11

Thoughts on the future

11.1 Introduction

Throughout the previous chapters we have examined a number of the most commonly used types of parameterization schemes within numerical weather prediction models. While individual parameterization schemes are constantly undergoing revision and new schemes appear in both the literature and operational models routinely, the underlying need for the parameterization of fundamental atmospheric processes has not changed. Indeed, the number of processes that are being parameterized has increased over the past 20 years to allow for more realism in both forecasts and climate simulations. These additional parameterizations may not be crucial to the model forecasts at all times and places, but they can make a significant difference regarding a particular event of importance to a specific user or community.

This evolution of parameterization highlights the fact that numerical models are becoming more capable (see Roebber *et al.* 2004). Models now can reproduce many of the phenomena that are observed in the atmosphere. As simple examples, moderate- and high-resolution models of today can reproduce mesoscale convective systems with their leading line of deep convection and trailing stratiform precipitation region as well as sea breezes and mountain-valley flows, while coupled ocean–atmosphere models can reproduce El Niño Southern Oscillation (ENSO) events. These phenomena could not be reproduced by any of the operational models (or even many research models) in use back in the 1970s, owing in part to their large grid spacing and in part to the parameterization schemes in use. This increased model capability often leads to higher expectations as well as higher perceived confidence in the forecasts. It can be very difficult for a human forecaster to challenge the prediction from a numerical model forecast at 1 km grid spacing that indicates the development of a severe thunderstorm, or from a seasonal climate model that predicts an ENSO event. These model forecasts often provide details that cannot be observed, while the structures and behaviors produced by the model appear very realistic.

Yet it is hoped that some degree of uncertainty or doubt has been created as the various parameterization schemes have been examined and their sensitivities explored. Parameterization schemes generally develop from a reasonable theoretical foundation, but require a number of simplifying assumptions and are tested using a limited data set over a specific range of environmental conditions. However, when incorporated into numerical models, these schemes are then used to make predictions over the entire globe and in environments for which they likely have never been tested fully. In addition, the behavior of an individual parameterization often depends upon the behavior of other parameterization schemes and interactions between schemes often occur. For example, planetary boundary layer schemes depend upon the behavior of the soil–vegetation–atmosphere transfer scheme that predicts ground temperature and soil moisture, yet the boundary layer scheme also influences the net radiation reaching the surface through the vertical mixing of moisture. Empirical tuning also comes into play as model developers and some users attempt to optimize model skill (however defined) for particular problems or scenarios. This empirical tuning often does not occur through a systematic approach and may instead be guided by case studies of important events or model intercomparison tests with a small sample size. The truly remarkable aspect of all this is that the resulting numerical predictions have value! And this value appears to be increasing over time, in part because the model parameterization schemes are becoming more and more accurate and realistic. This trend in the ever increasing realism in parameterization schemes is expected to continue. But this situation also should cause us to pause and think about how best to use these valuable, yet flawed, tools that we create.

It is important to recognize that it is impossible to test a parameterization scheme for all atmospheric conditions that may occur. The observational data do not exist for such a test, and the time needed to conduct such tests would be prohibitive. This is not to suggest that parameterization schemes should not be tested on large data sets. Indeed, many parameterization schemes are already tested on fairly large data sets and this testing is important to the improvement and validation of schemes. However, we need to recognize and appreciate that a truly comprehensive test of a parameterization scheme is impossible. And even if a comprehensive test indicated a weakness of a scheme under specific conditions, this does not imply that the scheme should be discarded. It is just that the scheme has limitations, which is true for all parameterization schemes. Numerical models are imperfect, so the key to success is how we deal with these imperfections.

We move at this point from the realm of what is known into a realm that mixes knowledge with conjecture and opinion. This is perhaps a bit unusual for a meteorological textbook, but it seems appropriate to discuss what the future may hold to stimulate discussions and debate in this important enterprise called numerical weather prediction. Topics briefly touched upon are ensemble predictions, ensembles and high-resolution single forecasts, statistical postprocessing, and the road forward.

