P2.30 WEATHER FORECAST UNCERTAINTY MANAGEMENT AND DISPLAY

Randy J. Lefevre*, Jonathan Pfautz, and Kenneth Jones

ApMet, Albuquerque, New Mexico;
Charles River Analytics Inc., Cambridge, Massachusetts

1. INTRODUCTION

This paper presents initial results from the Weather Prediction Uncertainty Management And Representation (PUMAR) project sponsored by the Army Research Laboratory, White Sands Missile Range in New Mexico. The goal of the PUMAR project was to cognitively engineer a user-friendly software tool to effectively communicate weather forecast uncertainty to a decision-maker assessing weather impacts.

2. BACKGROUND

One of the greatest challenges in meteorology is quantifying and communicating the uncertainty in weather forecasts, especially those forecasts based on numerical weather prediction (NWP) information. Weather decision aids, widely used in military operations, are often exploited to "translate" the NWP information into actionable weather intelligence. However, unrepresentative or erroneous environmental observations, errors in the data assimilation process, and inaccurate or imprecise microphysics or numerical analysis techniques can lead to uncertainty in the NWP output data. Extensive, on-going research on each segment of the NWP process is leading to the development of better analytical techniques, but the issue of quantifying the uncertainty remains a challenge. Ensemble techniques applied to NWP have proven to be valuable forecast tools but interpreting the output is often difficult, and they have demonstrated sensitivities to initial conditions. Bayes' Theorem and Bayesian Hierarchical Model (BHM) techniques have recently been applied to the NWP process with documented success at quantifying weather forecast uncertainty (Wikle et al. 2003, Berliner et al. 2003, Berliner et al. 1998).

The development of the PUMAR software began by establishing a representative scenario related to US Army operations—using a tactical unmanned aerial vehicle (T-UAV) for reconnaissance, surveillance, and target acquisition. The domain of the T-UAV reconnaissance scenario was limited to a portion of the White Sands Missile Range (WSMR) in central and southern New Mexico. The PUMAR demonstration focused on a specific 36-hour window in which specific targets needed to be imaged by the T-UAV.

The meteorological conditions used for the scenario covered a subset of the period of low clouds and heavy precipitation over New Mexico from 2-5 April 2004 (Figure 1). During this period, several cities in New Mexico recorded over 2.5 inches of precipitation. The Air Force Weather Agency (AFWA) provided the NWP information based on their operational implementation of the National Center for Atmospheric Research Pennsylvania State University Mesoscale Model (MM5). The PUMAR demonstration included several different NWP variables that could impact T-UAV operations, but for this paper we used the relative humidity (RH), cloud water (CW), cloud ice (CI), cloud snow (CS), and cloud rain (CR) mixing ratios to predict clouds—specifically cloud ceiling. Cloud ceiling, which dictates the ability of the T-UAV to image ground targets, most easily demonstrates the feasibility of the PUMAR application to military decision aids.

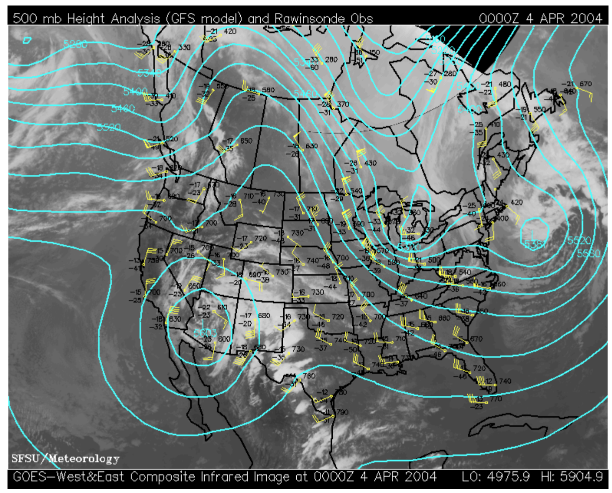

Figure 1. A representative meteorological infrared satellite image and 500 hPa geopotential height contours (courtesy of the California Regional Weather Server, http://www.squal.sfsu.edu) from 4 April 2004 at 0000 UTC depicting a cut-off low over Arizona and extensive clouds and precipitation over New Mexico.

The PUMAR demonstration incorporated the BHM approach using the commercially available BNet.Builder© software package, which allows the user to graphically interact with the NWP information to represent the types and sources of uncertainty in the forecast process (BNet.Builder© is available from Charles River Analytics Inc., http://www.cra.com/bnet). It was beyond the scope of this paper, however, to explore the use of BHM to address all the factors in the NWP process leading to uncertainty.

