NOTES AND CORRESPONDENCE

On Predicting Probability Distributions of Time-Averaged Meteorological Data

ROLAND A. MADDEN

National Center for Atmospheric Research, Boulder, Colorado*

3 October 1988 and 14 December 1988

1. Introduction

Leith (1973) offered some simple statistical arguments facilitating a quantitative assessment of the differences between climate model runs and of climate change itself. His suggestion of considering model, or real climate, time series as made up of "signal" and a constant "noise," makes it possible to easily determine conditional probability distributions as a function of the signal-to-noise ratio (his Fig. 4). It is these probabilities that are important for making practical decisions (e.g., Gilman 1985; Brown et al. 1986). This note complements Leith's by making first order estimates of how often specific probability distributions might be predicted. It is an extension of an aspect of a paper on a quantitative approach to long-range prediction (Madden 1981, hereafter M).

By assuming that an observed time series is made up of predictable signal and constant unpredictable noise, and that the signal can be predicted through a simple linear regression, the frequency with which certain probability distributions will be predicted can be determined exactly. We can use these exact results as first order approximations for the frequency with which we can predict certain probability distributions using more general prediction procedures. This frequency has been referred to as the predictive distribution (Brown et al. 1986). Also, considering estimates of "potential predictability" (Madden 1976; Madden and Shea 1978; M), we can get some idea of the best we can hope to do in predicting conditional probability distributions. "Best" implies the ability to frequently predict conditional probability distributions that are considerably different from the climatological distribution.

2. Example of a simple linear regression prediction

Figure 1 is taken from M. It is a scatter diagram of December–February (DJF) New Orleans temperatures versus June–August (JJA) Darwin pressures. In M it was assumed that the resulting linear regression equation

$$\hat{T} = 807.2 - 0.7841P \qquad (1)$$

could be used to predict DJF New Orleans temperature ($\hat{T}$ in degrees C) from the previous JJA Darwin pressures (P in mb). Here $\hat{T}$ is the expected value of the temperature given P. The physical connection was unspecified, but the correlation was shown to occur on time scales usually associated with the Southern Oscillation. The correlation between these data is −0.31, so that 9% of the total variance of DJF New Orleans temperatures can be explained by JJA Darwin pressures.

It was further assumed in M that both predictor (JJA Darwin pressures) and predictand (DJF New Orleans pressures) are normally distributed. The distributions on the right-hand side of Fig. 1 give an indication of the predictand's expected distribution given an observed average value in the predictor (solid); the dotted curves show the predictand's expected distribution given extremely low or high values of the predictor. The areas under the three curves give an idea of the resulting shifts in three-category probabilities [above (A), normal (N), below (B)]. The signal is defined as the shift in the mean of the predictand and the noise as the standard deviation associated with the unpredictable variations. In the case of New Orleans temperatures predicted by Darwin pressures the unpredictable temperature variations are 91% of the total variance.

To bring out the signal-to-noise ratio quantitatively, consider a predictor x and a predictand $\hat{y}$, which is the conditional expected value of y given x, that is,

$$\hat{y} = a + bx. \qquad (2)$$

The slope $b = r_{y,x}\sigma_y/\sigma_x$, where $r_{y,x}$ is the correlation coefficient between y and x and σ_x and σ_y are the standard deviations of x and y. Similarly,

$$\bar{y} = a + b\bar{x}, \qquad (3)$$

where $\bar{y}$ and $\bar{x}$ are the average values of y and x. Subtracting (3) from (2) and substituting for b gives one equation for the signal in the predictand:

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author address: Dr. Roland Madden, NCAR, P.O. Box 3000, Boulder, CO 80307-3000.

