Communicating probability with real-time calibrated forecasts

Patrick Tewson

Introduction

The UW Ensemble Bayesian Model Averaging website serves as a probabilistic forecast tool and a test bed for examining new techniques to create and visualize probabilistic weather forecasts for the Pacific Northwest, with the aim of promoting the use of probabilistic forecasts within forecast communities. Specifically, this tool uses model output from the UW MM5 core ensemble, the members of which are then combined using Bayesian Model Averaging to produce calibrated and sharp forecasts of surface parameters. The end result is a high resolution forecast that includes reasonable estimations of uncertainty and probability that a certain forecast event (rain, temperature below freezing) may occur. This work is funded by a MURI grant shared by the University of Washington's departments of atmospheric sciences, statistics, psychology, and the Applied Physics Lab (APL).

Features

The UW BMA web site provides probabilistic forecasts of afternoon surface temperature, maximum and minimum temperature over a 24-hour period, and accumulated precipitation over a 12-hour daytime period. These forecasts attempt to be calibrated, which is to say the probability accurately reflects the expected outcome, and sharp, meaning that confidence is high that the observed outcome will fall within a small range of the predicted outcome. These properties are significant because a forecast that follows these tenants will be most useful to a user concerned with certain fundamentals -- “Will it freeze tonight? Will it rain, and if so how much?” Verification information is calculated as part of the system and made available to the user via a companion web site.

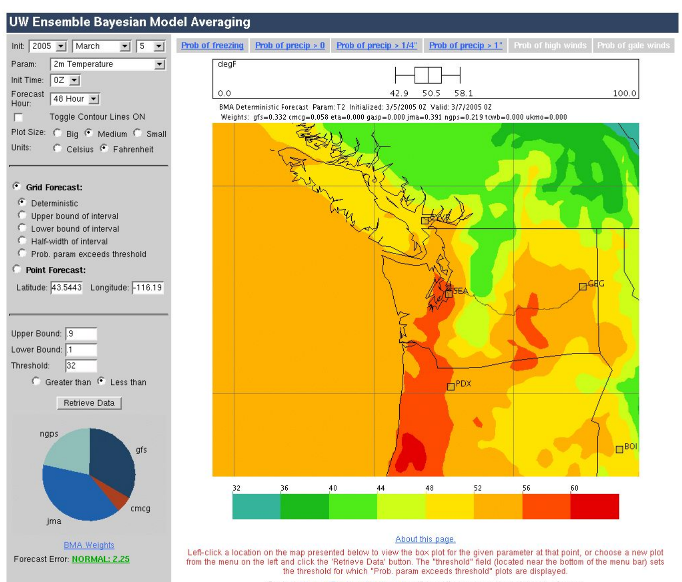

Figure: BMA page showing a deterministic forecast of surface temperature.

While numerous methods of post-processing models and ensembles have been discussed and researched, few products available to the end user make use of these techniques. Similarly, the raw output of an ensemble tends to be too diverse and complex to be useful to forecast consumers, and raw models include no information as to the uncertainty of the forecast. The UW BMA web page provides a summary of the UWME that include estimations of probability and uncertainty that are verifiably more accurate than those produced by the raw ensemble or climatology. In addition, straightforward visualizations of the parameters are provided to the end user. Since all of the utilized model data originates from the UWME 12-km domain, the BMA product is constrained to the same region, consisting of Washington, Oregon, much of Idaho, and Southwest Canada. However, the techniques used here should be broadly applicable to any forecast ensemble.

When arriving at the BMA web site, the user is greeted with a deterministic forecast of surface temperature, and several tools to specify new forecasts or understand the probabilistic nature of the current one. The forecast display on the web side is controlled from a menu located on the left side of the page. In addition to providing the most recent forecast, “forecasts” stretching into the past are also made available for comparison and general interest. The top elements of the menu control the date (defaulted to today's forecast), followed by UI elements to select model parameter (max/min temperature, accumulated precipitation, etc.), model initialization time, and forecast hour. The user can also control certain elements of the visualization with elements located beneath the former, such as display of contour lines, plot size, and units.

A line of buttons at the top of the page provides direct access to six forecasts deemed most useful to the user. These include **probability of freezing surface temperature, probability of measurable rain, probability of more than a 1/4” of rainfall,** and **probability of more than 1” of rainfall.** Two others relating to high and dangerous winds are not yet activated as wind prediction has not yet been implemented. Freezing temperatures apply to the overnight period, while rainfall applies to an 12-hour daytime accumulation period.

Five map plots of primary interest to the user are selectable from the mid-level of the left-hand menu. These include a deterministic forecast, and upper bound of the confidence interval (or forecast of highest likely outcome), lower bound of the confidence interval (lowest likely outcome), half-width of the predictive interval (an estimation of forecast uncertainty or predictive ability) and probability that a certain threshold is exceeded, such as the temperature drops below freezing or more than a quarter inch of rain will fall. Below these elements the user can control the levels of the upper and lower bound or threshold.

