CHAPTER 5

Computational Data Sciences for Actionable Insights on Climate Extremes and Uncertainty

Auroop R. Ganguly, Evan Kodra, Snigdhansu Chatterjee, Arindam Banerjee, and Habib N. Najm

CONTENTS

5.1 OVERVIEW AND MOTIVATION

5.1.1 Climate Extremes: Definitions and Concepts

The Intergovernmental Panel on Climate Change (IPCC) SREX (IPCC, 2011) summary for policymakers defines climate extremes as follows:

> The occurrence of a value of a weather or climate variable above (or below) a threshold value near the upper (or lower) ends of the range of observed values of the variable. For simplicity, both extreme weather events and extreme climate events are referred to collectively as "climate extremes."

Climate extremes in this chapter are defined as extreme weather events, or those that may last from several hours to several days. Thus, they may include heat waves and cold snaps, rainfall patterns in space and time potentially leading to floods and droughts, tropical cyclones, tornadoes, and storm surges. Figure 5.1 provides an overview.

There is evidence that statistical attributes of certain climate extremes have been growing steadily and significantly worse as a result of human influence, and these changes can be projected from analysis of physics-based computational climate model simulations as well as observations from remote or *in-situ* sensors. However, climate science cannot predict any particular event at decadal to centennial scales or assign a specific cause, and the confidence in statistical projections differs by the variable considered, the extent of spatial or temporal aggregation, regional and seasonal characteristics, and other considerations. Thus, we have relatively higher confidence in projections of temperature-related extremes, followed by extremes of precipitation and tropical cyclones. The climates of the extra-tropics are often relatively easier to project than that of the tropics, while statistical properties of extremes and change are typically better projected at aggregate space-time scales compared to finer resolutions.

5.1.2 Societal and Stakeholder Priorities

Stakeholder communities across multiple sectors such as water and food security, natural hazards preparedness and humanitarian aid, or

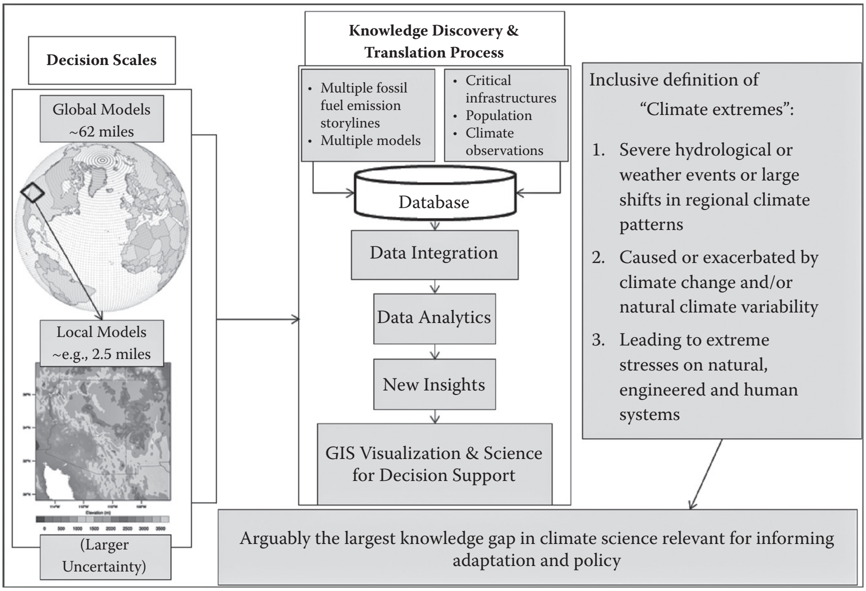

FIGURE 5.1 Uncertainty quantification for climate extremes, which are broadly construed in this context, represents one of the largest knowledge gaps in terms of translating the physical science basis of climate to information relevant for impacts assessments and adaptation decisions, and eventually to mitigation policy. However, the cascade of uncertainties is difficult to quantify. The societal costs of action and inaction are both potentially large for climate adaptation and mitigation policies; hence, uncertainties in climate are important to effectively characterize and communicate. Climate extremes may broadly include large shifts in regional climate patterns or severe weather or hydrological events caused or exacerbated by natural climate variability or climate change. This chapter primarily focuses on the statistical attributes of severe events, or, changes in tail behavior.

management of natural and engineered infrastructures, as well as policy makers dealing with urbanization, population growth, or migration, land use or energy and water sustainability, and energy or emissions control, require actionable insights about climate extremes at local to regional scales. The costs of action and inaction can both be large, as adaptation and mitigation strategies designed for climate extremes may be costly for the current generation and potentially even more costly for future generations. Specific examples of prior work that demonstrated the importance of climate extremes to the stakeholder community include the following: an international climate change policy related to global emissions negotiations (Engelhaupt, 2008; Tollefson, 2008a, b); national security decisions related to regional threat assessments and preparedness, for example, the 2010 Quadrennial Defense Review report of the United States Department of Defense (Ganguly et al., 2009a); and a war game related to the Arctic sea ice (cited in NRC, 2011).

5.1.3 Computational Data Sciences: Challenges and Opportunities

Computer science, informatics, and computational (or cyber) infrastructures have played a major role in our current understanding, projections, and attributions of climate extremes, primarily through physics-based models. The climate system is nonlinear, dynamical (and often chaotic, or very sensitive to initial conditions), subject to multiple feedback mechanisms (e.g., among ocean, atmosphere, and land processes), thresholds and intermittence (e.g., for precipitation and cloud physics), exhibits low-frequency (and even "*1/f*") variability and complex dependence structures (e.g., long-memory processes over time and long-range spatial dependence or "teleconnections" in space), as well as nonstationary (e.g., the relative dominance of processes generating extremes may change in a warmer world). Thus, purely data-driven extrapolation may not be adequate or even appropriate, especially for long lead time projections (e.g., decadal to centennial scales), where data assimilation methods may also have limited value. State-of-the-art physical climate models are based on fundamental physical laws (e.g., laws of motion and conservation of mass and momentum). Physical approximations and mathematical discretization techniques (e.g., strategically chosen finite difference equation system formulations) are applied to these laws, resulting in complex systems encapsulated in hundreds of thousands or millions of lines of low-level source code (Christensen et al., 2005). Current global climate models are composed of multiple interacting components, including atmospheric, oceanic, and often land and sea ice models (IPCC, 2007). Such physics-based models, whether global climate or Earth system models or regional climate models used to downscale the outputs of global models, are more credible for variables such as mean temperature at continental to global scales. The same models are less reliable for climate extremes; for example, they are inadequate for precipitation extremes and tropical cyclones, especially at the precision required for making informed decisions.

The research opportunities for computational data sciences are threefold: (1) improved characterization, projections, and attribution of climate extremes; (2) characterization of uncertainties, especially at local to regional scales for annual or seasonal projections over decades and centuries; and (3) enhanced predictive insights over and above what may be obtained from direct extrapolation of historical trends or analysis of climate model simulations. The volume of the data (e.g., hundreds of terabytes going on petabytes for archived climate model simulations, and gigabytes going on terabytes for remotely sensed observations) and the complexity of the methods (Lozano et al., 2009a, b; Steinhaeuser et al., 2011a, b) require data-intensive computational methods. A schematic is shown in Figure 5.2.

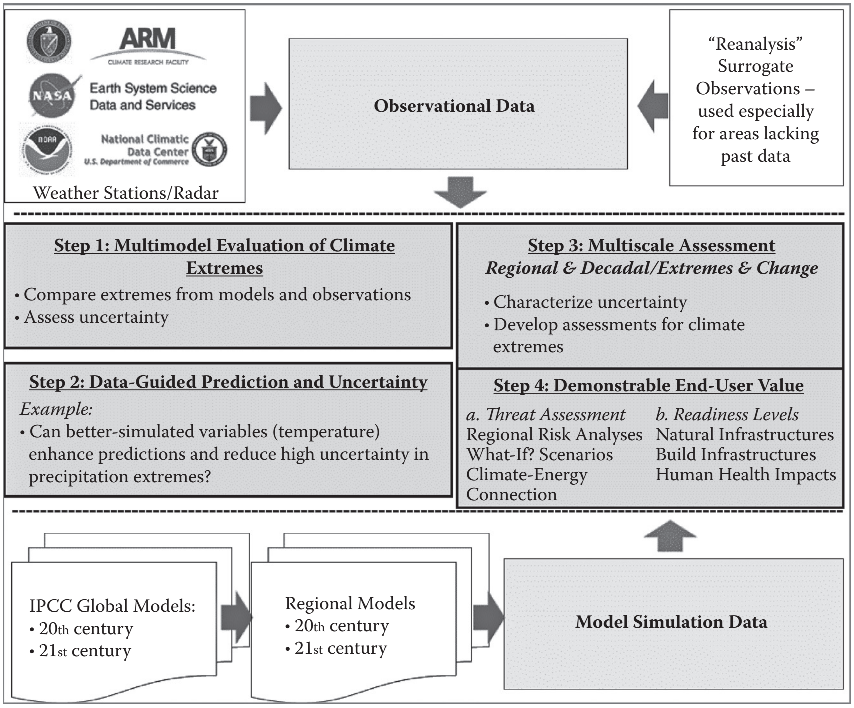

FIGURE 5.2 Remote or *in-situ* sensor observations and climate model simulations can be investigated through computational data science methods for multimodel evaluations, enhanced projections, and multiscale assessments to inform decisions and policy. The growth in climate data from models and observations is expected to grow exponentially over the next several decades (Overpeck et al., 2011), providing a vast set of challenges and opportunities for data science communities.

