

THE CONSTRUCTION OF A PROBABILITY
DISTRIBUTION FOR RAINFALL ON A
WATERSHED BY SIMULATION

by

Gary Williamson & Donald Ross Davis
Systems & Industrial Engineering
University of Arizona
Tucson, Arizona 85721

ABSTRACT

A raingage reading is a sample from the point rainfall population of an area. The actual average rainfall on the area (watershed) is a conditional probability distribution. For the case of thunderstorm rainfall this distribution is simulated by looking at all storms that could have produced the raingage reading. The likelihood of each storm is a function of its center depth. The amount of rain dumped on the watershed by each storm is weighted by the likelihood of its occurrence and the totality of such calculations is used to produce a probability distribution of rainfall on the watershed. Examples are given to illustrate the versatility of the program and its possible use in decision analysis.

INTRODUCTION

Areal rainfall, average rainfall over a watershed, is the consequence of a rainfall event and is estimated from raingage readings. In this paper we shall describe how simulation may be used to conduct experiments on a model of a watershed in lieu of direct

experimentation with the system itself. Such experiments will be used to construct an empirical probability distribution function (p.d.f.) on a watershed.

When the number of raingages on or near a watershed is small the estimate of areal rainfall may be in considerable error, especially if the rainfall event is a thunderstorm. Watershed models such as the Kentucky model (James, 1970) and the U.S. Geological Survey's flood discharge model (Dawdy and Bergman, 1969) give runoff predictions that become increasingly uncertain as the number of raingages on a watershed declines and as the location of the raingage moves away from the approximate center of the watershed.

A raingage reading may be viewed as a sample of the rainfall on the watershed. The amount of areal rainfall on a watershed can vary between extremely wide limits while the point rainfall over a single gage remains constant. Given a rainfall model this variability in the areal rainfall for a given raingage reading and location may be quantified as a p.d.f., by simulation. The quantification of the uncertainty in areal rainfall enables the use of decision theoretic methods (Davis et al, 1972) in the solution of water resource problems where areal rainfall, as estimated by a small number of raingages, is a factor.

In this model, we are going to determine the areal rainfall over a watershed by simulating the interaction among the variables of raingage reading, raingage location, and storm location.

The practice of simulation is rapidly growing and yet many people who rely on the results of simulation studies are not aware of exactly what

is involved. For this reason, we are going to describe the simulation approach to this model in some detail so that it may be seen how the individual equations and operations are built into a simulation model. This procedure must be understood to insure proper use of simulation results and is of course necessary if one wants to define their own model. Once the concept is understood, it may be readily applied to a system of most any complexity.

THE MODEL ENVIRONMENT

We will consider the relations between a raingage reading and its location, the storm location and height, and the areal rain on the watershed. The objective in doing so might be to most accurately determine runoff, determine peak flows, to minimize some economic criteria, or many other possibilities. The objective for this paper is to determine a probability distribution for the areal rainfall on the watershed, given the raingage reading at a particular location.

If this were to be done in the field first the watershed would have to be densely gaged in order to determine the areal rainfall from each storm. Then we would have to wait until a storm came along, take measurements and record the raingage reading and the areal rainfall. When we had enough data to represent all possible storms for all possible raingage readings an empirical p.d.f. of areal rainfall conditioned on raingage reading could be calculated. If this information was desired for a location that was not gaged the study would have to be repeated. Such an approach is clearly horrendous so we turn to the technique of simulation to approximate the watershed system.

At this point it must be stressed that simulation is only as good--meaning accurate, reliable, and useful--as the model and program behind it. Too many people are misled into believing that the techniques of simulation coupled with a large computer will somehow validate the results. The model and resultant simulation is only a conception of how the system behaves.

In our simulation we could have mimicked reality directly and allowed the simulation to proceed as follows:

1. Choose a random location and random height for a storm from some probability density function.
2. Determine the raingage reading given the distance from raingage to storm center.
3. Determine the average raingall on the watershed.

As in the field situation, we might repeat the process an indefinite number of times until the required statistics stabilized. Then we could move the raingage in our model and repeat the sequence. All that we would gain is the ability to examine one storm after another while remaining dry in the computer lab.

This approach is called Monte Carlo Simulation (Emshoff and Sisson, 1970). Values are chosen from probability density functions and are entered into the model. Repeated runs are analagous to repeated sampling in an experimental situation.

