recombination events between nonallelic copies. The finding of a recombinagenic motif within the repeats may therefore help to explain the observation that the breakpoints of nonallelic recombination events are often clustered (*12*). The overall influence of mobile DNA elements on recombination remains unclear, however, with some over- and some underrepresented within hotspots.

The seven-nucleotide motif is not among those previously associated with recombination in other species. However, its role in influencing recombination is supported by sperm-typing experiments, as is the role of another nine-nucleotide motif motif (CCCCACCCC) identified by the authors. Indeed, at a subset of hotspots in humans, mouse, and yeast, variation in hotspot intensity among individuals has been shown to depend on particular alleles, with recombination events occurring more often initiating on the background of the "hot" variant. When Myers *et al*. examined the sequence context of two human hotspots whose intensity has been shown to vary among alleles, they found that the "hot" alleles were their top-scoring seven and nine oligomer motifs and that in both cases, the "colder" alleles were a mutation away from that motif. This result strongly suggests that these sequences modulate hotspot activity (see the figure). Further evidence will come from sperm-typing studies of other hotspots polymorphic at the same motifs, as well as at other candidate sequences.

In light of recent reports that hotspot locations are largely discordant in humans and chimpanzees (*9*, *13*), the discovery of human motifs that appear to influence hotspot activity raises a number of additional questions: Can changes to sequence motifs explain most of the interspecies differences, or do other genomic features, such as chromatin accessibility or transposable element activity, explain their rapid evolution? Given that most recombination events take place within hotspots, and hotspot locations appear to be rapidly evolving, is there any constraint on recombination rates below that of a chromosomal arm? For example, are the density and intensity of hotspots constrained within circumscribed regions of the genome? With more sperm-typing experiments and extensive linkage disequilibrium data collection in close evolutionary relatives of humans, answers to these questions should no longer be elusive.

References

1. T. Hassold, P. Hunt, *Nat. Rev. Genet.* **2**, 280 (2001).
2. A. Lynn, T. Ashley, T. Hassold, *Annu. Rev. Genomics Hum. Genet.* **5**, 317 (2004).
3. S. Myers, L. Bottolo, C. Freeman, G. McVean, P. Donnelly, *Science* **310**, 321 (2005).
4. N. Arnheim, P. Calabrese, M. Nordborg, *Am. J. Hum. Genet.* **73**, 5 (2003).
5. L. Kauppi, A. J. Jeffreys, S. Keeney, *Nat. Rev. Genet.* **5**, 413 (2004).
6. T. D. Petes, *Nat. Rev. Genet.* **2**, 360 (2001).
7. D. C. Crawford *et al.*, *Nat. Genet.* **36**, 700 (2004).
8. G. A. McVean *et al.*, *Science* **304**, 581 (2004).
9. S. E. Ptak *et al.*, *Nat. Genet.* **37**, 429 (2005).
10. D. A. Hinds *et al.*, *Science* **307**, 1072 (2005).
11. P. Fearnhead, N. G. C. Smith, *Am. J. Hum. Genet.*, www.journals.uchicago.edu/AJHG/home.html.
12. J. R. Lupski, *Genome Biol.* **5**, 242 (2004).
13. W. Winckler *et al.*, *Science* **308**, 107 (2005).

10.1126/science.1120154

ATMOSPHERIC SCIENCE

Weather Forecasting with Ensemble Methods

Tilmann Gneiting and Adrian E. Raftery

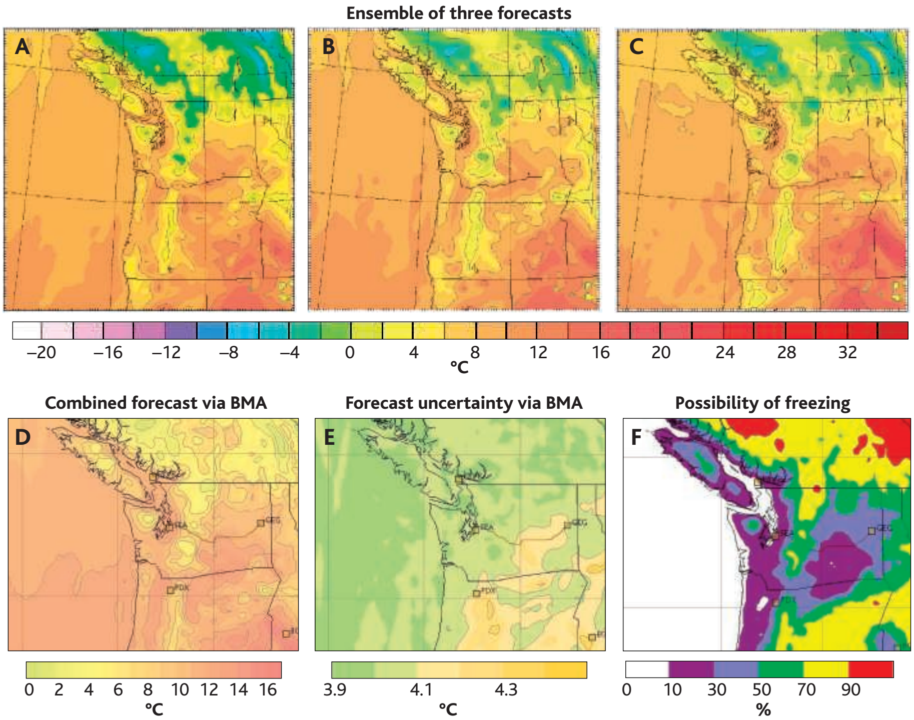
A radical change has occurred in the practice of numerical weather prediction over the past decade. Until the early 1990s, atmospheric scientists viewed weather forecasting as an intrinsically deterministic endeavor: For a given set of "best" input data, one "best" weather prediction is generated. Armed with sophisticated computing resources (including supercomputers), weather centers ran carefully designed numerical weather prediction models to produce deterministic forecasts of future atmospheric states. Although this is still the case today, weather prediction has been transformed through the implementation of ensemble forecasts. An ensemble forecast comprises multiple (typically between 5 and 100) runs of numerical weather prediction models, which differ in the initial conditions and/or the numerical representation of the atmosphere, thereby addressing the two major sources of forecast uncertainty.

