

The smoothing that is provided by the component-resampling procedure is illustrated by the sequences of time slices through the projection pdfs (in Figure 2) shown in Figure 3. The temperature-change pdfs spread and trend toward warmer conditions as the twenty-first century climate evolves. The spread is mostly a result of divergence between the models and divergence of the emissions scenarios, with relatively little contribution by increasing interannual variability within any given model's projections. Notice that, by as early as 2025, realizations that are cooler than the 1951–1980 "normal" are exceedingly rare.



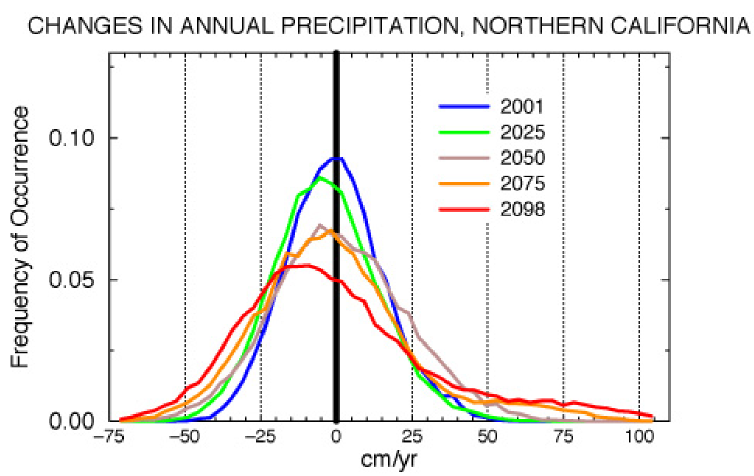


Figure 3. Time slices of the distributions of resampled ensemble realizations from Figure 2.

In contrast, the precipitation-change pdfs translate and spread much less than do the temperature pdfs (Fig. 3). Overall, the component-resampled realizations (as in the raw projections) most commonly exhibit only modest twenty-first century precipitation changes over California. The modes of the smoothed (resampled) pdfs in Fig. 3 trend toward drier conditions, which is much more difficult to perceive in the scattered red dots of Fig. 2 or in a corresponding "spaghetti" plot overlaying each ensemble member's projected time series. Thus, although no new information is introduced by the component-resampling procedure, its smoothing can nonetheless be very informative.

The general rate of expansion of the ensemble spread around this mean precipitation-change behavior is small, except for a distinct heavy tail spread towards substantially wetter conditions. That heavy tail spread reflects the contributions to the ensemble from the Canadian model's projections, the two outlying much wetter projections in the original 18-member ensemble. That model, under each of the emissions scenarios, evolves towards a much wetter California, as part of its tendency (unique among the models compiled here) to respond to increasing greenhouse forcing with enhanced El Niño conditions.

The component-resampling procedure applied here generates realizations of temperature and precipitation change that are jointly distributed. Thus, it is also possible to evaluate tendencies for correlated temperature and precipitation changes. The joint probabilities of precipitation and temperature change among the 20,000 resampled realizations is mapped in Figure 4 for several years during the twenty-first century. Notice that, as indicated previously, temperatures generally warm and precipitation changes little overall. However, the joint probability distribution is also somewhat bimodal in ways not obvious from either the unvariate pdfs or the spaghetti plots. The joint probabilities indicate that the warmest climate-change projections tend to also bring drier conditions; the cooler projections tend to be slightly wetter, most obviously by 2050. By 2100, when all the scenarios have warmed considerably, the same tendency still persists, but the warmer-drier scenarios dominate overall.

