MONTHLY
WEATHER
REVIEW

VOLUME 94, NUMBER 10

OCTOBER 1966

POINT AND AREA PRECIPITATION PROBABILITIES [1,2]

EDWARD S. EPSTEIN

Department of Meteorology and Oceanography, The University of Michigan

ABSTRACT

The relationship between point and area precipitation probabilities is examined on the basis of a simple model in which circular precipitation cells of uniform size are distributed at random over an area that is large compared to the forecast area. From knowledge of the cell size and the number of cells per unit area it is then possible to state both the point and area precipitation probabilities. Formulas and graphs of these relationships are shown. When the cells are large, point and area precipitation probabilities are almost equal, but they differ markedly when the cells are small. Joint and conditional probabilities of precipitation at two or more stations are also examined. An extension of the model is presented in which uncertainty regarding the density of cells is expressed as an elementary probability density, and the effects of this on the expected point and area precipitation probabilities are shown.

The present practice among forecasters is to express precipitation forecasts in probabilistic terms. These probabilities are specifically point probabilities, i.e., the probability that measurable precipitation will be observed during the forecast period at one, or any, given point in the forecast area. The probabilities are specifically *not* area probabilities, i.e., the probability that measurable precipitation will be observed at some point in the forecast area during the forecast period. The relationship between area and point probabilities has been a source of some confusion. The purpose of this note is to shed some light on this subject.

Let us consider the area for which the forecast is valid as our unit area, and suppose that during a given period there are N precipitation cells per unit area distributed at random (i.e., all locations are equally likely) over an area large compared to the forecast area. We will consider that each precipitation cell covers an area Q. For simplicity we will assume that both the forecast area and the precipitation cells are circular, thus having radii of $(1/\pi)^{1/2}$ and $(Q/\pi)^{1/2}$, respectively. The area probability of precipitation is then equivalent to the probability that one or more cells occur within an area of

$$\pi[(1/\pi)^{1/2}+(Q/\pi)^{1/2}]^2=[1+Q^{1/2}]^2$$

(see fig. 1). The point probability of precipitation is the probability that one or more cells occur within an area Q surrounding the observation point.

The probability that k cells fall within the area Q is, from the Poisson probability distribution,

$$e^{-NQ}(NQ)^k/k!.$$

The probability that one or more cells fall in that area, i.e., the point probability of precipitation, is then

$$P_p=1-e^{-NQ}. \quad (1)$$

Similarly, the area probability of precipitation is

$$P_a=1-e^{-N[1+Q^{1/2}]^2}. \quad (2)$$

Furthermore, one can see that the area and point probabilities are related by

$$1-P_a=(1-P_p)^{[1+(1/Q)^{1/2}]^2}.$$

[1] This research was supported in part by the U.S. Department of Commerce, Environmental Science Services Administration, Weather Bureau, under Contract Cwb-10847.

[2] Publication Number 107 from the Department of Meteorology and Oceanography. The University of Michigan.

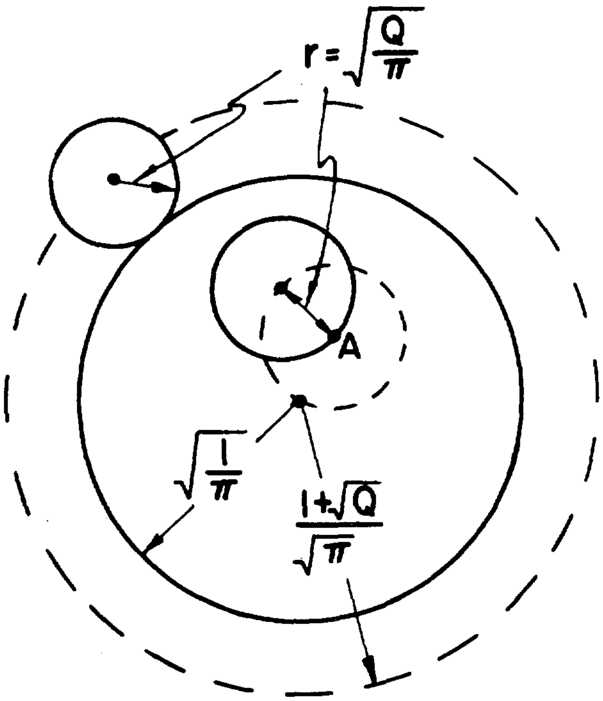

FIGURE 1.—Any cell located within the large dashed circle produces precipitation in the forecast area (large solid circle). Any cell located within the small dashed circle produces precipitation at station A. The small solid circles represent precipitation cells.

