4.2 BAYESIAN PROCESSOR OF OUTPUT: A NEW TECHNIQUE FOR PROBABILISTIC WEATHER FORECASTING

Roman Krzysztofowicz*
University of Virginia, Charlottesville, Virginia

1 INTRODUCTION

Rational decision making by industries, agencies, and the public in anticipation of heavy precipitation, snow storm, flood, or other disruptive weather phenomenon, requires information about the degree of confidence that the user can place in a weather forecast. It is vital, therefore, to advance the meteorologist's capability of quantifying forecast uncertainty to meet the society's rising expectations for reliable information.

The long-term goal of our research is to lay down a methodological foundation for the next generation of probabilistic forecasting systems. The specific objective is to develop and test a coherent set of theoretically-based techniques for probabilistic forecasting of weather variates. The basic technique, called *Bayesian Processor of Output* (BPO), will process output from a numerical weather prediction (NWP) model and optimally fuse it with climatic data in order to quantify uncertainty about a predictand. The extended technique, called *Bayesian Processor of Ensemble* (BPE), will process an ensemble of the NWP model output (or multiple model outputs).

As is well known, Bayes theorem provides the optimal theoretical framework for fusing information from different sources and for obtaining the probability distribution of a predictand, conditional on a realization of predictors, or conditional on an ensemble of realizations. The challenge is to develop and test techniques suitable for operational forecasting. This paper provides an overview of the research plan, the Bayesian theory of forecasting, the development strategy, and the testing strategy.

* *Corresponding author address*: Roman Krzysztofowicz, University of Virginia, P.O. Box 400747, Charlottesville, VA 22904-4747; e-mail: rk@virginia.edu.

2 RESEARCH PLAN

The research will cycle through two phases: the development phase and the testing phase.

2.1 Development Phase

The goal is to develop the BPO (and the BPE) that has a solid theoretical foundation, is suited to all types of predictands, is amenable to supervised-automatic estimation, and is computationally efficient. The development will proceed in three stages.

1. The *theoretic structures* of the BPO will be derived from the laws of probability theory. In particular, the principles of Bayesian forecasting and fusion will be followed (Krzysztofowicz, 1983, 1999; Krzysztofowicz and Long, 1990). There will be three structures, for

- binary predictands (e.g., indicator of precipitation occurrence),
- multi-category predictands (e.g., indicator of precipitation type),
- continuous predictands (e.g., precipitation amount conditional on precipitation occurrence, temperature, visibility, ceiling height, wind speed).

2. Within each theoretic structure, *parametric models* for the needed multivariate and conditional probability distributions will be developed. These developments will be based on the meta-Gaussian family of multivariate distributions (Kelly and Krzysztofowicz, 1994, 1995, 1997), which allows for (i) any form of marginal distributions of predictors and predictand, (ii) a stochastic dependence structure that is pairwise and that for each pair of variates may be nonlinear (in the conditional mean) and heteroscedastic (in the conditional variance), and (iii) an analytic forecasting equation for the probability distribution of the predictand.

3. Given the parametric models, procedures for *parameter estimation, model validation,* and *predic-*

tors selection will be designed.

2.2 Testing Phase

The BPO will be tested by producing *probability of precipitation* (PoP) occurrence, *probabilistic quantitative precipitation forecast* (PQPF), and *quantitative precipitation forecast* (QPF) derived from the PQPF. The primary benchmark for evaluation of the BPO will be the Model Output Statistics (MOS) technique (Glahn and Lowry, 1972) used in operational forecasting by the National Weather Service (NWS). In the currently deployed AVN-MOS system (Antolik, 2000), the predictors for the MOS forecasting equations are based on output fields from the Global Spectral Model run under the code name AVN. The samples for the estimation and verification of the BPO will be retrieved from the same archive that the Meteorological Development Laboratory of the NWS utilized to develop the AVN-MOS system.