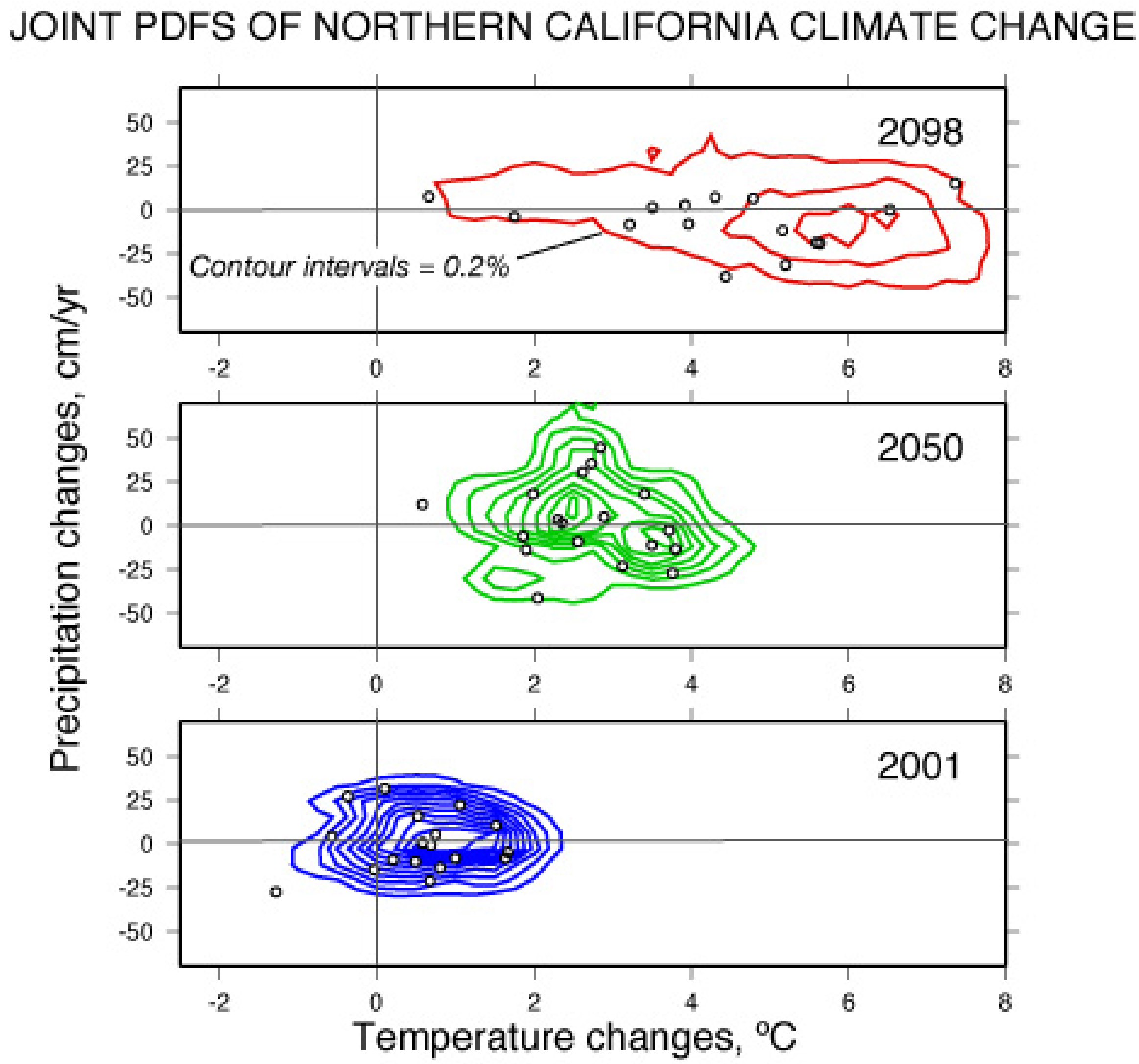


Figure 4. Time slices of the joint temperature-precipitation distributions of resampled ensemble realizations from Figure 2. Circles indicate values in the original 18-member ensemble of projections.

