U.S. DEPARTMENT OF COMMERCE
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NATIONAL WEATHER SERVICE
OFFICE OF SYSTEMS DEVELOPMENT
TECHNIQUES DEVELOPMENT LABORATORY

TDL OFFICE NOTE 98-2

A PROPOSAL FOR A NEW GENERATION OF CENTRALIZED STATISTICAL QUANTITATIVE PRECIPITATION FORECAST GUIDANCE PRODUCTS

Mark S. Antolik

February 1998

A PROPOSAL FOR A NEW GENERATION OF CENTRALIZED STATISTICAL QUANTITATIVE PRECIPITATION FORECAST GUIDANCE PRODUCTS

Mark S. Antolik

1. INTRODUCTION

At least partly in response to the scientific recommendations of the United States Weather Research Program, the National Weather Service (NWS) Office of Meteorology (OM) has placed higher priority on the study of atmospheric processes related to precipitation production in order to effect improvements in quantitative precipitation forecasts (QPF). Concurrent with this effort, greater emphasis also will be placed on the enhancement of related technologies and support and service functions (National Weather Service 1996). Work at the Techniques Development Laboratory (TDL) consistently has shown that interpretive statistical guidance can enhance the utility of information already available from numerical weather prediction (NWP) models, and precipitation prediction has proven to be no exception. In recent studies, TDL statistical QPF guidance utilizing output from the Nested Grid Model (NGM; Hoke *et al.* 1989) has been shown to substantially improve upon the performance of the gridpoint precipitation forecasts available directly from the model. Specifically, the current guidance system based on the Model Output Statistics (MOS; Glahn and Lowry 1972) technique exhibited superior Critical Success Index scores when compared to the raw NGM gridpoint precipitation for all forecast projections and accumulated amounts (Antolik 1995, 1996). These performance characteristics were apparent during both the warm and cool seasons of a year-long verification of operational forecasts. Clearly, MOS-based QPF systems can provide useful support to forecasters wishing to evaluate the information supplied by NWP models. The added value inherent in MOS QPF can, if properly used, be a major contributor to a guidance package which enables the operational meteorologist to improve precipitation forecasts.

Presently, however, the output from the NGM MOS QPF system is disseminated to the forecaster in categorical form only, even though probabilistic forecasts are its principal components. During the past few years, however, the field has begun to move toward the production of probabilistic QPF (PQPF), driven at least partly by lessons learned from the Midwest floods of 1993 (U.S. Department of Commerce 1994). Probabilistic forecasts of precipitation amounts produced in the modernized NWS Forecast Office (WFO) would, in theory, be used to facilitate probabilistic river stage forecasting at the River Forecast Centers (RFC). Probabilistic information should be more advantageous to individual users of river forecasts than deterministic stage forecasts, enabling users to take appropriate actions according to their own unique cost/benefit analyses. The Eastern Region of the NWS, in cooperation with the University of Virginia, has taken the lead in the development of a pilot PQPF program at the Pittsburgh WFO and the Ohio RFC starting in late 1990 (Krzysztofowicz *et al.* 1993). More recently, the NWS Hydrologic Information Working Group (HIWG) has recommended that this pilot program be expanded into a complete risk reduction exercise prior to the phase-in of PQPF programs at all WFO's and RFC's (National Weather Service 1997). While the forecasters involved in the Pittsburgh project eventually were able to produce reasonably skillful subjective PQPFs, at least during the cool season (Krzysztofowicz and Drake 1992), they lacked objective probabilistic guidance which was specifically tailored to their needs. As these probabilistic forecast procedures are expanded into other WFOs, TDL stands in a unique position to provide meaningful PQPF guidance which will facilitate these new forecast tasks.

Since the current MOS QPF guidance was not intended originally to support a PQPF methodology, the format and content of the guidance as now disseminated are probably not well suited to the anticipated needs of such a program. Thus, we see any proposed TDL response to the problem of providing synoptic-scale statistical guidance to this effort as twofold: (1) adaptation of current NGM MOS QPF guidance to a format more compatible with the proposed NWS PQPF initiative as risk reduction exercises commence; and (2) development of new products which are mathematically and scientifically more appropriate to the long-term needs of the PQPF program. Given the immediacy of NWS needs associated with PQPF risk reduction, we propose to address the former largely by post-processing current NGM MOS QPF guidance. For the latter, we will proceed with research and development of new MOS products which rigorously satisfy the requirements of the new forecast methodology. To be sure, any short-term solutions arrived at in support of the NWS pilot PQPF program will require further evaluation and modification before becoming a standard part of future MOS guidance. However, we fully expect that the lessons learned from providing support to the NWS pilot program will guide our efforts to develop the next generation of statistical QPF guidance based on the high-resolution NWP models of the future. The remainder of this document addresses specific steps which TDL plans to take to address the immediate and long-term requirements for centralized statistical guidance associated with a comprehensive NWS PQPF program.

2. SHORT-TERM FOCUS AND PRODUCT REFINEMENT

A. Development of NGM MOS PQPF Guidance

Appropriate Statistical Parameters

The current timetable for NWS PQPF risk reduction exercises calls for experimental field forecasts to commence at selected WFO's in 1998. The plans also anticipate the development of centralized PQPF guidance intended for the support of program participants prior to that time (National Weather Service 1997). Given these severe time constraints, it will be necessary to adapt current TDL QPF guidance to formats which are more compatible with these efforts. Fortunately, as indicated above, probabilistic information is already available from the current NGM MOS QPF system. The NGM MOS QPF guidance is derived from the output of statistical equations which yield the probability that accumulated liquid precipitation will meet or exceed several specified cutoff amounts over various 6- and 12-h time periods (Antolik 1995). The probabilistic forecasts are available for a wide range of forecast valid periods at numerous locations throughout the contiguous United States and Alaska. There is, therefore, great likelihood that a prospective user of MOS PQPF products would quickly be overwhelmed by this vast amount of information. Clearly, if we are to make MOS probabilistic QPF guidance available to the field on a routine basis, it is imperative that the guidance be packaged both usefully and concisely.

When working with probabilistic information, it is often helpful to consider the characteristics of the probability distribution or density function associated with a given event. By definition, the probability density function of a random variable is that function whose integral is exactly unity over the entire range of possible values of the variable. The likelihood that a range of values of the random variable (in our case, accumulated liquid-equivalent precipitation amount) will be observed is given by the associated area under the probability density function bounded by the two endpoints. The information contained in the probability distribution can be conveniently summarized in a number of ways. The theory of statistics tells us that it is a relatively simple exercise to calculate the mean or expected value of a given distribution, and that the variance of the distribution measures its breadth. Taken together, these summary measures give an idea of which values are most likely to occur and an indication of the uncertainty inherent in the random process. The uncertainty represented by the probability distribution can also be expressed in other ways, such as through the use of confidence intervals about the mean or median. For instance, forecasters involved in the Pittsburgh WFO pilot PQPF project have used <u>exceedance fractiles</u> to delineate a range of the most likely precipitation amounts given a fixed degree of uncertainty about the median of the distribution (Krzysztofowicz *et al.* 1993; hereafter K93).