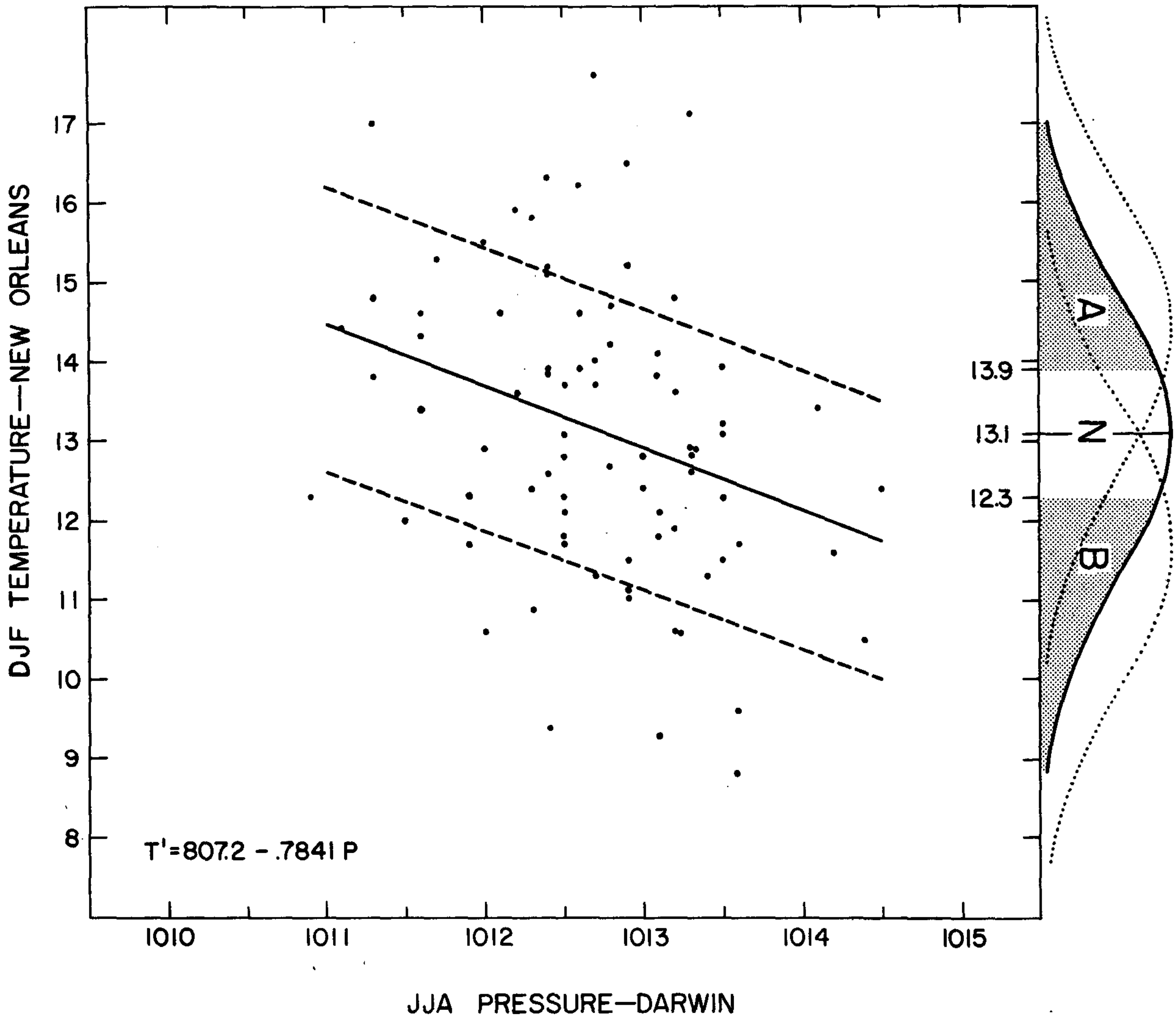

FIG. 1. Scatter diagram of JJA-averaged Darwin pressures and the following DJF-averaged New Orleans temperatures. Correlation between these data is −0.31. Sloping solid line is a regression line given by the equation in lower left. Sloping dashed lines indicate plus or minus one standard deviation (±1.8°C). Long-term DJF average temperature at New Orleans is 13.1°C. There is a 1/3 probability that the actual DJF temperature will be <12.3°C (B), between 12.3° and 13.9°C (N), or >13.9°C (A) given a previous JJA pressure at Darwin equal to the long-term mean (1012.7 mbar). Predicted distribution for this case is indicated by solid, normal curve and regions marked B, N, and A along the right-hand side. Shifted distributions predicted by extremely low or high JJA Darwin pressures are indicated by the dotted curves.

$$\hat{y} - \bar{y} = r_{y,x} \frac{\sigma_y}{\sigma_x} (x - \bar{x}). \qquad (4)$$

The noise in the predictand is the standard deviation of the remaining unpredictable variations, or $\sigma_y(1 - r_{y,x}^2)^{1/2}$. The signal-to-noise ratio is then

$$\frac{\hat{y} - \bar{y}}{\sigma_y(1 - r_{y,x}^2)^{1/2}} = \frac{r_{y,x}}{(1 - r_{y,x}^2)^{1/2}} \frac{x - \bar{x}}{\sigma_x}. \qquad (5)$$

Equation (5) gives the predicted signal-to-noise ratio (left-hand side) for an observed standardized deviation away from the mean of the predictor $[(x - \bar{x})/\sigma_x]$.