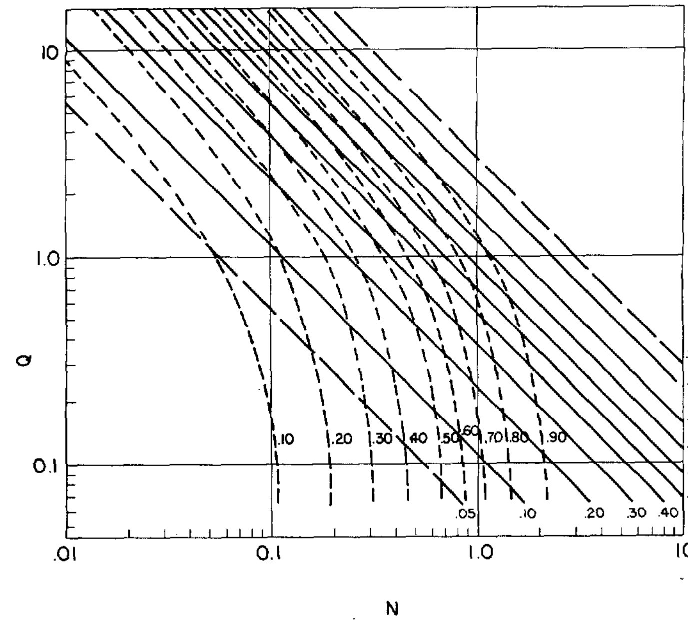

FIGURE 2.—The straight solid lines are lines of equal point probability of precipitation. The curved dashed lines are lines of equal area probability of precipitation. The ordinate Q is the area of each precipitation cell. $Q=1$ indicates a cell size equal to the forecast area. N is the number of cells per forecast area.

These relationships are shown in figure 2. For large Q, P_a and P_p approach each other asymptotically. In other words, when the precipitation cells are large, if precipitation occurs at any point in an area it is likely to occur at all points in the area.

It is reasonable to expect a seasonal variation in cell size. During summer, when convective type storms predominate, Q will be small and area probabilities near 1 may correspond to quite small point probabilities of precipitation (cf. McDonald, [1]). During winter, when large cyclonic systems produce much of the precipitation, Q will be large and area and point probabilities will not differ much.

These arguments can be extended to probabilities of precipitation at two or more stations. If the separation between the stations, δ, is greater than the diameter of the precipitation cells, then from the earlier assumption of randomness, the probabilities of precipitation at the two stations are independent.[3] If events A and B represent, respectively, the occurrence of precipitation at stations A and B, then, for $\delta \geq 2(Q/\pi)^{1/2}$,

$$\text{Prob}\ \{A\} = \text{Prob}\ \{B\} = P_p,$$

$$\text{Prob}\ \{A, B\} = P_p^2,$$

$$\text{Prob}\ \{A+B\} = 2P_p - P_p^2 = 1-(1-P_p)^2.$$

[3] If we had not assumed circular cells, but instead elongated cell tracks, we could not have made this statement. In that case the orientation of the tracks and stations becomes important and leads to a far more complicated analysis.

However if $\delta < 2(Q/\pi)^{1/2} = 2r$, then the area in which a cell must occur to give precipitation is that enclosed by the intersecting circles in figure 3. The area of this region is

$$\gamma = 2Q - 2(Q/\pi)\cos^{-1}(\delta/2r) + \delta[r^2 - (\delta^2/4)]^{1/2}$$

and the probability of precipitation at either A or B is

$$P_2 = \text{Prob}\ \{A+B\} = 1 - e^{-N\gamma},$$

$$= 1-(1-P_p)^{2-(2/\pi)\{\cos^{-1}(\delta/2r) - (\delta/2r)[1-(\delta/2r)^2]^{1/2}\}}.$$

Curves of P_2 versus $\delta/2r$ are shown in figure 4 as solid curves for various values of P_p. Also plotted in this figure are the conditional probabilities, $P_c = 2-(P_2/P_p)$, of precipitation at one station given the occurrence of precipitation at another station at a distance δ.