So, when we examine the collection of MOS probabilistic precipitation forecasts valid over the same time period, it is perhaps most advantageous to consider the underlying probability distribution which they describe. After appropriate post-processing, the MOS cumulative probabilities define areas underneath the forecast probability density function (Antolik 1996). This information then can be used to estimate the mathematical form of the density function. What remains to be determined is how best to solve the integral relationship posed by the MOS forecasts or, in other words, to fit an appropriate curve to the forecast probabilities. The fitting process can be simplified if certain assumptions are made about the shape of the curve. Selker and Haith (1990) have shown that the Weibull distribution matches well with the skewed nature of observed precipitation distributions, particularly in the modeling of rare events. This conclusion also is supported by the Krzysztofowicz and Sigrest (1997) examination of historical precipitation records in support of the Pittsburgh PQPF project. Once an appropriate Weibull curve has been fit to the MOS forecasts, the distribution can be specified by the scale and shape parameters of the Weibull model, and any additional desired statistics of the distribution (mean, fractiles, etc.) can be readily calculated.

Expected values of the MOS probabilities also may be calculated without first resorting to a curve fitting procedure (Antolik 1995). This can be done by performing the appropriate subtractions such that the MOS probabilities are valid for discrete precipitation amount intervals. These discrete probabilities are then multiplied by the mean observed precipitation amount within each interval, and the results are summed over the entire range of MOS categorical predictand values. Spatial plots of NGM MOS QPF expected values have been shown (Antolik 1995, 1996) to be useful for identifying the shape and orientation of forecast precipitation patterns and in pinpointing the locations of expected precipitation maxima. Often, these details appeared in the MOS QPF expected value fields before they became evident in the numerical model forecasts. Thus, there are potential benefits to be realized if forecasters are able to examine MOS QPF expected values on a routine basis. Currently, TDL produces experimental estimates of the NGM MOS QPF expected values twice daily. These expected-value forecasts are estimates in the sense that algebraic midpoints of the discrete categories are used in the calculations in lieu of the actual observed mean precipitation amounts within each category. These forecasts are gridded and transmitted to the Heavy Precipitation Branch (HPB) of the National Centers for Environmental Prediction (NCEP) for display on the National Centers Advanced Weather Interactive Processing System (N-AWIPS).

Requirements of Eastern Region PQPF Risk Reduction Exercise

Since the K93 project was intended to examine the feasibility of forecast support for future probabilistic river stage forecasts at RFC's, PQPF was produced to serve as guidance for the hydrologic forecast cycle. These are daily river stage forecasts for 6-h periods beginning at 1200 UTC, with the first forecasts valid at 1800 UTC. Depending upon the particular RFC in question, the first of these stage forecasts may actually be issued with as little as 3-4 hours lead time with subsequent forecasts commonly extending out to 3 or 4 days. Nonetheless, the nominal valid period for this package of forecasts begins anew each day at 1200 UTC. The efforts of K93 focused on providing guidance for the first 24-h period so that updated probabilistic guidance would be available from project participants as each new daily hydrologic forecast package was being prepared at the Ohio RFC. In addition to being compatible with the hydrologic forecast cycle, a 24-h valid period is more likely to completely encompass a given precipitation event. This somewhat mitigates the need to consider issues regarding the statistical dependence of probabilistic forecasts for a single meteorological event occurring over portions of two successive forecast periods.

During the forecast experiment described in K93, participants attempted to forecast the nature of the probability distribution for the 24-h expected mean areal precipitation (MAP) amount. This probability distribution was specified via the assessment of three exceedance fractiles for each forecast valid period. The values forecast were those amounts which participants expected to be exceeded 25%, 50%, and 75% of the time. The X_{25} and X_{75} fractiles define a 50% confidence interval about the median of the distribution (X_{50}), and provide an assessment of its shape. As forecasters become more confident in the median amount forecast, the probability distribution becomes narrower and the difference between X_{25} and X_{75} decreases. As forecaster confidence declines, the reverse holds true.

Current river forecast models, however, require input of a deterministic QPF at 6-hr timesteps. K93 addressed this problem by the temporal disaggregation of 24-h expected MAP amounts implied by the probability distributions forecast in the above fashion, this time conditioned on the assumption that measurable precipitation would fall sometime during the period. Forecasters assessed the fractional amount of the anticipated 24-h total mean areal rainfall which they expected to occur within each of the four component 6-h subperiods over which their PQPFs were valid. The 6-h expected fractions determined in this fashion can be assessed without prior knowledge of the total rainfall that eventually occurs over the complete 24-h period or the nature of the anticipated 24-h probability distribution, provided that the fractions are stochastically independent of the 24-h precipitation amount. This stochastic independence has been found to hold in situations where the timing of the precipitation can be estimated with reasonable accuracy (Krzysztofowicz and Pomroy 1997). This disaggregative procedure may also be more compatible with the natural forecast thought process where field meteorologists tend to think more readily in terms of forecasting a total precipitation amount expected to occur during an entire weather event (i.e., a "storm total").

The upcoming NWS Eastern Region PQPF risk reduction exercise essentially will attempt to implement the concepts of K93 on a broader scale, albeit using procedures which have been modified somewhat, based on previous experience and subsequent consideration of what is feasible in a quasi-operational setting. In one of these modifications to K93, forecasters in the risk reduction exercise will concentrate on forecasts of conditional amounts valid for 24-h time periods. Participants also will be asked to estimate only two of the three exceedance fractiles, namely X_{50} and X_{25}, since only two points are required to specify the probability density curve if certain assumptions are made regarding its shape. Exactly as in K93, however, the risk reduction PQPF methodology assumes that the precipitation amounts conform to a Weibull distribution. Once the density function has been specified, the 24-h expected MAP amount will be determined and disaggregated into 6-h conditional amounts by using the forecasters' best estimate of the 6-h expected fractions.

A word of caution is perhaps in order regarding the suitability of the experimental risk reduction forecasts to the problem of providing 6-h probabilistic guidance for hydrologic forecasts. Once the expected fractions have been determined, the conditional 24-h expected MAP amounts may be partitioned (under the assumption of disaggregative invariance) into 6-h subtotals for input into existing RFC models. While this satisfies the current requirements of streamflow prediction models, it does not satisfy the eventual goal of probabilistic input for the 6-h forecast periods. The problem lies with the dependence of events expected to occur across one or more of the 6-h sub-periods. Probability theory allows for the combination and conditional disaggregation of mean values of the expected density functions, but the same is not true for other statistics of the distribution (e.g., the exceedance fractiles). While it is a valid exercise to ingest the three 24-h QPF exceedance fractiles into existing RFC models in order to obtain confidence limits on 24-h river stage forecasts, it is not mathematically sound to obtain 6-h probabilistic river stage forecasts by disaggregating these fractiles to generate companion sets of 6-h confidence intervals. Thus, PQPF in support of future 6-h probabilistic river stage forecasts will need to be generated in a different fashion.