5.1.3.1 Overview of Research Areas: 1. Extremes Characterization

Extremes may be characterized based on their relevance to impacts (e.g., heat waves based on nighttime minima, which relate to human discomfort and possible loss of lives: Meehl and Tebaldi, 2004) or through statistical distributions (e.g., extreme value theory for precipitation extremes: Kharin et al., 2007). Thus, our analysis (Ganguly et al., 2009b) of model simulations and surrogate observations (reanalysis data generated by assimilating weather data from disparate sensors into a numerical weather prediction model) pointed to higher trends but larger uncertainty in heat waves in this century based on a plausible but high emissions scenario, which in turn implies greater urgency but caution in adaptation or mitigation decisions. On the other hand, our analysis (Kodra et al., 2011a; see report by Tollefson, 2011) of multiple model simulations and reanalysis data revealed that while extreme cold events may grow less frequent, the intensity and duration of the ones that do occur may often persist at current levels. Our analysis of intense precipitation events (Kao and Ganguly, 2011) suggested an amplification of extremes, especially over the extratropics and in an aggregate sense at continental to global scales. Major community-wide efforts are necessary to comprehensively characterize the statistical attributes of gradual or sudden changes in extremes over space and time, including less well-defined or predictable climate extremes such as droughts. A combination of state-of-the-art methods, new methodological adaptations, and novel approaches in spatial or spatiotemporal statistics and data mining are motivated.

5.1.3.2 Overview of Research Areas: 2. Uncertainty Assessments

Uncertainty assessments for extremes (Wehner, 2010) need to consider knowledge-gaps in model physics (e.g., based on statistical methods to balance model skill in the past and multimodel convergence in the future through extensions of approaches such as Smith et al., 2009), and the applicability and statistical validity of the definitions or distributions of extremes, as well as uncertainties in parameter estimation processes (e.g., through the bootstrap as in Kharin et al., 2007 or Kao and Ganguly, 2011). The potential differences in the nature of the insights and uncertainties based on definitions of extremes become obvious by comparing our recent work (Ghosh et al., 2011) with a previous approach (Goswami et al., 2006). Current methods for attribution of extremes, e.g., for intense precipitation events, include statistical techniques (Min et al., 2011) or numerical simulations (Pall et al., 2011): these methods can benefit from rigorous uncertainty quantification approaches. New mathematical methods for uncertainty are critically needed in these areas.

5.1.3.3 Overview of Research Areas: 3. Enhanced Predictions

Large gaps continue to exist in our scientific understanding and projections of certain crucial variables, often related to climate extremes, and fine-scale or aggregate processes that drive the extremes. There have been claims that *the sad truth of climate science is that the most crucial information is the least reliable* (Schiermeier, 2010). One question relevant for enhanced projections of climate extremes is the extent to which the variables that are relatively better predicted (e.g., sea surface temperature or ocean meteorology in general, atmospheric temperature or humidity profiles over land) may have information content about the variables that may be more crucial (e.g., precipitation extremes or intensity and frequency of hurricanes), and whether such information can be utilized for developing predictive models. There is evidence of information content; for example, Liu et al. (2009) and others have reported on temperature dependence of precipitation extremes. And there is literature attempting to develop physics-based relations; for example, O'Gorman and Schneider (2009), Sugiyama et al. (2010), as well as Muller and O'Gorman (2011), have sought to develop a better understanding of precipitation processes related to extremes based on atmospheric covariates, while Emanuel et al. (2008) have attempted to produce projections of the statistical attributes of hurricanes based on climate model-simulated oceanic variables. While data-driven methods should be explored to develop novel and actionable predictive insights, the methods have to be able to handle nonlinear processes as well as complex dependence patterns, yet remain physically interpretable and able to generalize to nonstationary conditions. This area may represent a steep challenge for computational data sciences, and perhaps motivate truly interdisciplinary approaches conceived from traditionally disparate methods ranging from computational statistics, signal processing, and econometrics, to nonlinear dynamics, graph-based methods, data mining, and machine learning. Our recent efforts (Steinhaeuser et al., 2011a, b and Chatterjee et al., 2012) for improved regional predictions over land based on ocean meteorological variables or Kawale et al., 2011 for understanding climate dipoles) are only initial steps in an area that may well represent a grand challenge for data-intensive sciences.

5.2 EXTREMES CHARACTERIZATION

There are several challenges in characterizing and analyzing data related to climate extremes. One of the first challenges is the nature of the data: observational data are of relatively short duration and typically do not allow for many important extreme conditions to be manifest, they are unevenly spread spatially, and data quality is also uneven. Climate model outputs and reanalysis data do not have several of these problems, but Mannshardt-Shamseldin et al. (2011) demonstrate that the nature of extremes from gridded data differ considerably from observed data. Moreover, as Wehner (2010) observes:

> …to the extent that climate models can be tuned to reproduce the recent past, model developers focus on the mean values of climatic observations, not the tails of their distributions.

Several other studies (O'Gorman and Schneider, 2009; Sugiyama et al., 2010; Wehner, 2010; Min et al., 2011) have pointed out that our current understanding of precipitation extremes has room for improvement and that the current generation of climate models probably fails to reflect reality. Recent studies (O'Gorman and Schneider, 2009; Sugiyama et al., 2010) suggest a deficiency in our understanding of the relationship between precipitation extremes and atmospheric moisture content. Wehner (2010) suggests that climate models might actually be able to emulate extreme events if they were run at sufficiently high resolution; that is not the case for models run at the typical resolution level adopted for the "International Panel on Climate Change's Fourth Assessment Report" (AR4, a landmark assessment report on the state of climate change science) (IPCC, 2007), and higher-resolution runs are computationally expensive. Min et al. (2011) point to the possible underestimation of trends in precipitation from a potential lack of accounting for anthropogenic effects on future precipitation extreme events. All of these suggest that there is room for improvement in the quality of data, and in the development of methodology to analyze available extreme data.

Another challenge for statistical and data-driven analysis of climate extremes is that the definition of *what* is extreme should be guided by various stakeholders and users. For example, in the case of rainfall, (1) the amount of rainfall in a given unit of time, (2) the total amount of rainfall, (3) the duration of rainfall, (4) the time gaps between rainfall events,

(5) the spatial pattern of rainfall, and several other variables can be of interest. Accordingly, the definition and the notion of an extreme may be different. The trifecta of intensity, duration, frequency (IDF), which is often characterized using extreme value theory (Kao and Ganguly, 2011) is useful in many cases, but not all. Another example is that of cold temperatures, which are important for crop production and food security. The variables of interest in this example could be the number of days of a certain level of frost, consecutive frost days, and time spent below a temperature threshold (Kodra et al., 2011a). Not all such definitions of "extremes" lend themselves to common, theoretically satisfying statistical analysis (Coles, 2001).

Another potential problem is that of identification of *extreme* events versus *rare* events, which are not always the same; in other words, an event might be extreme in impact but not rare, and vice versa. The definition of an extreme event may often be determined by its impact, and this definition will, in turn, often determine its rarity. The rarity of the defined events, along with other data properties, will dictate which statistical inference approaches may be appropriate. In some cases, summary measures have been used to obtain conclusions about extreme events (Goswami et al., 2006), although subsequent uses of the extreme-value model have provided different conclusions on similar data (Ghosh et al., 2011). Also, as Ferro and Segers (2003) observe, extremes can be clustered, which may present additional challenges related to the independence assumed by some extreme value statistical approaches.