3. BAYESIAN HIERACHICAL MODEL (BHM)

The BHM formalism, based on probability theory developed by Thomas Bayes published posthumously in 1763 (http://www.bayesian.org/bayesian/bayes.html), is expressed as:

* Corresponding author address: Randy J. Lefevre, Apmet, 1215 Diamondback Drive NE, Albuquerque, NM 87113-2294; e-mail: randy.lefevre@apmet.net.

$$[H \mid E,c] = \frac{[E \mid H,c] * [H \mid c]}{[E \mid c]} \qquad (1)$$

where [H|E,c] is the POSTERIOR probability distribution of a hypothesis (H) given the evidence (E) and the background context (c), [E|H,c] is the LIKELIHOOD probability distribution of the evidence given the hypothesis, [H|c] is the PRIOR probability distribution alone, and [E|c] is a normalizing distribution to allow the POSTERIOR probability to be between zero and one. The hierarchical aspect of the BHM is the application of Bayes' Theorem to several variables, arranged hierarchically, that express some causal influence on the forecast variable (and on each other).

The POSTERIOR distribution for this paper is the probability of cloud given the MM5 NWP evidence (i.e. RH, CW, CI, CS, and CR). The shape of the probability distribution quantifies the uncertainty assessment. The LIKELIHOOD distributions are the probability distributions explicit to the MM5 NWP model based on data assimilation of observations, microphysical parameters, and numerical analysis techniques. The PRIOR distribution as captured in the Conditional Probability Table (CPT) of BNet.Builder© is based on expert knowledge of cloud formation. The PRIOR "probability distribution" is indicative of the inherent uncertainty in the expert's knowledge.

4. BNET.BUILDER© ANALYSIS

The BNet.Builder application, graphically depicting the hierarchical nature of the cloud analysis (Fig. 2), was executed on each grid point (spatially and temporally) throughout the PUMAR domain. The PRIOR distribution (Conditional Probability Table independent of NWP) shown in the lower portion of Figure 2 related the *a priori* knowledge of the cloud variables (e.g. CW, CR, CS, and CI) to the cloud condensate, and the cloud condensate distribution combined with the relative humidity distribution (thus the hierarchy) determined the overall PRIOR cloud distribution. The PRIOR distributions were based on expert knowledge and did not include NWP intelligence, nor did the PRIOR distribution account for increased uncertainty with forecast hour. Since the focus of this paper was on the application of the BHM information to influence graphical interpretation of the data, the PRIOR cloud condensate and relative humidity distributions were discretized to high, medium, low, or none. For example, the cloud condensate PRIOR distribution was skewed to the high end if all the cloud variables were considered high, and it was skewed to the low end if all the cloud variables were considered low.

The LIKELIHOOD cloud condensate and relative humidity distributions (evidence) were likewise discretized to high, medium, low or none. The evidence for each cloud variable was provided by AFWA's MM5 data. The discrete category of each variable was based on AFWA's threshold value for each mixing ratio. For example, the cloud water (CW) mixing ratio was considered high if the grid point value was greater than or equal to 0.025 g kg^{-1}, medium if CW was between 0.025-0.015 g kg^{-1}, low if CW was between 0.015-0.001 g kg^{-1}, and none if CW was equal to zero.

The LIKELIHOOD distributions (evidence), based on the NWP information of cloud condensate and relative humidity, combined with the PRIOR distributions of cloud condensate and relative humidity provided the input to BNet.Builder©'s inference engine (based on Jensen's Junction Tree Algorithm [2001]). The output was the POSTERIOR distribution of cloud (simplified to true or false) at each grid point and a quantitative assessment of uncertainty. The minimum height of cloud (a value of "true") determined the cloud ceiling.

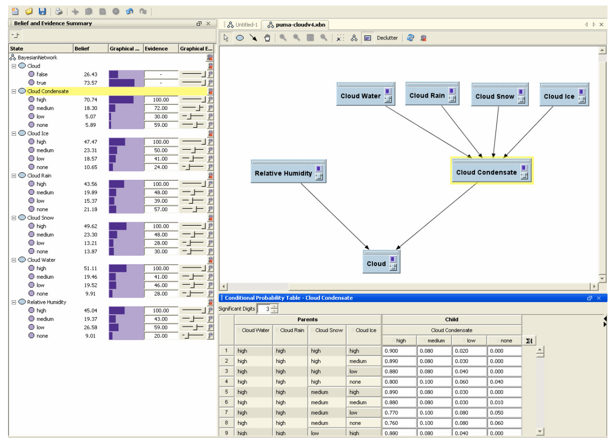

Figure 2. Graphic sample of the BNet.Builder software as applied to the determination of cloud and forecast uncertainty. The Conditional Probability Table in the lower panel of the image is the PRIOR distribution. The values on the left panel to the right of the bar graphs represent the LIKELIHOOD distributions (evidence), and the values to the left of the bar graphs represent the POSTERIOR distribution.