3. Frequency of occurrence of given signal-to-noise ratios

Figure 2 shows the predicted signal-to-noise ratio as a function of observed standardized deviations from the mean of the predictor, assuming $|r_{y,x}| = 0.31$, the case depicted in Fig. 1. For example, if JJA Darwin pressure was observed to be one standard deviation above or below its average value, the signal-to-noise ratio for the following DJF New Orleans temperature would be slightly larger than 0.31. From the normal distribution, a signal-to-noise ratio of 0.31 would change two-category probabilities from a 50% chance

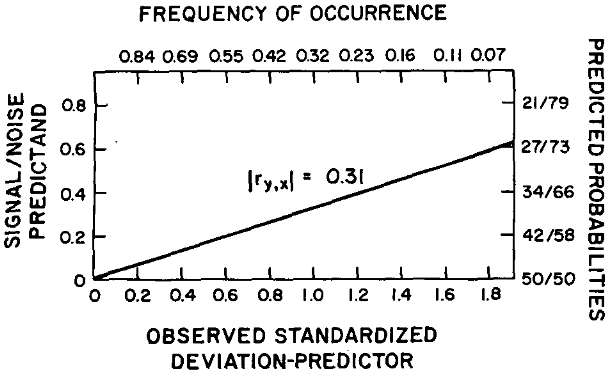

FIG. 2. Predictand signal-to-noise ratio, or number of standard deviations that predictand is predicted to deviate from its mean (left-hand ordinate) for a correlation of |0.31| as a function of the number of standard deviations that the predictor is observed to deviate from its mean (bottom abscissa). Right-hand ordinate is labeled with the predicted, two-category probabilities associated with the signal-to-noise ratios on the left. Labels across the top indicate the frequency that standardized deviations in the predictor indicated at the bottom will occur. It is assumed that both predictor and predictand are normally distributed.

of being below or above the average value to a 62% chance of being below or above the average. If the JJA Darwin pressures are normally distributed, then we can determine how often we might predict a 62% or greater chance that the DJF New Orleans temperatures will be below or above average. Again, from the normal distribution the Darwin pressures will exceed plus or minus one standard deviation from average about 32% of the time. The upper scale on Fig. 2 is the frequency of occurrence of observed standardized deviations in the predictor.

The right-hand scale of Fig. 2 indicates two-category probabilities associated with the left-hand signal-to-noise ratios. With a correlation of .31 we can explain 9% of the variance of the predictand, and, as a result, could predict a probability 42/58 (favoring above or below average depending on the sign of the signal) about 55% of the time. If, to be of value we must predict shifts of at least 30/70, then we can only expect to do that about 11% of the time.

Of course if we can explain more of the variance through a larger correlation, then prediction of large shifts in probability will be more frequent. Figure 3 extends results for other correlations. Rather than label lines for varying $r_{y,x}$'s, they are labeled in percent variance predictable ($100 \times r_{y,x}^2$). If we can explain none of the variance ($r_{y,x} = 0$), then we must always predict a 50/50 chance of being above or below normal. In the other extreme, if we were to be able to predict 99% of the variance then we would be able to predict probabilities of 10/90 or more extreme about 90% of the time. Table 1 is included to facilitate easy conversion of the abscissa to three category probabilities.

4. Implications for long-range prediction

While Fig. 3 is exact only for the special case of a linear relationship between normally distributed predictands and predictors, it can be used as a first order approximation for more general prediction methods. For example, operational experience of the U.S. National Weather Service (NWS) suggests that their long-range forecasts explain less than 2% of the variance of temperature when considering all of the contiguous United States (Epstein 1988). Predicting only 2% of variance allows predicting shifts in probability from 50/50 to 40/60 about 10% of the time (Fig. 3).