Proposed NGM MOS PQPF Guidance

While the NGM MOS QPF system does not explicitly forecast MAP amounts, expected-value amounts (and expected fractions derived from them) interpolated to an appropriate grid from MOS station locations should serve as acceptable guidance to forecasters attempting the prediction of areal means. MOS QPF probabilities and expected-value estimates currently used by the NCEP HPB are interpolated via a simple Cressman analysis to the 92.5-km AWIPS grid 213 for display. Case studies examined at TDL (Antolik 1995, 1996) have suggested that the NGM MOS QPF may have an effective information content which is somewhat below the resolution of the numerical model output on which it is based. This finer-scale information is contained in the long-term climatological record of station data used for equation development. The exact "resolution" of the NGM MOS QPF guidance is dependent upon a number of factors related to the distribution and location of the MOS sites contained in the dependent data sample. These factors, all of which vary considerably over different geographical areas of the United States, include the average station spacing, the size of the regions over which predictand data have been combined, and the characteristics of the terrain. The manner in which the information content of the MOS PQPF is transferred to the analysis grid undoubtedly also plays a role in determining the ultimate resolution of any interpolated, station-oriented MOS QPF product. In general, the effective resolution of such a product is directly proportional to the MOS station density and inversely proportional to the magnitude of the terrain gradient between sites, as any inverse-distance interpolation procedure is likely to be most valid over flat terrain. Use of a climatologically-weighted interpolation procedure such as that used by Charba *et al.* (1997) in the production of locally-generated very short-range QPF would likely optimize the information content of the gridded product, and is an enhancement which should be built into future centralized, gridded MOS QPF products derived from station-oriented predictand data. For the purpose of PQPF risk reduction, we propose to make probabilistic NGM MOS QPF guidance available for all 718 MOS forecast locations within the contiguous U.S.; for NCEP use, we will also interpolate these data to the 20-km AWIPS 215 grid in order to make the guidance compatible with requirements already specified for RFC analysis of precipitation data on the scale of individual river basins. It is doubtful, however, that the NGM MOS QPF has much information content at this resolution.

Since expected-value estimates are already available at station locations from the NGM MOS QPF system, these 6- and 12-h expected amounts will be used to construct MOS guidance for the expected fractions at station locations and at the AWIPS 215 gridpoints. Ratios between the 6-h NGM MOS QPF expected values and a 24-h expected amount constructed from the two 12-h MOS QPF expected values covering the concurrent period will be used to determine the four expected fractions at each station or gridpoint as required for the 24-h forecast period of the risk reduction exercise. Again, the 24-h period in question runs from 1200 UTC to 1200 UTC the following calendar day (i.e., between the initial valid times for successive versions of the hydrologic forecast package). According to Krzysztofowicz and Pomroy (1997), this disaggregative procedure is proper as long as the expected fractions are statistically independent of the total 24-h precipitation amount. This condition generally holds to the extent that the occurrence of precipitation within a given period is more predictable than the actual expected amount. Once obtained in this fashion, the expected fractions will then be input as guidance into software designed to assist forecaster assessment of 6-h precipitation amounts for the RFC.

As indicated in Antolik (1995), the current experimental NGM MOS QPF expected-value product makes use of the midpoints between predictand amount cutoff values. In order to produce a strictly unbiased and reliable expected value, mean observed precipitation amounts should be used in place of the midpoints for each of the discrete intervals between QPF predictand breakpoints. The means in question are those dictated by the sample of NGM MOS developmental data for each forecast region and diurnal period. Due to the skewness of the distribution of observed precipitation amounts, approximation of the means by midpoints should result in overestimation of the true expected value. However, as noted in Antolik (1995), the value used for the "mean" of the uppermost (unbounded) category in the 6- and 12-h expected value estimates has been deliberately chosen to be somewhat less than the true climatic mean suggested by cursory examination of the dependent dataset. This has been done to partially compensate for the likely overestimation of the contribution from categories dealing with lesser precipitation amounts. Given the time constraints involved with providing meaningful guidance for the start of risk reduction, TDL will continue to use the current expected-value estimates in the production of guidance for the expected fractions. However, the use of seasonal, regionalized mean values may be incorporated into future versions of TDL PQPF guidance.

In addition to the four 6-h expected fractions, we also will provide guidance for the shape of the 24-h forecast probability distribution by post-processing existing NGM MOS QPF probability forecasts. The modification of NGM MOS probabilities to comply with risk reduction requirements is a considerably more challenging task. MOS probabilities are available for 6- and 12-h periods at projections ranging from 6-60 hours after the NGM cycle times of 0000 and 1200 UTC. Theoretically, this enables TDL to produce an "outlook" forecast in support of hydrologic forecast activities for the 24-48 h projection after the 1200 UTC NGM run on the preceding day. A second, updated forecast package can be generated 12 hours later based on the 12-36 h MOS guidance from the 0000 UTC run. Before the functional form of the desired 24-h probability distribution may be determined, however, an acceptable method must be found for the construction of 24-h probabilities from the existing 6- and 12-h MOS forecasts. Combining probabilities is not a straightforward exercise due to the statistical dependence of events expected to occur during one or more of the shorter MOS forecast periods which make up the 24-h intervals desired for hydrologic PQPF. In order to arrive at proper estimates of the 24-h MOS probabilities, the dependence of events spanning one or more of the MOS QPF valid periods must somehow be quantified.

Wilks (1990) has devised a procedure that combines probability of precipitation (PoP) forecasts valid for consecutive time intervals into forecasts applicable to the entire period. The dependence of events occurring during two consecutive time intervals is handled by parameterization of the probability of precipitation occurring in the second time interval conditional upon occurrence of precipitation in the first. The Wilks parameters were empirically derived and account for seasonal variability in the dependence of precipitation events. TDL currently uses the Wilks procedure to produce 24-h medium-range MOS PoP forecasts from 12-h probabilities. Unfortunately, this methodology is not applicable to the combination of probabilities for events defined by the accumulation of precipitation amounts greater than the minimum measurable. The problem lies in the fact that the Wilks model does not allow for the possibility that an event occurring over a 24-h period may be realized by the occurrence of "partial events" in each of the two 12-h subperiods, a situation which is clearly not possible in the case of PoP forecasts but has a nonzero probability of occurrence with specific precipitation amounts. In mathematical terms, we must respect the distinction between discrete and continuous random variables. In the case of combining PoP forecasts for consecutive 12-h time intervals, the resulting 24-h probability is for the union of two discrete random variables; in the case of combination of probabilities for precipitation amounts, we desire an expression for the probability of occurrence of the sum of two continuous random variables (i.e., the two 12-h precipitation amounts).

Krzysztofowicz (personal communication) has suggested that an acceptable methodology for combination of 12-h QPFs into forecasts valid over 24-h time periods involves estimation of the first and second moments of the 24-h probability density function, given the known 12-h distributions and our assumption of a Weibull model for the 24-h probability distribution. As in the preceding section, the 24-h mean can be calculated by adding the means of the distributions for the component 12-h periods. Estimating the variance of the 24-h distribution from 12-h statistics is somewhat more complicated and requires knowledge of the covariance between events expected to occur during each of the 12-h time periods. There are basically three approaches that could be taken to deal with the unknown covariance and arrive at an acceptable estimate of the second moment. First of all, we could calculate the variance based upon an assumption of complete independence between the two 12-h periods (i.e., a covariance of zero). Secondly, climatological information and/or a historical archive of MOS forecasts could be used to make empirical estimates of the covariance. These estimates would probably work best if stratified seasonally, diurnally, and perhaps also by synoptic regime or other characteristics of precipitation events. Thirdly, it is also possible that the nature of the 6-h MOS QPF probability distributions might lend some case-specific insight into the nature of the correlation between the 12-h forecasts. That is, we could use information already contained within the MOS QPF system to forecast the dependence between events expected to occur during each of the 12-h periods. For the purposes of the risk reduction exercise, we plan either to assume independence or to estimate the covariance empirically via examination of a 3-year archive of operational NGM MOS QPF and observed precipitation. As the risk reduction proceeds, we will evaluate the validity of these two methods as well as further investigate methods which utilize information from the MOS QPF's themselves.