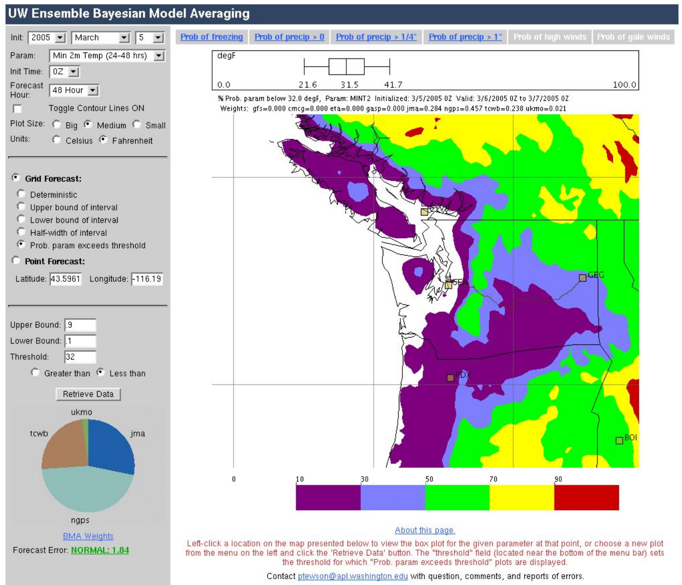

Figure: BMA page showing a probability of surface temperature falling below freezing.

Other elements of the display help the user understand the probabilistic nature of the forecast on display. Located directly above the map, the **box plot** gives information about a single map location, selected by the user by clicking on the map display. The box plot displays a mean or median forecast along with 10%, 25%, 75%, and 90% confidence bounds, conveying the likely outcome along with the range of expected possibilities within certain quantities of confidence. In addition, the **model weights pie chart** at the bottom of the left-hand menu displays the contribution of each of the core ensemble members to the forecast on display. Beneath the pie chart, a small readout gives an indication of recent verification error. The latitude and longitude fields display the location of the current box plot on display, or allow the user to manually enter a location. Selecting the "point forecast" button brings up a forecast probability distribution function for the selected location instead of the map.

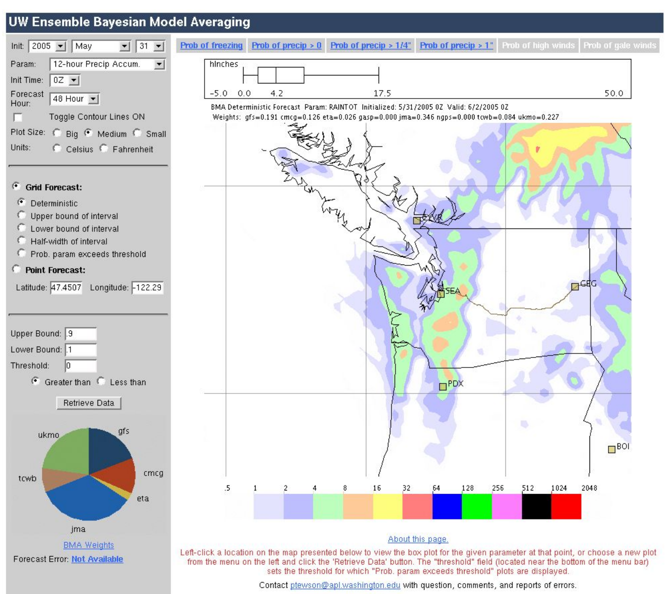

Figure: BMA page showing a deterministic precipitation forecast.

Description of BMA algorithm

The UW BMA web site uses a Bayesian Model Averaging technique to combine the output of the core UW Mesoscale Ensemble (UWME). The core UWME is the combined output of eight runs of the Mesoscale Model 5 (MM5), using identical versions of the model but differing initial conditions, in attempt to capture a broader degree of possible outcomes then would a single model. Ensembles tend to exhibit a relationship between model agreement and predictive ability knows as the **spread-skill relationship.** However, while ensemble predictions are sharp they are uncalibrated.

BMA describes a method of fitting a probability distribution to every individual ensemble member, after which these distributions can be combined by a weighted sum to produce a single distribution based on the ensemble output for every point within the model domain. Model variance and weights are computed by applying the algorithm to a training data set consisting of past set of model output and observations (nominally 25 days worth), after which these outputs are applied to generate the BMA forecast from the latest set of model data.