11.2 Ensemble predictions

Ever since the pioneering study of Lorenz (1963), it has been recognized that small errors can grow rapidly in non-linear models. However, even if the model is perfect, there is a finite time limit to the predictability of the atmosphere, since it is impossible to observe the atmosphere perfectly due to both sampling and instrument errors. This predictability time limit becomes shorter as the scales of interest become smaller (Lorenz 1969). Model simulations starting from ever so slightly different initial conditions diverge and eventually have little relationship to one another (Fig. 11.1). Since the true atmospheric state at any point in time can only be known approximately, the atmosphere prediction problem needs to be formulated in terms of the time evolution of a probability distribution function (PDF) for the atmosphere. This realization that the atmosphere is chaotic and has this sensitive dependence upon initial conditions led to the development of ensemble forecasting systems that explicitly attempt to predict the evolution of the atmospheric PDF (see the historical review by Lewis 2005). Since their first operational use in the early 1990s, ensembles have become a critical component of both operational numerical weather prediction and climate studies, and remain an important research topic. Ensembles are now used for climate, seasonal, medium-range, and short-range forecasting. As the models used within ensemble forecasting systems improve, the ensembles improve as well. Thus, model and ensemble forecast improvements go hand-in-hand.

Ensembles are simply groups of forecasts that are valid over the identical time period. Typically, each forecast member of the ensemble differs in its initial conditions in order to provide an initial sample of the atmospheric PDF. Differences in model characteristics may also be included as part of the ensemble to account for model error. A numerical weather prediction model is used to provide a forecast from each of these different initial conditions, and the properties of the atmospheric PDF are assumed to be determined by the statistics calculated from the ensemble members at any selected forecast time (Epstein 1969; Leith 1974; Molteni *et al.* 1996). The ensemble statistics

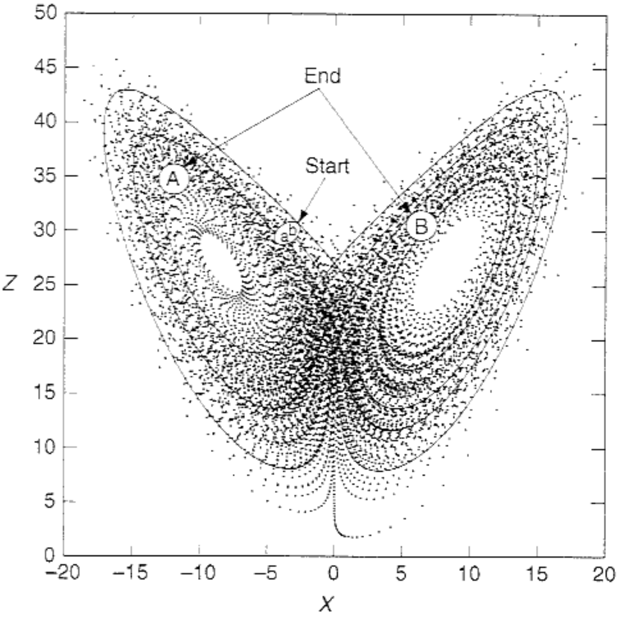

Figure 11.1. Depiction of the Lorenz (1963) attractor (dots), with trajectories *a* and *b* indicated that start from an initial small difference and evolve into a large difference at their end points *A* and *B*.

approximate the true atmospheric PDF closely if the initial perturbations accurately represent the PDF of analysis errors and if the numerical model provides a very good approximation of the atmosphere. Thus, instead of a single (deterministic) forecast that only provides one picture of the evolution of the atmosphere, an ensemble of forecasts is created and the forecast is now inherently probabilistic in nature. Murphy and Winkler (1979) strongly argue that forecasts cannot be used to their best advantage unless the forecast uncertainty is quantified and expressed in a useful manner to the end users. Ensemble forecasting systems are one way to express forecast uncertainty (Fig. 11.2).