From a purely statistical perspective, there is a gap between finite sample data-based extreme events and the general asymptotic theory that is used for extreme event analysis. Classic extreme-value statistical approaches attempt to extrapolate the extreme behavior of variables by fitting distributions to tail observations, such as annual maxima or exceedances above or below some predetermined (quantile) threshold (Coles, 2001; Kharin et al., 2007). Note that the generalized extreme value distribution or the generalized Pareto distribution, which have been used in the climate extremes literature (Kharin and Zwiers, 2000; Perkins et al., 2009; Kao and Ganguly, 2011), are asymptotic limits of probabilities relating to finite-sample size extreme events, and need not be exact characterizations. Also, most real data are temporally and spatially correlated, a fact that is often ignored in computing return-level characteristics, quantifying uncertainty, or making inference. There is no consensus about the best parameter estimation and inference technique for extreme-value distributions (Hosking and Wallis, 1997; Kharin and Zwiers, 2000; Coles and Dixon, 1999; Kharin and Zwiers, 2005), and approaches for including information from covariables are still in development (Hall and Tajvidi, 2000).

The bootstrap, broadly speaking, is a class of resampling techniques that can be employed to quantify sampling variability (uncertainty) in parameter estimation, among other uses (Efron, 1979). Parametric bootstrap and the traditional nonparametric bootstrap approaches of Efron (1979) were used in conjunction with the L-moments method and the maximum likelihood method for studying climate extremes in Kharin and Zwiers (2000; 2005; 2007), who also compared various estimation techniques and listed several caveats. Inference for climate extremes may benefit from a better understanding of the limits and appropriateness of popular statistical inference procedures (such as extreme value theory), as well as the application and/or creation of other approaches that relax assumptions or are robust to limitations of available extreme data.

5.3 UNCERTAINTY ASSESSMENTS

5.3.1 Statistical Modeling of Uncertainty in Multimodel Ensembles

Here we discuss the state-of-the-art in uncertainty quantification (UQ) for situations where ensembles of global climate models or Earth system models (GCMs/ESMs) are used to assess regional climate change. While statistical and dynamical (regional climate models) downscalings are often used for regional assessments, they are in turn driven by ESMs, and hence UQ in ESMs remains an important challenge. UQ is often inundated with a sense of urgency, and ensembles of ESMs are tools from which practical and timely uncertainty assessments can be readily formed.

Structural uncertainty, or that which arises from variations in the mathematical mechanics of climate models, is the principal focus of UQ in approaches discussed in this section; it has been studied in several forms with multimodel ensembles where weights are assigned to individual models as a measure of their reliability. We distinguish the ensemble approaches discussed here from other UQ methodologies—for example, physics perturbed ensembles—that have been used to explore parameter uncertainty within single climate models (Stainforth et al., 2005), and approaches based on or similar to polynomial chaos expansion (see Section 5.3.2). It is important to be aware of all approaches for

UQ to understand the scope and limitations of multimodel ensembles for climate UQ: to date, no statistical multimodel ensemble UQ methodology explicitly incorporates uncertainty within climate models (e.g., to understand the uncertainties contributed by parameterizations of climate processes and the propagation of these uncertainties along the rest of the model components). The ensemble methods discussed here, however, provide a basis for exploring inter-model (i.e., structural) uncertainty.

While the use of multimodel ensembles for prediction has been studied extensively in data science disciplines (Seni and Elder, 2010), an important distinction must be made for the climate science domain. In many typical time series and classification applications, for example, a forecast horizon of interest is often one or two (or a few) periods ahead, or the binary classification is for the next observation (Seni and Elder, 2010). In such cases, there is the potential for a predictive model to learn from an ensemble of predictions and recalibrate its next prediction upon validation (Fern and Givan, 2003). Several key challenges, however, distinguish climate forecasting from more typical problem settings: long prediction lead times (multidecadal to centennial scales), potential nonstationarity (where the validity of processes embedded within a GCM may change), the difficulty in selecting metrics that are meaningful for model training (Knutti, 2010), and finally the impossibility of true validation for the prediction horizons of interest. Many of the methods developed in data sciences, while valuable on their own and across applications, do not directly translate well to climate prediction. For example, even in the case where past data are segmented into multiple training samples in order to rigorously develop a multimodel ensemble prediction formula founded on well-chosen, physically meaningful error metrics, there is no guarantee that nonstationarity will not invalidate the prediction formula in the future, given the lead time. Thus, while mature data science ensemble methodologies may be valuable as foundational approaches, novel and creative methods are needed to solve the problem of UQ in climate with multimodel ensembles.

One persisting notion in the climate literature is that the multimodel average (MMA), or the average of spatial, temporal, or spatiotemporal fields of climate outputs from multiple GCMs, is a robust approach for making "most likely" projections; this robustness is largely based on alleged bias (noise) cancellation and orthogonal skills of GCMs (Knutti et al., 2010). The concept of its potential utility in climate may have followed from success in weather forecasting (Krishnamurti et al., 1999; Hagedorn et al.,

2005) and has been empirically justified in climate attribution studies (Pierce et al., 2009; Santer et al, 2009); in fact, the latter studies implicitly suggest that the MMA of a random selection of an adequate number of GCMs will form a reliable projection, at least for anthropogenic attribution. Equal-weighted MMAs may represent a more conservative approach, as optimum model weighting may hold less potential benefit than risk compared to equal weighting (Weigel et al., 2010). The MMA has been a default in some climate stakeholder assessment reports (Karl et al., 2009), where they are sometimes displayed visually without clear reference to uncertainty. This may be a questionable practice, as averaging of dynamically consistent spatial fields or time series may lead to physically meaningless signals, and the exclusion of individual model results may serve to obscure plausible lower and upper bounds of climate change (Knutti et al., 2010). A recent case study (Perkins et al., 2009) has also implicitly questioned the notion of *a priori* use of MMAs and inclusion of demonstrably poor GCMs. Work with historical (20th century) observations and simulations of Indian monsoon climatology (Kodra et al., 2012) may suggest that the MMA should not be a default choice, and that all models within an ensemble should be evaluated carefully.

Given this debate surrounding MMAs, the display of worst and best cases as derived from archived GCM outputs may be advisable as the bare minimum requirement for communicating uncertainty. However, because multimodel ensembles are not true random samples of independent GCMs from a larger population, they should not be considered formal probability distributions (Tebaldi and Knutti, 2007). More rigorous and statistically grounded approaches may be desired; recently, several notable probabilistic approaches have been developed: Giorgi and Mearns (2002) proposed the Reliability Ensemble Average (REA) method for assigning reliability to simulations of regional mean temperature from GCMs; this method has since been expanded and developed into a more formal Bayesian framework that has become perhaps the most prominent method for UQ using multimodel ensembles. Essentially, the REA method attempts to weight GCMs based on their alleged reliability, which is a balance of (1) historical model bias relative to observations and (2) future multimodel consensus (convergence), or model distance from the center of the ensemble spread. Giorgi and Mearns (2002) define bias and convergence as the following, respectively:

$$\lambda_{B,j} = \min\left(1, \frac{1}{|x_j - \mu|}\right), \tag{5.1}$$

$$\lambda_{C,j} = \min\left(1, \frac{1}{|y_j - \tilde{Y}|}\right), \tag{5.2}$$

where X_j is a historical (20th century) temperature output from climate model j, and μ is observed ("true") temperature from the same time period; Y_j and $\tilde{Y}$ are, respectively, the same for a future time period. If the arguments

$$\frac{1}{|X_j - \mu|} \text{ or } \frac{1}{|Y_j - \tilde{Y}|}$$

are more than 1, then 1 is chosen as $\lambda_{B,j}$ or $\lambda_{C,j}$, with the notion that $|X_j - \mu|$ or $|Y_j - \tilde{Y}|$ could have been small just by chance. "$\tilde{Y}$" is an unknown and thus must be estimated using the following weights:

$$\lambda_j = \left(\lambda_{B,j}^m \lambda_{C,j}^n\right)^{1/mn}, \tag{5.3}$$

$$\tilde{Y} = \frac{\Sigma_j \lambda_j Y_j}{\Sigma_j \lambda_j}. \tag{5.4}$$

In practice, Giorgi and Mearns (2002) arbitrarily set $m = n = 1$ so that bias and consensus received equal favor. Because $\tilde{Y}$ and $\lambda_{C,j}$ are both unknowns but depend on each other, a recursive procedure was used to compute both (Giorgi and Mearns, 2002).

These two criteria suggest that information on the credibility (weights) of models can be estimated by performance compared to observed data, as well as degree of convergence; if a model falls far from the center of the ensemble, it may be treated more like an outlier than a likely outcome. The consensus criterion may have been borne at least partially from the ideas of bias cancellation and orthogonal skills of MMAs discussed previously (Krishnamurti et al., 1999; Hagedorn et al., 2005), and strong criticisms of the criterion have been acknowledged (Tebaldi and Knutti, 2007). Additionally, the criterion of GCM skill (*bias* in most recent work) is difficult to define and evaluate; in most cases, it is difficult to determine whether metrics measuring past GCM skill will translate to the future (Tebaldi and Knutti, 2007; Knutti et al., 2010).