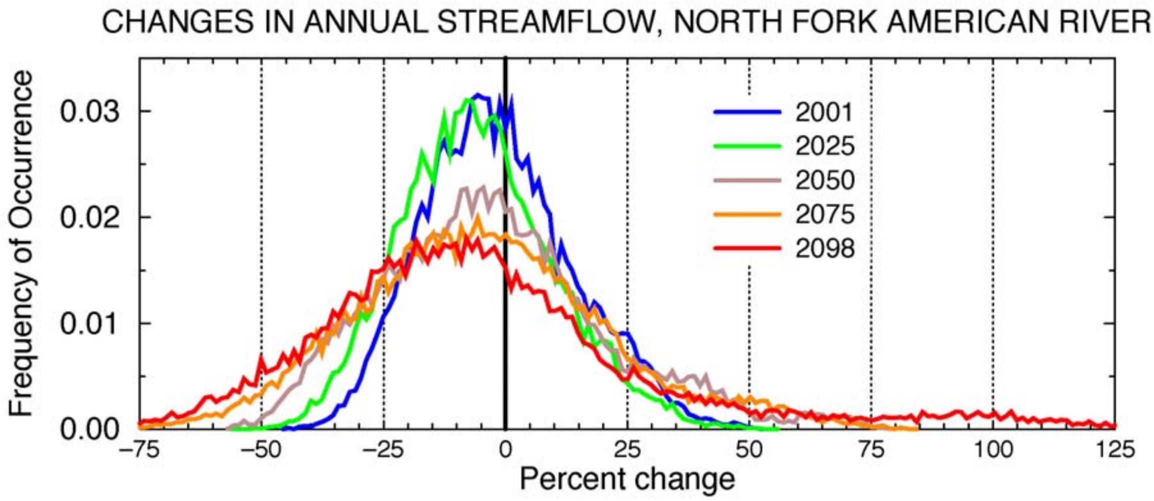
The component-resampled realizations of the projections also provide a ready supply of examples of coordinated temperature and precipitation changes for use in evaluating climate-change impacts. As a simple example, the 20,000 temperature-and-precipitation-change realizations generated for Figures 2 and 3 were introduced to the streamflow amount and timing response surfaces mapped by Jeton et al. (1996) for the North Fork American River in the central Sierra Nevada. Those response surfaces (Figs. 16b and 17c in Jeton et al. 1996) show the mean simulated changes in annual streamflow amounts and in the median-flow dates (days of year by which half the year's flow is past), in response to 100-year-long synthetic climate series with arbitrarily specified mean-climate changes ranging from cooler to warmer, and from drier to wetter. The mean streamflow changes mapped by Jeton et al. (1996)—corresponding to the temperature and precipitation changes in each of the 20,000 resampled ensemble realizations (from each of the time slices in Fig. 3)—were accumulated, and the resulting pdfs of streamflow amount and timing are shown in Figure 5.

The pdfs of annual streamflow changes in Fig. 5 are similar to the pdfs of precipitation change in Fig. 3, reflecting the strong control that precipitation change exerts on total streamflow amount, as well as the nearly complete buffering of streamflow amounts against responses to temperature changes, discussed at length by Jeton et al. (1996). By the end of the twenty-first century, streamflow amounts are significantly biased towards a drier mean and mode, although the much wetter Canadian climate models ensures a heavy tail of significantly wetter streamflow-amount realizations.

The corresponding projections of streamflow timing (Fig. 5, bottom panel) mostly reflect the warmer temperatures projected by all the models, although concurrent precipitation changes in the realizations couple nonlinearly with the temperature effects in the Jeton et al. (1996) response surfaces to yield much broader and more multimodal timing distributions. Some of the multimodal character of the timing pdfs presumably derives from the bimodal character of the joint temperature-precipitation distributions (Fig. 4). By 2025, years with earlier than normal median-flow dates (1951–1980) are all but eliminated among the resampling-driven realizations. By the end of the twnety-first century, the most common median-flow date projections are over a month earlier than the 1951–1980 norms; see Stewart et al. (2004) for a more comprehensive and geographically far-reaching discussion of this phenomenon.

Now, consider the differences between the messages and information content of Fig. 1 and Fig. 3 (or 5). How attractive does the bookending strategy look, once the pdfs have been examined? From the spaghetti of Fig. 1, we concluded mostly that projected temperature changes are most unanimous than are the projections of precipitation change, and that very wet futures are a significant threat (or opportunity).

