6.1

An Experiment in Probabilistic Quantitative Precipitation Forecasting

Steve Amburn* and James Frederick
National Weather Service Forecast Office, Tulsa, Oklahoma

1. Introduction

Jorgensen and Klein (1969) demonstrated a method to produce probabilistic quantitative precipitation forecasts (PQPFs) based on seasonal average climatological precipitation amounts and precipitation probabilities. Based on 15 years of data, they computed tables of the probabilities to exceed certain seasonal precipitation amounts at 108 stations across the conterminous United States. Those tables display unconditional probabilities that include both rain and no rain cases and also conditional probabilities that include data for only days when rain occurred. The latter is of most interest since it can be combined with the forecast probability of rain to produce an unconditional Probability of Exceedance (POE) for arbitrarily selected rainfall amounts.

Wilks (1990) used a similar approach to produce PQPFs (or POEs) using probabilities of precipitation (PoPs) together with climatological distributions of precipitation amounts, also based on the condition that rain occurred. Wilks concluded, "PQPFs could be more skillful than previous experience has indicated if appropriate climatological distributions of conditional precipitation amounts were to be consulted as one element of the guidance."

Additional information and work on probabilistic quantitative precipitation forecasts can be found in: Hashemi and Decker, 1969; Hughes, 1980; Wilks, 1990; Krzysztofowicz, 1993; Sigrest, 1998; Applequist, 2002; Cope, 2004 and others.

Unfortunately, climatologies are averages of a diverse spectrum of rainfall events, with varying precipitation means and distributions. This diversity is observed not only on an annual time scale but also on a seasonal scale. Therefore, a single climatology, even for a season, frequently does not represent the variety of events that occur. It should be better to forecast the appropriate distribution for the expected event, and then compute the PQPFs from those distributions. The National Weather Service Forecast Office (WFO) in Tulsa, Oklahoma is analyzing this method.

Corresponding author address: Steven A. Amburn, National Weather Service, 10159 E. 11th Street, Suite 300, Tulsa, OK 74128;
Email: Steve.Amburn@noaa.gov

This paper describes an attempt to further the work of both Jorgensen and Klein (1969) and also Wilks (1990). Assuming the distribution of rainfall observations across a given area, from a typical precipitation event, can be represented reasonably well by an exponential probability density function (PDF) with a given mean, then the probability to exceed a particular rainfall amount can be calculated. This Probability of Exceedance (POE) is the probabilistic QPF produced at WFO Tulsa.

The Tulsa method produces POEs using four terms: 1) the probability density function (PDF) of the exponential distribution; 2) WFO generated QPFs that are used for the mean of the PDF; 3) WFO generated PoP; and 3) software in the Gridded Forecast Editor (GFE, Global Systems Division, 2005). A description of the Tulsa method is discussed along with comparisons to the work of Jorgensen and Klein.

Examples of graphic products are described and shown below. An example of a text product is also shown, which includes specific average POEs for one of the 32 counties across the Tulsa County Warning and Forecast Area (TSA CWFA).

2. Mathematical Background

This formula-based method of producing POEs is based on the assumption that the distribution of given precipitation amounts can be approximated by the gamma distribution. Wilks (1995) states, "the versatility in shape of the gamma distribution makes it an attractive candidate for representing precipitation data, and it is often used for this purpose." The gamma function is shown in Equation (1). A brief mathematical explanation is provided here.

$$\Gamma(\alpha) = \int x^{(\alpha-1)} e^{-x} \, dx \qquad (1)$$

,for α > 0, integrated from 0 to ∞.

The gamma distribution takes on several different shapes, as shown in Figure 1, depending on the values of the shape parameter alpha (α). However, based on climatology, an appropriate distribution for most precipitation events is where alpha is equal to 1.0. In this case, the frequency of small rainfall amounts is highest, with a rapid decrease in frequencies of higher amounts. Where α= 1, the gamma distribution simplifies to a special distribution called the exponential distribution which can be used in producing the POEs. The density function of the exponential distribution is defined by Equation (2), where the mean value of the distribution is given by μ (Mu).

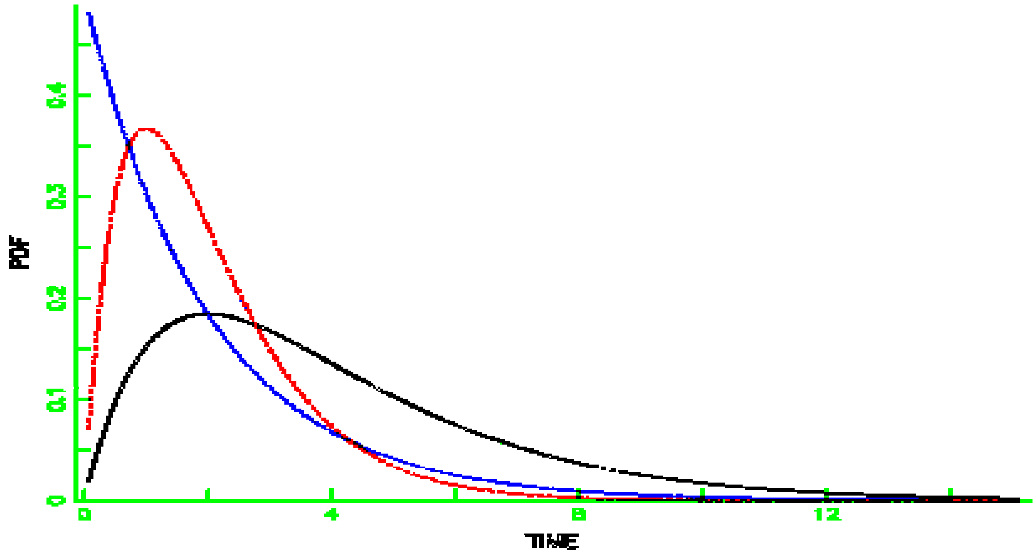

Figure 1. Examples of gamma PDF, where alpha = 1 (blue line), 2 (red line), 3 (black line), from Engineering Statistics Handbook (2005)