The REA was admittedly ad hoc; however, its two notions of GCM skill and consensus (Giorgi and Mearns, 2002) have formed the foundation for a prominent line of work, beginning with Tebaldi et al., (2004), that formalized them in a Bayesian model. One of the most recent versions of this statistical model can be found in Smith et al. (2009), which also allows for the joint consideration of multiple regions. Specifically, using this model, a posterior distribution for past or current temperature μ and future temperature υ can be simulated from a Markov Chain Monte Carlo (MCMC) sampler using a weight λ_j for each GCM j. Next, each λ_j can be simulated from a posterior by considering the bias and consensus of GCM j. The weights λ_j then inform a new estimate of μ and υ, which informs a new estimate of each λ_j, and so on. Specifically, λ_j follows a Gamma posterior distribution with the following expectation:

$$E\left[\lambda_j \middle| \bullet\right] = \frac{a+1}{b+0.5\left(X_j-\mu\right)^2+0.5\theta\left(Y_j-\upsilon-\beta\left(X_j-\mu\right)\right)^2} \tag{5.5}$$

where a and b are prior parameters (usually uninformative), β is an unknown quantity representing the correlation between historical and future model outputs, and θ is a random quantity that allows future climate variance to differ from that of the past. The conditionality serves to illustrate the fact that the expectation of λ_j is contingent upon the values of all other random parameters and data. Thus, the value of λ_j is a function of the random quantities μ, υ, θ, and β, which in turn have their own posterior conditional distributions. In general, it is notable that under this formulation all of these random quantities are conditionally independent and cannot be readily analyzed marginally.

An initial post-hoc analysis (Kodra et al., 2011b) of simulation results from the latest model (Smith et al., 2009) suggested it may rely more on consensus and less on bias (i.e., skill), and that the bias criterion may not be adequate in representing model skill. Conceptually, in the state-of-the-art statistical model, holding all else constant, the posterior distribution for $\upsilon-\mu$ will "shrink" toward the multimodel mean, even if all GCMs exhibit poor skill with respect to past data; in such circumstances, it might

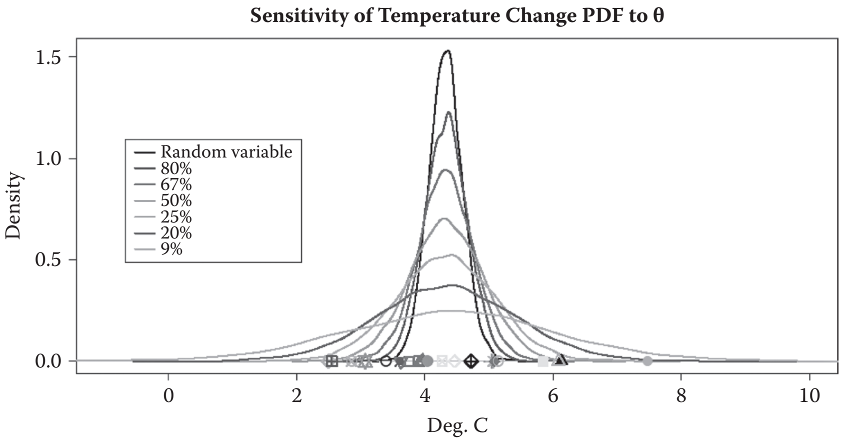

FIGURE 5.3 The univariate (one region) Bayesian model from Smith et al. (2009) illustrates the importance of the parameter θ in dictating the spread of the probability density function (PDF) for change in regional mean temperature. This particular set of PDFs is obtained for spatiotemporally averaged Greenland temperature change from 1970 to 1999 to 2070 to 2099. The horizontal axis measures change in degrees Celsius, while the vertical axis measures frequency. The legend indicates the condition of θ and from top to bottom corresponds to PDFs from narrow to wide: "random" is the narrowest density, and treats θ as a random unknown quantity as in Smith et al. (2009). For the remaining PDFs, values of θ are fixed at different quantities that come to represent the relative importance of convergence versus bias (where the importance of bias is 100% minus that of θ). Notably, treating θ as a random quantity yields a result where convergence is favored much more than bias.

make sense that that uncertainty should actually increase. The degree of shrinkage toward the mean is dictated by the parameter θ. Its importance is stressed in Figure 5.3 and clear from Equation 5.1: if $\theta \gg 1$, then holding all else constant, consensus is favored by the statistical model. Indeed, an earlier work by Tebaldi et al. (2004) featured a slight variant of their statistical model with a provision that, through a prior distribution, restricted the influence of θ; this provision was not included later by Smith et al. (2009). While the above represents the most prominent line of work on combining multimodel ensembles for quantifying uncertainty in regional climate, a few other initial approaches have been developed. These approaches have extended the broad multimodel UQ line of work by integrating methodology for quantifying inter-model covariance (Greene et al., 2006) and spatial variability structure (Furrer et al., 2007; Sain et al., 2008), Bayesian model averaging as a technique for model weighting (Min and Hense, 2007), and methodology for quantifying GCM biases (Buser et al., 2009).

Monteleoni et al. (2010) proposed an online learning approach where weights for GCMs can change with new data. More specifically, α-experts predict annual global mean temperature; the experts learn from the mistakes (i.e., squared errors) of GCMs and can actually predict one–year-ahead temperature better than the best prediction of all GCMs or the multimodel mean. While this is an intriguing approach for combining GCMs and it seems to handle one-step-ahead nonstationarity quite well, in its present form it does not allow for long lead time prediction, which is the primary goal of climate models in general. Additionally, uncertainty bounds have not yet been developed within the algorithm.

The regional climate UQ research area is still relatively nascent and may benefit from new approaches (Knutti et al., 2010). Methods must be developed that simultaneously accommodate long lead times as well as potential nonstationarity, where conditions (e.g., relationships between variables) could change. In addition, there may be value in considering physical relationships between multiple variables to encourage robustness and interpretation in the GCM weighting process, as is discussed in Section 5.4. Recent UQ work (Tebaldi and Sanso, 2009), along the same line as Smith et al. (2009), developed a hierarchical Bayesian model for joint projections of mean temperature and precipitation; this model attempted to statistically utilize correlation between the two variables. Extending this type of statistical model to one that considers climate extremes would be valuable to the climate community. For instance, O'Gorman and Schneider (2009) developed a relatively simple conceptual physical model for an increase in precipitation extremes under anthropogenic climate change, while Kao and Ganguly (2011) explored conceptual physical bases in their characterization of 21st century precipitation extremes. Careful statistical utilization of such insights could lead to novel multimodel ensemble UQ methodology for extremes, which could be useful for informing various impact sectors.

5.3.2 Parametric Uncertainties in Individual Climate Models

While approaches for quantifying structural uncertainty were discussed in Section 5.3.1, accounting for parametric uncertainty is important for understanding intrinsic model variability, especially because archived global models do not traverse the entire space of plausible prediction space

(Stainforth et al., 2005). Climate models employ a wide range of parameterizations in land, ocean, atmosphere, and ice sub-model components. These parameters are typically calibrated based on data. Bayesian methods are useful for parameter estimation with quantified uncertainty, as they provide estimation of the joint posterior density of the parameters. While Bayesian methods have been used for parameter estimation in climate model components (Jackson et al., 2003, 2004, 2008; Annan and Hargreaves, 2006; Govaerts and Lattanzio, 2007; Villagran et al., 2008), much more remains to be done in general for estimating the wide variety of climate model parameters. One key challenge is the construction of effective climate model surrogates for facilitating the Bayesian/inverse problem solution. Gaussian processes, as well as other surrogate models, have been used for representing climate data (Jackson et al., 2004; Banerjee et al., 2008; Cressie and Johannesson, 2008; Sanso et al., 2008; Villagran et al., 2008; Drignei et al., 2009; Furrer and Sain, 2009). High dimensionality, in the present context largely embodied in the large number of uncertain parameters, is a significant challenge for surrogate construction and for uncertainty quantification in general. Global sensitivity analysis methods (Morris, 1991; Saltelli et al., 2000) have been used to identify a small subset of parameters that are critical to the climate system outputs of interest. Alternatively, compressed sensing (CS) methods have been developed to provide means of constructing sparse representations of high-dimensional information (Donoho, 2006; Candès and Wakin, 2008). Bayesian CS methods have been used for discovering sparsity in land models (S. Babacan et al. 2010).