A comparative verification of the BPO forecasts and the MOS forecasts will be conducted in three stages. To wit, the implementation of the BPO will progressively diverge from the implementation of the MOS technique in order to evaluate both the *magnitude* and the *source* of the expected improvement of forecasts. There will be three sources (stages) of improvement:

- Improvement due to the theoretic structure of the BPO.
- Improvement due to the use of all climatic data and the optimal fusion of information according to Bayes theorem.
- Improvement due to the optimal selection of predictors according to a criterion of informativeness (or sufficiency, in the sense of Blackwell (1951)).

3 BAYESIAN THEORY

The theoretic structure of the BPO is outlined below for two types of predictands, binary and continuous. Both types of predictands are involved in the production of a PQPF.

3.1 Binary Predictand

3.1.1 Variates and Samples

Let V be the *predictand*, a binary indicator of some future event, such that $V = 1$ if and only if the event occurs, and $V = 0$ otherwise; its realization is denoted v, where $v \in \{0, 1\}$. Let $\mathbf{X} = (X_1, ..., X_I)$ be the vector of I predictors; its realization is denoted $\mathbf{x} = (x_1, ..., x_I)$. Each X_i $(i = 1, ..., I)$ is assumed to be a continuous variate — an assumption that simplifies the presentation but can be relaxed if necessary.

Suppose the forecasting problem has already been structured, and the task is to develop the forecasting equation in a setup similar to that of the MOS technique (Antolik, 2000). For example, let the event be the occurrence of precipitation (practically, the accumulation of at least 0.254 mm of water) at a station during the 12-h period beginning 12 h after the run of the NWP model, at 0000 UTC, in August. Let $\{v\}$ denote the climatic sample of the predictand. Say, a homogeneous record of precipitation observations over 10 years exists, so that one can extract a sample of size J_1 = 10 years × 31 days = 310 realizations. Let $\{(\mathbf{x}, v)\}$ denote the sample of joint realizations of the predictor vector and the predictand. Say, a homogeneous record of the NWP model output over the last 4 years exists, so that one can form a sample of size J_2 = 4 years × 31 days = 124 joint realizations.

The point of the above example is that typically the *joint sample* is much shorter than the *climatic sample*: $J_2 << J_1$. Classical statistical methods, such as the MOS technique, deal with this *sample asymmetry* by simply ignoring the long climatic sample. In effect, these methods throw away vast amount of information about the predictand. In contrast, the BPO will use effectively both samples; it will extract information from each sample and then optimally fuse information according to the laws of probability. (Pooling of samples from different months and stations in order to increase the sample size is a separate issue.)

3.1.2 Inputs

With P denoting the probability and p denoting a generic density function, define the following objects.

$g = P(V = 1)$ is the *prior probability* of event $V = 1$; it is to be estimated from the climatic sample $\{v\}$. Probability g quantifies the uncertainty about the predictand V that exists before the NWP model output is available. Equivalently, it characterizes the natural variability of the predictand.

$f_v(\mathbf{x}) = p(\mathbf{x}|V = v)$ for $v = 0, 1$; function f_v is the I-variate density function of the predictor vector $\mathbf{X}$, conditional on the hypothesis that the event is $V = v$. The two conditional density functions, f_0 and f_1, are to be estimated from the joint sample $\{(\mathbf{x}, v)\}$. For a fixed realization $\mathbf{X} = \mathbf{x}$, object $f_v(\mathbf{x})$ is the *likelihood* of event $V = v$. Thus (f_0, f_1) comprises the *family of likelihood functions*. This family quantifies the stochastic dependence between the predictor vector $\mathbf{X}$ and the predictand V. Equivalently, it characterizes the informativeness of the predictors with respect to the predictand.