$$f(x) = (1/\mu) * e^{-x/\mu} \quad (2)$$

, where μ is the mean

Integrating (2) yields the cumulative density function (Equation 3), where the POE can be computed for any selected rainfall amount, x. (A more rigorous explanation can be found in a number of statistics books, such as Wilks (1995)).

$$POE(x) = e^{-x/\mu} \quad (3)$$

Examples of the exponential PDF are shown in Figure 2 for a variety of means. Note, as the mean increases, the PDF becomes “flatter” with a larger area under the right tail of the PDF. This indicates that events with higher average rainfalls will have a higher frequency of larger individual rainfall amounts, and therefore higher POEs for large rainfall amounts.

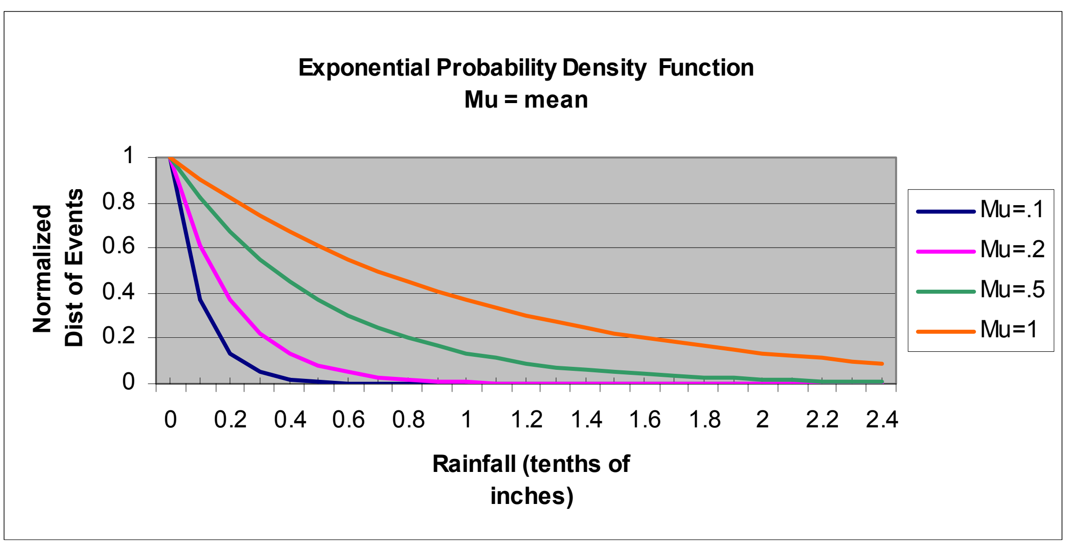

Figure 2. Exponential density functions for varying means (Mu). For larger means, the decline in the number of rainfall amounts is less dramatic, indicating a greater frequency of heavier amounts.

Table 1 contains examples of POEs, using Equation (3), for certain threshold QPFs with different mean values. POEs in the table are conditional upon the occurrence of rain. The reader will notice that the probability for attaining the mean value is less than 50%. This is a characteristic of the exponential distribution. The third line in Table 1 provides a good example. The mean on that line is 0.50 inches, yet the POE for 0.50 inches is only 36.8%. Since the exponential distribution is skewed strongly toward lower values, the mean and median are not the same. The mean value will be higher than the median. More robust explanations can be found in statistics textbooks, such as Walpole and Myer (1978).

Table 1. The table shows examples of POEs for different mean values, given a PoP of 100%.

Mean QPF(in.)	POE(.10)	POE(0.25)	POE(0.50)	POE(1.00)	POE(2.00)
0.10	0.368	0.082	0.007	0.000	0.000
0.20	0.607	0.287	0.082	0.007	0.000
0.50	0.819	0.607	0.368	0.135	0.018
0.75	0.875	0.717	0.513	0.264	0.069
1.00	0.905	0.779	0.607	0.368	0.135
1.50	0.936	0.846	0.717	0.513	0.264
2.00	0.951	0.882	0.779	0.607	0.368
2.50	0.961	0.905	0.819	0.670	0.449

3. Examples of Rainfall Frequency Distributions

Plots of precipitation data for a ten-year period from 1995 through early 2005 for sites in the TSA CWFA match the shape of the exponential distribution rather well. Figure 3 shows12-hour rainfall data plots for two sites in the TSA forecast area. It can be seen that the distribution of rainfall amounts in each 0.05-inch category bin decreases rapidly as the amounts increase.