Ensembles initially consisted of the same model with the same parameterization schemes, but using different initial conditions for each ensemble member. A number of different techniques have been developed to perturb the model initial conditions around a control, or best estimate, analysis of the atmospheric state in order to sample the analysis error. One Monte Carlo approach mimics the differences between the global analyses of different operational centers, which is an estimate of analysis error, and so draws the different initial conditions from this specified distribution (Errico and Baumhefner 1987; Mullen and Baumhefner 1994). However, it takes a very

Figure 11.2. Spaghetti diagram of 5820 m, 500 hPa geopotential height isolines from the NCEP short-range ensemble forecasting system started at 0900 UTC 22 August 2005. Geopotential height isolines valid at the 39 h forecast time, or 0000 UTC 24 August 2005. Notice the variability over central North America, with some members indicating that a ridge or anticyclone will develop and others indicating zonal flow across the United States. Courtesy of Dr. Jun Du of the National Centers for Environmental Prediction.

large number of ensemble members to randomly sample the analysis error well and it is prohibitively expensive at this time to produce forecasts from hundreds of ensemble members. Thus, several initial condition perturbation techniques sample the analysis uncertainty space more strategically. Both singular vectors (Buizza and Palmer 1995; Molteni *et al.* 1996; Buizza 1997) and the breeding of growing modes (Toth and Kalnay 1993, 1997) select perturbations to the control analysis that are the most unstable and grow the fastest. It is hoped that these fastest growing perturbations lead to a reasonable sampling of the true atmospheric PDF with fewer ensemble members, since they should dominate the ensemble variability. Results further indicate that these two perturbation techniques, initially developed for medium-range ensemble systems, focus upon perturbations to synoptic-scale weather systems over the midlatitudes that are associated with baroclinic instability (Toth and Kalnay 1997). This temporal and spatial scale is well suited for numerical weather prediction, since baroclinic instability is inherent within the equations of motion. Yet another perturbation approach uses perturbed observations in data assimilation systems to produce a set of representative analysis errors (Houtekamer *et al.* 1996; Houtekamer and Lefaivre 1997), instead of focusing upon growing modes that differ systematically from analysis errors (Houtekamer 1995). Regardless of the perturbation method, after 5 days of forecast time, it is difficult if not impossible to determine which perturbations originated from which technique.

As the spatial and temporal scales get smaller, more instabilities and physical processes are known to play a role in the evolution of important atmospheric features, at least on an intermittent basis. This also suggests that forecasts of these smaller-scale features will be useful over commensurately shorter times. This situation highlights the need for improved and increased observations and analysis techniques (see Daley 1991 for a summary of many data assimilation techniques) in order to provide an accurate depiction of the initial atmospheric state for models to use. New sensing capabilities, such as *in situ* observations from commercial aircraft and remote sensing observations from satellite and radars, have helped to provide greater information to use in specifying the atmospheric state. But the information available on mesoscale and cloud-scale atmospheric features is still woefully inadequate. Sophisticated data assimilation techniques, such as three- and four-dimensional variational assimilation (e.g., Derber 1989; Županski and Mesinger 1995; Gauthier *et al.* 1999; Lorenc *et al.* 2000; Rabier *et al.* 2000; Barker *et al.* 2004; Županski *et al.* 2005) and the ensemble Kalman filter (Evensen 1994, 1997; Mitchell and Houtekamer 2000; Houtekamer and Mitchell 2001; Snyder and Zhang 2003; Dowell *et al.* 2004) help to make the most use of the available observations. Yet it is clear that many uncertainties remain in the initial conditions provided to models.

The uncertainties present in specifying the atmospheric state at any given time also influence the perturbation strategies designed for creating the ensemble members. For example, perturbation techniques designed for the medium-range forecast problem may not be well suited to the short-range forecast problem. While baroclinic instability is still important at short ranges, many initial state uncertainties and short-range forecast concerns have little to do with synoptic-scale features. The use of optimization periods of 12 h to 2 days for error growth may define the perturbation types that are generated and these perturbations may not be especially meaningful for the short range. Alternative approaches to generating initial condition perturbations that use input from human forecasters appear worthy of further exploration for predicting unlikely events (e.g., Xu *et al.* 2001; Homar *et al.* 2006). Yet it may be that combining various approaches, such as Monte

Carlo and the breeding of growing models, yields the best results. More research is needed on how to generate perturbations for short-range ensembles. The same can probably be said of ensembles used for seasonal and climate simulations and forecasts.