To save computer time and to make the problem more tractable, we have constrained the simulation to perform in a predefined manner. Given the same parameters, equations, and criteria, the simulation will always

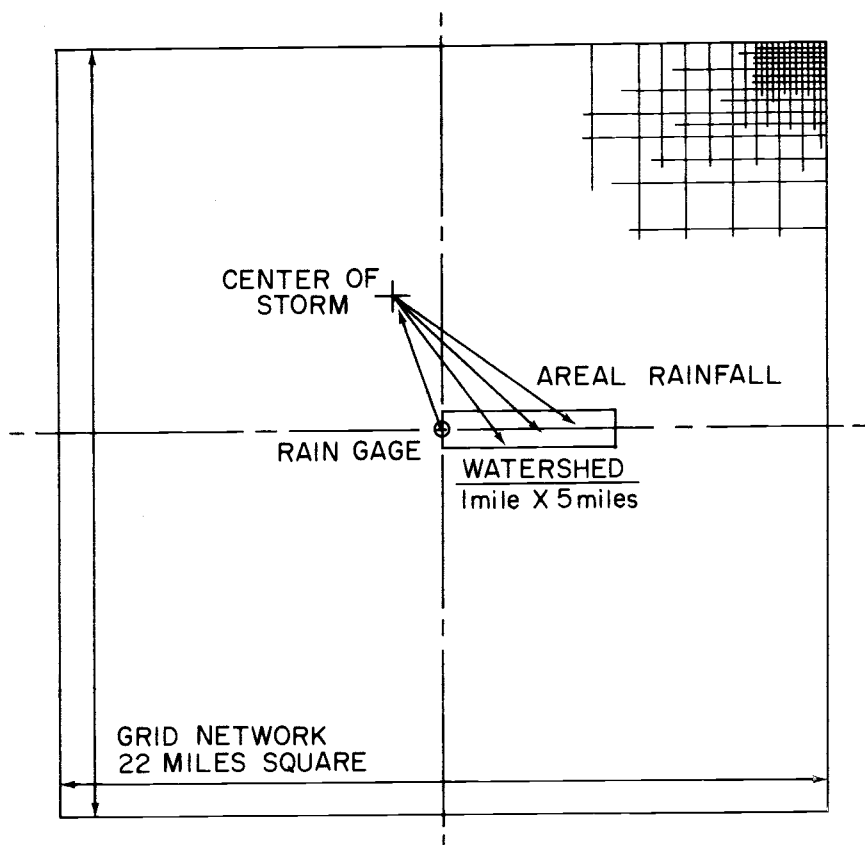
give the same output. This is not Monte Carlo Simulation but what is sometimes referred to as deterministic simulation.

THE SIMULATION

We examined a rectangular watershed of dimensions one mile by five miles. First the raingage was located on the long axis of the watershed at its western boundary. Storms were allowed to occur in a square 22 miles on a side centered on the raingage.

To reduce the effort involved we constrained the sampling. First we arbitrarily chose a raingage reading. For this reading there are an infinite number of storm heights and locations that may produce the raingage reading, each of which has its own effect on the watershed of interest. To keep calculating time under control we reduced the number of allowable storm locations. The 22 mile square was divided into a grid of 64 units on a side, as shown in Figure 1. For calculating purposes storm centers were located only on grid points. It was now easy to calculate the distance from the raingage to the storm center, the height of the storm at the center and the areal rainfall on the watershed. The first calculations were made for an 8 x 8 grid. Second calculations were made for a 16 x 16 grid. The number of equally distant grid lines on the side of the 22 mile square was doubled until the variance of the empirical p.d.f. obtained from the calculations had a variance within one percent of the variance calculated with the previous grid, or until a 64 x 64 grid was used.

For each grid point we calculated the height of the storm necessary to produce the raingage reading. The areal rainfall on the watershed was also



LAYOUT OF WATERSHED
FIGURE 1

calculated. The empirical p.d.f. of areal rainfall is not a histogram of the areal rainfalls so calculated, because each storm is not equally likely. The empirical p.d.f. is a histogram of areal rainfalls weighted by the likelihood of storm from which they came.

In the calculations presented here we have assumed the watershed to increase in height from 2,500 feet at the western end to 7,500 feet at eastern. The surrounding area in the east-west and north-south directions is level with the watershed.