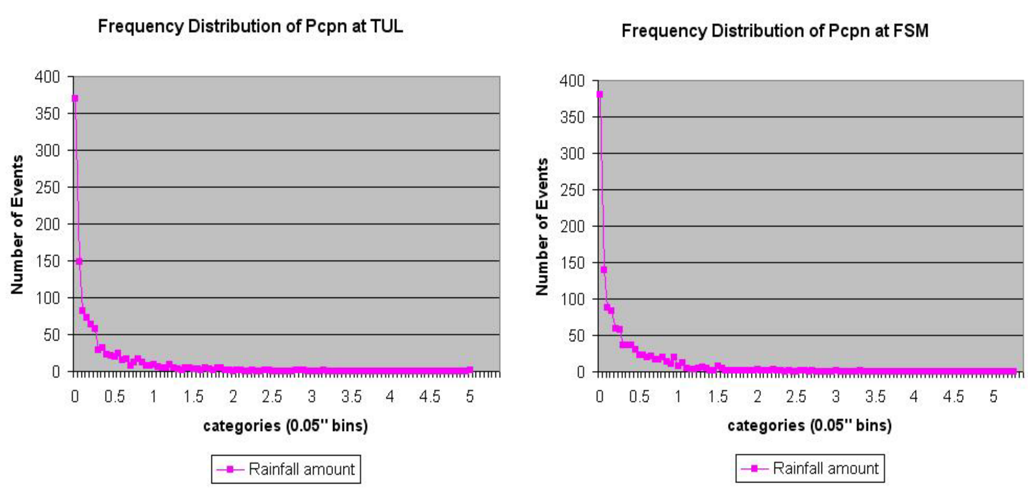

Figure 3. Rainfall frequency distributions from 1995 through early 2005 for Tulsa, OK (TUL), Fort Smith, AR (FSM). Frequencies are for 0.05 inch categories.

While the exponential distribution is valid as a composite of events, is it also valid for individual rainfall events? Figure 4 shows 86 individual 12-hour rainfall events in the Tulsa WFO forecast area from mid summer to early autumn 2005. (A 12-hour rainfall is defined here as any 12-hour period where measurable rain occurred anywhere in the WFO TSA forecast area. For the period described, 140 events were possible.) The consistency in the shape of the plots would indicate that the exponential distribution also applies to individual events. Even larger and heavier rainfall events, as shown in Figure 5, are reasonably represented by the exponential distribution. However, further analysis and study of a wider variety of events, particularly winter events, would be prudent, as more appropriate distributions may be required for some events, depending on the synoptic scale influences.

The data for Figure 4 are human quality controlled quantitative precipitation estimates from the NWS Arkansas Red Basin River Forecast Center in Tulsa, Oklahoma. These data analyses are performed hourly on a 4x4km grid that covers the Tulsa WFO forecast area. Each 4x4km grid is effectively considered a rain gage. The hourly analyses are summed over 12-hour periods and grouped into 0.05 inch categories to create the frequency distributions in the figure.

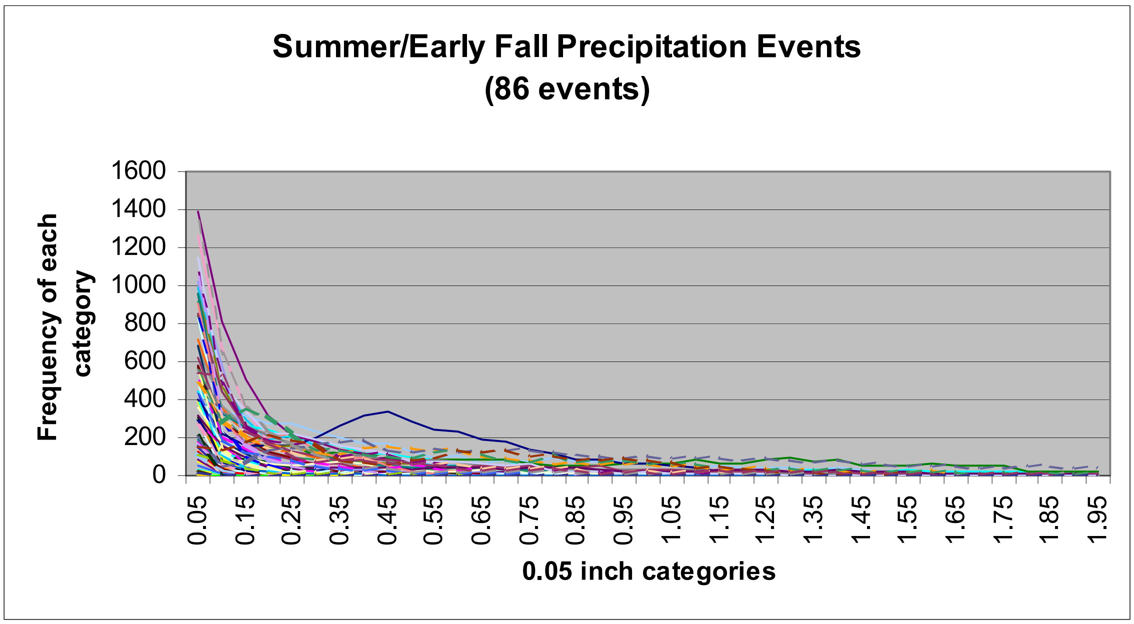

Figure 4. Rainfall distributions for 86 individual 12-hour rainfall events in the TSA forecast area.