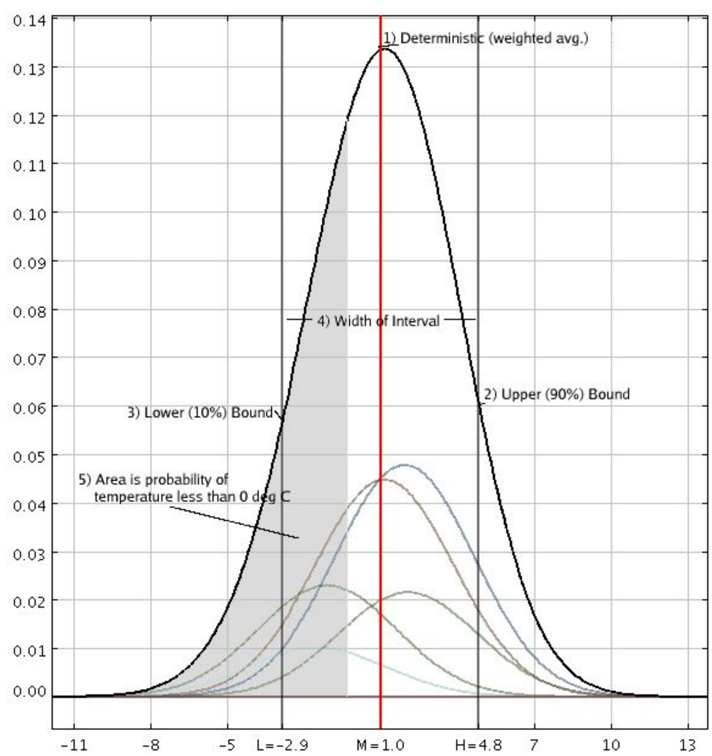

Figure: BMA PDF

The five forecast parameters are derived from the BMA forecast PDF (seen here for a normally-distributed parameter such as temperature) – (1), the **deterministic forecast**, (2), the **upper bound** of predictive interval (90% of the area of the curve below this point), (3) the **lower bound** of the predictive interval (90% of the curve's area list above this point, (4) the **width of the predictive interval** (spread between the upper and lower bounds and estimation of uncertainty, and (5) the **probability of exceeding a threshold,** in this case temperature dropping below freezing.

Technical details

The BMA web application provides the user managed access to an archive of model data that extends for over four years and uses over two terrabytes of disk space. This data is accessed both at request time, when plots are dynamically generated and returned to the client browser, and as part of background processes that update the BMA training data set and calculate new sets of BMA weights. This separation between tasks that occur in the background on a scheduled basis, such as the retrieval of new data and calculation of the weights, and those that are triggered by user interaction with the web site forms the fundamental division of work for the application.

Background tasks involve gathering the latest observational and model data, combining these into the file that will be used to train the BMA, and calculating a set of BMA outputs (weights, model variance, and bias correction terms). These generally run as small Java applications, or as in the case of the BMA calculation itself, scripts written for the free software package R.

The web-based application operates under the Catalina / Tomcat web server, a free web server that runs Java technologies. When a user choses to view a plot on the web site, BMA parameters are combined with the appropriate ensemble forecast to produce the requested visualization. This may involve evaluation of the CDF at every point in the model grid, totaling to almost 100,000 Gaussian integrations to create a single plot. However, with modern processor capabilities and optimized software, the calculation requires only a brief second or two.

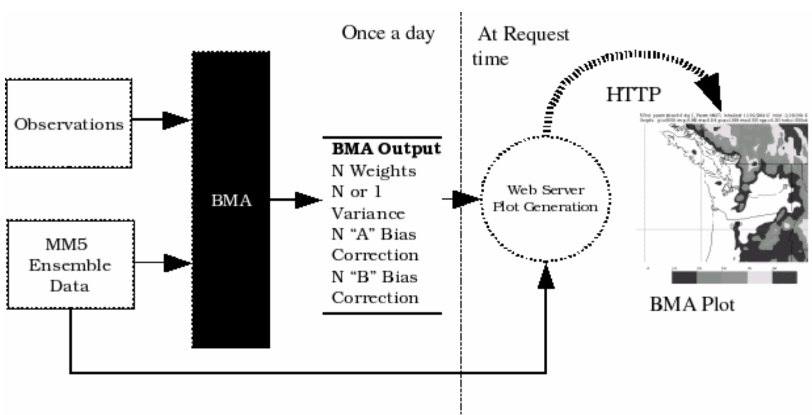

Figure: Simple data flow diagram

Collaborators

The BMA web application represents the combined effort of several different groups within and without the University of Washington, among them atmospheric sciences, statistics, psychology, and the Applied Physics Laboratory. This project represents a unique opportunity to bring together groups of experts from many different fields and unify their talents with the goal of shifting thought on weather forecasting towards adoption of probability and uncertainty information.

Future developments

In addition to bringing forecasts of surface winds into the application, the BMA web site will soon offer dramatically sharper forecasts through improvements to the use of the BMA technique that introduce spacial variability to the BMA parameters. Further improvements in usability and simplification are ongoing and will be manifest in future product iterations.

Conclusion

The BMA website is a prototype that demonstrates the potential for real-time forecast products that include probability and uncertainty, and represents the combined efforts of the MURI team. The design of the site advertises a large set of capability that a final product could include, and therefore serves to illustrate these features more than act as a streamlined and targeted application for a more specific use. Future applications can reduce and distill this set of capabilities to improve the experience of a known user demographic.

References

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

Eckel, F. A. and Mass, C. F. (2005). Effective mesoscale, short-range ensemble forecasting. Accepted to *Wea. Forecasting*.

Raftery, A.E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Mon. Wea. Rev.,* **133,** 1155-1174.