Similar extensions to any network of stations are possible. All that is necessary is a knowledge of the geometry of the station network. Some results of $k \times k$ square arrays of stations are shown in figure 5.

From the point of view of forecasting precipitation probabilities, direct estimation of point probabilities may well be the forecaster's most effective procedure. The forecaster could then infer, from equation (1), the product NQ. Given additional information (climatological, synoptic, etc.) on the expected cell size, the forecaster could then also infer from his estimate of P_p such things as N and P_a.

On the other hand, the forecaster may ultimately find it easier to forecast N, or at least its expected value. If

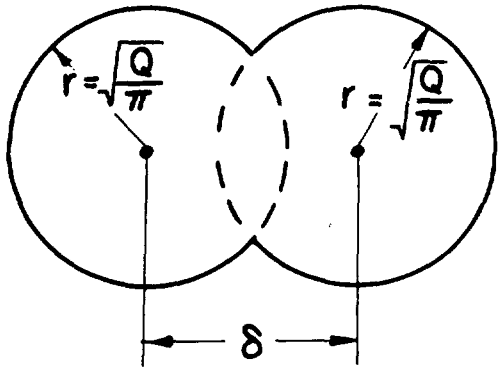

FIGURE 3.—Geometry for determining the probability of precipitation at station A or station B when δ, the distance between the stations, is less than the cell diameter $2r$.

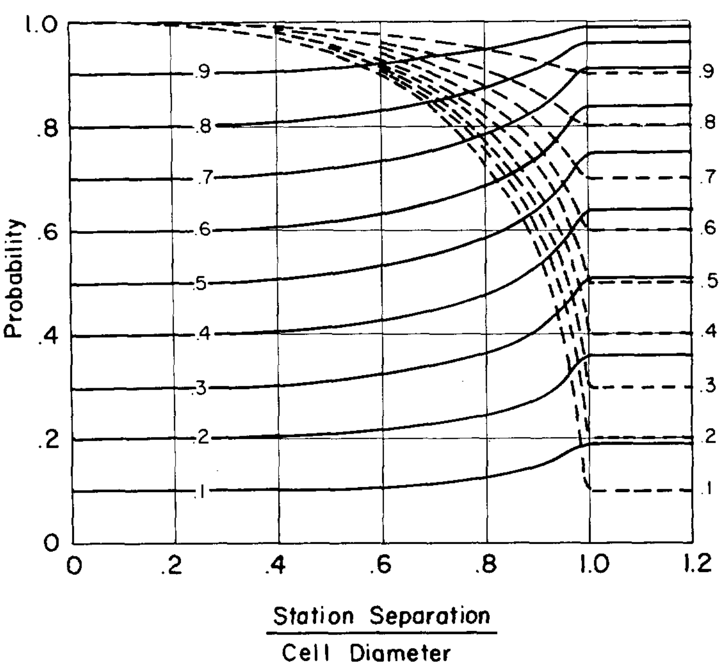

FIGURE 4.—The probability of precipitation at either of two stations (solid curves) and the conditional probability of precipitation at one station given the occurrence of precipitation at another station (dashed curves). Both sets of curves are plotted against the ratio of the distance between the stations to the cell diameter and are labeled according to P_p, the point probability of precipitation.

one assumes that Q is known, and treats N as a random variable, assuming some simple probability distribution, it becomes possible to infer, from the forecast of the expectation of N, $E(N)$, the expected values of P_p, and also P_a.