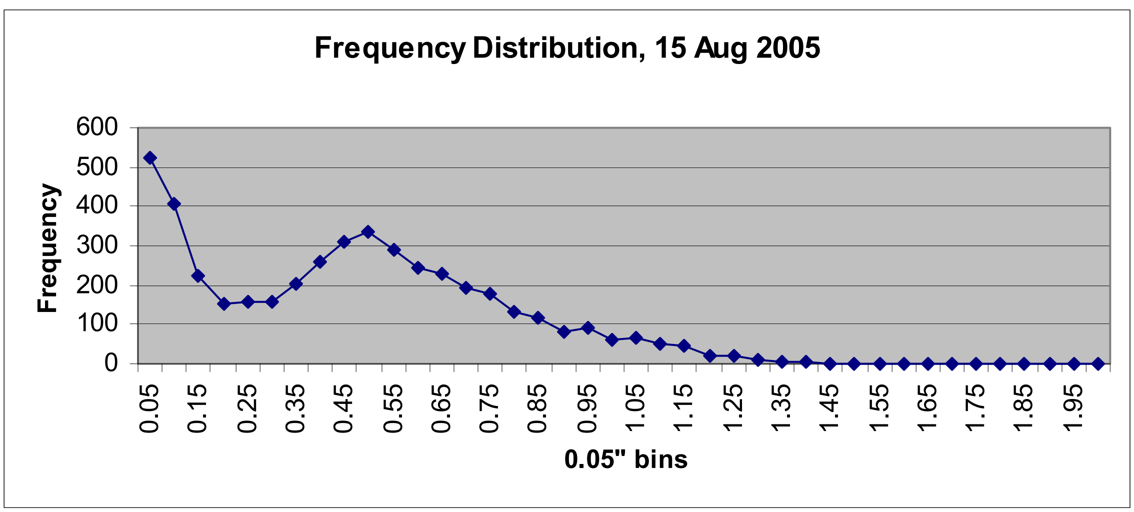

Figure 5. Non-typical exponential distribution of rainfall for an event in eastern Oklahoma.

4. The Forecast Process

WFO meteorologists already forecast all the necessary parameters for the production of POEs. No additional workload is required. In fact, the process of producing the grid files, products and graphics is automated within the GFE. Forecasters create their PoP and QPF forecasts, as they currently do. Then, GFE calculates the POE grid fields for threshold rainfall amounts of 0.10, 0.50, 1.00 and 2.00 inches.

During the development phase of the Tulsa Method, the type of POE had to be determined. Would it be conditional (based on the occurrence of rain), or unconditional (independent of the occurrence of rain). Following the lead of Jorgenson and Klein, it was decided to produce the unconditional POE (uPOE). Equation (3) is then used to calculate the POE based on the mean (μ) rainfall amount based on the condition that rain occurs. That results in a conditional POE. The uPOE is then simply the product of the conditional POE and the NWS PoP as shown in Equation (4).

$$\text{uPOE}(x) = (e^{-x/\mu}) * \text{PoP} \qquad (4)$$

This solved an interesting problem. When calculated, conditional POEs frequently exceeded the standard PoP for that same period, since the value of μ is based on the condition that rain occurs. As an example, given that rain occurs, the conditional POE for 0.10 inches of rain may be 80%. However, the NWS PoP to measure 0.01 inches of rain may only be 30%. This might be confusing to the less sophisticated user.

One last obstacle had to be overcome. NWS QPF forecasts are unconditional, i.e., they are areal average amounts a forecaster expects when <u>all</u> gages are considered, including the ones that recorded no rain. Therefore, the NWS QPF needs to be converted to a conditional QPF, which is the value of μ used in Equation (4). Within the GFE, Equation (5) accomplishes this. This is also consistent with the work of Jorgenson and Klein.

$$\begin{aligned}\mu &= \text{Conditional QPF}\\ &= (\text{unconditional QPF}) / \text{PoP} \qquad (5)\end{aligned}$$

The critical element to the entire process is the QPF supplied by the forecasters. That QPF is substituted for μ and changes the shape or "steepness" of the exponential PDF, thereby changing the resultant POEs. This step in the process takes advantage of a forecaster's expertise to identify events that may not match the "average" for that season. This should provide for much more accurate POEs than can be computed by simply using the seasonal mean as offered by Jorgenson and Klein. Events not typical for the season will likely depend on the nature of the event (convective or non-convective).

After GFE performs the calculations using Equations (4) and (5), all output products are generated. The forecasters may then choose to alter the POEs, although that is not expected to happen very often. Verification and feedback to the forecaster should help determine if and when these adjustments will add value.

5. Comparisons to Previous Work

Jorgenson and Klein (1969) derived Equation (6), which defines the unconditional probability to exceed a certain amount of rainfall (r), for a given event. $P_t(r/0.01)$ is the conditional probability that an amount greater than "r" will occur, and is provided in the tables they compiled for 108 stations across the conterminous United States. P(0.01) is the probability of measurable rain (0.01 inches), which is the standard NWS PoP. An excerpt is provided in Appendix A.

$$P_t(r, 0.01) = P_t(r/0.01) * P(0.01) \qquad (6)$$

Jorgenson and Klein provided the following example where they compute the unconditional probability to exceed 0.50 inches of rain, based on a PoP of 60%. "Consider, for example, the problem of determining the probability of .50 inches or more of rain in the 'tonight' period for Atlanta during the spring months. Assume that the public probability forecast has assigned a .60 probability to the event of measurable precipitation for 'tonight' (00Z-12Z for Atlanta), so that P(.01) is .60. The data in table 1 provide the conditional probability P(.50/.01) = .27. Substituting into equation ([6]):

$$\begin{aligned}P_t(0.50/0.01) &= P_t(0.50/0.01) * P(0.01)\\ &= 0.27 * 0.60 = 0.16 \qquad (7)\end{aligned}$$

The desired probability is then 0.16."