Once the 24-h probability distribution has been specified in the above fashion at each point, additional statistics required for the risk reduction exercise can be readily calculated. Namely, the risk reduction methodology will require the estimation of at least two exceedance fractiles (the X_{50} and X_{25} values) from the 24-h conditional probability distribution. For completeness, TDL will also provide MOS estimates of X_{75}. The 24-h conditional distribution may be obtained from the unconditional MOS QPF distribution as specified above with the additional knowledge of the 24-h PoP. We will generate estimates of the 24-h PoP from MOS QPF 12-h PoP forecasts combined by using the Wilks (1990) method or, alternatively, by using more general procedures as outlined in Krzysztofowicz (1997). Once the 24-h conditional probability distribution has been specified, the required exceedance fractiles will be calculated. In summary, then, TDL will provide forecasts of eight parameters to be used as guidance for the risk reduction exercise: the four 6-h expected fractions, the 24-h PoP, and three exceedance fractiles of the conditional 24-h distribution of precipitation amount.

B. Evaluation of NGM MOS PQPF Products

The experience gained from the production of MOS PQPF for risk reduction exercises should provide practical guidance as to which types of probabilistic statistical products will be of most benefit to future meteorological and hydrological applications. At this point, it should be evident that a number of the procedures proposed to provide guidance to the upcoming project are pragmatic in design. Our primary goal is to provide usable probabilistic guidance from the NGM MOS QPF in its current operational form, without resorting to lengthy development of a completely new system. Thus, it is possible that the NGM MOS PQPF guidance as formulated and disseminated to project participants may not be the most suitable for the ultimate long-term requirements of a NWS-wide PQPF program.

Assessing the skill and appropriateness of products developed for this effort will not be a straightforward task, however. First of all, we expect that the most important information regarding the practicality of MOS PQPF will come from the participants in the risk reduction exercise. Every effort should be made to solicit the opinions of the forecasters regarding the perceived information content and usefulness of the products as formatted for the exercise. It does the operational forecaster little good if the information contained in probabilistic guidance is either too voluminous or complex, or is not well packaged. No matter how technically rigorous such a system might be, it likely will be of little practical use. Worse still, unsuitably formatted guidance may actually have a detrimental effect on operational performance and skill if the guidance requires the forecaster to devote a large block of time to its examination and interpretation.

The second part of the evaluation process should consist of formal forecast verification studies. Verification, even of deterministic forecasts, presents many challenges associated with how forecast skill is defined and measured. Many different skill measures or scores are often used in attempts to highlight the characteristics of a given forecast system. Widely used skill measures such as the critical success index work well for deterministic systems, but these measures are not suited to probabilistic forecasts. Many of the parameters which will be forecast as part of the risk reduction exercise, however, deal with the character of the expected probability distribution and hence will not be adequately verified by the examination of a few individual events. Nevertheless, if the probability forecasts are well calibrated, the marginal distribution of the observations given the forecasts should equal the <u>relative frequency</u> of observed precipitation over all possible probability values. Given the unbiased nature of MOS forecasts, one would expect that as the forecast exercise proceeds the mean of the MOS expected value forecasts should approach the average precipitation observed during the period. In addition, estimates of the exceedance fractiles should demarcate the appropriate quartiles of the observed precipitation frequency when they are examined piecewise over the entire range of precipitation amounts. If there are problems with the assumptions relied upon in the construction and parameterization of the precipitation distributions, then the statistics generated from the Weibull model should differ substantially from the observations and from the bulk statistics of the MOS forecasts. Thus, the MOS probability forecasts could serve as one check on the suitability of the Weibull fit if the Weibull model were to be applied to the 6- and 12-h forecasts, as these distributions are not subject to combination via the moment-estimation procedure of the preceding subsection.

As outlined in Antolik (1996), TDL has used the Ranked Probability Score (RPS; Epstein 1969) to evaluate the probability forecasts comprising the NGM MOS QPF system. This score accounts for the relative magnitudes of errors in a multi-category probabilistic forecast system (Murphy 1970), and is most useful when comparing a given system against some type of benchmark (a long-term climatology, for example). A similar procedure could be used to assess the skill of the 24-h probabilities as given by the Weibull curve constructed according to the moment-estimation procedure. If the combination procedure is a useful one, one might expect that 24-h forecasts could be produced with minimal degradation in skill relative to the component 12-h probabilities.

The Bayesian Correlation Score (BCS; Krzysztofowicz 1992) is a new measure of the economic value or utility of forecasts from the standpoint of end users. In order for a given forecast method to earn a higher BCS than a competitive system, the new method must offer greater potential benefit to <u>every</u> user of the information. K93 employed the Bayesian correlation score in an attempt to quantify the economic value of QPF from the WFO Pittsburgh pilot project. The BCS may be extended to multivariate systems (Krzysztofowicz, personal communication), which allows estimation of the value added to the K93 forecasts as a result of introducing probabilistic information. At no time during the Pittsburgh project did the BCS for probabilistic forecasts drop below that of the corresponding deterministic forecasts valid over the same time period, indicating the potential benefit of including probabilistic information in QPF products. Identical scoring procedures could be used to help assess the utility of the PQPF guidance produced for the upcoming risk reduction exercise.

C. Interim MOS QPF Packages Based on Other NWP Models

Since the implementation of the NGM MOS guidance package, NCEP has continued to develop and expand the suite of NWP models, and TDL has collected output data from these new models. TDL currently possesses an archive of output from the Aviation run of the global forecast model (AVN; Kanamitsu 1989), as well as output from the 'early' run of the eta-coordinate model (Black *et al.* 1993). MOS QPF guidance based on output from these models has yet to be developed, however. This is due, in part, to the frequent design changes which these models have undergone. (See Kanamitsu *et al.* 1991 for a discussion of recent changes in the AVN analysis and forecast system, and Rogers *et al.* 1995a and 1995b for a similar treatment of eta-model developments). Statistical techniques work best when long, stable samples of data are available for development. Thus, there is a risk that the use of a traditional MOS approach with either of these models may result in sub-optimal guidance.

Nevertheless, operational forecasters faced with the task of interpreting an ever-increasing stream of output from various NWP models probably would benefit from new statistical guidance based on model output. If nothing else, the existence of AVN or eta MOS QPF guidance would provide a "first-guess" interpretation of AVN and eta model precipitation forecasts, representing a potential savings of time in active weather situations. To be sure, there has been a persistent call from the field for MOS packages based on these newer models, especially for guidance from the eta model. With these considerations in mind, we developed and tested an experimental package of MOS QPF guidance based on output from the eta model (Antolik 1998). By first concentrating on the eta model, we are taking a step toward satisfying the operational demand for statistical guidance which is more mesoscale in nature. During development of this experimental guidance, we used the basic procedures which have traditionally been in place at TDL, without modification. Consequently, the use of a limited (less than 3-year) sample of developmental data may impact the stability and robustness of the system to future changes in eta model design.

Adverse impacts from the limited sample of data may have two origins. Not only has the structure of the eta model itself undergone a number of changes during the data collection period, but the introduction of observations from the Automated Surface Observing System (ASOS) also may have changed the statistical properties of the predictand data. It remains to be seen whether a blend of precipitation observations from ASOS and manual sites will lead to useful statistical relationships. In order to develop a future short-sample eta statistical system with optimum skill, extensive quality control of ASOS observations will be necessary, especially in dealing with liquid-equivalent measurements during the winter months. In addition, in order to combat the effects of any changes which may have occurred to the basic structure of the eta model during data collection, care must be taken to utilize only those output fields which exhibit reasonable stability over the entire period of record. Other steps, such as the pooling of predictand data over larger areas during regionalization and the imposition of strict minima on predictor variance, would help enhance the stability of eta guidance developed with limited data samples.