3.1.3 Theoretic Structure

The probability g and the family of likelihood functions (f_0, f_1) carry information about the prior uncertainty and the informativeness of the predictors into the Bayesian revision procedure. The *predictive density function* κ of the predictor vector $\mathbf{X}$ is given by the total probability law:

$$\kappa(\mathbf{x}) = f_0(\mathbf{x})(1-g) + f_1(\mathbf{x})g, \tag{1}$$

and the *posterior probability* $\pi = P(V = 1|\mathbf{X} = \mathbf{x})$ of event $V = 1$, conditional on a realization of the predictor vector $\mathbf{X} = \mathbf{x}$, is given by Bayes theorem:

$$\pi = \frac{f_1(\mathbf{x})g}{\kappa(\mathbf{x})}. \tag{2}$$

By inserting (1) into (2), one obtains an alternative expression:

$$\pi = \left[1 + \frac{1-g}{g}\frac{f_0(\mathbf{x})}{f_1(\mathbf{x})}\right]^{-1}, \tag{3}$$

where $(1-g)/g$ is the *prior odds* against event $V = 1$, and $f_0(\mathbf{x})/f_1(\mathbf{x})$ is the *likelihood ratio* against event $V = 1$. Equation (3) defines the theoretic structure of the BPO for a binary predictand.

3.2 Multi-Category Predictand

The BPO for a multi-category predictand will be a generalization of the BPO for a binary predictand. It will be tested for forecasts of the precipitation type that are of import to the transportation industry during the winter but that are currently rather poor for some regions (e.g., east of the Blue Ridge Mountains in Virginia).

3.3 Continuous Predictand

3.3.1 Variates and Samples

Let W be the *predictand*, which is a continuous variate; its realization is denoted ω. Let $\mathbf{X}$ be the vector of I continuous *predictors*; its realization is denoted $\mathbf{x}$.

As before, suppose the forecasting problem has already been structured, and the task is to develop the forecasting equation. Let $\{\omega\}$ denote the climatic sample of the predictand, and let $\{(\mathbf{x}, \omega)\}$ denote the sample of joint realizations of the predictor vector and the predictand.

3.3.2 Inputs

$g(\omega) = p(\omega)$ is the *prior density function* of the predictand W; it is to be estimated from the climatic sample $\{\omega\}$. The corresponding distribution function is denoted G.

$f(\mathbf{x}|\omega) = p(\mathbf{x}|W = \omega)$; function $f(\cdot|\omega)$ is the I-variate density function of the predictor vector $\mathbf{X}$, conditional on the hypothesis that the actual value of the predictand is $W = \omega$. The family $\{f(\cdot|\omega) : \textit{all}\ \omega\}$ of the density functions of $\mathbf{X}$ is to be estimated from the joint sample $\{(\mathbf{x}, \omega)\}$. For a fixed realization $\mathbf{X} = \mathbf{x}$, object $f(\mathbf{x}|\cdot)$ is the *likelihood function* of W. Symbol f will denote the *family of likelihood functions*.

3.3.3 Theoretic Structure

The density function g and the family of likelihood functions f carry information about the prior uncertainty and the informativeness of the predictors into the Bayesian revision procedure. The *predictive density function* κ of the predictor vector $\mathbf{X}$ is given by the total probability law:

$$\kappa(\mathbf{x}) = \int_{-\infty}^{\infty} f(\mathbf{x}|\omega)g(\omega)\, d\omega, \tag{4}$$

and the *posterior density function* $\phi(\omega) = p(\omega|\mathbf{X} = \mathbf{x})$ of predictand W, conditional on a realization of the predictor vector $\mathbf{X} = \mathbf{x}$, is given by Bayes theorem:

$$\phi(\omega) = \frac{f(\mathbf{x}|\omega)g(\omega)}{\kappa(\mathbf{x})}. \tag{5}$$

The corresponding *posterior distribution function* $\Phi(\omega) = P(W \le \omega|\mathbf{X} = \mathbf{x})$ is given by

$$\Phi(\omega) = \frac{1}{\kappa(\mathbf{x})}\int_{-\infty}^{\omega} f(\mathbf{x}|u)g(u)\, du. \tag{6}$$

Equations (4)–(6) define the theoretic structure of the BPO for a continuous predictand.

4 DEVELOPMENT

4.1 Modeling Issues

To operationalize a BPO, it is necessary to formulate a model for the family of likelihood functions and to derive a method (analytic and/or numerical) for the effective calculation of the posterior probability π or the posterior distribution function Φ. With regard to modeling, there are two key issues.