The rainfall mean for a 12-hour, spring event in Atlanta, obtained from the Jorgensen and Klein table is 0.36 inches. By substituting 0.36 for μ and using a PoP of 60% in equation 4, the Tulsa method yields the following:

$$\begin{aligned}\text{uPOE}(x) &= (e^{-x/\mu}) * \text{PoP} \qquad (8)\\ &= \text{Exp}(-.50/.36) * 0.60 = 0.15\end{aligned}$$

The results of equations (7) and (8) are remarkably close. Table 2 shows other examples, given a PoP of 100%. Not all values are as close as the above example, but the results are probably well within the forecast errors of both the QPF and the PoP.

Table 2. Sample POEs for 0.25 and 0.50 inches as taken from Jorgenson and Klein (J&K) and also calculated from uPOE equation (6) using 100% for the PoP. Average difference between methods was 3.38%. A maximum difference was 8% at Detroit and Fort Worth.

City	Mean	J&K(.25)	uPOE(.25)	J&K(.50)	uPOE(.50)	Avg Diff
Detroit (winter)	0.11	13%	10%	4%	1%	3.00%
Detroit (spring)	0.14	22%	16%	6%	3%	4.50%
Detroit (summer)	0.25	29%	37%	16%	14%	5.00%
Detroit (autumn)	0.20	26%	28%	11%	8%	2.50%
Fort Worth (winter)	0.19	24%	27%	11%	7%	3.50%
Fort Worth (spring)	0.39	47%	53%	27%	28%	3.50%
Fort Worth (summer)	0.32	38%	46%	21%	21%	4.00%
Fort Worth(autumn)	0.30	38%	43%	18%	19%	3.00%
Atlanta (winter)	0.30	37%	43%	19%	19%	3.00%
Atlanta (spring)	0.36	44%	50%	27%	25%	4.00%
Atlanta (summer)	0.34	43%	48%	24%	23%	3.00%
Atlanta (autumn)	0.25	34%	37%	19%	14%	4.00%
Sacramento (winter)	0.24	32%	35%	14%	12%	2.50%
Sacramento (spring)	0.19	28%	27%	8%	7%	1.00%
Sacramento (summer)	0.11	14%	10%	7%	1%	5.00%
Sacramento (autumn)	0.26	36%	38%	18%	15%	2.50%
					Avg Diff	3.38%

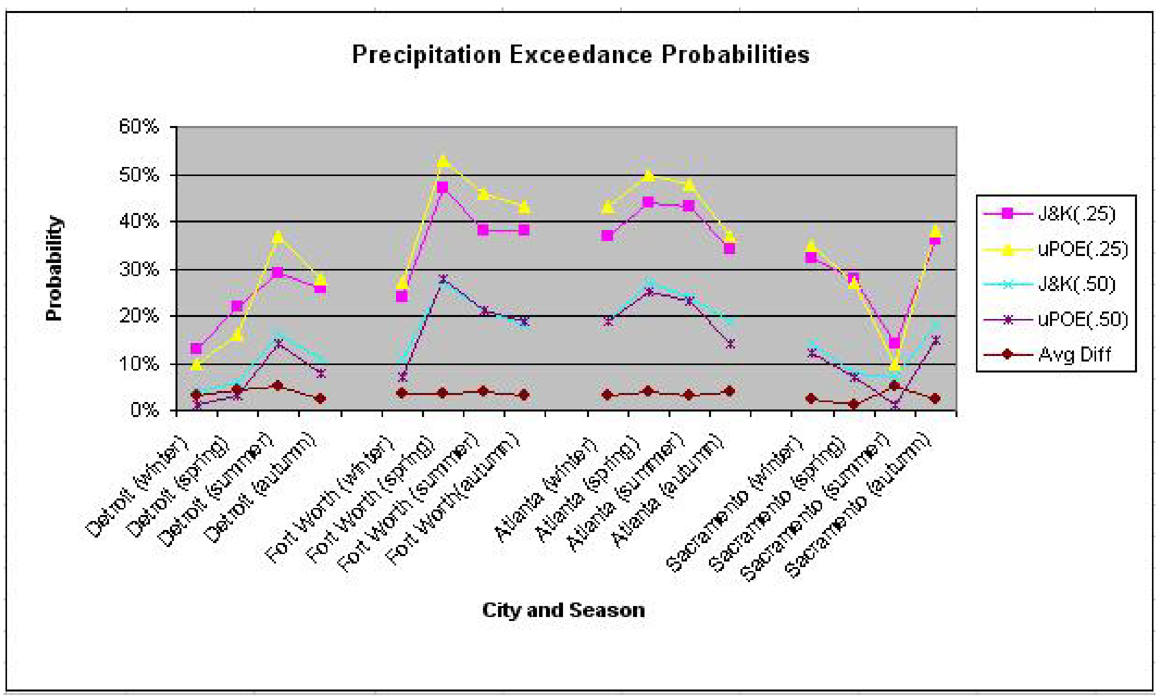

Figure 6. Plot of data in table 2.

6. Justification for the Method

Simply put, accuracy is the justification for using formula-based POEs. Once the decision is made to provide probabilistic QPFs to the user community, it is incumbent upon the NWS to provide the best ones possible. Using a forecaster's mean QPF for μ in the exponential PDF rather than the climatological mean should result in much better POEs for those events.