The pdfs, in contrast, suggest that the envelope of (most likely) temperature projections spreads more through time than does the envelope of precipitation changes. The less-than-obvious tendency for the mode of precipitation changes to drift towards drier conditions is also much clearer in the pdfs. In fact, no new information has been added to the ensemble by the component-resampling procedure, but our understanding of the potentialities that the ensemble represents is arguably much clearer. In addition to this clarification, the users of such an ensemble have much more freedom to select their own levels of risk aversion when ensemble



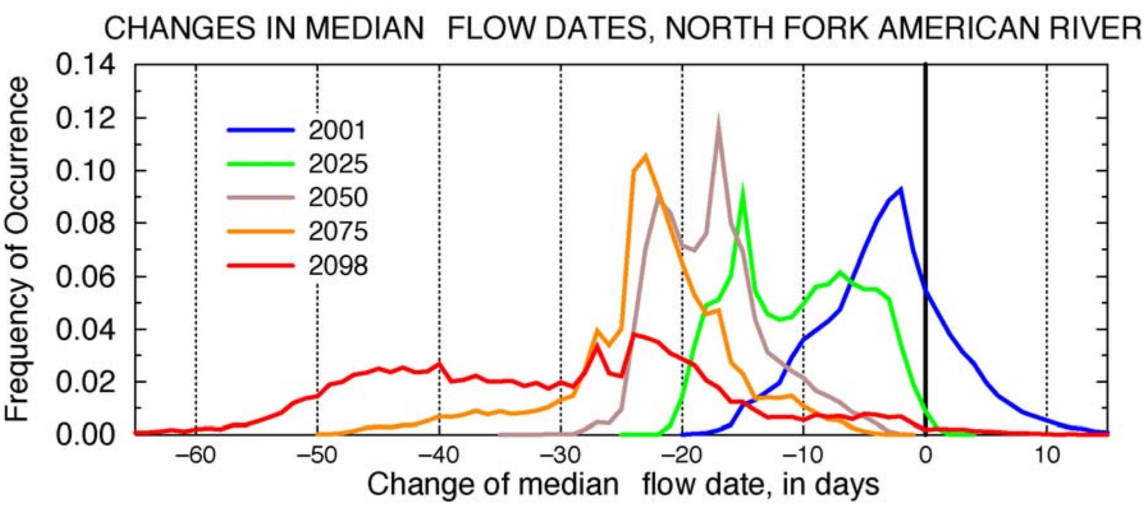


Figure 5. Distributions of annual streamflow amounts and median-flow dates (i.e., date by which half of a year's flow is past) in response to 20,000 resampled climate-change projections (illustrated in Fig. 3). Streamflow responses were estimated from response surfaces mapped in Jeton et al. (1996).

results are quantified by pdfs rather than by spaghetti. That freedom is needed, because risk is not simply the likelihood of an adverse impact; rather, risk is essentially a product of likelihood and cost of that impact. Consequently, in applications, each newly discovered potential impactbrings with it its own unique requirements from the projection ensembles. A more pdf-centric approach is the more proactive approach.

Although the resampling procedure used in this section added no real information to the ensembles, the procedure can readily be extended to add crucial information in clear and helpful ways. For example, the component-resampling procedure used above treated each model's projections as equally likely and each emissions scenario as equally likely.

However, the procedure can be modified to reflect any assumed weighting of the various models and scenarios. For example, if the accuracies of each model were quantitatively indexed by a measure of the likelihood that its projections were the most accurate (among all the models considered), then that index could be used to weight the fraction of samples that each model would contribute to the resampling procedure. This would mean that the most accurate models would contribute the most to the resampled distributions, and the least accurate models would contribute the least. Similarly, if the likelihood of emissions scenarios could likewise be ranked quantitatively, then the resampling probabilities could be adjusted to reflect those outcomes as well. This opportunity to hone the pdfs generated from the 18-member climate-change ensemble considered is an important motivation for our ongoing efforts to characterize the historical accuracies of the models with respect to a combination of local and global simulation-skill scores.

5.0 Summary

In current climate-change applications, the availability of ensembles of predictions that contain very large numbers of members and ensembles that evenly mix model uncertainties with emissions uncertainties are rare. The availability of such ensembles would substantially ease statistical analyses and interpretations, and could be used to judge simulation skill. This study describes and demonstrates the clarifications that are possible when projection-distribution functions can be estimated quantitatively by resampling much smaller ensembles.