Even if it is possible to mitigate the effects of model changes during collection of the developmental data sample, it is impossible to anticipate the effects of changes which are made to the eta model <u>after</u> a statistical system has been implemented. One major question will center around the impact of continuing increases in model resolution on a MOS system which might already be in place. For instance, NCEP has already introduced a newer, mesoscale version of the eta model with 48-km resolution (Black 1994), and in February 1998, the 'early' eta model was replaced operationally by a 32-km version of the mesoscale eta model. Changes in model resolution will likely alter the statistics of certain output fields used as predictor variables, although the particular change documented in Black (1994) appeared to have minimal impact on our experimental eta MOS QPF system (Antolik 1998). This suggests that it may be possible to formulate predictors for MOS systems whose statistics may be somewhat resistant to a change in resolution. In order to compensate completely for other types of changes to NWP models, however, adaptive statistical procedures may be necessary. (For a further discussion of this resolution enhancement and other recent changes to eta model structure, see Rogers *et al.* 1996).

For the above reasons, we recognize that the performance of any interim eta MOS QPF package might not be as good as the current NGM MOS system. Such a package should still exhibit some measure of improvement over climatology and the raw precipitation forecasts from the eta model itself, however (see Dallavalle 1996), and might be the preferred guidance in situations where the eta model is deemed to have the best grasp of the synoptic situation at hand. Indeed, recent research suggests that the performance of a "consensus" system whereby MOS forecasts based on two or more different models are used is superior, on average, to a single set of statistical forecasts (Vislocky and Fritsch 1995a). Thus, the experimental eta guidance may fill a role as a valuable second statistical "opinion" with regard to daily forecast problems.

3. ISSUES AFFECTING FUTURE STATISTICAL PQPF DEVELOPMENT

It has become clear that future interpretive statistical guidance will be developed and tested in an era where growing computer power will permit NWP models with greater spatial resolution. Not only will model data be available on finer grids, but forecasts of output variables also will be available at shorter time intervals, perhaps even down to a few minutes. Furthermore, this expansion of NWP capability promises to occur at an ever-quickening pace. New models will come on line even more rapidly than in the past, and changes to operational models will occur with greater frequency.

As a result, the sheer volume and variety of NWP model output and other meteorological data available to the field forecaster are likely to increase markedly, but without corresponding growth in the time available for analysis. Field forecasters also will be hard pressed to adapt to rapid changes in model performance characteristics. That is, new models are likely to be introduced and changed before forecasters are able to develop sufficient "hands on" experience to make qualitative adjustments for model biases. Thus, the need for automated interpretive guidance will only become more critical.

A. New Statistical Approaches

One of the goals of an extensive effort currently underway to modernize TDL development and operational software is to arrive at a system which is more adaptable to the rapid changes anticipated in the future. If future MOS developmental procedures are not made more efficient, then the MOS technique could well be rendered obsolete. Future development procedures must turn out statistical equations in timely fashion after new models are introduced and must minimize the need for complete redevelopment of the statistical system when changes are introduced to model architecture.

The rapid pace of NWP advancement further complicates the prospects for developing sound statistical guidance by limiting the time available for collection of model output and historical observations prior to development. Thus, statistical guidance based on future NWP models will likely need to be developed with small samples of dependent data. The NGM MOS QPF system was developed with a stable 7-year sample of model and precipitation data and, as shown in Antolik (1995, 1996), performs well. As we have emphasized, however, traditional statistical techniques tend to work best when long samples of data are available for development. It is difficult to envision a dependent data sample of this length for development of MOS QPF guidance in the future. Thus, performance of statistical PQPF guidance based on newer NWP models may be jeopardized. The performance of future statistical systems will hinge upon two competing effects: the skill gained via improvements to the structure of the NWP models will be offset to some degree by the effects of having less-than-ideal datasets available for statistical development. There is some degree of risk, then, that future statistical QPF guidance may not significantly outperform the current NGM MOS QPF system.

Nevertheless, there are certain steps which can be taken to mitigate the negative effects of short dependent data samples. Future statistical PQPF systems will be designed such that their relationships may be quickly re-derived or "updated" over time as additional model output and observed precipitation data are collected. As long as a model remains in operation without any significant changes in its structure, the skill of its associated statistical guidance package should increase with increasing sample size. If the model has changed significantly, however, and new data are merely added to that upon which the existing statistical system is based, undesirable effects may result. If significant changes to predictor/predictand relationships have occurred, the performance of the guidance would be compromised. Statistically, the crucial issue is how well the dependent sample reflects (or can be adjusted to reflect) the relationships between NWP output and observed precipitation on the entire population of possible weather events.

In an environment where NWP models are frequently modified, it may be advisable to employ predictor variables which have been "normalized" by their means and variances over the developmental sample. Such a procedure may be useful whether the NWP model changes occur after collection of dependent data, or while data collection is in progress prior to MOS equation development. If the NWP model is changed during collection of the dependent sample, then the dependent data represent a mixture of the statistical properties of predictor variables both before and after the change. In this case, the performance of "unadjusted" statistical guidance on independent data might be somewhat better than that of a system developed solely with data collected before the model changes were implemented, but it would be quite difficult to compensate numerically for the effects of the model changes after the guidance is operational. If normalized predictors were employed in the MOS equations, however, the dependent sample could be partitioned and predictor statistics calculated separately before and after each change to the underlying NWP model. On the other hand, if NWP model changes occur subsequent to MOS development, then the statistics of predictor variables are homogeneous throughout the dependent sample, but these "pre-change" predictor statistics might be very different from those encountered operationally after implementation of MOS guidance. Therefore, statistical guidance implemented prior to NWP model changes may initially perform poorly, but may be more easily "corrected" for differences in statistical relationships between the dependent and independent data. In this latter case, normalized predictors could be used to compensate operationally for model changes made after MOS system implementation. By monitoring the statistics of model output over a specified, sliding time interval, predictor means and variances could thus be gradually adjusted without completely redeveloping the statistical guidance. When changes to the model are frequent and their seasonal effects unknown, multiple running means of statistics calculated over several, varying time periods might even be employed to "tune" the response of the statistical system. The values of the moments which result in optimal performance on a sample of independent test data could then be chosen for use in daily operations.

Other techniques which allow updates of predictor/predictand relationships also exist. For example, Krzysztofowicz (personal communication) has suggested the use of a meta-Gaussian model to construct joint density functions for multivariate systems. In this technique, relationships involving each of the predictors in question are specified in terms of (univariate) marginal distributions. The multivariate likelihood function for the predictand is then decomposed into a product of univariate conditional densities, and can be revised by using a sequential Bayesian procedure. The marginal distributions can also be updated as new data are added to the dependent sample, making it possible to include this new information without completely redeveloping the entire system.

Finally, nonparametric regression techniques might improve the performance of MOS-type systems, especially when short samples make it difficult to specify properly the nonlinear relationships in predictor data. For instance, the use of Generalized Additive Models has been shown to increase the performance of regression-based forecasts, especially when the forecast problem in question suggests that it is meteorologically advantageous to develop a regression-based statistical system with a number of nonlinear predictor fields. Techniques of this type also promise to reduce the labor-intensive interaction currently necessary to linearize important meteorological variables prior to their introduction to the regression process (Vislocky and Fritsch 1995b). All in all, there are many reasons to be optimistic that the problem of dealing with short samples of output from rapidly changing models can be solved in a mathematically rigorous, yet operationally feasible, way.