We have seen in earlier chapters that parameterization schemes play a large role in forecasts of sensible weather – low-level temperatures, dewpoints, winds, and rainfall about which the public is most concerned. And, therefore, one would think that parameterization schemes may contribute to forecast sensitivities and to important forecast errors (Fig. 11.3). This realization that model imperfections may contribute substantially to forecast error has led to the inclusion of different models or different physical process parameterization schemes or stochastic errors within ensemble forecast systems for both the medium and short range (Houtekamer *et al.* 1996; Atger 1999; Buizza *et al.* 1999; Harrison *et al.* 1999; Stensrud *et al.* 1999, 2000; Fritsch *et al.* 2000; Evans *et al.* 2000; Ziehmann 2000; Wandishin *et al.* 2001; Hou *et al.* 2001; Stensrud and Weiss 2002). Results from these studies clearly indicate that ensembles containing different models or different parameterization schemes are more skillful than ensembles that do not contain some aspect of model uncertainty. It is anticipated that as we explore the optimum balance of ensemble member grid spacing and the number of ensemble members, the probabilistic forecasts from multimodel ensemble systems will only further improve.

The value of multimodel ensembles also is seen in seasonal to interannual climate prediction. Nine-member ensembles are created from each of seven different coupled ocean–atmosphere models that use nearly the same control initial conditions and the results from all nine ensembles pooled into one large multimodel ensemble as part of the Development of a European Multimodel Ensemble system for seasonal-to-inTERannual prediction (DEMETER; Palmer *et al.* 2004) project. Results indicate that coupled ocean–atmosphere multimodel ensembles are more reliable, have better resolution, and have consistently better performance than single-model ensembles (Hagedorn *et al.* 2005).

For climate simulations, Stainforth *et al.* (2005) use widely distributed desktop computers to conduct over 2500 simulations of future climate in a world with doubled CO_2 levels. These simulations differ in that parameters within the model physical process parameterization schemes are varied over a range of values deemed plausible by the scheme developers, and in that they have variations in the initial conditions. The ensemble mean forecasts are in agreement with other predictions, indicating a global warming of nearly 3.4 K for doubled CO_2 conditions. However, the range of warming is much greater than seen in previous studies, with some simulations indicating warming of