With parameters calibrated using Bayesian methods, and described using a posterior density, one can estimate the forward propagation of uncertainty in model observables resulting from the input parametric uncertainty. While this can be done via random sampling, employing the parameter posterior density, model complexity, and computational cost render this an infeasible task due to the large number of random samples needed for convergence of Monte Carlo methods (Caflisch, 1998). In recent years, Polynomial Chaos (PC) UQ methods have been developed and used to great advantage in representing uncertain/random variables, and in accelerating forward uncertainty propagation in computational models (Ghanem and Spanos, 1991; Xiu and Karniadakis, 2002; Debusschere et al., 2004; Le Maitre et al., 2004; Soize and Ghanem, 2004). PC methods rely on a representation of random variables as truncated expansions in terms of orthogonal functions of a given random basis. Thus, an uncertain model parameter χ can be written as a PC expansion (PCE) in terms of an n-dimensional basis:

$$\chi \approx \sum_{k=0}^{P} \alpha_k \Psi_k \left(\xi_1, \xi_2, \ldots, \xi_n\right) \tag{5.6}$$

The key task is then the propagation of uncertainty from parameter χ to climate model output Z, where $Z = \mathcal{H}(\chi)$ (Najm, 2009). Generally, PC methods come in two variants (Najm, 2009): (1) intrusive methods, where the governing equations are transformed employing Galerkin projection, arriving at a new set of equations for the PC mode coefficients; and (2) nonintrusive methods, where the original model is employed as a black box in a sampling context. The former approach requires changing the source code/forward solvers of the underlying physical model. The latter is the more practical, given established legacy codes. In this approach, the mode coefficients in the PCE for a model output of interest Z are evaluated employing a chosen set of N samples of the basis ξ, $\{\xi^j\}_{j=1}^{N}$ and associated model output values Z^j. Adaptive anisotropic deterministic sampling methods, employing sparse-quadrature evaluations of projection integrals for the PC coefficients, are highly effective in this regard (Nobile et al., 2008).

The outcome of the PC forward UQ problem is a PCE for the uncertain model outputs of interest, which can be employed to generate corresponding probability density functions, or moments of interest. Note that the PCE, being a functional representation of model outputs over the range of uncertainty in input parameters, is also useful simply as a surrogate for the dependence of model outputs on the parameters. PC surrogates have been used, very much like GP surrogates, for accelerating Bayesian inference (Marzouk and Najm, 2009).

The use of PC UQ methods in the global scale climate-modeling context is feasible in principle employing a nonintrusive formalism, although it is severely challenged by the high dimensionality of the stochastic input space, particularly given the computational expense of a single model run. Presuming known uncertainties of model parameters, targeted studies exploring small subsets of parameters are quite feasible. Sparsification via sensitivity analysis or CS methods, as well as utilization of known parameter correlations, and the hierarchical structure of climate models, is useful to reduce the effective dimensionality of the uncertain input space, thereby rendering the overall problem more feasible.

Finally, given that the most damaging consequences of climate change are associated with climate extremes, it is worthwhile to discuss the prediction of extreme behavior under uncertainty in model parameters. This is a more challenging pursuit than that of estimating means/moments of uncertain climate predictions when extreme events of interest are also rare. Capturing the tails of probability distributions accurately in a computational setting is challenging when there is small probability of sampling the tail region, and hence implies the need for a very large number of samples overall. While PC methods can avoid random sampling, the accuracy of the PC representation for modeling the tail-behavior of the underlying random variable requires particular attention. In principle, using high-order expansions may help; however, that renders the problem even more computationally challenging, requiring, for example, a large number of samples in a nonintrusive setting. It is also feasible to use a PC basis that is tailored to achieve higher accuracy in the tail region than conventional constructions. This can be done, for example, by choosing a basis that has a density with fat-tail behavior. This approach, however, still requires further development to be implemented practically with climate models.

5.4 ENHANCED PREDICTIONS

The key desiderata from predictive models in the context of extremes include accurate and uncertainty-quantified projections of crucial variables related to extremes as well as succinct characterizations of covariates and climate processes collectively influencing extremes. Such characterizations must be cognizant of complex and possibly nonlinear dependency patterns while staying physically interpretable, thereby yielding scientific understanding of extremes and the processes driving them. While uncertainty quantification methods were discussed in Section 5.3, we now briefly introduce some methodology that could be useful in enhancing predictions and perhaps reducing uncertainty in crucial climate variables that are not captured well by current-generation physical models.

Standard predictive models, such as least squares or logistic linear regression, fall short of such desiderata in multiple ways. The number p of covariates and fine-scale climate processes potentially influencing key variables such as extreme precipitation far surpass the number of examples n of such extreme events. In the $n \ll p$ regime for regression, consistency guarantees from standard theory breakdown, implying that the model inferred is not even statistically meaningful (Girko, 1995; Candès and Tao, 2007). Moreover, such standard models will assign nonzero regression coefficients to all covariates and processes provided as input without any model selection (Tibshirani, 1996; Zhao and Yu, 2006). The corresponding model is not likely to have a meaningful physical interpretation or to aid hypothesis generation by identifying a small number of covariates or processes potentially influencing the key response variable(s).

Sparse regression, a relatively recent development in statistics and machine learning, may be a promising approach toward building enhanced prediction models. Preliminary studies in the context of multivariate autoregressive models and Granger causality have shown promise for modeling both normal as well as extreme behavior (Lozano et al., 2009a, b; Liu et al., 2009). Sparse regression models attempt to identify a small set of predictors out of a large space of candidates by balancing model fit and parsimony. Regularization (penalty) terms are combined with error minimization criteria to encourage predictive models that are accurate while choosing a relatively small subset of covariates with nonzero coefficients. Penalty terms could enforce temporal, spatial, and/or covariate sparsity (Lozano et al., 2009a, b), which may be needed to obtain physically sensible solutions. The models can be shown to be statistically consistent even in the $n \ll p$ regime, along with rigorous finite sample rates of convergence (Meinshausen and Bühlmann, 2006; Zhao and Yu 2006; Candès and Tao, 2007; Meinshausen and Yu, 2009; Negahban et al. 2009; Ravikumar et al. 2010; Negahban and Wainwright, 2011; Obozinski et al. 2011). While the number of samples n for extremes will be rather small, sparse regression methods will still be able to do model selection in a statistically consistent way. There are, however, a few key assumptions made while developing statistical consistency guarantees, which need not be true for climate data, especially extremes. A typical linear model of the form $y = X\theta + w$ assumes that (1) y has a linear dependence on the covariates, (2) the noise w is Gaussian or sub-Gaussian, and (3) samples are independently drawn from a fixed but unknown distribution. Such assumptions can often be violated in the context of extremes. In particular, (1) the extremes may have a nonlinear dependence on the covariates (Koltchinskii and Yuan, 2008; Raskutti et al., 2010), (2) the noise component may be non-Gaussian and may be even heavy tailed (Falk et al., 2010; Embrechts et al., 2011), and (3) the samples (y_i, X_i) may be dependent or even from a nonstationary distribution, violating the independence assumptions (Meir, 2000; Mohri and Rostamizadeh, 2010). The ability to accommodate such nonlinear, non-Gaussian, nonstationary behavior using advanced predictive models will determine the success of enhanced prediction for extremes.

An alternative promising approach for regression of unreliable variables, especially extremes, can be based on correlated but more reliable variables, or those that are currently better predicted. For example, although rainfall extremes may be difficult to characterize directly in terms of a set of covariates, given that rainfall extremes may be correlated with temperature in some way (O'Gorman and Schneider, 2009; Sugiyama et al., 2010), and temperature can be effectively characterized using suitable (sparse) regression models, one then obtains an indirect but reliable way to model rainfall extremes. The above idea is a prominent theme in multivariate or multiple output regression, especially in the context of spatial and geostatistics, where such models are studied as Linear Models of Coregionalization (LMC) (Wackernagel, 2003; Gelfand and Banerjee, 2010) and related spatiotemporal models (Higdon, 2002; Mardia and Goodall, 1993). In recent years, correlated multiple-output, nonlinear, nonparametric regression models have been studied in statistical machine learning using Gaussian Processes (Agovic et al., 2011; Álvarez and Lawrence, 2011). In particular, our recent work on Probabilistic Matrix Addition (Agovic et al., 2011) has two additional capabilities that may be useful in this context: (1) the ability to utilize a nonlinear covariance function (kernel) K_x among the covariates, as well as a nonlinear covariance function K_y among the multivariate output (e.g., precipitation and temperature); and (2) the ability to handle data matrices with missing entries. Recent literature has shown evidence that reliable auxiliary variables may contain valuable information on crucial variables that are currently difficult to physically model (e.g., O'Gorman and Schneider, 2009; Sugiyama et al., 2010), and the ideas outlined above may lead to a systematic way to leverage such signals. Such an accomplishment might not only enhance predictability of such key variables, but also augment physical understanding, which could inform future efforts in the climate modeling community.