B. MOS QPF for High-Resolution Numerical Models and New Precipitation Datasets

An additional challenge as we develop statistical forecasts based on new NWP models lies in the changing characteristics of observed precipitation data. The statistical characteristics of precipitation data collected on the synoptic scale likely have changed significantly since the development of the NGM MOS QPF package. The U.S. network of manned reporting stations which comprised the backbone of the TDL hourly predictand database is currently being replaced by a network of ASOS sites. Though observations may be on a comparable spatial scale, these automated sensors have properties which are very different from the instruments which have been in use for the past few decades. In addition, certain operational problems inherent in the ASOS raingauge threaten the stability of the precipitation database (Kuligowski 1997). For example, underestimation of amounts has been seen in heavy rainfall events, and erroneous reports of measurable amounts (i.e., 0.01 inches) have occurred when heavy accumulations of dew collect in the tipping bucket. Furthermore, the ASOS tipping bucket gauge has proven to be inadequately heated, making it difficult to obtain accurate liquid-equivalent measurements during wintertime precipitation events. Lastly, although not a sensor performance issue, the placement of future ASOS sites also may affect the database if they are moved significant distances from the locations at which manual observations were taken.

As we consider the development of guidance based on mesoscale models, we also confront the choice of appropriate predictand datasets to help define and verify events predicted on an ever-shrinking spatial scale. The properties of ASOS precipitation observations notwithstanding, use of any synoptic network of gauges will probably not be adequate when developing QPF guidance based on these models. For precipitation estimation on finer scales, hydrometeorologists have been placing increased emphasis on the use of remotely-sensed data from both satellite and radar sensors. Satellite precipitation estimates have yet to reach levels of accuracy and reliability sufficient to stand alone as predictand ground truth. Radar-based estimates are not without their shortcomings, either. In the case of radar data, what would ordinarily be collected by a gauge at a point location is estimated in an indirect fashion from a quantity (i.e., reflectivity) which is not strictly proportional to rainfall rate, but is more closely related to dropsize distribution. Radar precipitation measurements also suffer from other limitations due to their estimation from a remote location on the earth's surface (i.e., radar horizon limitations and beam blockage). Nevertheless, analysis and quality control of hourly accumulated precipitation estimates from WSR-88D data are routinely being performed at various RFC's. RFC's also generate regional composites, or mosaics, of gridded precipitation estimates from a number of WSR-88D sites. These composites have become known as the WSR-88D Stage III precipitation product. Plans exist for the assembly of a national composite of Stage III data, commonly referred to as the Stage IV product (Kuligowski 1997). Provided that the limitations associated with radar measurement of rainfall can be appropriately addressed, a product of this type seems ideally suited to the development of centralized statistical guidance.

Once reliable precipitation estimates are routinely available nationwide at these finer scales, it should no longer be necessary to confine TDL efforts to station-specific MOS PQPF equations. It would then be possible to contemplate development of a grid-oriented MOS PQPF system where Stage IV precipitation observations are interpolated to a fine-scale grid, and the statistical development is done on that grid. Such a gridded MOS PQPF product would be better suited to hydrologic applications which require basin MAP estimates. A 2-year archive of Stage IV precipitation data is a likely prerequisite for nationwide gridded MOS PQPF, however. In order that these data are available in timely fashion for development of MOS guidance from the newest NWP models, TDL should begin the collection and archival of Stage IV precipitation estimates as soon as possible.

C. Probabilistic Guidance for Flexible Time Periods

Finally, one of the major issues central to the discussion of PQPF requirements for hydrologic needs is that of the time periods over which forecast guidance must be available. Current NWS public forecast requirements call for PoP guidance at 6- and 12-h intervals following the 0000 and 1200 UTC cycle times, whereas the current hydrologic forecast schedule requires QPF valid for 24-h periods. In the preceding section, we proposed a set of 24-h QPF guidance products to be generated from the existing NGM MOS probabilities. While this configuration of guidance products should satisfy the short-term needs of the PQPF effort, we intend to investigate other methods of addressing the issue of probabilistic QPF over longer valid periods.

From the earlier discussion, it should be apparent that the specification of a mathematically proper system of 24-h QPF guidance from component 6- or 12-h products is not straightforward. We have already seen that combination of probability forecasts for "events" other than the occurrence of measurable precipitation represents, in reality, an estimate of the probability of occurrence of the sum of two random variables. As such, this problem will not yield to simple probability combination techniques such as those of Wilks (1990) or Krzysztofowicz (1997). These two techniques combine probability forecasts for discrete random variables (such as PoP) over successive time intervals to arrive at an estimate of the probability of the union of the two discrete random variables. Nevertheless, the Eastern Region risk reduction methodology makes use of both 24-h unconditional precipitation amount forecasts and a forecast of the 24-h PoP in order to convert these unconditional amounts to amounts conditional upon precipitation occurrence. Thus, there are two component parts to the problem of providing complete 24-h MOS PQPF guidance for hydrologic use, that of specifying probabilistic forecasts of the 24-h total amount, and of specifying an estimate of the 24-h PoP. In light of the foregoing, we now examine methods by which 24-h MOS PQPF guidance might be constructed:

Aggregation of Short-term PQPF

Since we require estimates of the probability of occurrence of the sum of two random variables, these estimates must be constructed by consideration of the complete probability distribution of precipitation amount for each of the component subperiods. According to probability theory, estimates of the distribution of the sum of two random variables may be made if the moments of the distributions for each of the variables to be summed are known. As discussed in section 2a, our interim procedure for aggregation of PQPF will estimate the first and second moments of the 24-h probability distribution given the (known) moments of the 12-h probability distributions and the assumption of a Weibull model. However, information regarding the covariance of the precipitation amounts is also required in the process of estimating the variance of the sum. At first, we will attempt simple specification (complete independence, or climatic relationships) of the unknown dependence structure of the 12-h forecasts. As we have indicated, it may be possible to derive information about the dependence of 12-h events directly from 6-h probability information already produced by the MOS system.

In addition to this moments-based approach, alternate methods of arriving at the 24-h probability distribution may also be possible. For instance, we may be able to solve the problem by first specifying the functional form of the 12-h marginal distributions via fitting of Weibull curves to each set of 12-h MOS QPF probabilities. Once the form of the 12-h density functions has been specified, it might be possible to arrive at the complete 24-h distribution through the known marginals.

Aggregation of Short-term PoP Forecasts

Obviously, estimates of the 24-h PoP are readily available from procedures which estimate the functional form of the 24-h distribution of precipitation amount. If, however, there is an operational need for a 24-h PoP forecast without detailed guidance regarding the complete distribution of expected 24-h precipitation amounts, the PoP may be estimated by a number of other methods. For instance, the Wilks (1990) procedure might be used to combine two PoP forecasts for the component 12-h subperiods, or the more general Krzysztofowicz (1997) methodology might be employed. Alternatively, the 24-h PoPs could be constructed with the help of MOS equations developed specifically to forecast the conditional relationships between precipitation occurrence in each subinterval. (Development of MOS equations of this type can be accomplished by appropriate stratification of the dependent data.) This would enable forecasters to "build up" a set of probability forecasts for a desired time period by combination of a number of basic 6-h MOS forecasts, with MOS-derived conditional probability forecasts available for any combination of component 6-h time intervals. Climatic information regarding the correlation between events in the various 6-h periods might also be used to assist forecasters with the combination process. Even so, if a 24-h PoP forecast were to be assembled in strictly rigorous fashion from a complete set of basic 6-h MOS unconditional probabilities and forecasts of all possible conditional relationships, 15 separate equations would be required. Upon further examination, it is likely that many of these relationships would prove to be quite weak, especially among nonconsecutive time intervals. As a result, judicious assumptions of independence may greatly reduce the number of equations required.