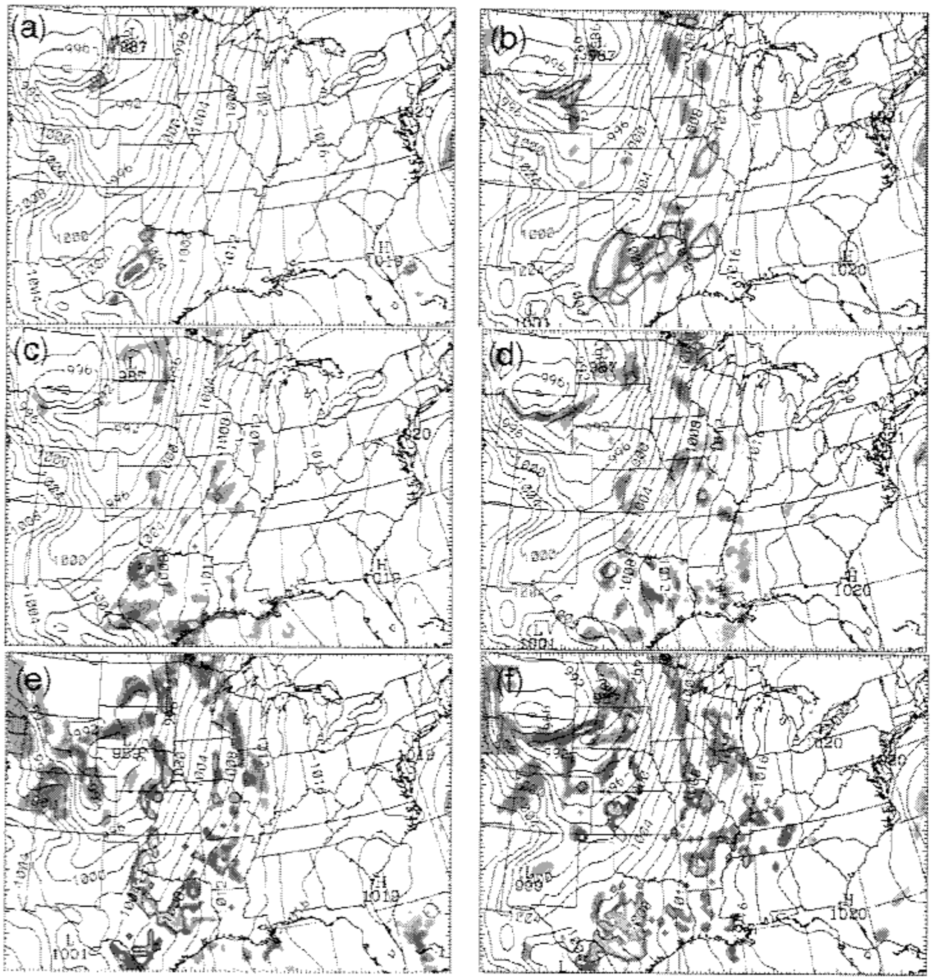

Figure 11.3. Sea-level pressure (contoured every 2 hPa) and 3 h accumulated convective rainfall (mm shaded) at the 24 h forecast time from six different mesoscale short-range ensemble members with parameterization scheme diversity valid at 0000 UTC 4 May 1999. While the sea-level pressure patterns are very similar in the six members, the rainfall patterns and amounts are very different. This highlights the importance of model parameterizations to the resulting forecasts and the need for parameterization diversity in ensemble forecasting systems. From Stensrud and Weiss (2002).

over 10 K is possible and others indicating very little warming at all. This result highlights the importance of the model physical process parameterizations to climate modeling, the need for improved parameterizations, and the need for a better understanding of how to create ensembles to respond to the ever increasing demand for more accurate and skillful weather information across a range of temporal scales.

In retrospect, this situation should not be surprising. We have seen that parameterization schemes can provide very different answers under the same environmental conditions. It is also possible that some parameterization schemes are entirely locked out of some environmental conditions and will never be able to reproduce the observed atmospheric behaviors in these environments. One example of this situation is the Betts–Miller convective parameterization being unable to develop convection in "loaded-gun" soundings (Fawbush and Miller 1954) that are typical of severe storm environments. This situation occurs because loaded-gun soundings have very dry mid-levels, such that the scheme fails to activate and hence is not able to produce rainfall. This is not to say that the Betts–Miller scheme is without value; it does an incredibly good job in many locations, but on the relatively rare occasion when these severe storm environments occur it is not able to activate. One can easily think of other examples in microphysical parameterizations, where some interactions between the various microphysics species are not included in the scheme, thereby locking out some behaviors. Another example is radiation parameterizations that do not allow for partly cloudy skies.

With the realization that parameterization schemes are imperfect and may not even function realistically in some environments, the idea of ensembles with model or model parameterization diversity in addition to initial condition uncertainty becomes very appealing and intuitive. The value of this type of approach is that the ensemble is more likely to be capable of reproducing the observed atmospheric phenomena over a broader range of environmental conditions, owing to the variety of parameterization schemes being used. It maximizes the chances that at least one of the parameterization schemes is capable of producing a realistic forecast for a given environment. It may be that not all members are equally plausible in a given environment, but over the entire range of environments visited by the atmosphere many parameterizations are equally skillful. The only argument for ensembles without model or model physics diversity is that all the parameterization schemes function well in all environmental conditions, a hypothesis that appears questionable at best.

11.3 Ensembles and high-resolution single forecasts

Computer resources are a finite quantity. Even with the incredibly rapid increases in computer processor speed, model developers and users are able to consume all the available processor cycles. This is particularly true in operational centers, where one also wants to make the best use of these precious resources. This situation has led occasionally to a perceived competition between high-resolution single deterministic forecasts and ensembles,