Figure 7 is a bar chart showing the mean precipitation from 86 12-hour rainfall events from the end of July 2005 to the end of September 2005 (140 events possible during this time). The climatological mean for that same season, based on the 10-year Tulsa data and also the 15-year Jorgenson and Klein data, is approximately 0.36 inches, as shown by the line through the middle of the chart. POEs calculated using 0.36 for μ would clearly have been too high. However, POEs based on a forecaster's best judgment should more closely match the actual means for each of those events and therefore should result in more accurate POEs. Verification feedback and training is expected to make that true.

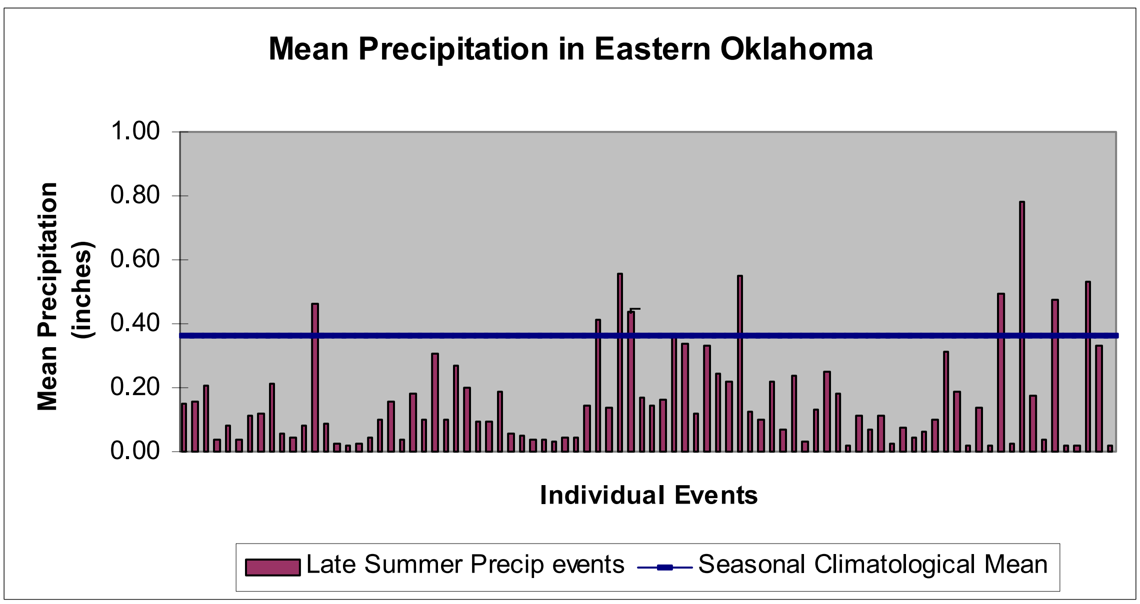

Figure 7. Late summer mean precipitation for 86 events in the WFO Tulsa forecast area. The line shows the climatological mean for any event.

7. Examples of POE Output

POE products are both graphic and text. Gridded POE within the GFE are available for 6-hour, 12-hour, or 24-hour periods. Figure 8 is a GFE depiction of the POE(0.10) for a specific 6-hour period. Those grids can be output directly to the TSA web page or used to generate other graphics, such as the bar graph shown in Figure 9. Finally, a text product is shown in Figure 10, depicting the average POEs one of the 32 counties in the Tulsa forecast area.