Direct Development of Long-term Guidance and "Disaggregative" Approaches

Complications with aggregative approaches for both long-term PoP and QPF will arise as forecast output from new models becomes available at more frequent intervals. Obviously, the number of equations involved in an aggregative approach using probabilities valid for less than 6-h periods increases accordingly. Also, the subdivision of precipitation records into 6-, 3-, or even 1-h segments presents a problem with the development of stable forecast equations for the rare events. It is difficult to observe a sufficient number of the heaviest accumulations over extremely short time periods, especially if the length of the dependent data sample is limited by frequent model changes.

Instead of combining short-term QPF's, one could, of course, develop MOS QPF guidance explicitly for 24-h periods. This approach may require some refinement of current QPF development procedures, and forecasting for 24-h events might require an unacceptable degree of time-averaging in the specification of appropriate predictors. Since most fields available from current NWP models are valid for periods of 6 hours or less, development of appropriate predictors from these fields for use in MOS equations valid over long time periods could smooth out important signals in the model output.

As we have seen, however, there are serious technical difficulties associated with the "disaggregation" of long-term probabilities (i.e., for 24-h periods) into forecasts valid for specific events expected to occur over shorter time intervals (6-h periods, for example). Without adequate specification of the dependence structure for precipitation occurrence over every component subperiod, disaggregation of probabilistic forecasts is impossible. Rather than completely disaggregating a single 24-h forecast, for instance, one might wish to develop 24-h probabilities which are valid for staggered, overlapping 24-h periods in an attempt to discount events anticipated to occur completely within one of the 6-h periods. Again, without information regarding the stochastic dependence between the 6-h interval in question and either of the two, sliding 24-h valid periods, this approach is not mathematically sound.

Given the near-impossibility of decomposing long-term probability forecasts without violating the fundamental tenets of probability theory, aggregation of basic short-term probabilities is likely to be the method of choice. Indeed, when theoretical issues are considered in tandem with the difficulty of producing stable statistical forecast equations for rare events, it appears that the optimal combination of MOS QPF probabilities might be comprised of several short-term probability forecasts for the more common amounts taken in conjunction with a single probability forecast of the rare event(s) valid over the entire period of interest.

4. PLANNED TDL GUIDANCE BASED ON CURRENTLY OPERATIONAL NWP MODELS

The above issues notwithstanding, TDL will continue to provide operational meteorologists with the best possible statistical QPF given the constraints imposed by the available data. Specifically, we plan to move forward with the production of new statistical QPF guidance based on operational versions of the AVN and eta models. At the outset, these new statistical systems primarily will be developed with a traditional MOS approach and the station-oriented predictand data already archived at TDL. Undoubtedly, greater care will need to be taken than in past developments to employ predictor fields which are largely insensitive to changes in the NWP models. In order to produce guidance with an acceptable level of skill, some type of procedure to compensate for intra-sample changes in the statistics of predictor variables also may be required with these initial versions of AVN- and eta-based MOS.

We already have expanded the experimental eta-based QPF system discussed in section 2 to produce forecasts nationwide, and this system awaits implementation. It is only through daily operational experience that we will begin to fully understand the level of skill that realistically can be expected from MOS QPF developed in the traditional fashion, but with limited samples of data from a changing numerical model. Preliminary tests indicated that the performance of this experimental eta-based system was only slightly inferior to that of the current, operational NGM MOS QPF system which was developed with 7-year dependent data samples (Antolik 1998). Clearly, some of the skill that may be sacrificed by the need to employ limited data samples in MOS development may be compensated by the improved performance of the eta model itself. Shortly after the end of the data collection period for this experimental guidance, however, the eta model underwent yet another series of design changes (Rogers *et al.* 1996). Changes which occur after the collection of dependent data are potentially as damaging to the integrity of a statistical system as those which occur while data collection is in progress. Therefore, an "updatable" MOS procedure utilizing predictor variables which have been normalized by their moments (such as that described more fully in section 3.1) may be needed to realize optimal performance from any operational eta MOS QPF system. In light of the Antolik (1998) result, another alternative to the development of updatable MOS guidance also seems plausible. That is, if acceptable results can be achieved with dependent data samples of 1-2 seasons in length, then it should be feasible to generate a stable dependent sample consistent with the traditional MOS approach simply by re-running any new versions of the operational NWP models on a consistent, limited collection of historical data. If so, then it would be possible to avoid altogether the use of data samples with varying statistical characteristics of both predictor variables and observations.

After implementation of an eta MOS QPF system, we will proceed with the development of a similar, station-oriented system based on AVN data. All of the above considerations regarding the need for careful consideration of unstable predictor/predictand relationships will apply as well to any AVN MOS QPF developed by using traditional methods. We expect to begin development of this AVN MOS QPF guidance sometime during FY98.

For the more distant future, we plan to shift our emphasis to the development of grid-oriented MOS QPF whereby predictand values are specified over a grid rather than at individual station locations. This means that future MOS QPF should be more compatible with hydrologic applications which require MAP guidance. As stated in the above section, a nationwide Stage IV precipitation archive should be well suited to the development of grid-oriented MOS QPF systems. Given that attempts to develop statistical systems with less than two years of dependent data are likely to be met with questionable results, we propose to begin archiving Stage IV data at TDL once the data become routinely available on a nationwide basis. It is likely that the development of such gridded statistical QPF will be attempted for future versions of both the AVN and eta models. We expect that additional details regarding the structure of such guidance and statistical methods employed will be resolved by ongoing development and testing. Furthermore, we anticipate that the format of most operational products from future QPF systems will be probabilistic in nature, regardless of the methods employed during development. The post-analysis and verification of NGM MOS PQPF risk-reduction forecasts will guide our choices regarding the type and composition of future eta and AVN PQPF products made available.

5. SOME FINAL THOUGHTS

We have entered an era in which new numerical models and improvements to existing models are introduced at a rate which far exceeds that which was common when the MOS technique was first introduced in the early 1970's. Continuing advances in the field of NWP now make it imperative that the approach of the interpretive statistical guidance evolve as well. In short, it will no longer be possible to produce MOS with the techniques of the past on models of the future. In recognition of this, TDL already has embarked upon an effort to redesign the software and data structures used in its statistical forecast development. The goal of this effort is a software and data storage system which will enable TDL meteorologists to update statistical forecast systems more efficiently and more rapidly when changes are made to the underlying NWP models.