Realizing the full potential of an ensemble forecast requires statistical postprocessing of the model output, in that model biases, insufficient representations of forecast uncertainty, and the differing spatial scales of model gridboxes and observations need to be addressed. In concert with statistical postprocessing, ensembles provide flow-dependent probabilistic forecasts in the form of predictive probability distributions over future weather quantities or events. Probabilistic forecasts allow one to quantify weather-related risk, and they have greater economic value than deterministic forecasts in a wide range of applications, including electricity generation, aircraft and ship routing, weather-risk finance, and disease modeling (*1*).

A maturing area is that of medium-range probabilistic forecasting at prediction horizons up to 10 days, which involves ensembles of global numerical weather prediction models (*1*, *2*). Three operational methods for the generation of medium-range initial condition ensembles have been developed. The U.S. National Centers for Environmental Prediction (NCEP) and the European Centre for Medium-Range Weather Forecasts (ECMWF) seek directions of rapid error growth in selective sampling procedures, known as the bred-vector perturbation method (*3*) and the singular-vector technique (*4*), respectively. The Meteorological Service of Canada (MSC) uses the Monte Carlo–like perturbed-observation approach (*5*), in which the model physics parameterizations vary as well. Ensemble forecasting and atmospheric data assimilation (the melding of weather observations into a numerical model) can mutually benefit from each other, and there are promising options for a linked system (*6*). A recent comparative study suggests that the ECMWF data assimilation, numerical modeling, and ensemble generation system has the best overall performance, with the NCEP system being competitive during the first few days, and the MSC system during the last few days, of the 10-day forecast period (*7*). The successful operation of forecast ensembles on the global scale has motivated the development of limited-area short-range ensembles driven by initial and boundary conditions supplied by different weather centers, such as the University of Washington ensemble system (*8*–*10*) over the North American Pacific Northwest (see the figure).

Probabilistic forecasting has become an integral part of seasonal prediction as well (*1*, *11*). Forecasts on seasonal to interannual time scales rely on comprehensive global coupled ocean-atmosphere models and have become feasible with an improved understanding of the coupling between sea surface temperature anomalies and atmospheric circulation patterns. A recent special issue of *Tellus* (*12*) is dedicated to results from the European Union–spon-

The authors are in the Department of Statistics, University of Washington, Box 354322, Seattle, WA 98195, USA. E-mail: tilmann@stat.washington.edu



An improved forecast. Ensemble forecast of surface temperature over the North American Pacific Northwest, with postprocessed probabilistic forecast products derived by Bayesian model averaging (BMA) (*14*). (**A** to **C**) The ensemble consists of nine 48-hour forecasts (of which three are shown) valid at 4 p.m. local standard time on 2 April 2005, using the MM5 mesoscale model with initial conditions provided by different weather centers (*8–10*). (**D**) The BMA combined forecast is a weighted average of the bias-corrected ensemble members (*10, 14*). (**E**) The uncertainty plot is a map of the half-width of the BMA forecast intervals. Higher values correspond to more uncertainty. (**F**) The BMA probability of freezing refers to the 24-hour period ending at the valid time.

sored DEMETER (Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction) project. A single supercomputer hosted seven independent state-of-the-art models, which produced a series of 6-month ensemble reforecasts with common archiving systems and diagnostics. Each model was run nine times with different initial conditions, resulting in global multimodel, multi–initial condition ensemble reforecasts over the past 50 years. The DEMETER ensemble improved both deterministic and probabilistic forecast skill when compared to the single-model ensembles, in ways that cannot be attributed to the increase in ensemble size only. Applications to malaria incidence and crop yield prediction have shown the benefits of linking seasonal forecast ensembles to end-user models that are also run in ensemble mode. Building on the success of the DEMETER project, an operational real-time seasonal ensemble prediction system has been established at ECMWF.