5.5 CONCLUSIONS AND FUTURE RESEARCH

Computational data sciences may offer a path forward to one of the key science challenges relevant for stakeholders, resources managers, and policy makers—specifically, the consequences on statistical attributes of extreme events as a consequence of climate change. This chapter presented several critical challenges in the science of climate extremes that are not handled by the current generation of climate models, represent long-standing challenges in scientific understanding, may not be solved in the near future by improvements in physics-based modeling, and where data-driven computational methods may offer novel solutions. The chapter presented specific details on three interrelated problem areas—*extremes characterization*, *uncertainty quantification*, and *enhanced prediction*—as well as several potential conceptual and methodological directions for the interdisciplinary computational and data sciences communities to advance the science in these areas.

The close integration between physical understanding, or physics-based modeling, and data-driven insights is emphasized from three interrelated perspectives.

1. Data-driven insights from observations inform model diagnostics and uncertainty quantification.
2. Enhanced projections rely on data-guided functional mappings (e.g., between precipitation extremes and temperature), which in turn may be derived from both observations and model-simulations but remain conditioned on physics-based, model-simulated results in the future (e.g., future data-driven insights on precipitation extremes may be conditioned on projected temperatures in the future from models).
3. New insights from massive data on multivariate associations derived from observed or model-simulated data not only improve our understanding of relevant processes but also may inform the physical formulation and parameter choices within global or regional climate models.

This chapter made no attempt to comprehensively address two important research directions. The first area, on attribution of climate extremes, has been primarily motivated by prior and ongoing work led by the climate science community but has witnessed recent progress by the computer science community. The second area, on graphical models and complex networks in climate, has been motivated by nonlinear dynamics communities within the geosciences as well as more recently by the interdisciplinary data science communities. Attribution of climate extremes to global warming and anthropogenic emissions, regional changes in urbanization and land use, as well as other plausible causal factors, is a key concern for policy makers. The corresponding methods include data-driven approaches such as Granger causality and optimal fingerprinting, that, for example, may be used for attribution of global and regional changes in rainfall extremes (Min et al., 2011). Numerical modeling techniques have attempted to generate a range of simulations at specific locations and time periods to delineate what may be caused by natural variability versus what may have to be attributed to climate change, for example, with application to location-based precipitation extremes in a given season (Pall et al., 2011). While data science communities have developed (Lozano et al., 2009b) innovative approaches, new developments are motivated. Nonlinear dynamics methods have been used in climate for a while (e.g., to relate climate oscillators to the variability of river flows: Khan et al., 2006, 2007). Correlation-based complex networks have been used in climate to capture multivariate dependence (Steinhaeuser et al., 2011b), diagnose model performance through their ability to capture ocean-based oscillators (Kawale et al., 2011), relate ocean-based oscillators to regional climate (Steinhaeuser et al., 2011a; Chatterjee et al., 2012), as well as for abrupt change (Tsonis et al., 2007) and extremes processes (Malik et al., 2011). Further developments may make these emerging approaches important tools for climate extremes.

ACKNOWLEDGMENTS

This research was primarily supported by the National Science Foundation through their Expeditions in Computing program; grant number 1029166. ARG and EK were supported in part by a Northeastern University grant, as well as by the Nuclear Regulatory Commission. HNN acknowledges support by the U.S. Department of Energy (DOE), Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

REFERENCES

Agovic, A., Banerjee, A., and Chatterjee, S. 2011. Probabilistic Matrix Addition. *International Conference on Machine Learning* (*ICML*).

Álvarez, M.A., and Lawrence, N.D. 2011. Computationally efficient convolved multiple output Gaussian processes. *Journal of Machine Learning Research,* 12: 1425–1466.

Annan, J.D., and Hargreaves, J.C. 2006. Using multiple observationally-based constraints to estimate climate sensitivity. *Geophysical Research Letters*, 33, L06704, doi: http://dx.doi.org/10.1029/2005GL025259.

Babacan, S., Molina, R., and Katsaggelos, A., 2010. Bayesian comprehensive sensing using LaPlace priors. *IEEE Transactions on Image Processing*, 19(1): 53–63.

Banerjee, S., Gelfand, A.E., Finley, A.O., and Sang, H. 2008. Gaussian predictive process models for large spatial data sets. *Journal of the Royal Statistical Society B,* 70(4): 825–848.

Buser, C.M., Künsch, H.R., Lüthi, D., Wild, M., and Schär, C. 2009. Bayesian multi-model projection of climate: Bias assumptions and interannual variability. *Climate Dynamics,* 33: 849–868, doi:10.1007/s00382-009-0588-6.

Caflisch, R. 1998. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica,* 7: 1–49.

Candès, E., and Tao, T. 2007. The Dantzig selector: Statistical estimation when *p* is much larger than *n*. *Annals of Statistics,* 35(6): 2313–2351.

Candès, E.J., and Wakin, M.B., 2008. An introduction to compressive sampling. *IEEE Signal Processing Magazine,* 25(2), 21–30.

Chatterjee, S., Banerjee, A., Chatterjee, S., Steinhaeuser, K., and Ganguly, A.R. 2012. Sparse group lasso: Consistency and climate applications. *Proceedings of the SIAM International Conference on Data Mining* (in press).

Christensen, C., Aina, T., and Stainforth, D. 2005. The challenge of volunteer computing with lengthy climate model simulations. *First International Conference on e-Science and Grid Computing.* doi: 10.1109/E-SCIENCE.2005.76.

Coles, S.G. 2001. *An Introduction to Statistical Modeling of Extreme Values*. New York: Springer, 208 pp.

Coles, S.G., and Dixon, M.J. 1999. Likelihood-based inference for extreme value models. *Extremes,* 2(1): 5–23.

Cressie, N., and Johannesson, G. 2008. Fixed rank kriging for very large spatial data sets. *Journal of the Royal Statistical Society B*, 70(1): 209–226.

Debusschere, B., Najm, H., Pebay, P., Knio, O., Ghanem, R., and Le Maitre, O. 2004. Numerical challenges in the use of polynomial chaos representations for stochastic processes. *SIAM Journal of Scientific Computing*, 26: 698–719.

Donoho, D.L. 2006. Compressed Sensing. *IEEE Transactions on Information Theory,* 52(4): 1289–1306.

Drignei, D., Forest, C.E., and Nychka, D. 2009. Parameter estimation for computationally intensive nonlinear regression with an application to climate modeling. *Annals of Applied Statistics,* 2(4): 1217–1230.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. *Annals of Statistics,* 7: 101–118.

Emanuel, K., Sundararajan, R., and Williams, J. 2008. Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bulletin of the American Meteorological Society,* 89(3): 347–368.

Embrechts, P., Klüppelberg, C., and Mikosch, T. 2011. *Modelling Extremal Events for Insurance and Finance.* New York: Springer.

Engelhaupt, E. 2008. Climate change: A matter of national security. *Environmental Science & Technology,* 42(20): 7458–7549.

Falk, M., Hüsler, J., and Reiss, R. 2010. *Laws of Small Numbers: Extremes and Rare Events.* New York and Berlin: Springer.

Fern, A., and Givan, R. 2003. Online ensemble learning: An empirical study. *Machine Learning,* 53: 71–109.

Ferro, C.A.T., and Segers, J. 2003. Inference for clusters of extreme values. *Journal of the Royal Statistical Society, Series B*, 65(2): 545–556.

Furrer, R., and Sain, S.R. 2009. Spatial model fitting for large datasets with applications to climate and microarray problems. *Statistics and Computing,* 19: 113–128, doi: http://dx.doi.org/10.1007/s11222-008-9075-x.

Furrer, R., Sain, S.R., Nychka, D., and Tebaldi, C. 2007. Multivariate Bayesian analysis of atmosphere-ocean general circulation models. *Environmental and Ecological Statistics,* 14: 249–266.

Ganguly, A.R., Steinhaeuser, K., Erickson, D.J., Branstetter, M., Parish, E.S., Singh, N., Drake, J.B., and Buja, L. 2009b. Higher trends but larger uncertainty and geographic variability in 21st century temperature and heat waves. *Proceedings of the National Academy of Sciences of the United States of America,* 106(37): 15555–15559.

Ganguly, A.R., Steinhaeuser, K., Sorokine, A., Parish, E.S., Kao, S.-C., and Branstetter, M. 2009a. Geographic analysis and visualization of climate extremes for the Quadrennial Defense Review. *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems,* p. 542–543.

