Statistical postprocessing for ensembles of numerical weather prediction models

Tilmann Gneiting
University of Heidelberg, t.gneiting@uni-heidelberg.de

Abstract: The past fifteen years have witnessed a radical change in the practice of weather forecasting, in that ensemble prediction systems have been implemented operationally. An ensemble forecast comprises multiple runs of numerical weather prediction models, which differ in initial and lateral boundary conditions, and/or the parameterized representation of physical processes. However, ensemble forecasts are subject to biases and dispersion errors, and thus statistical postprocessing is required, with Bayesian model averaging and ensemble model output statistics being state of the art approaches. Future work is called for to ensure that the postprocessed forecast fields show physically realistic and coherent joint dependence structures across meteorological variables, geographic space and look-ahead times.

Keywords: Bayesian model averaging; ensemble model output statistics; numerical weather prediction; statistical postprocessing

1 Introduction

A major human desire is to make forecasts for an uncertain future. Consequently, forecasts ought to be probabilistic in nature, taking the form of probability distributions over future quantities or events (Dawid 1984; Gneiting 2008). That said, weather forecasting has traditionally been viewed as a deterministic exercise, drawing on highly sophisticated numerical models of the atmosphere. The advent of ensemble prediction systems in the 1990s marks a radical change (Palmer 2002; Gneiting and Raftery 2005). An ensemble forecast comprises multiple runs of numerical weather prediction models, which differ in initial conditions, lateral boundary conditions, and/or the parameterized representation of the atmosphere being used. An example from the University of Washington Mesoscale Ensemble (Grimit and Mass 2002) over Western North America and the Northeast Pacific Ocean is shown in Figure 1.

2 Statistical postprocessing of ensemble weather forecasts

Realizing the full potential of an ensemble forecast requires statistical postprocessing of the model output, to address model biases and dispersion errors.

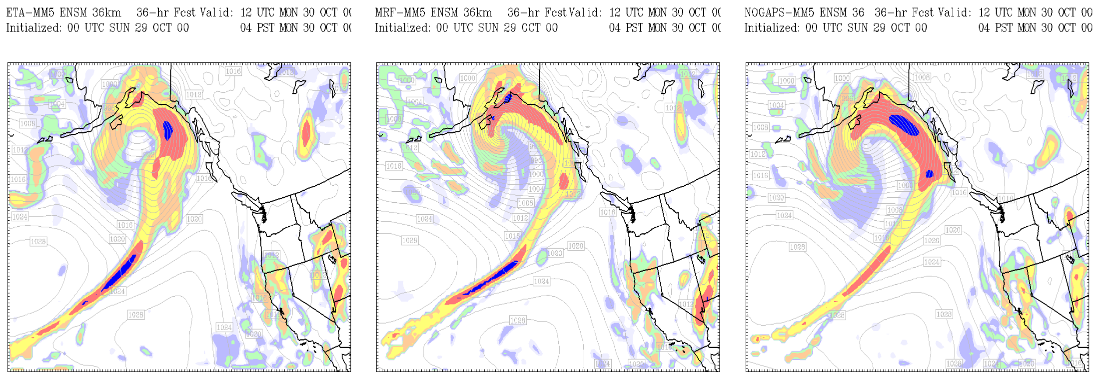

Figure 1: 36-hour ahead ensemble forecast valid October 30, 2000 over Western North America and the Northeast Pacific Ocean, with color representing precipitation amounts. Three members of the University of Washington Mesoscale Ensemble (Grimit and Mass 2002) are shown.

Popular approaches for doing this include the Bayesian model averaging (BMA) method developed by Raftery et al. (2005) and the ensemble model output statistics (EMOS), or heterogeneous regression, technique introduced by Gneiting et al. (2005). The BMA approach employs a mixture distribution, where each mixture component is a parametric probability density associated with an individual ensemble member, with the mixture weight reflecting the member's relative contributions to predictive skill over a training period. In contrast, the EMOS predictive distribution is a single parametric distribution.

To fix the idea, consider an ensemble of NWP forecasts, $f_1, \ldots, f_k$, for temperature, x, at a given time and location. Let $\phi(x; \mu, \sigma^2)$ denote the normal density with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$ evaluated at $x \in \mathbb{R}$. The BMA approach of Raftery et al. (2005) employs Gaussian components with a linearly bias-corrected mean. The BMA predictive density for temperature then becomes

$$p(x \,|\, f_1, \ldots, f_k) \;=\; \sum_{i=1}^{k} w_i \, \phi(x; a_i + b_i f_i, \, \sigma^2),$$

with BMA weights, $w_1, \ldots, w_k$, that are nonnegative and sum to 1, bias parameters $a_1, \ldots, a_k$ and $b_1, \ldots, b_k$, and a common variance parameter, σ^2, all of which being estimated from training data over a rolling training period that consists of the recent past. The EMOS approach of Gneiting et al. (2005) employs a single Gaussian predictive density, in that

$$p(x \,|\, f_1, \ldots, f_k) = \phi(x; a + b_1 f_1 + \cdots + b_k f_k, \, c + d s^2),$$

with regression parameters a and $b_1, \ldots, b_k$, and spread parameters c and d, where s^2 is the variance of the ensemble values. The EMOS technique thus is more parsimonious, and the BMA method is more flexible.

