Reply

C. A. DOSWELL III, R. DAVIES-JONES AND D. L. KELLER*
NOAA/ERL National Severe Storms Laboratory, Norman, Oklahoma
21 March 1991

Generally speaking, we agree with most of Allan Murphy's comments and welcome the opportunity to clarify the points in question. Indeed, some of the things we said are subject to misinterpretation and Murphy's comments offer us the chance to try to prevent readers from drawing inappropriate conclusions.

Our statement about the "illusory" distinction between categorical (dichotomous) and probabilistic (polychotomous) forecasts should not be taken too literally. Our intent, albeit imperfectly accomplished, was to suggest that it is possible to adapt forecast methods designed for dichotomous forecasts to polychotomous forecasts, and vice-versa. This is not always commonly recognized and when forecasts of one type are being verified, that verification is not restricted only to the verification measures associated directly with that type. We agree fully with Murphy's caveat that potentially useful information about polychotomous forecasts will be ignored when using methods designed for dichotomous forecasts. However, it seems useful for comparison purposes to evaluate forecasts both ways, so we wanted to emphasize that it could, in fact, be done. At the most elementary level, our idea is that dichotomous forecasts really are polychotomous forecasts using only two categories (the minimum number).

The same sort of possible misinterpretation is at the root of Murphy's concern for our comparison between the dichotomous and polychotomous forecasts via the linear regression model. The intent of our comparison was to reveal the differences between the implicit, subjective "thresholds" used by different forecasters (and, possibly, at different times by the same forecaster). Indeed, there was no effort to derive the dichotomous forecasts from the probabilities during the experiment—had we done so using some prespecified procedure, then "consistency" would have been assured. If forecasters are free to make subjective thresholding decisions from one forecast to another, without any guidelines about the procedure, then the "consistency" between their dichotomous and polychotomous forecasts should be of some interest to Murphy and others, as it was to us. As before, the idea was that had we been unable to use dichotomous verification measures for polychotomous forecasts, it would have been difficult to make the comparisons we made. There may well be better ways to do this comparison and we welcome any suggestions that Murphy might make; we do not consider ourselves to be verification "experts." Again, we agree that our words may not have expressed our ideas as well as we wished.

Concerning the problems Murphy has mentioned about our attempts at "reducing the dimensionality" of verification problems, we concur once again with Murphy's concerns about this. However, he has credited us with far more than we deserve; what we wanted was to produce comparable measures for both the 2×2 and the $k \times k$ problems. Our goal certainly was not to reduce the problem's dimensionality. The fact that such a reduction may involve loss of information should be fairly obvious, since we certainly recognized it, but that fact deserves being brought to the attention of potential users.

Murphy's final general comment about our generalization of the TSS seems to represent no more than a slight amplification of the points we made about it in our paper. In fact, we did note that the TSS and S measures were quite closely related, with only a subtle difference involving how the expected values for the contingency table are derived. Thus, there doesn't appear to be any important contention between Murphy and us about this point.

Regarding Murphy's enumerated "specific comments," these also seem to require little or no response from us. Basically, we agree fully with Murphy's contributions and clarifications. The only thing we can add concerns his specific comment #5: we had some initial difficulty with Murphy and Daan's (1985) use of "proper" in the quoted context. This was cleared up in informal discussions with Murphy prior to publication of our paper, but we still had some trouble crafting the verbiage to reflect the correct interpretation. We hope this open discussion with Murphy makes it clear once and (literally) for all what the intentions of Murphy and Daan were in that particular passage.

* Also affiliated with the Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma.

Corresponding author address: Dr. Charles A. Doswell III, National Severe Storms Laboratory, Norman, OK 73069

REFERENCE

Murphy, A. H., and H. Daan, 1985: Forecast evaluation. *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A. H. Murphy and R. W. Katz, Eds., Westview Press, 379–437.