NWS skill in predicting temperature is a function of location and season (see maps of skill prepared by D. Gilman in Kalnay and Livezey 1985), and at specific locations and during specific seasons current capabilities exceed the 2% value in Fig. 3. The actual variance that NWS predicts at individual locations is not readily available, but consideration of lag correlations between months and seasons (e.g., van Loon and Jenne 1975; Namias 1978; van den Dool et al. 1986) and other forecasting techniques (e.g., Barnett 1981) suggest that at a few places we might currently predict about 10%–25% of the variance depending on the time of year. Development work underway at the NWS that combines analog and persistence information in a bivariate linear regression is able to explain 10%–30% of the variance of winter temperatures in

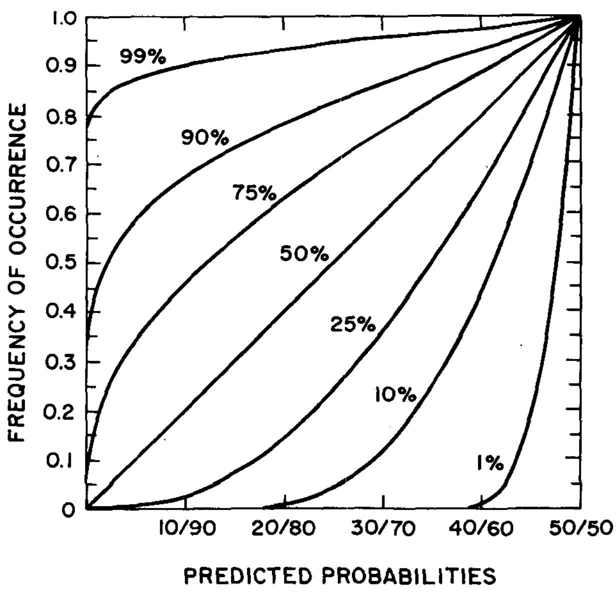

FIG. 3. Frequency with which one can expect to predict various two-category probabilities when the prediction scheme allows indicated percent of total variance of the predictand to be predicted, and both predictand and predictor are normally distributed. For example, if a prediction scheme can predict 10% of the variance, then it would allow predicting two-category probabilities of 30/70 or more extreme a little more than 10% of the time.

TABLE 1. Conversion between two- and three-category probability distributions.

Signal/noise	0	.25	.52	.84	1.28
Two-category probabilities	50/50	40/60	30/70	20/80	10/90
Three-category probabilities	33/33/33	24/34/42	17/24/54	10/24/66	4/16/80

parts of the eastern United States (R. Livezey, personal communication).

As a result, we take the area between the 10% and 25% lines on Fig. 3 to be a first order approximation to current capabilities in forecasting monthly or seasonal mean temperatures at a few favorable locations. Studies of potential long-range predictability of temperature indicate that, because of the restrictions imposed by sampling variability of finite time-averaged weather fluctuations, the best we could ever predict is 50%–60% of the variance in limited regions (Madden and Shea 1978; Madden 1981). That means the 50% line in Fig. 3 represents an approximate upper limit to the performance of future prediction schemes even in the most favored locations and during the most favored seasons.

REFERENCES

Barnett, T. P., 1981: Statistical prediction of North American air temperatures from Pacific predictors. *Mon. Wea. Rev.*, **109**, 1021–1041.

Brown, B. G., R. W. Katz and A. H. Murphy, 1986: On the economic value of seasonal-precipitation forecasts: The fallowing/planting problem. *Bull. Amer. Meteor. Soc.*, **67**, 833–841.

Epstein, E. S., 1988: Long-range weather prediction limits of predictability and beyond. *Wea. Forecasting*, **3**, 69–75.

Gilman, D. L., 1985: Long-range forecasting: The present and the future. *Bull. Amer. Meteor. Soc.*, **66**, 159–164.

Kalnay, E., and R. Livezey, 1985: Weather predictability beyond a week: An introductory review. *Proceedings of the International School of Physics, Course LXXXVIII*, M. Ghil, R. Benzi and G. Parisi, Eds., Elsevier, 311–346.

Leith, C. E., 1973: The standard error of time-averaged estimates of climate means. *J. Appl. Meteor.*, **12**, 1066–1069.

Madden, R. A., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Mon. Wea. Rev.*, **104**, 942–952.

——, and D. J. Shea, 1978: Estimates of the natural variability of time-averaged temperature over the United States. *Mon. Wea. Rev.*, **106**, 1695–1703.

——, 1981: A quantitative approach to long-range prediction. *J. Geophys. Res.*, **86**, 9817–9825.

Namias, J., 1978: Persistence of U.S. seasonal temperatures up to one year. *Mon. Wea. Rev.*, **106**, 1557–1567.

Van den Dool, H. M., W. H. Klein and J. Walsh, 1986: The geographical distribution and seasonality of persistence in monthly mean air temperature over the United States. *Mon. Wea. Rev.*, **114**, 546–560.

van Loon, H., and R. L. Jenne, 1975: Estimates of seasonal mean temperature, using persistence between seasons. *Mon. Wea. Rev.*, **103**, 1121–1128.