Gelfand, A.E., and Banerjee, S. 2010. Multivariate spatial process models. In A.E. Gelfand, P. Diggle, P. Guttorp, and M. Fuentes (Eds.), *Handbook of Spatial Statistics*. Boca Raton, FL: CRC Press.

Ghanem, R., and Spanos, P. 1991. *Stochastic Finite Elements: A Spectral Approach.* New York and Berlin: Springer.

Ghosh, S., Das, D., Kao, S.-C., and Ganguly, A.R. 2011. Lack of uniform trends increasing spatial variability in observed Indian rainfall extremes. *Nature Climate Change*. (In press—available online). 2, 86–91, DOI: 1038/nclimate1377.

Giorgi, F., and Mearns, L.O. 2002. Calculation of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "Reliability Ensemble Averaging" (REA) method. *Journal of Climate,* 15: 1141–1158.

Girko, V.L. 1995. *Statistical Analysis of Observations of Increasing Dimension*. New York: Kluwer Academic.

Goswami, B.N., Venugopal, V., Sengupta, D., Madhusoodanan, M.S., and Xavier, P.K. 2006. Increasing trend of extreme rain events over India in a warming environment. *Science,* 314(5804): 1442–1445.

Govaerts, Y.M., and Lattanzio, A. 2007. Retrieval error estimation of surface albedo derived from geostationary large band satellite observations: Application to Meteosat-2 and Meteosat-7 data. *Journal of Geophysical Research,* 112 D05102, doi: http://dx.doi.org/10.1029/2006JD007313.

Greene, A.M., Goddard, L., and Lall, U. 2006. Probabilistic multimodel regional temperature projections. *Journal of Climate*, 19: 4326–4343.

Hagedorn, R.F., Doblas-Reyes, F.J., and Palmer T.N. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting. I. Basic concept. *Tellus A,* 57: 219–233.

Hall, P., and Tajvidi, N. 2000. Nonparametric analysis of temporal trend when fitting parametric models to extreme value data. *Statistical Science,* 15(2): 153–167.

Higdon, D.M. 2002. Space and space-time modelling using process convolutions. In *Quantitative Methods for Current Environmental Issues*, p. 37–56. Berlin: Springer-Verlag.

Hosking, J.R.M., and Wallis, J.R. 1997. *Regional Frequency Analysis: An Approach Based on L-Moments*. Cambridge: Cambridge University Press.

IPCC. 2007. Fourth Assessment Report (AR4), Working Group I, Chapter 8.

IPCC. 2011. Summary for Policymakers. In Intergovernmental Panel on Climate Change Special Report on Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation, Field, C.B., Barros, V., Stocker, T.F., Qin, D., Dokken, D., Ebi, K.L., Mastrandrea, M.D., Mach, K.J., Plattner, G.-K., Allen, S., Tignor, M., and Midgley, P.M. (Eds.). Cambridge (UK) and New York: Cambridge University Press.

Jackson, C., Sen, M., and Stoffa, P. 2004. An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions. *Journal of Climate,* 17(14): 2828–2841.

Jackson, C., Xia, Y., Sen, M., and Stoffa, P. 2003. Optimal parameter and uncertainty estimation of a land surface model: A case example using data from Cabauw. *Netherlands Journal of Geophysical Research,* 108 (D18): 4583 http://dx.doi.org/10.1029/2002JD002991.

Jackson, C.S., Sen, M.K., Huerta, G., Deng, Y., and Bowman, K.P. 2008. Error reduction and convergence in climate prediction. *Journal of Climate,* 21: 6698–6709, doi: 10.1175/2008JCLI2112.1.

Kao, S.-C., and Ganguly, A.R. 2011. Intensity, duration, and frequency of precipitation extremes under 21st-century warming scenarios. *Journal of Geophysical Research,* 116(D16119): 14 pages. DOI: 10.1029/2010JD015529.

Karl, T.R., Melillo, J.M., and Peterson, T.C. Eds. 2009. *Global Climate Change Impacts in the United States.* Cambridge (UK) and New York: Cambridge University Press, 196 pp.

Kawale, J., Liess, S., Kumar, A., Steinbach, M., Ganguly, A., Samatova, N.F., Semazzi, F., Snyder, P., and Kumar, V. 2011. Data guided discovery of dynamic climate dipoles. *Proceedings of the 2011 NASA Conference on Intelligent Data Understanding (CIDU).*

Khan, S., Bandyopadhyay, S., Ganguly, A.R., Saigal, S., Erickson, D.J., Protopopescu, V., and Ostrouchov, G. 2007. Relative performance of mutual information estimation methods for quantifying the dependence among short and noisy data. *Physical Review,* E 026209.

Khan, S., Ganguly, A.R., Bandyopadhyay, S., Saigal, S., Erickson, D.J., Protopopescu, V., and Ostrouchov, G. 2006. Nonlinear statistics reveals stronger ties between ENSO and the tropical hydrological cycle. *Geophysical Research Letters,* 33: L24402.

Kharin, V.V., and Zwiers, F.W. 2000. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *Journal of Climate,* 13: 3760–3788.

Kharin, V.V., and Zwiers, F.W. 2005. Estimating extremes in transient climate change simulations. *Journal of Climate,* 18: 1156–1173.

Kharin, V.V., Zwiers, F.W., Zhang, X., and Hegerl, G.C. 2007. Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate,* 20: 1419–1444.

Knutti, R. 2010. The end of model democracy? *Climatic Change,* 102(3-4): 395–404.

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G.A. 2010. Challenges in combining projections from multiple climate models. *Journal of Climate,* 23(10): 2739–2758.

Kodra, E., Chatterjee, S., and Ganguly, A.R. 2011b. Challenges and opportunities toward improved data-guided handling of global climate model ensembles for regional climate change assessments. *ICML Workshop on Machine Learning for Global Challenges.*

Kodra, E., Ghosh, S., and Ganguly, A.R. 2012. Evaluation of global climate models for Indian monsoon climatology. *Environmental Research Letters,* 7, 014012, 7 pp., doi: 10.1088/1748-9326/7/1/014012.

Kodra, E., Steinhaeuser, K., and Ganguly, A.R. 2011a. Persisting cold extremes under 21st-century warming scenarios. *Geophysical Research Letters,* 38(L08705): 5 pages. DOI: 10.1029/2011GL047103.

Koltchinskii, V., and Yuan, M. 2008. Sparse recovery in large ensembles of kernel machines. *Conference on Learning Theory (COLT).*

Krishnamurti, T.N. et al. 1999. Improved weather and seasonal climate forecasts from multimodel superensemble. *Science,* 285(5433): 1548–1550.

Le Maitre, O., Najm, H., Ghanem, R., and Knio, O. 2004. Multi-resolution analysis of Wiener-type uncertainty propagation schemes. *Journal of Computational Physics,* 197: 502–531.

Liu, S.C., Fu, C., Shiu, C.-J., Chen, J.-P., and Wu, Fu. 2009. Temperature dependence of global precipitation extremes. *Geophysical Research Letters,* 36(L17702): 4 pages. DOI: 10.1029/2009GL040218.

Lozano, A.C., Abe, N., Liu, Y., and Rosset, S. 2009b. Grouped graphical Granger modeling methods for temporal causal modeling. *ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2009*, p. 577–586.

Lozano, A.C., Li, H., Niculescu-Mizil, A., Liu, Y., Perlich, C., Hosking, J., and Abe, N. 2009a. Spatial-temporal causal modeling for climate change attribution. *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,* p. 587–596.

Malik, N., Bookhagen, B., Marwan, N., and Kurths, J. 2011. Analysis of spatial and temporal extreme monsoonal rainfall over South Asia using complex networks. *Climate Dynamics* 39(3–4), 971–987, DOI: 10,1007/500382-011-1156-4.

Mannshardt-Shamseldin, E.C., Smith, R.L., Sain, S.R., Mearns, L.O., and Cooley, D. 2011. Downscaling extremes: A comparison of extreme value distributions in point-source and gridded precipitation data. *Annals of Applied Statistics*, 4(1), 486–502.

Mardia, K.V., and Goodall, C.R. 1993. Spatial-temporal analysis of multivariate environmental monitoring data. In *Multivariate Environmental Statistics*, pp. 347–386. Elsevier Science Publishers B.V., North Holland.

Marzouk, Y.M., and Najm, H.N. 2009. Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems. *Journal of Computational Physics,* 228(6): 1862–1902.

Meehl, G.A., and Tebaldi, C. 2004. More intense, more frequent, and longer lasting heat waves in the 21st-century. *Science,* 305(5686): 994–997.