While the original methodological development of Raftery et al. (2005) and Gneiting et al. (2005) was addressed at temperature and surface pressure, more recent work aims at the statistical postprocessing of ensemble forecasts for quantitative precipitation (Sloughter et al. 2007), wind speed (Sloughter et al. 2010; Thorarinsdottir and Gneiting 2010) and wind direction (Bao et al. 2010). For a fully Bayesian alternative to the BMA approach of Raftery et al. (2005), see Di Narzio and Cocchi (2010).

3 Challenges for future work

Even though Bayesian model averaging and ensemble model output statistics are state of the art methods, they treat distinct weather variables at distinct geographic locations and distinct look-ahead times independently of each other. This conflicts with key applications such as air traffic control, flood management or winter road maintenance, where it is critically important that the postprocessed forecast fields show physically realistic and coherent joint dependence structures across meteorological variables, geographic space and look-ahead times.

Perhaps the most advanced technique in these directions is the Spatial BMA approach of Berrocal, Raftery and Gneiting (2007), who merged the traditional BMA approach of Raftery et al. (2005) with the geostatistical output perturbation (GOP) technique of Gel, Raftery and Gneiting (2004) to obtain probabilistic temperature field forecasts that honor the spatial structure of observations. Similarly, the Bernoulli-Gamma BMA approach of Sloughter et al. (2007) could be merged with the two-stage spatial method of Berrocal, Raftery and Gneiting (2008), which uses Gaussian copulas, to yield spatially and/or temporally coherent postprocessed forecast fields for quantitative precipitation. Variants of the Schaake shuffle (Clark et al. 2004) provide nonparametric alternatives. Work along these lines is a critical research need in the statistical postprocessing of ensemble weather forecasts, and there is ample scope for continued methodological development, using nonparametric tools, methods of spatial and spatio-temporal statistics, and/or copula techniques.

References

Bao, L., Gneiting, T., Grimit, E. P., Guttorp, P. and Raftery, A. E. (2010). Bias correction and Bayesian model averaging for ensemble forecasts of surface wind direction. *Monthly Weather Review*, **138**, 1811–1821.

Berrocal, V. J., Raftery, A. E. and Gneiting, T. (2007). Combining spatial statistical and ensemble information for probabilistic weather forecasting. *Monthly Weather Review*, **135**, 1386–1402.

Berrocal, V. J., Raftery, A. E. and Gneiting, T. (2008). Probabilistic quantitative precipitation field forecasting using a two-stage spatial model. *Annals of Applied Statistics*, **2**, 1170–1193.

Clark, M., Gangopadhyay, S., Hay, L., Rajagopalan, B. and Wilby, R. (2004). The Schaake shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, **5**, 243–262.

Dawid, A. P. (1984). Statistical theory: The prequential approach (with discussion and rejoinder). *Journal of the Royal Statistical Society Series A: General*, **147**, 278–292.

Di Narzo, A. F. and Cocchi, D. (2010). A Bayesian hierarchical approach to ensemble weather forecasting. *Journal of the Royal Statistical Society Series C: Applied Statistics*, **59**, 405–422.

Gel, Y., Raftery, A. E. and Gneiting, T. (2004). Calibrated probabilistic mesoscale weather field forecasting: The Geostatistical Output Perturbation (GOP) method (with discussion). *Journal of the American Statistical Association*, **99**, 575–587.

Gneiting, T. (2008). Editorial: Probabilistic forecasting. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **171**, 319–321.

Gneiting, T. and Raftery, A. E. (2005). Weather forecasting with ensemble methods. *Science*, **310**, 248–249.

Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Monthly Weather Review*, **133**, 1098–1118.

Grimit, E. P. and Mass, C. F. (2002). Initial results of a mesoscale short-range ensemble system over the Pacific Northwest. *Weather and Forecasting*, **17**, 192–205.

Palmer, T. N. (2002). The economic value of ensemble forecasts as a tool for risk assessment: From days to decades. *Quarterly Journal of the Royal Meteorological Society*, **128**, 747–774.

Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, **133**, 1155–1174.

Sloughter, J. M., Raftery A. E., Gneiting, T. and Fraley, C. (2007). Probabilistic quantitative precipitation forecasting using Bayesian model averaging. *Monthly Weather Review*, **135**, 3209–3220.

Sloughter, J. M., Gneiting, T. and Raftery, A. E. (2010). Probabilistic wind forecasting using ensembles and Bayesian model averaging. *Journal of the American Statistical Association*, **105**, 25–35.

Thorarinsdottir, T. L. and Gneiting, T. (2010). Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society Series A: Statistics in Society*, **173**, 371–388.