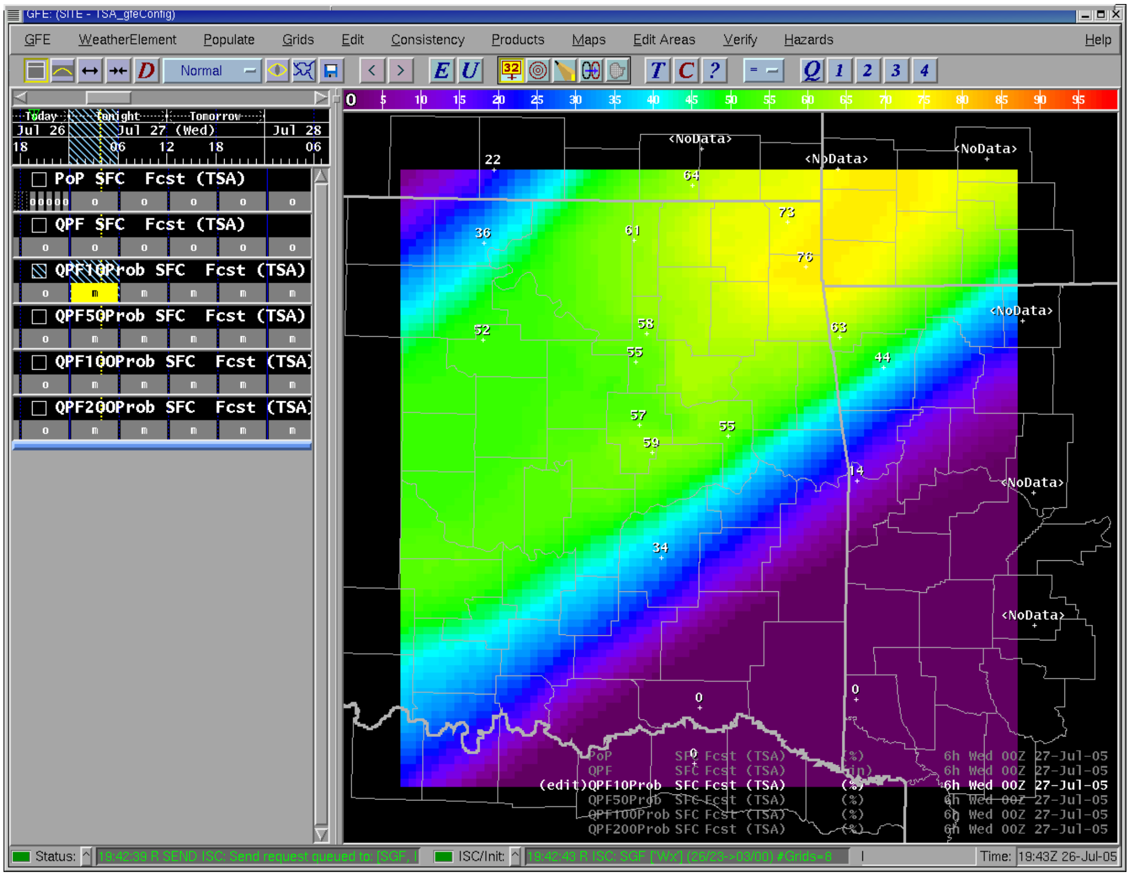

Figure 8. POE for unconditional probability to exceed 0.10 inches as shown in GFE graphics.

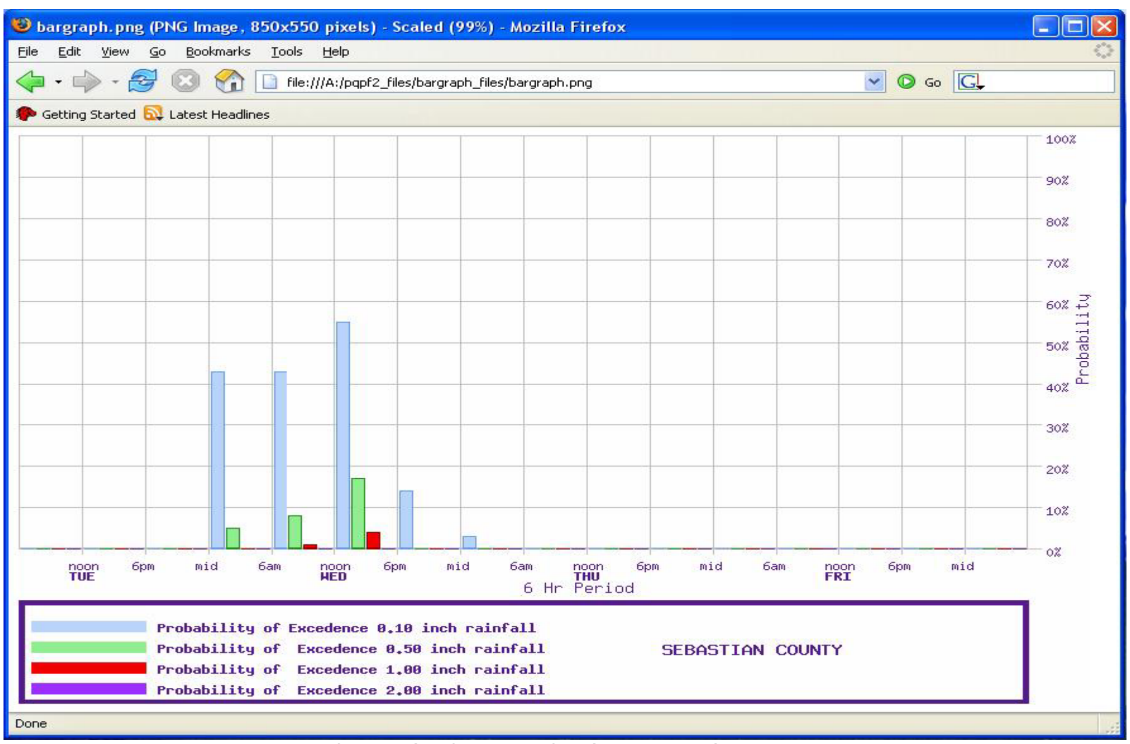

Figure 9. Bar graph output of the POE forecast for Sebastian County, Arkansas.

```
OKZ060-271445-
TULSA-
INCLUDING THE CITY OF...TULSA
937 PM CDT TUE JUL 26 2005

DATE         TUE 07/26/05     \       WED 07/27/05     \       THU 07/28/05     \
CDT          06 09 12 15 18 21 00 03 06 09 12 15 18 21 00 03 06 09 12 15 18 21 00

POP 6HR         30    40    50    70    50    20    10    10     5     5     5
QPF 6HR          0     0  0.08  0.20  0.16  0.04     0     0     0     0     0
X 0.10           0     0    17    49    29     0     0     0     0     0     0
X 0.50           0     0     0     9     2     0     0     0     0     0     0
X 1.00           0     0     0     1     0     0     0     0     0     0     0
X 2.00           0     0     0     0     0     0     0     0     0     0     0
```

Figure 10. Text output for Tulsa County, Oklahoma, showing the PoP and QPF for each 6-hour period, along with the POEs for 0.10, 0.50, 1.00, and 2.00 inches of rainfall.