But that is only part of the story. In order to cope in an era when short samples of model data are all that is routinely available to the development process, some fundamental changes in the techniques used to develop statistical forecast systems are also likely to be necessary. This is not to say, a priori, that the basic MOS technique of multivariate linear regression should be abandoned; it may perhaps work best with modification of the methods used to define functional relationships between predictors and the predictand or in concert with other statistical techniques. Indeed, preliminary tests on data from the "early" eta model have shown promising results from a procedure which is essentially the same as that used to develop the current operational NGM MOS QPF system. In this case, the only differences involved modifications to the regionalization of predictand data and judicious selection of candidate predictors to help control the effects of reduced sample size. If new NWP models bring attendant increases in the skill of NWP precipitation forecasts, then the reduction in skill of statistical forecasts owing to development on smaller samples of data should at least be partially compensated. We also anticipate that an increased reliance on station climatic data will help to improve the performance of site-specific statistical forecasts based on short samples of model data. This should especially hold true for circumstances in which data must be pooled over large geographic areas during MOS development in order to insure statistical stability of rare-event forecast equations. In these cases, station-specific climatic information should compensate for some of the skill at individual point locations which may be sacrificed by the need to combine data from diverse geographic regions.

Lastly, it is clear that most future MOS QPF guidance issued to forecasters will be probabilistic. Probabilistic formulation not only makes it possible for the forecaster to convey information regarding the uncertainty of the forecasts, but allows individual end users to decide how best to react, given this uncertainty. Different scenarios likely mean that a variety of actions should be taken by various end users. PQPF allows each user to tailor the NWS forecast to their own unique needs. MOS QPF fits quite naturally in this environment in that probabilistic information can be gleaned from a single run of the underlying model. Furthermore, the MOS QPF presently represents the primary means of obtaining probabilistic information from short range NWP models. Thus, every effort should be made to expand the suite of MOS PQPF products according to the strategies set forth herein. At the same time, we recognize that a NWS-wide shift to a PQPF forecast methodology will require a major commitment of resources--not only involving the restructuring of NWS products and procedures, but also in the training of forecasters and the public (National Weather Service 1997). With its expertise in the generation of probabilistic model-based guidance products, TDL is in a position to make major contributions to these efforts.

REFERENCES

Antolik, M. S., 1995: NGM-based quantitative precipitation forecast guidance: Performance tests and practical applications. Preprints 14th Conference on Wea. Analysis and Forecasting, Dallas, Amer. Meteor. Soc., 182-187.

_____, 1996: Toward an objective NGM-based expected-value quantitative precipitation forecast system. Preprints 13th Conference on Probability and Statistics in the Atmospheric Sciences, San Francisco, Amer. Meteor. Soc., 57-64.

_____, 1998: Eta-based MOS quantitative precipitation forecasts in an era of rapidly changing numerical models: How well can we do? Preprints 16th Conference on Wea. Analysis and Forecasting, Phoenix, Amer. Meteor. Soc., 446-449.

Black, T. L., 1994: The new NMC mesoscale eta model: Description and forecast examples. Wea. Forecasting, **9**, 265-278.

_____, D. G. Deaven, and G. J. DiMego, 1993: The step-mountain eta-coordinate model: 80 km 'Early' version and objective verifications. NWS Technical Procedures Bulletin No. 412, NOAA, U.S. Department of Commerce, 31 pp.

Charba, J. P., Y. Liu, M. H. Hollar, and B. E. Exley, 1997: The LAMP QPF product. Part I: The model development. Wea. Forecasting, in press.

Dallavalle, J. P., 1996: A perspective on the use of Model Output Statistics in objective weather forecasting. Preprints 15th Conference on Weather Analysis and Forecasting, Norfolk, Amer. Meteor. Soc., 479-482.

Epstein, E. S., 1969: A scoring system for probability forecasts of ranked categories. J. Appl. Meteor., **8**, 985-987.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. J. Appl. Meteor., **11**, 1203-1211.

Hoke, J. E., N. A. Phillips, G. J. DiMego, J. J. Tucillo, and J. G. Sela, 1989: The regional analysis and forecast system of the National Meteorological Center. Wea. Forecasting, **4**, 323-334.

Kanamitsu, M., 1989: Description of the NMC Global Data Assimilation and Forecast System. Wea. Forecasting, **4**, 335-342.

_____, J. C. Alpert, K. A. Campana, P. M. Kaplan, D. G. Deaven, M. Iredell, B. Katz, H.-L. Pan, J. G. Sela, and G. H. White, 1991: Recent changes implemented into the Global Forecast System at NMC. Wea. Forecasting, **6**, 425-435.

Krzysztofowicz, R., 1992: Bayesian correlation score: A utilitarian measure of forecast skill. Mon. Wea. Rev., **120**, 208-219.

_____, 1997: Probabilities for a period and its subperiods: Theoretical relations for forecasting. (draft manuscript submitted to Mon. Wea. Rev.).

_____, and T. R. Drake, 1992: Probabilistic quantitative precipitation forecasts for river forecasting. Preprints, Symposium on Weather Forecasting, 72nd Annual Meeting of the American Meteorological Society, Atlanta, Amer. Meteor. Soc., 66-71.

_____, W. J. Drzal, T. R. Drake, J. C. Weyman, and L. A. Giordano, 1993: Probabilistic quantitative precipitation forecasts for river basins. Wea. Forecasting, **8**, 425-439.

_____, and T. A. Pomroy, 1997: Disaggregative invariance of daily precipitation. J. Appl. Meteor., **36**, 721-734.

_____, and A. A. Sigrest, 1997: Local climatic guidance for probabilistic quantitative precipitation forecasting. Mon. Wea. Rev., **125**, 305-316.

Kuligowski, R. J., 1997: An overview of National Weather Service quantitative precipitation estimates. TDL Office Note 97-2, National Weather Service, NOAA, U.S. Department of Commerce, 26 pp.

Murphy, A. H., 1970: The ranked probability score and the probability score: A comparison. Mon. Wea. Rev., **98**, 917-924.

National Weather Service, 1996: Office of Meteorology 1996-2005 strategic operating plan. NOAA, U.S. Department of Commerce, 56 pp.

_____, 1997: The end-to-end forecast process for quantitative precipitation information: Hydrometeorological requirements, scientific issues, and service concepts. NOAA, U.S. Department of Commerce, in press.

Rogers, E., T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, Y. Lin, N. W. Junker, and M. E. Baldwin, 1995a: Changes to the NMC operational eta model analysis/forecast system. NWS Technical Procedures Bulletin, No. 423, NOAA, U. S. Department of Commerce, 60 pp.

_____, D. G. Deaven, and G. J. DiMego, 1995b: The Regional Analysis System for the Operational 'Early' eta model: Original 80-km configuration and recent changes. Wea. Forecasting, **10**, 810-825.

_____, T. L. Black, D. G. Deaven, G. J. DiMego, Q. Zhao, M. E. Baldwin, N. W. Junker, and Y. Lin, 1996: Changes to the operational "Early" eta analysis/forecast system at the National Centers for Environmental Prediction. Wea. Forecasting, **11**, 391-413.

Selker, J. S., and D. A. Haith, 1990: Development and testing of single-parameter precipitation distributions. Water Resources Research, **26**, 2733-2740.

U.S. Department of Commerce, 1994: The Great Flood of 1993. Natural Disaster Survey Report, National Weather Service, NOAA, 281 pp.

Vislocky, R. L., and J. M. Fritsch, 1995a: Improved Model Output Statistics forecasts through model consensus. Bull. Amer. Meteor. Soc., **76**, 1157-1164.

_____, and _____, 1995b: Generalized Additive Models versus linear regression in generating probabilistic MOS forecasts of aviation weather parameters. Wea. Forecasting, **10**, 669-680.

Wilks, D. S., 1990: On the combination of probabilities for consecutive precipitation periods. Wea. Forecasting, **5**, 640-650.