First, among the most informative predictors of precipitation occurrence and precipitation amount is an estimate of the total precipitation amount output from the NWP model. Typically, this estimate takes on value zero on some days, and positive values on other days. Thus it should be modeled as a binary-continuous variate. This adds complexity to the structure of the conditional density functions, f_v for $v = 0, 1$, and $f(\cdot|\omega)$ for all w, because each of them becomes a two-component mixture.

Second, a flexible and convenient model is needed for the multivariate conditional density functions of the (absolutely) continuous predictor vectors. For that purpose we employ the meta-Gaussian model developed by Kelly and Krzysztofowicz (1994, 1995, 1997) and applied successfully to probabilistic river stage forecasting (Krzysztofowicz and Herr, 2001; Krzysztofowicz, 2002).

4.2 Meta-Gaussian Model

The meta-Gaussian model belongs to the class of multivariate distributions constructed from marginals. It offers these properties.

1. The marginal conditional distribution function of each predictor X_i can take any form; it can be parametric or nonparametric.

2. The transform for each predictor X_i is uniquely specified once its marginal distribution function has been estimated.

3. The conditional dependence structure among the predictors $X_1, ..., X_I$ is pairwise; the degree of dependence is quantified by the conditional correlation matrix.

4. The conditional dependence structure between any two predictors X_i and X_j, $i \neq j$, can be nonlinear and heteroscedastic.

5. The probabilistic forecast (the posterior probability π, the posterior density function ϕ, or the posterior distribution function Φ) is specified by an analytic expression.

Properties 1 and 4 imply the flexibility in fitting the model to data. Properties 2 and 3 imply the simplicity of estimation. Property 5 implies the computational efficiency — an important attribute for operational forecasting.

5 TESTING

5.1 PQPF Production

Comprehensive testing and evaluation will be performed of the BPO for two kinds of forecasts: PoP (binary predictand) and PQPF (binary-continuous predictand).

The PoP will be produced from (3). The PQPF that meets requirements of hydrological models (Krzysztofowicz, 2002) will be produced from (3) and (6) as follows. With W defined as the precipitation amount accumulated at a station (or at a grid point) during a specified period, (3) will produce a probability

$$\pi = P(W > 0|\mathbf{X} = \mathbf{x}), \tag{7}$$

and (6) will produce a conditional distribution function

$$\Phi(\omega) = P(W \leq w|\mathbf{X} = \mathbf{x}, W > 0), \quad w \geq 0, \tag{8}$$

such that $\Phi(\omega) > 0$ if $w > 0$, and $\Phi(\omega) = 0$ if $w = 0$. Then the PQPF will be specified in terms of the (unconditional) distribution function of W:

$$P(W \leq w|\mathbf{X} = \mathbf{x}) = \pi\Phi(\omega) + (1 - \pi), \quad w \geq 0. \tag{9}$$

This distribution function will be continuous because Φ will be continuous.

The evaluation will include a *comparative verification* of the BPO forecasts and the MOS forecasts. Two basic questions will be asked during the evaluation: (i) Does the BPO produce better probabilistic forecasts than the MOS technique does? (ii) Does the BPO overcome the statistical deficiencies and the operational limitations of the MOS technique?

5.2 Verification Measures

Verification of forecasts will be performed according to the Bayesian decision theory. This theory prescribes two verification measures: a measure of *calibration* and a measure of *informativeness* (DeGroot and Fienberg, 1982, 1983; Vardeman and Meeden, 1983; Krzysztofowicz, 1996). These verification measures characterize the necessary and sufficient attributes of forecast from the viewpoint of a rational decision maker. Both measures are based on the family of likelihood functions which characterizes the stochastic dependence between the forecast and the predictand, and which must be estimated from a joint sample of realizations.

For the PoP, the primary verification measures will be the *calibration function* (Krzysztofowicz and Long, 1991) and the *receiver operating characteristic* (Krzysztofowicz and Long, 1990). For the PQPF, the primary verification measures will be the *calibration score* (Krzysztofowicz and Sigrest, 1999) and the *informativeness score* (an extension of the Bayesian correlation score of Krzysztofowicz (1992)).