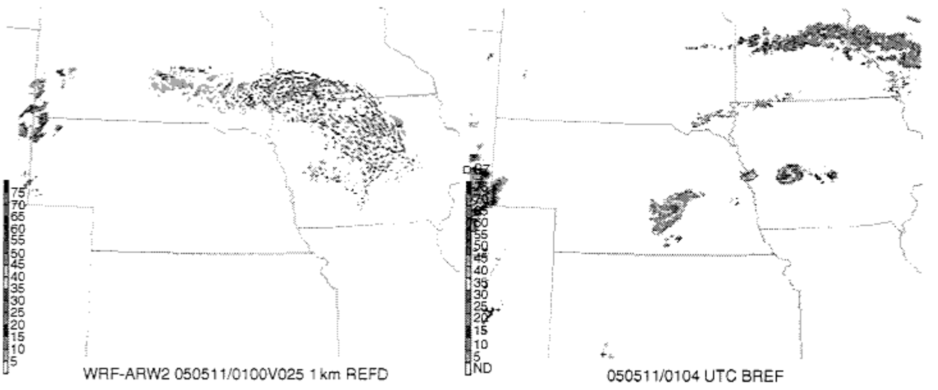

Figure 11.4. Cloud-scale model prediction (left) and observed (right) radar reflectivity fields valid 0100 UTC 11 May 2005. While the cloud-scale model clearly captures the typical types of reflectivity structures seen in the observations, the model has clearly missed the thunderstorms active in Nebraska and Iowa and misplaced the convection in Minnesota (in the upper right portion of the figure) southwestward. Image created as part of the Storm Prediction Center 2005 spring forecasting experiment.

since both ensembles and high-resolution (i.e., small grid spacing) single forecasts are very computationally demanding. The argument for high-resolution deterministic forecasts derives from the desire to have forecast models that are capable of predicting the observed atmospheric phenomena that are deemed most important. Without argument, high-resolution forecasts may be quite valuable for forecasting commonly observed, small-scale features with large societal or economic impact, such as severe thunderstorms and sea breezes. However, participation in several spring forecasting experiments that evaluated high-resolution (2 and 4 km grid spacing) operational forecasts of deep convection suggests that the lack of uncertainty information hampers the best of these forecasts (Fig. 11.4). It is not clear that some of the behaviors that are explicitly seen in high-resolution forecasts cannot be anticipated from the environmental conditions of lower-resolution forecasts with just as much skill (see discussions of Brooks *et al.* 1992 and Roebber *et al.* 2004). It is generally easier to disagree with lower-resolution forecasts than it is to disagree with forecasts that provide data on the same scales as the best of our observational systems.

In forecasts over western Washington state, Colle *et al.* (2000) show significant forecast improvement as the model grid spacing is decreased from 36 to 12 km. However, little forecast improvement is seen as the grid spacing is further decreased from 12 to 4 km. Over the northeastern United States, Colle *et al.* (2003) show little improvement at all when going from 36 to 12 km grid spacing, likely owing to the less sharp terrain features in the northeastern as compared to the northwestern United States. Gallus (1999) further indicates little or no increase in forecast skill when reducing from 30 to 10 km grid spacing in simulations of several midwestern convective systems. These results clearly highlight that model grid spacing by itself is not necessarily the answer to forecast improvements. While the optimal grid spacing for a given model likely depends upon a host of factors, including data assimilation systems, observational density, and model parameterization schemes, any general assumption that a reduction in the grid spacing automatically leads to improved forecasts must be suspect.

Instead of a competition between ensembles and high-resolution deterministic forecasts, there may be ways to merge high-resolution forecasts that are capable of reproducing the smaller-scale features of significant societal and economic interest with an ensemble forecasting system. Stensrud *et al.* (2000b) and Leslie and Speer (2000) discuss this situation, and suggest a combined forecast system in which an ensemble using larger grid spacing is run first and the results evaluated using clustering methods to define the most likely forecast scenarios of the day (e.g., Alhamed *et al.* 2002; Yussouf *et al.* 2004). These most likely forecast scenarios are then used to provide boundary conditions and a first-guess field for assimilating data into a small grid spacing forecast system. Using this type of approach, ensembles are used to ensure that the high-resolution forecast is actually a likely scenario and not one that is outside of the set of ensemble solutions. With a bit more ingenuity, many other possible ways to merge ensembles with high-resolution forecasts are certainly possible to provide both detailed forecasts and information on forecast uncertainty. As the forecast time increases past a few days, the value of single (deterministic) forecasts decreases rapidly and ensembles or other statistical approaches become the only useful prediction approach.