8. Summary

Probabilistic QPFs, or probabilities of exceedance (POEs), are being produced at each forecast cycle at the WFO in Tulsa, Oklahoma. These POEs are generated in the Gridded Forecast Editor and are unique for each grid point across the TSA forecast area for threshold amounts of 0.10, 0.50, 1.00 and 2.00 inches. The meteorologist’s unconditional QPF grid fields are used as input to the probability density function of the exponential distribution. Those QPFs effectively change the shape of the distribution so it will more closely match the expected distribution for the rainfall event. Conditional POEs are then generated for the specified threshold precipitation amounts. These conditional POEs are then multiplied by the PoPs at each grid point to arrive at the final unconditional POEs. This method is automated and requires no additional effort from the forecasters.

This Tulsa method of issuing a PQPF is experimental and still needs to be evaluated for accuracy and reliability. However, it does compare well with the previous results of Jorgenson and Klein (1969). There is some concern that this PQPF method may not be entirely appropriate for stratiform precipitation events. Examples of TSA PQPF output can be found at http://www.srh.noaa.gov/tsa/pqpf.htm, in both graphic and text modes.

9. Acknowledgements

We are grateful to Bill Lawrence of the Arkansas Red Basin River Forecast Center for providing the rainfall data for use in this study. Without his help, this study would not have been possible. We would also like to acknowledge the assistance of Dr. Peyton Cook, professor of statistics at the University of Tulsa, for his help with the mathematics. Finally, we would like to thank Ms. Diane Cooper, Mr. Bernard Meisner, and Ms. Diana Riddle for their comments and suggestions for this manuscript.

10. References

Applequist, S., et.al., 2002: Comparison of Methodologies for Probabilistic Quantitative Precipitation Forecasting. *Wea. Forecasting*, **17**, pp. 783–799.

Cope, A. M. and M. P. Delisi, 2004: An Experiment in Probabilistic Quantitative Snowfall Forecasting, Preprints, *20th Conference on Weather Analysis and Forecasting / 16th Conference on Numerical Weather Prediction*, 84th, Seattle, WA, Amer. Meteor. Soc., P2.18.

Engineering Statistics Handbook, 2006: Section 8.1.6.5 Gamma, http://www.itl.nist.gov/div898/handbook/apr/section1/apr165.htm

Global Systems Division, 2006: Interactive Forecast Preparation System (IFPS), GFE Suite Technical Review, http://www.fsl.noaa.gov/projects/development.html.

Hashemi, F. and W. Decker, 1969: Using Climatic Inforamtion and Weather Forecasts for Decisions in Economizing Irrigation Water. *Agricul. Meteor.*, **6**, 245-257.

Hughes,Lawrence A., 1980: Probability Forecasting – Reasons, Procedures, Problems. NOAA Technical Memorandum NWS FCST 24, Meteorological Services Division, Silver Spring, MD, 84 pp.

Jorgensen, D.L. and W.H. Klein, 1969: Conditional Probabilities of Precipitation Amounts in the Conterminous United States. ESSA Tech. Memo, WBTM TDL 18, U.S. Dept. Commerce, 89 pp. [NTIS PB 183144]

Krzysztofowicz, R., W. J. Drzal, T. R. Drake, J. C. Weyman and L. A. Giordano, 1993: Probabilistic Quantitative Precipitation Forecasts for River Basins. *Wea Forecasting*, **8**, pp. 424-439.

Sigrest, 1998: Spatially Averaged versus Point Precipitation in Monongahela Basin: Statistical Distinctions for Forecasting. *Wea. Forecasting*, **13**, pp.1063-1077.

Walpole, R.E. and R.H. Myers, 1978: Probability and Statistics for Engineers and Scientists, 2nd Edition, Macmillan Publishing Co., Inc. New York, New York, 580 pp.

Wilks, Daniel S., 1990: Probabilistic Quantitative Precipitation Forecasts Derived from PoPs and Conditional Precipitation Amount Climatologies**.** *Monthly Weather Review*: Vol. 118, No. 4, pp. 874–882.

Wilks, Daniel S., 1995: Statistical Methods in the Atmospheric Sciences, Academic Press, San Diego, CA, 467 pp.

Appendix A
An Excerpt from Jorgenson and Klein (1969)

This excerpt provides the statistical basis Jorgenson and Klein used for making probabilistic quantitative precipitation forecasts. Their Equation (4) defines the unconditional probability to exceed a selected rainfall amount, r. Table 1, to which they refer, is their tabulated data that gives the conditional probabilities of precipitation occurrence in seven quantitative ranges for 108 stations combined by seasons. The Tulsa Method to compute POEs uses Equation (4) below.

"To obtain the probability of a precipitation event consisting of any fixed amount of rain falling in a given time period, we can make use of the definition of conditional probability. The conditional probability of an event A given that event B will occur is

$$\mathbf{P(A/B) = P(A,B) / P(B)} \qquad \textbf{(2)}$$

Where P(A/B) is the conditional probability of A, the condition being that B occurs, P(A,B) is the probability for the joint occurrence of A and B, and P(B) is the probability of B.

Applying this definition to a rain amount in excess of r in a period t, we write

$$\mathbf{Pt(r/0.01) = Pt(r,0.01) / P(0.01)} \qquad \textbf{(3)}$$

Or

$$\mathbf{Pt(r,0.01) = Pt(r/0.01) * P(0.01)} \qquad \textbf{(4)}$$

The conditional probability of an amount greater than r, Pt(r/0.01), is given in table 1 for time periods of 6, 12, and 24-hours. The probability of measurable rain, P(0.01), is obtained from the public probability forecast. The product of these two gives the desired probability."