5. COGNITIVE ENGINEERING AND VISUALIZATION

A critical aspect of the approach taken during the PUMAR project was the use of cognitive engineering techniques such as Cognitive Task Analysis (CTA) to aid in the system's design (Shraagen et al. 2000). This approach encourages the design to focus on supporting the user's tasks through relevant information in an appropriate method. Several recent case studies have demonstrated the power of CTA in driving the design of novel visualizations of abstract domain concepts and the subsequent improvement in performance (Roth et al. 2001, Bisantz 2002, Potter et al. 2003). In PUMAR, we used CTA to identify key types and sources of uncertainty information impacting the decision-maker and focused our efforts on quantifying that uncertainty.

An area of active research in the cognitive engineering community is the development of visualizations of uncertainty (Basapur 2003, Finger & Bisantz 2002). Presenting supplemental uncertainty information to the decision-maker can improve their situational awareness, and improve their trust in the display system and underlying predictive models. While some methods for representing uncertainty have been developed, they need to be carefully tied, via CTA techniques, to the specific decision-making domain. In the PUMAR project, we developed specific visualizations for understanding the impact of uncertain weather forecast information on military operations.

6. RESULTS

The benchmark for the results of the PUMAR demonstration was the current capability of the Department of Defense Integrated Weather-Effects Decision Aid (IWEDA) as described by Sauter (2000). The IWEDA-like spatial (not shown) and temporal display (Fig. 3) of information gives the decision-maker an adequate indication of the NWP situation. The often-used red (severe impact), yellow (marginal impact) and green (minimal impact) “stop-light” displays, however, lack the critical meta-information about the underlying uncertainty of the NWP forecasts. The PUMAR demonstration, on the other hand, expanded upon the “stop-light” displays by experimenting with different hues (colors), transitions between hues, and transparency to represent NWP forecast uncertainty (Fig. 4).

The IWEDA-like temporal display (Fig. 3) indicates severe impact to T-UAV operations at the particular location during the early periods of the forecast, and then conditions improve to marginal on the way to minimal impact. The “cause” of the severe or marginal impact can be identified by looking at the impacts of the individual components (IWEDA rules). The “low cloud ceiling” rule fired from the initial time period through 27 forecast hours, as depicted by the yellow “pixels”, indicating the cloud ceiling was below a certain threshold. In addition, the “very low cloud ceiling” rule fired from the initial time period until the 12-hour forecast, as depicted by the red “pixels”, indicating the cloud ceiling was well below operational thresholds. The overall system impact is simply the worst-case component (rule). In this IWEDA-like display the operator is not able to decipher the uncertainty or the magnitude of the impact (e.g. how red is red).

The PUMAR temporal display of the same NWP information (Fig. 4) does allow the operator to assess the uncertainty (either by hue or transparency) of the specific component (rule). In this display the aggregate or system “pixels” are colored (hue) based on the number of component rules that fired, and the transparency is based on the uncertainty. The color (hue) of the cloud ceiling “pixels” are the same as the IWEDA-like display, but the transparency decreased with time indicating the uncertainty in the low and very low cloud ceiling forecasts. Since the PRIOR distribution used in this demonstration did not include temporal variability, the transparency (uncertainty) was likely the result of a higher cloud base forecast. The overall system impacts were marginal in the IWEDA-like display from 21 through 27 forecast hours (Fig. 3), but using the PUMAR display the operator could see that only a single rule fired during this period (the low cloud ceiling rule), and due to the uncertainty of the forecast the overall system impacts are green (actually light green indicating lower confidence in the minimal impact).

A similar discussion is also applicable for the spatial displays of IWEDA-like products (Fig. 5) and the PUMAR display of the same data (Fig. 6). The cloud ceiling severely impacted the T-UAV operations at this particular time, but the operator did not have access to the uncertainty information in Figure 5. If the operator was given the uncertainty information as in Figure 6 the outcome of the decision could be different. The light red or pink “pixels” indicate the very low ceiling rule fired for those locations, but there was a significant level of uncertainty in the forecast. Seeing the light red next to the yellow at several locations, the operator could assume some risk and execute the mission.