11.4 Statistical postprocessing

Arguably one of the most overlooked aspects of numerical weather prediction is the postprocessing of the forecast data to remove or reduce obvious and persistent errors. Most of the original operational postprocessing schemes used multivariate linear regression to relate the model forecast variables to observations (Glahn and Lowry 1972; Jacks *et al.* 1990). These techniques have provided improved forecast guidance to human forecasters for many decades. The downside to this type of approach is that it requires a lengthy data archive of both observations and an unchanged model, making it difficult to use this type of approach when models are changing frequently. Modifications to this approach to allow for updates have been developed

(Ross 1989; Wilson and Vallée 2002). Other approaches are also possible, such as using a Kalman filter (Homleid 1995; Roeger *et al.* 2003) and other regression techniques (Mao *et al.* 1999; Hart *et al.* 2004).

With the advent of ensembles, other approaches to postprocessing have been developed. Krishnamurti *et al.* (2000) show how a simple bias correction approach can improve the precipitation forecasts in a global ensemble. Stensrud and Yussouf (2003, 2005) and Woodcock and Engel (2005) show that a bias correction approach when applied to near-surface variables yields results that improve upon model output statistics (MOS) and also provide reliable probability forecasts for the short-range predictions of sensible weather (Fig. 11.5). A different approach, called reforecasting, produces an ensemble of retrospective reforecasts from a fixed model over a long time period (15 25 years) in order to diagnose the operational model bias and to provide improved precipitation forecasts using a regression approach (Hamill

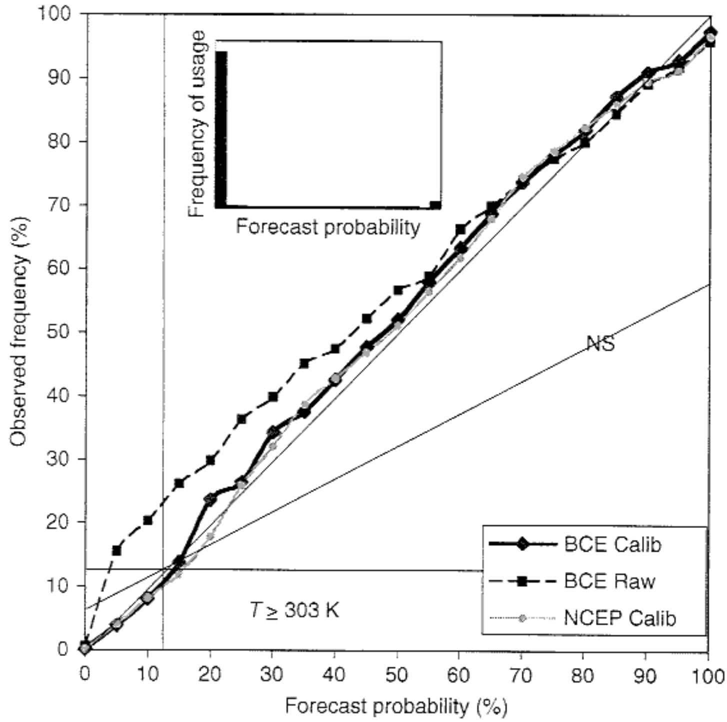

Figure 11.5. Attribute diagram for 2 m temperatures equal to or exceeding 303 K created from a 31-member ensemble over North America. Dashed line indicates results from the raw ensemble, while solid lines indicate results from two versions of postprocessed ensemble data with the bias removed from each ensemble member. Note the improved reliability in the postprocessed ensemble data. From Stensrud and Yussouf (2005).

et al. 2004, 2006). Although producing these reforecasts takes months of computer time, the process is not difficult and leads to significantly improved operational forecasts (Hamill *et al.* 2006). These types of efforts maximize the benefits of the numerical prediction systems and often lead to forecast improvements that are equivalent to decades worth of model improvement alone.

11.5 The road forward

A recurring theme in this final chapter is that there is a lot of uncertainty in how to best use our numerical weather prediction models and their associated computational resources. The meteorology community has spent over 50 years learning how to develop deterministic numerical weather prediction models. The successes arising from this investment are clear and evident every day. Numerical models handle a wide variety of weather scenarios on a vast range of spatial and temporal scales very well and influence operational and investment decisions in a large number of industries. Numerical models and parameterizations will continue to improve as new observations, theories, and ideas are converted into algorithms and studied. Without doubt, numerical weather prediction is an incredibly active and exciting field that will continue to yield forecast improvements for years to come, and one key component of these forecast improvements will continue to be the physical process parameterization schemes.