Current challenges include the representation of forecast uncertainty due to the use of imperfect numerical models. Model uncertainty can be addressed through the use of multimodel ensembles (in which each single model run is deterministic), or through stochastic representations of parameterized physical processes, as implemented in the ECMWF medium-range ensemble, thereby introducing randomness into the model runs (*13*). Both options link flow-dependent forecast uncertainty and model-related errors, and it remains to be seen whether they are superior in any way to approaches based purely on statistical postprocessing (*7, 14*). Nor has the debate on selective versus Monte Carlo sampling of initial condition uncertainty been resolved, although it may evolve in novel directions as operational experience with various methods of sequential data assimilation accrues.

From daily to seasonal time scales, probabilistic forecasts based on ensembles have become a prominent part of numerical weather prediction. The ability of ensemble systems, in concert with statistical postprocessing, to improve deterministic forecasts—in that the ensemble mean forecast outperforms the individual ensemble members—and to produce probabilistic and uncertainty information to the benefit of weather-sensitive public, commercial, and humanitarian sectors has been convincingly established. More work needs to be done to routinely provide fully reliable, flow-dependent probabilistic forecast distributions, particularly of weather fields, as opposed to forecasts at individual sites. In keeping with the remarkable pace of progress since the early 1990s, we anticipate notable improvements in deterministic and probabilistic forecast skill through the continued development of multimodel, multi–initial condition ensemble systems and advanced, grid-based statistical postprocessing techniques.

Additional effort is required in the communication, visualization, and evaluation of probabilistic forecasts, and differing interpretations of probability need to be reconciled, to avoid the risk of perfecting ensemble methodologies without a clear aim (*15, 16*). To this end, the paradigm of maximizing the sharpness of the probabilistic forecasts under the constraint of calibration may offer guidance. Calibration refers to the statistical consistency between the probabilistic forecasts and the observations; sharpness refers to the spread of the predictive distributions and is a property of the forecasts only. The goal is to increase sharpness in the forecasts, without compromising the validity of the probability statements.

References and Notes

1. T. N. Palmer, *Q. J. R. Meteorol. Soc.* **128**, 747 (2002).
2. For real-time ensemble forecasts, see NCEP Operational Ensemble Forecasts (www.cdc.noaa.gov/map/images/ens/ens.html) and MSC Ensemble Forecasts (http://weatheroffice.ec.gc.ca/ensemble/index_e.html).
3. Z. Toth, E. Kalnay, *Bull. Am. Meteorol. Soc.* **74**, 2317 (1993).
4. F. Molteni, R. Buizza, T. N. Palmer, T. Petroliagis, *Q. J. R. Meteorol. Soc.* **122**, 73 (1996).
5. P. L. Houtekamer, L. Lefaivre, J. Derome, H. Ritchie, H. L. Mitchell, *Mon. Weather Rev.* **124**, 1225 (1996).
6. G. Evensen, *J. Geophys. Res.* **99**, 10143 (1994).
7. R. Buizza *et al.*, *Mon. Weather Rev.* **133**, 1076 (2005).
8. E. P. Grimit, C. F. Mass, *Weather Forecast.* **17**, 192 (2002).
9. F. A. Eckel, C. F. Mass, *Weather Forecast.* **20**, 328 (2005).
10. Real-time forecast products are available at the University of Washington Mesoscale Ensemble (www.atmos.washington.edu/~ens/uwme.cgi) and Bayesian Model Averaging (http://isis.apl.washington.edu/bma/index.jsp) sites.
11. T. N. Palmer *et al.*, *Bull. Am. Meteorol. Soc.* **85**, 853 (2004).
12. T. N. Palmer, Ed., Special issue on the DEMETER project, *Tellus A* **57** (May 2005), pp. 217–512.
13. R. Buizza, M. Miller, T. N. Palmer, *Q. J. R. Meteorol. Soc.* **125**, 2887 (1999).
14. A. E. Raftery, T. Gneiting, F. Balabdaoui, M. Polakowski, *Mon. Weather Rev.* **133**, 1155 (2005).
15. I. T. Jolliffe, D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science* (Wiley, New York, 2004).
16. R. de Elía, R. Laprise, *Mon. Weather Rev.* **133**, 1129 (2005).

17. While preparing this Perspective, T.G. was on sabbatical leave at the Soil Physics Group, Universität Bayreuth, Germany. We thank C. Mass for comments; J. Baars, E. Grimit, and P. Tewson for assistance in preparing the figure; and the U.S. Department of Defense Multidisciplinary Research University Initiative (MURI) for financial support.

10.1126/science.1115255