The uncertainty information available during the PUMAR project allowed us to experiment with several different NWP display techniques to try to communicate additional forecast intelligence and eliminate information saturation. Using the display techniques demonstrated in the PUMAR study and limiting the display to two variables (for example magnitude and uncertainty) seemed to work best.

7. CONCLUSION

The goal of the Weather Prediction Uncertainty Management and Representation (PUMAR) project was to develop a technique to quantify the uncertainty of numerical weather prediction (NWP) output and to display weather impact information in a cognitively congruent manner for easy interpretation by an operator. We quantified the NWP uncertainty of cloud forecasts using a simplified Bayesian Hierarchical Model (BHM) technique (BNet.Builder). We demonstrated that the existing Integrated Weather-Effects Decision Aid (IWEDA) display could be enhanced by including intelligence about the level of uncertainty of the input information.

For this paper we simplified the distribution of cloud variables by quantizing the cloud constituents and relative humidity. An obvious extension of this work would be to use the complete/continuous distributions of cloud variables and to modify the PRIOR distribution temporally to account for variable forecast skill in the BHM process. In fact, the strength of the BHM process is the ability to account for uncertainty in all stages of the NWP process. The focus of this paper was on the post-processed variables, but the BHM technique could be applied in the data assimilation process as well as exploring the uncertainty of sub-grid scale processes (e.g. turbulent mixing, cloud microphysics, etc.). Similarly, additional uncertainty visualization techniques could be developed for additional types of uncertainty in NWP process. These additional visualizations may require additional cognitive engineering efforts to understand the impact of this information on the decision-making process.

Figure 3: Current IWEDA-like temporal display showing the predicted T-UAV capabilities (top), and the expanded tree view showing the results of the individual rules (bottom) extracted from one particular grid point within the PUMAR domain. The NWP forecasts extended for 36 hours (the intervening hours were persisted for display purposes). The red, yellow and green "pixels" were assigned based on operational T-UAV weather sensitivities.

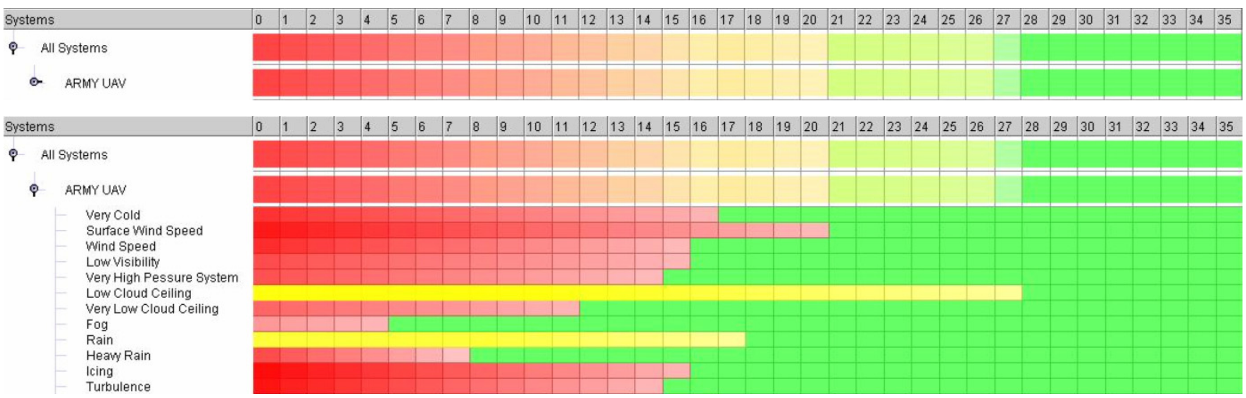

Figure 4: The same display as in Fig. 3, except for the application of cognitive engineering techniques to change the hue (color) and transparency of the mission impacts based on the uncertainty of the predicted NWP variables.