The improvement of many parameterization schemes often requires collaborations with scientists outside of traditional meteorology and atmospheric science. Not only are meteorologists working with computer scientists, but also with biologists, plant physiologists, remote sensing scientists, foresters, engineers, hydrologists, statisticians, ecologists, economists, and oceanographers. It is clear that the atmospheric sciences are already a multidisciplinary effort and this trend is only going to continue and probably accelerate. We must continue to learn the languages of other sciences in order to learn from their expertise and continue to improve model forecasts.

Beyond the models themselves, we also need to be concerned with how best to use and support our computational and human resources. Some scientists strive to improve finite-differencing or parameterization schemes, while others are working in data assimilation or basic research. Few live in the transition zone between research and operations, and fewer still have the time to step back and look more broadly at how we use these tools we build. For example, not many meteorologists examine how these numerical tools are used by human forecasters to provide guidance products to the public, or to produce forecasts tailored to specific end user communities. Many studies show that human forecasters routinely improve upon numerical guidance in a variety of ways (Olson *et al.* 1995; Leftwich *et al.* 1998; McCarthy *et al.* 1998; Reynolds 2003), yet it is not clear that we are providing the forecast data to forecasters in ways that would allow for the best use of this information or that we are even providing the correct information for all forecast concerns. It also may be that the output from weather and climate models can provide sufficient information to improve human and environmental conditions in ways we never thought possible. An increased emphasis on collaboration between model developers and a broad spectrum of the model user community could be very beneficial.

It is particularly encouraging that seasonal numerical weather prediction systems are starting to be linked directly to models that predict specific human impacts such as malaria incidence (Morse *et al.* 2005) and crop yield (Cantelaube and Terres 2005; Challinor *et al.* 2005; Marletto *et al.* 2005). These types of linkages could have a substantial impact on planning activities that lead to disease prevention and crop selection. Connections between the predictions from global climate models and ecosystem responses are also being explored (Higgins and Vellinga 2004). In short-range forecasting, numerical weather prediction models are being linked to emission and chemistry models to predict air quality (Otte *et al.* 2005), and are used to alert the public to poor air quality conditions that affect human and ecosystem health. These types of activities need to continue and to increase over time to link weather and climate predictions to other quality of life and quality of environment concerns, and may end up profoundly changing the way in which weather and climate model predictions are valued and used by the public.

A need also exists to determine more accurately the economic value of forecasts (e.g., Morss *et al.* 2005) in order to strengthen support for these activities within government budgeting agencies often looking to cut programs (Doswell and Brooks 1998). While there are a number of case studies illustrating the value of weather forecasts (e.g., Katz and Murphy 1997), there has yet to be a sector-wide evaluation of the value of forecasts. This information will be difficult to obtain, but the need for this activity increases from year to year.

Perhaps we have come to the point where the success of numerical weather prediction and the forecast enterprise often is taken for granted by many outside of meteorology. Forecast failures are highlighted, while forecast successes are simply viewed as routine. The success of the numerical weather prediction enterprise breeds increased expectations that model forecasts can provide information on not only the large-scale weather pattern, but also the sensible local weather elements (2 m temperatures, rainfall, precipitation type, and turbulence) that are much more difficult to predict. There also is increased expectation that models can provide information on unlikely events, such as damaging windstorms, tornadoes, tropical cyclones and floods, in order to help with emergency management activities and disaster planning. On longer timescales, the global society is looking to the atmospheric sciences for guidance regarding the best crops to plant for the upcoming growing season and help in understanding how our societies influence global climate change and the outcomes of any mitigation activities. At this point in time, it should not be surprising that there is a need to take the time both to learn how to use these numerical tools most effectively and perhaps even to defend our desires for further forecast improvement. In some ways, parts of numerical weather prediction are a victim of their own success. Continued improvements will occur, but we need to be more careful to illustrate and quantify the value of numerical model improvements to our constituents – the public – and educate them in how to use the output from these tools to their best advantage. If this education occurs, then the future of numerical weather prediction will be brighter than ever.