A pdf, of the simple form used here, simply provides more information of direct relevance to resource managers, engineers, utilities, farmers, and others, and a clearer depiction of central tendencies and the risks at the extremes than does the typical spaghetti plot. Although little or no actual new information was introduced by the component-resampling procedure in the example shown here, it already provides an objective method for developing reproducible estimates of detailed distribution functions from small ensembles that clarifies the implications of an available climate-change ensemble considerably. Information describing the historical skills of contributing models, and probabilities of various forcing scenarios, can readily be added to improve the uncertainty estimates. The result is a bridge between subjective interpretations of spaghetti plots of small ensembles and the kinds of visualizations and calculations that could be accomplished with much larger ensembles. Thus, the methods and ensemble explored here suggest that:

1. Depicting climate-change ensembles in terms of the density distributions in the projection ensembles can provide new insights into the projections that are not obvious or directly measured in the more common spaghetti diagrams/listings—even when no

new information is added in the process of estimating those distributions. In the example presented here, spaghetti diagrams had fueled the idea that precipitation projections were more scattered than temperatures, and that a very wet California was a strong possibility. An objective depiction of the distribution of projections indicated instead that the ensemble distribution of temperature projections spreads more (in relative terms) than does the corresponding precipitation distribution, and that the wet projections are true outliers with much smaller changes in precipitation being much more common (likely) among current projections.

2. The process of estimating projection distributions from the ensemble spaghetti offers a natural opportunity for actually adding information to the interpretation of ensembles. In the resampling procedure used here, skill scores for the models and any outside information about the relative likelihoods of various emissions scenarios can easily be used to condition the resampling probabilities, so that the resulting estimates of projection probabilities more nearly reflect the strengths and weaknesses of each contributing ensemble member. Such weighting is not an option unless the step from spaghetti to distribution is taken.

3. Uncertainties from both model differences and emissions scenarios cloud our view of the future climate. The even-handed mixture of projections from both different models and different forcings is an ideal that should be pursued as much as possible, and that should be brought to California applications at the earliest feasible date.

For the future, besides working to develop and use model skill scores in the resampling procedure, it is worth noting that the procedure demonstrated here is currently only suited for use with a handful of projected variables. Estimating the joint pdfs of simultaneous projections of many variables or many locations will require modifications and extensions of the procedure. However, the benefits of replacing ensemble spaghetti with projection distributions (densities) make that extension a worthwhile goal.

References

Allen, M. R. 2003. Possible or probable? *Nature* 425: 242.

Allen, M. R., and W. J. Ingram. 2002. Constraints on future changes in climate and the hydrologic cycle. *Nature* 419: 224–232.

Barnett, T., R. Malone, W. Pennell, D. Stammer, B. Semtner, and W. Washington. 2004. The effects of climate change on water resources in the West: Introduction and overview. *Climatic Change* 62: 1–11.

Dai, A., T. M. L. Wigley, B. A. Boville, J. T. Kiehl, and L. E. Buja. 2001. Climates of the Twentieth and Twenty-first Centuries Simulated by the NCAR Climate System Model. *J. Clim.* 14: 485–519.

Dettinger, M. D., D. R. Cayan, M. K. Meyer, and A. E. Jeton. 2004. Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900–2099. *Climatic Change* 62: 283–317.

Ghil, M., M. R. Allen, M. D. Dettinger, K. Ide, D. Kondrashov, M. E. Mann, A. W. Robertson, A. Saunders, Y. Tian, F. Varadi, and P. Yiou. 2002. Advanced spectral methods for climatic time series. *Rev. Geophys.* 40(1), 1003, doi:10.1029/2000RG000092, 1–41.

Ghil, M., and S. Childress. 1987. *Topics in geophysical fluid dynamics – Atmospheric dynamics, dynamo theory, and climate dynamics.* Springer-Verlag: New York, 485 p.

Houghton, J. T., et al. (eds.). 2001. *Climate Change 2001: The scientific basis: Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, 525–582.

Jeton, A. E., M. D. Dettinger, and J. L. Smith. 1996. *Potential effects of climate change on streamflow, eastern and western slopes of the Sierra Nevada, California and Nevada.* U.S. Geological Survey Water Resources Investigations Report 95–4260, 44 p.

Krishnamurti, T. N., C. M. Kishtawal, T. LaRow, D. Bachiochi, Z. Zhang, C. E. Williford, S. Gadgil, and S. Surendran. 2000. Multimodel superensemble forecasts for weather and seasonal climate. *J. Climate* 13:4196–4216.