Meinshausen, N., and Bühlmann, P. 2006. High-dimensional graphs and variable selection with the Lasso. *Annals of Statistics,* 34: 1436–1462.

Meinshausen, N., and Yu, B. 2009. Lasso-type recovery of sparse representations for high-dimensional data. *Annals of Statistics,* 37(1): 246–270.

Meir, R. 2000. Nonparametric time series prediction through adaptive model selection. *Machine Learning,* 39: 5–34.

Min, S.-K., and Hense, A. 2007. Hierarchical evaluation of IPCC AR4 coupled climate models with systematic consideration of model uncertainties. *Climate Dynamics,* 29: 853–868.

Min, S.-K., Zhang, X., Zwiers, F.W., and Hegerl, G.C. 2011. Human contribution to more-intense precipitation extremes. *Nature,* 470: 378–381.

Mohri, M., and Rostamizadeh, A. 2010. Stability bounds for stationary φ-mixing and β-mixing processes. *Journal of Machine Learning Research,* 11: 789–814.

Monteleoni, C., Schmidt, G., and Saroha, S. 2010. Tracking climate models. *NASA Conference on Intelligent Data Understanding.*

Morris, M. 1991. Factorial sampling plans for preliminary computer experiments. *Technometrics,* 33(2): 161–174.

Muller, C.J., and O'Gorman, P.A. 2011. An energetic perspective on the regional response of precipitation to climate change. *Nature Climate Change,* 1: 266–271.

Najm, H.N. 2009. Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics. *Annual Reviews of Fluid Mechanics,* 41: 35–52.

Negahban, S., Ravikumar, P.D., Wainwright, M.J., and Yu, B. 2009. A unified framework for high-dimensional analysis of M-estimators with decomposable regularizers. *NIPS,* 2009: 1348–1356.

Negahban, S., and Wainwright, M.J. 2011. Simultaneous support recovery in high-dimensional regression: Benefits and perils of L_1-L_∞-regularization. *IEEE Transactions on Information Theory,* 57(6): 3841-3863, June 2011.

Nobile, F., Tempone, R., and Webster, C. 2008. A sparse grid stochastic collocation method for partial differential equations with random input data. *SIAM Journal of Numerical Analysis,* 46(5): 2411–2442.

NRC. 2011. *National Security Implications of Climate Change for U.S. Naval Forces.* Committee on National Security Implications of Climate Change for U.S. Naval Forces. National Research Council of the National Academies. National Academies Press. 172 pp.

Obozinski, G., Wainwright, M.J., and Jordan, M.I. 2011. Union support recovery in high-dimensional multivariate regression. *Annals of Statistics,* 39(1): 1–47.

O'Gorman, P.A., and Schneider, T. 2009. The physical basis for increases in precipitation extremes in simulations of 21st-century climate change. *Proceedings of the National Academy of Sciences of the United States of America,* 106(35): 14773–14777.

Overpeck, J.T., Meehl, G.A., Bony, S., and Easterling, D.R. 2011. Climate data challenges in the 21st century. *Science,* 331: 700-702. doi: 10.1126/science.1197869.

Pall, P., Aina, T., Stone, D.A., Stott, P.A., Nozawa, T., Hilberts, A.G.J., Lohmann, D., and Allen, M.R. 2011. Anthropogenic greenhouse gas contribution to flood risk in England and Wales in autumn 2000. *Nature,* 470: 382–385.

Perkins, S.E., Pitman, A.J., and Sisson, S.A. 2009. Smaller projected increases in 20-year temperature returns over Australia in skill-selected climate models. *Geophysical Research Letters,* 36: L06710, doi:10.1029/2009GL037293.

Pierce, D.W., Barnett, T.P., Santer, B.D., and Gleckler, P.J. 2009. Selecting global climate models for regional climate change studies. *Proceedings of the National Academy of Sciences USA,* 106(21): 8441–8446.

Raskutti, G., Wainwright, M.J., and Yu, B. 2010. Minimax-Optimal Rates for Sparse Additive Models over Kernel Classes via Convex Programming. Technical report, http://arxiv.org/abs/1008.3654, UC Berkeley, Department of Statistics, August.

Ravikumar, P., Wainwright, M.J., and Lafferty, J. 2010. High-dimensional Ising model selection using L1 regularized logistic regression. *Annals of Statistics,* 38(3): 1287–1319.

Sain, S., Furrer, R., and Cressie, N. 2008. Combining Ensembles of Regional Climate Model Output via a Multivariate Markov Random Field Model. Technical report, Department of Statistics, The Ohio State University.

Saltelli, A., Chan, K., and Scott, E.M. 2000. *Sensitivity Analysis.* Wiley Series in Probability and Statistics New York.

Sanso, B., Forest, C.E., and Zantedeschi, D. 2008. Inferring climate system properties using a computer model. *Bayesian Analysis*, 3(1): 1–38.

Santer, B.D. et al. 2009. Incorporating model quality information in climate change detection and attribution studies. *Proc. Natl. Acad. Sci. USA,* 106(35): 14778–14783.

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

Seni, G., and Elder, J.F. 2010. Ensemble methods in data mining: Improving accuracy through combining predictions. *Synthesis Lectures on Data Mining and Knowledge Discovery.*

Smith, R.L., Tebaldi, C., Nychka, D., and Mearns, L.O. 2009. Bayesian modeling of uncertainty in ensembles of climate models. *Journal of the American Statistical Association,* 104(485): 97–116. doi:10.1198/jasa.2009.0007.

Soize, C., and Ghanem, R. 2004. Physical systems with random uncertainties: Chaos representations with arbitrary probability measure. *SIAM Journal of Scientific Computing*, 26: 395–410.

Stainforth, D.A. et al. 2005. Uncertainty in predictions of the climate response to rising levels of greenhouse gases. *Nature,* 433: 403–406.

Steinhaeuser, K., Chawla, N.V., and Ganguly, A.R. 2011a. Complex networks as a unified framework for descriptive analysis and predictive modeling in climate science. *Statistical Analysis and Data Mining,* 4(5): 497–511.

Steinhaeuser, K., Ganguly, A.R., and Chawla, N. 2011b. Multivariate and multiscale dependence in the global climate system revealed through complex networks. *Climate Dynamics,* DOI: 10.1007/s00382-011-1135-9. 39(3–4), 889–895

Sugiyama, M., Shiogama, H., and Emori, S. 2010. Precipitation extreme changes exceeding moisture content increases in MIROC and IPCC climate models. *Proceedings of the National Academy of Sciences of the United States of America,* 107(2): 571–575.

Tebaldi, C., and Knutti, R. 2007. The use of the multimodel ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A*, 365, 2053–2075.

Tebaldi, C., Mearns, L.O., Nychka, D., and Smith, R.L. 2004. Regional probabilities of precipitation change: A Bayesian analysis of multimodel simulations. *Geophysical Research Letters,* 31, L24213, doi:10.1029/2004GL021276.

Tebaldi, C., and Sanso, B. 2009. Joint projections of temperature and precipitation change from multiple climate models: A hierarchical Bayesian approach. *Journal of the Royal Statistical Society A,* 172(1): 83–106.

Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58(1): 267–288.

Tollefson, J. 2008a. Climate war games. *Nature,* 454(7205): 673. August 7.

Tollefson, J. 2008b. Climate war games: The "Angry Red Chart." Nature Blogs. 30 July: http://blogs.nature.com/news/blog/events/climate_war_game/.

Tollefson, J. 2011. Climate change: Cold spells in a warm world. *Nature,* 139. Research Highlights. 14 April. DOI:10.1038/472139d.

Tsonis, A.A., Swanson, K., and Kravtsov, S. 2007. A new dynamical mechanism for major climate shifts. *Geophysical Research Letters,* 34: L13705, 5 pp., doi:10.1029/2007GL030288.

Villagran, A., Huerta, G., Jackson, C.S., and Sen. M.K. 2008. Computational methods for parameter estimation in climate models. *Bayesian Analysis,* 3: 823–850, doi: http://dx.doi.org/10.1214/08-BA331.

Wackernagel, H. 2003. *Geostatistics: An Introduction with Applications*. Berlin: Springer-Verlag.

Wehner, M. 2010. Source of uncertainty in the extreme value statistics of climate data. *Extremes,* 13(2): 205–217.

Weigel, A.P., Knutti, R., Linger, M.A., and Appenzeller, C. 2010. Risks of model weighting in multimodel climate projections. *Journal of Climate,* 23: 4175–4191.

Xiu, D., and Karniadakis, G. 2002. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing,* 24: 619–644.

Zhao, P., and Yu, B. 2006. On model selection consistency of lasso. *Journal of Machine Learning Research,* 7: 2541–2567.