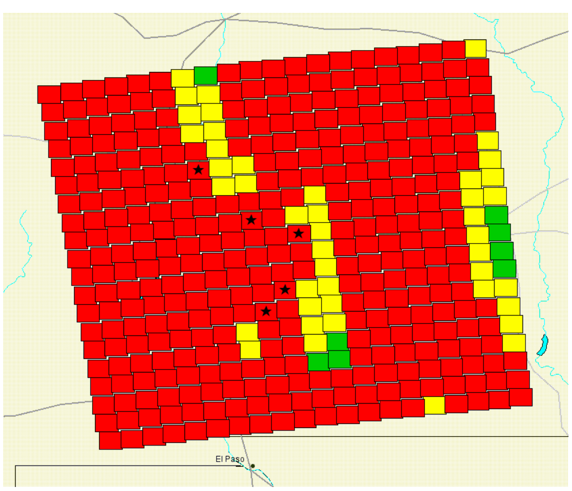

Figure 5: Current IWEDA-like spatial display showing the cloud ceiling at one particular time over the PUMAR domain. The black stars were points of interest during the T-UAV demonstration scenario. The red, yellow, and green "pixels" were assigned based on operational T-UAV cloud ceiling sensitivities.

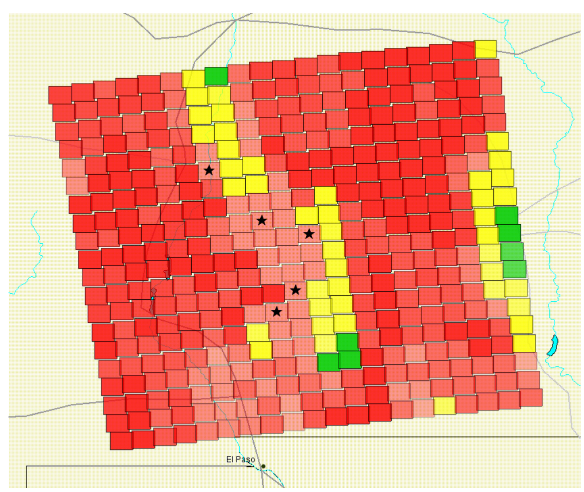

Figure 6: The same display as in Fig. 4, except for the application of cognitive engineering techniques to change the hue (color) and transparency of the cloud ceiling impacts based on the uncertainty of the predicted cloud ceiling.

8. REFERENCES

Basapur, S., A.M. Bisantz, and K. Kesavadas, 2003: The effect of display modality on decision-making in uncertainty, In Proceedings of *Human Factors and Ergonomics Society Annual Meeting.*

Berliner, L.M., J.A. Royle, C.K. Wikle, and R.F. Milliff, 2003: *Bayesian methods in the Atmospheric Sciences*. Bayesian Statistics 6, Oxford University Press, London, **1**, 341-359.

Berliner, L.M., R.F. Milliff, and C.K. Wikle, 2003: Bayesian hieirachical modeling of air-sea interaction, *J. Geophys. Res*., **108**, 3104-3122.

Bisantz, A.M., 2002: *Holistic Battlespace Visualization: Advanced Concepts in Information Visualization and Cognitive Studies.* Univ. Buffalo Technical Report.

Finger, R. and A.M. Bisantz, 2002: Utilizing graphical formats to convey uncertainty in a decision-making task, *Theoretical Issues in Ergonomic Science,* **3**(1), 1-25.

Jensen, F.V., 2001: *Bayesian Networks and Decision Graphs.* Springer-Verlag, New York.

Patterson, E.S., E.M. Roth, and D.D. Woods, 2001: Predicting vulnerability in computer-supported inferential analysis under data overload, *Cognition, Technology and Work*, 3224-3237.

Roth, E.M., L. Lin, S. Kerch, S.J. Kenney, and N. Sugibayashi, 2001: *Designing a first-of-a-kind group view display for team decision making: A case study*. In *Linking Expertise and Naturalistic Decision Making*, Salas, E. and Klein, G. (Eds.). Mahwah, NJ: Lawrence Erlbaum Associates.

Sauter, D., 2000: An interactive information and processing system to assist the military with command and control decision making, Proc. 16th Conf. on Interactive Info. and Proc. Sys for Met., Ocean, and Hydro. Long Beach, Amer. Met. Soc.

Schraagen, J.M., S.F. Chipman, and V.L. Shalin, 2000: *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wikle, C.K., L.M. Berliner, and R.F. Milliff, 2003: Hierarchical Bayesian approach to boundary value problems with stochastic boundary conditions, *Mon. Wea. Rev*., **131**, 1051-1062.

Acknowlegements

This work was performed under Office of the Secretary of Defense Contract number W911QX-04-C-0063 with the U.S. Army Research Laboratory's Computational and Informational Science Directorate. The authors thank Ms. Barb Sauter for her support and direction.