Miller, N. L., K. E. Bashford, and E. Strem. 2003. Potential impacts of climate change on California hydrology. *J. Amer. Water Resources Assoc.* 39: 771–784.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski. 2003. Breaking the cloud parameterization deadlock. *Bull. Am. Meteorol. Soc.* 1547–1564.

Richardson, D. S. 2001. Measures of skill and value of Ensemble Prediction Systems, their interrelationship and the effect of ensemble size. Q. J. R. Meteorol. Soc. 127:2473–2489.

Stewart, I., D. R. Cayan, and M. D. Dettinger. 2004. Changes in snowmelt runoff timing in western North America under a 'Business as Usual' climate change scenario. *Clim. Change,* 15 p.

Toth, Z., O. Talagrand, G. Candille, and Y. Zhu. 2003. Probability and ensemble forecasts, *in* I. T. Jolliffe and D. B. Stephenson (eds.), *Forecast Verification: A practitioner's guide in atmospheric science.* San Francisco: John Wiley & Sons, 137–164.

Van Rheenen, N. T., A. W. Wood, R. N. Palmer, and D. P. Lettenmaier. 2004. Potential implications of PCM climate change scenarios for Sacramento-San Joaquin River basin hydrology and water resources. *Clim. Change* 62: 257–281.

Wigley, T. M. L., and S. C. B. Raper. 2001. Interpretation of high projections for global-mean warming. *Science* 293: 451–454.

Zhu, Y., Z. Toth, R. Wobus, D. Richardson, and K. Mylne. 2002. The economic value of ensemble-based weather forecasts. *Bull. Am. Meteorol. Soc.* 83:73–83.

Appendix

Component-Resampling Method

Consider an ensemble of n forecasts of, say, temperature at a given model grid cell, each m days (or years) long, and each containing elements $\{x_{ij},\ i=1,\ m\}$ where j indicates the ensemble member. A PCA of the n forecast vectors, with all expectations calculated across the ensemble members, will decompose the original ensemble into m loading patterns $\{e_i,\ i=1,\ m\}$ and m corresponding sets of amplitudes $\{p_{ij},\ j=1,\ n\}$. By construction, the loading patterns (like empirical orthogonal functions) and the amplitudes (like principal component series) both form orthogonal bases for describing the original ensemble. The amplitudes measure the projections of the original ensemble members onto the set of independent normal modes of the variations (the loading patterns), and, by construction, the amplitude of one of the ensemble members projected onto a particular loading pattern is statistically independent of its projection onto any of the others. The k-th original (properly standardized) ensemble member can be recovered completely by:

$$x_{ik} = \sum_{j=1}^{j=n} e_{ij}\ p_{jk}$$

Another prediction vector that is indistinguishable from the original ensemble elements, to second order, can be obtained by scrambling the amplitudes (picking the k indices in equation (1) at random, with replacement, from $k=1,\ n$, at each step in the summation). Because the amplitudes are independent from loading pattern to loading pattern, statistically the first and second-order statistics do not depend on which one is chosen at each step.

The procedure is as follows:

1. Calculate the ensemble mean values μ_i of x_{ij} at each time i, with expectation taken across the ensemble:

$$\mu_i = \sum_j x_{ij} / n$$

and subtract these means from the forecast vectors to obtain an ensemble of centered (zero-mean at each lead time) forecast vectors. Any mean or trend shared by all the ensemble members is removed by this step. When the ensemble is resampled later, this temporally varying ensemble mean can readily be added again.

2. Calculate the ensemble standard deviations σ_i of the centered forecasts at each time *i*, again with expectation taken across the ensemble, and divide the centered forecast vectors at each lead time by the corresponding standard deviation to obtain a standardized forecast ensemble (zero mean and unit variance at each lead time). This ensures that any temporal evolution of the spread of the ensemble is captured and can be reintroduced after resampling. Removing the temporally varying standard deviation at this point in the analysis ensures that inter-ensemble variations in the early part of the forecasts (when the ensemble typically has not spread much) are treated in the same detail as those later in the forecasts.

3. Compute the cross correlations of the standardized forecasts at each time and lag, with expectations taken across the ensemble. The resulting cross-correlation matrix is m x m, and summarizes the covariation of the day-1 forecasts in each ensemble member with the day-2, day-3 (and so on) forecasts in the same ensemble member.