6 CLOSURE

In a nutshell, the research program outlined herein is expected to advance (i) the science of the NWP model output processing (including ensemble processing), and (ii) the capability of forecasting weather probabilistically.

A coherent set of Bayesian techniques will harness recent advances in Bayesian statistical theory, multivariate distributions, estimation methods, and ensemble forecasting. For each of the three types of predictands (binary, multi-category, continuous), there will be:

- the BPO, a technique for processing output from a NWP model into a probabilistic forecast;
- the BPE, a technique for processing an ensemble of outputs from a NWP model into a probabilistic forecast; and
- the Bayesian verification measures for probabilistic forecasts produced via either BPO or BPE.

In addition, this research program is expected to advance the capability of producing

- the PQPF, which is based on either a single output or an ensemble of outputs from a NWP model, and which meets requirements of hydrological forecasting.

Acknowledgments. This material is based upon work supported by the National Science Foundation under Grant No. ATM-0135940, "New Statistical Techniques for Probabilistic Weather Forecasting".

References

Antolik, M.S., 2000: An overview of the National Weather Service's centralized statistical quantitative precipitation forecasts. *Journal of Hydrology,* **239**(1–4), 306–337.

Blackwell, D., 1951: Comparison of experiments. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, J. Neyman (ed.), University of California Press, Berkeley, pp. 93–102.

DeGroot, M.H., and S.E. Fienberg, 1982: Assessing probability assessors: Calibration and refinement. In *Statistical Decision Theory and Related Topics,* III, J.O. Berger, and S.S. Gupta, (eds.), Academic Press, New York, vol. 1, 291–314.

DeGroot, M.H., and S.E. Fienberg, 1983: The comparison and evaluation of forecasters. *The Statistician*, **32**, 12–22.

Glahn, H.R., and D.A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology,* **11**(8), 1203–1211.

Kelly, K.S., and R. Krzysztofowicz, 1994: Probability distributions for flood warning systems. *Water Resources Research,* **30**(4), 1145–1152.

Kelly, K.S., and R. Krzysztofowicz, 1995: Bayesian revision of an arbitrary prior density. *Proceedings of the Section on Bayesian Statistical Science,* American Statistical Association, 50–53.

Kelly, K.S., and R. Krzysztofowicz, 1997: A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrology and Hydraulics,* **11**(1), 17–31.

Krzysztofowicz, R., 1983: Why should a forecaster and a decision maker use Bayes theorem. *Water Resources Research,* **19**(2), 327–336.

Krzysztofowicz, R., 1992: Bayesian correlation score: A utilitarian measure of forecast skill. *Monthly Weather Review,* **120**(1), 208–219.

Krzysztofowicz, R., 1996: Sufficiency, informativeness, and value of forecasts. *Proceedings,* Workshop on the Evaluation of Space Weather Forecasts, Space Environment Center, NOAA, Boulder, Colorado, 103–112.

Krzysztofowicz, R., 1999: Bayesian forecasting via deterministic model. *Risk Analysis,* **19**(4), 739–749.

Krzysztofowicz, R., 2002: Bayesian system for probabilistic river stage forecasting. *Journal of Hydrology,* **268**(1–4), 16–40.

Krzysztofowicz, R., and H.D. Herr, 2001: Hydrologic uncertainty processor for probabilistic river stage forecasting: Precipitation-dependent model. *Journal of Hydrology,* **249**(1–4), 46–68.

Krzysztofowicz, R., and D. Long, 1990: Fusion of detection probabilities and comparison of multisensor systems. *IEEE Transactions on Systems, Man, and Cybernetics,* **20**(3), 665–677.

Krzysztofowicz, R., and D. Long, 1991: Forecast sufficiency characteristic: Construction and application. *International Journal of Forecasting,* **7**(1), 39–45.

Krzysztofowicz, R., and A.A. Sigrest, 1999: Calibration of probabilistic quantitative precipitation forecasts. *Weather and Forecasting,* **14**(3), 427–442.

Vardeman, S., and G. Meeden, 1983: Calibration, sufficiency, and domination considerations for Bayesian probability assessors. *Journal of the American Statistical Association,* **78**(384), 808–816.