For example, let us consider that N has an exponential distribution with mean ν, i.e.,

$$\text{Prob}\,\{N_1<N<N_2\}=\int_{N_1}^{N_2}(1/\nu)e^{-N/\nu}dN.$$

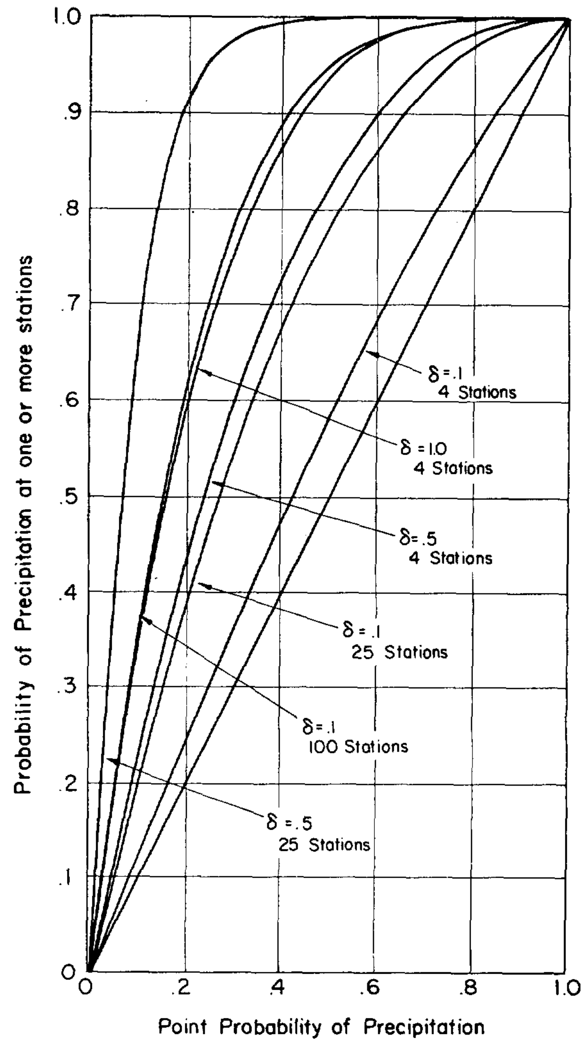

FIGURE 5.—The probability of precipitation at one or more stations in a square array, as a function of the point probability of precipitation. The curves are labeled according to the ratio of the station spacing to the cell diameter and the number of stations in the network.

Then

$$E(P_p)=\int_0^\infty(1-e^{-NQ})(1/\nu)e^{-N/\nu}dN,$$
$$=Q\nu/(Q\nu+1) \tag{3}$$

and

$$E(P_a)=\int_0^\infty\{1-e^{-N[1+Q^{1/2}]^2}\}(1/\nu)e^{-N/\nu}dN,$$
$$=[1+Q^{1/2}]^2\nu/\{[1+Q^{1/2}]^2\nu+1\}. \tag{4}$$

Graphs of EP_p and EP_a against ν and Q are shown in figure 6.

If the probability distribution for N were somewhat more general, specifically a gamma distribution with mean ν and variance σ^2, then