4. This cross-correlation matrix is decomposed into loading patterns and their attendant amplitude series by a simple PCA. The loading patterns describe the temporal evolution of the ensemble members in the most economical form. For example, perhaps most ensemble members trend throughout from warmer towards cooler, while a few might increase for a while and then decrease like the others. The two behaviors would tend to be captured by two distinct loading patterns, and those ensemble members in the former category would be weighted more heavily in the former loading pattern; whereas, the latter ensemble members would be weighted more heavily on the latter pattern. These weights are measured by the respective amplitude series. For a given loading pattern, the amplitude series measures the weight (similarity) of each ensemble member in turn to that pattern, and a given ensemble member's amplitude of any of the loading patterns is statistically independent of its weight on any other loading pattern.

5. Randomly resample the PCA results to generate as many additional "forecast" realizations as necessary. Because the amplitudes for the various loading patterns are, by construction, independent of each other, by construction it does not matter which ensemble member's amplitude for a given loading pattern is mixed with which other ensemble member's amplitude for another loading pattern. With m loading patterns, each of which can take on any of the n amplitudes, the number of distinct resamples that can be constructed is m^n; e.g., in a 10-member ensemble of 14-day forecasts, 14^{10}

resamples can be generated. If, as in many PCA, only about 20% of the amplitude series contributed much variance to the recontructions, the effective candidates for independent samples would drop to perhaps 3^{10} or about 60,000 possible independent samples, which is still a useful expansion of the apparent size of the ensemble.

6. Having reconstructed a "new" member of the standardized forecast ensemble by resampling (in Step 5), rescale the result by the time-varying ensemble standard deviations and then add the time-varying ensemble means. By this rescaling, the stationary and shared variability are restored, and the large numbers of results can be ranked and summarized in detailed histograms to obtain pdfs as fine as desired.

The method ensures that features shared by all members are shared by the component-resampled ensemble members, that variations shared by subsets of the ensemble members are reproduced realistically and in proportion to their occurrence in the original ensemble, and even that the noisy (unshared) variations are faithfully captured and reproduced in the component-resampled ensemble. Because the method is based on PCA, the component-resampled ensemble is described mostly in terms of its first and second statistical moments, so that the resulting smoothed pdfs tend toward Gaussian shapes; however, that tendency is relatively weak. When present, forecast bifurcations should be captured in the PCA loading patterns and weighted appropriately (in both amplitude and numbers of participating ensemble members) by the corresponding amplitude series. Then, when the amplitude series are resampled randomly, both the shapes and relative frequencies of the bifurcations are naturally reproduced in a satisfying way. Because of this property of the components that are resampled, even while the general divergence attributable to model and scenario differences, along with sensitive dependences upon initial conditions, any occasional, temporary confluences of ensemble members associated with visits near ghost limit points and cycles (Ghil et al. 2002), and in homoclinic orbits (Ghil and Childress 1987), shared by the ensemble members are also captured in a natural way.

One way to picture the method is to imagine that the original ensemble has been filtered into a large number of narrow and nonoverlapping frequency bands. The result is that an ensemble member has power A in the first frequency bin, has power B in the second frequency bin, and so on. Another ensemble member would have a different set of powers in each frequency bin. Now, assuming that an ensemble's power in the first frequency bin has little bearing on its power in the second, and so on, one can imagine generating new ensemble surrogates with the same statistical properties as the original ensemble, by taking the filtrate from one ensemble member (at random) from the first frequency bin, adding to it the filtrate from an ensemble member (at random) the second bin, and so on, until samples from all the frequency bins have been incorporated. The sum of the frequency components constitutes a new time series with statistical properties that are derived strictly from the ensemble's overall power spectrum. For example, if the 10-day periodicities in the ensemble members were most powerful and 8-day periodicities notably lacking, the resampling would still yield members with powerful 10-day periodicities and weak 8-day periodicities, because the resampling only uses observed values from each frequency bin. The present method improves on such a hypothetical frequency-binned resampling by (1) guaranteeing—by the construction of the PCA decomposition of the ensemble members—that the various elements resampled (the loading patterns, which would correspond to sine waves of given frequency in the hypothetical) are always independent of each other, and (2) allowing more flexibility of loading-pattern shape than is offered by a simple frequency-domain approach.