$$E(P_p)=1-[(\sigma^2Q/\nu)+1]^{-(\nu/\sigma)^2}$$

and

$$E(P_a)=1-\{(\sigma^2/\nu)[1+Q^{1/2}]^2+1\}^{-(\nu/\sigma)^2}.$$

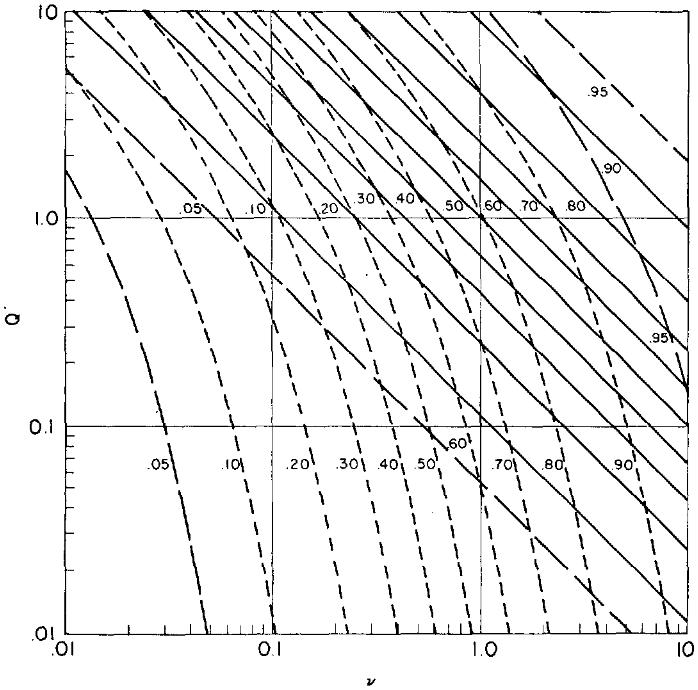

FIGURE 6.—The expected point probability (straight solid lines) and area probability (curved dashed lines) of precipitation when the number of cells per forecast area has an exponential distribution with expected value ν, and Q is the area of each cell.

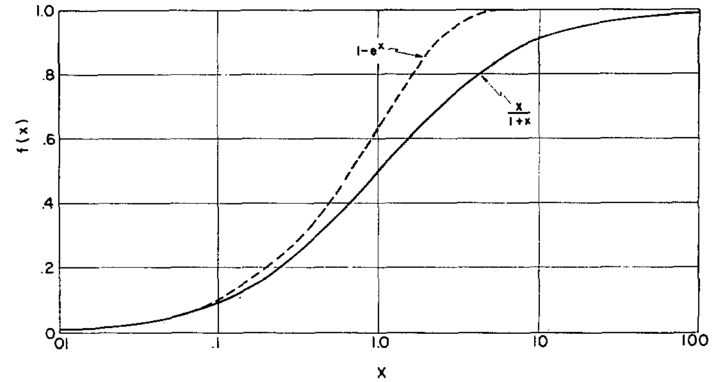

FIGURE 7.—The dashed curve gives the point probability of precipitation, P_p, if x is taken as QN, and the area probability of precipitation, P_a, if x is taken as $[1+Q^{1/2}]^2 N$. The solid curve gives $E(P_p)$ and $E(P_a)$ when x is interpreted as $Q\nu$ and $[1+Q^{1/2}]^2 \nu$, respectively, assuming N has an exponential distribution with $E(N)=\nu$. If N has a gamma distribution with $E(N)=\nu$ and $\text{Var}(N)<\nu^2$, the curve of $E(P_p)$ and $E(P_a)$ would lie between the two curves shown. For Var $(N)>\nu^2$, the appropriate curve would lie below those shown. CORRECTION: Change dashed curve label to $1-e^{-x}$

This reduces to the case for the exponential distribution when $\sigma=\nu$. It also approaches equations (1) and (2) as σ approaches zero, implying $N=\nu$ with certainty.

Finally, consider a comparison of the results for known N (equations (1) and (2)) with those when N is a random variable with an exponential distribution and expected value ν (equations (3) and (4)). The functional forms of these equations are similar, making it possible to show these relationships in a single graph. Figure 7 is basically a graph of two functions, $1-e^{-x}$ and $x/(x+1)$, versus x. For point probabilities, one can interpret x as QN or $Q\nu$. Then the two curves give the point probabilities of precipitation when N is known and when N is a random variable, respectively. For area probabilities, the same is true if x is interpreted as $[1+Q^{1/2}]^2N$ or $[1+Q^{1/2}]^2\nu$.

The difference between the ordinates represents the penalty for uncertainty, at least when the uncertainty is expressed in terms of an exponential distribution for N. For example, if QN is known to be 1.0, $P_p=0.632$. However, if N has an exponential distribution such that $QE(N)=Q\nu=1.0$, then $E(P_p)=0.5$. One is unable to state as high a probability of precipitation because of the strong implied probability that $N<\nu$ which outweighs the effect of the less probable occurrence of $N>>\nu$.

REFERENCE

1. J. E. McDonald, "'It Rained Everywhere But Here!'—The Thunderstorm Encirclement Illusion," *Weatherwise*, vol. 12, No. 4, Aug. 1959, pp. 158–160, 174.

[*Received July 20, 1966*]