We have used the Fogel-Duckstein (1969) model of thunderstorm rainfall:

$$R = R_0 e^{-b[(x - x_0)^2 + (y - y_0)^2]}$$

where R is the rainfall amount at the raingage coordinates (x, y) and R_0 is the rainfall at the storm center (x_0, y_0) and $b = 0.27 e^{-.67R_0}$

The model is modified to reflect the higher relative likelihood of storm occurring above the base level of the watershed. Assuming a base level of 2500 feet, the likelihood previously calculated was multiplied by the factor

$$1.0 + 0.0002727(h - 2500)$$

where h is the height as defined by a function intersecting the western boundary at 2500 feet and the eastern boundary at 7500 feet.

Otherwise, the depth of storm centers was assumed to be Gumbel distributed with a mode at 0.9 inches. Rainfall at the raingage assumed to be distributed exponentially with a mean of 0.25 inches.

These assumptions are made to model a typical small watershed in the foothills of Tucson, Arizona.

For the raingage reading, the height of a storm on the grid point is determined by the numerical analysis technique of Newton-Raphson. Double integration with Simpson's Rule is used to establish the area rainfall. The rainfall at each point within the watershed is calculated by substituting the distance to the storm center in the rain equation. A schematic representation of this sequence is shown in Figure 2. Two additional steps which are discussed later are also shown.

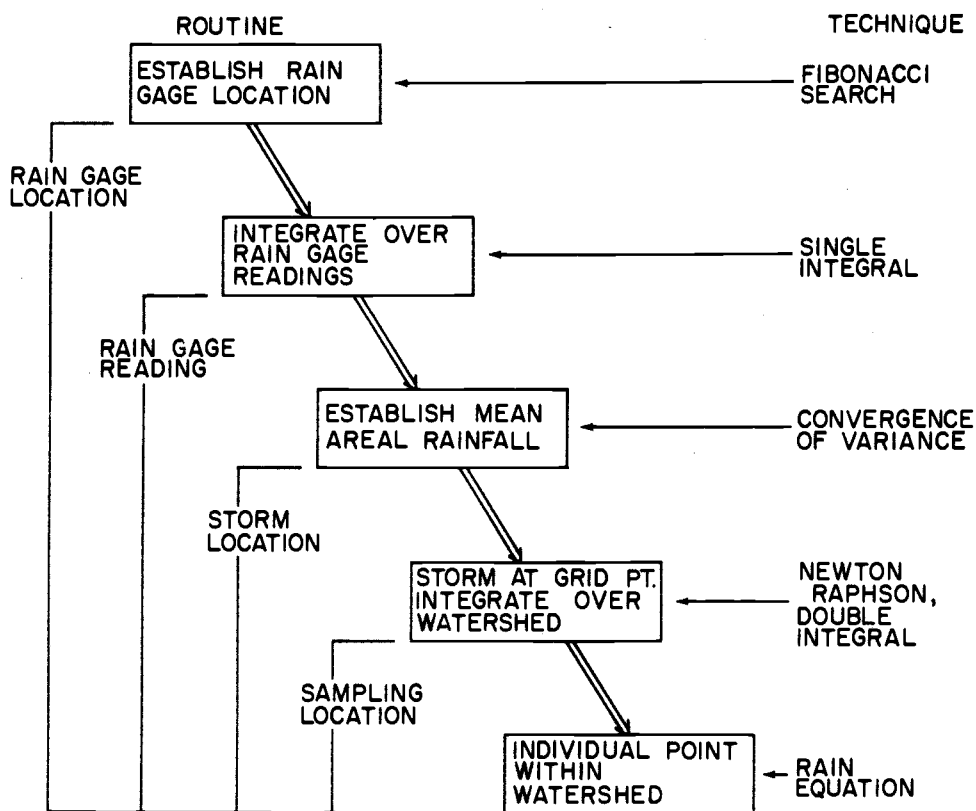
DATA DISPLAYS

For each point on the grid that is processed, we can record and plot the storm height, the likelihood, and the areal rainfall on the watershed. This is for a single reading on the raingage at a specific location.

We arbitrarily limited the maximum storm height to 4.0 inches. Given a raingage indicating 0.4 inches located at the western boundary of the watershed a plot of the storm heights would produce a convex surface with the raingage in the center. The maximum height of 4.0 inches occurs at a distance of just under six miles from the gage.

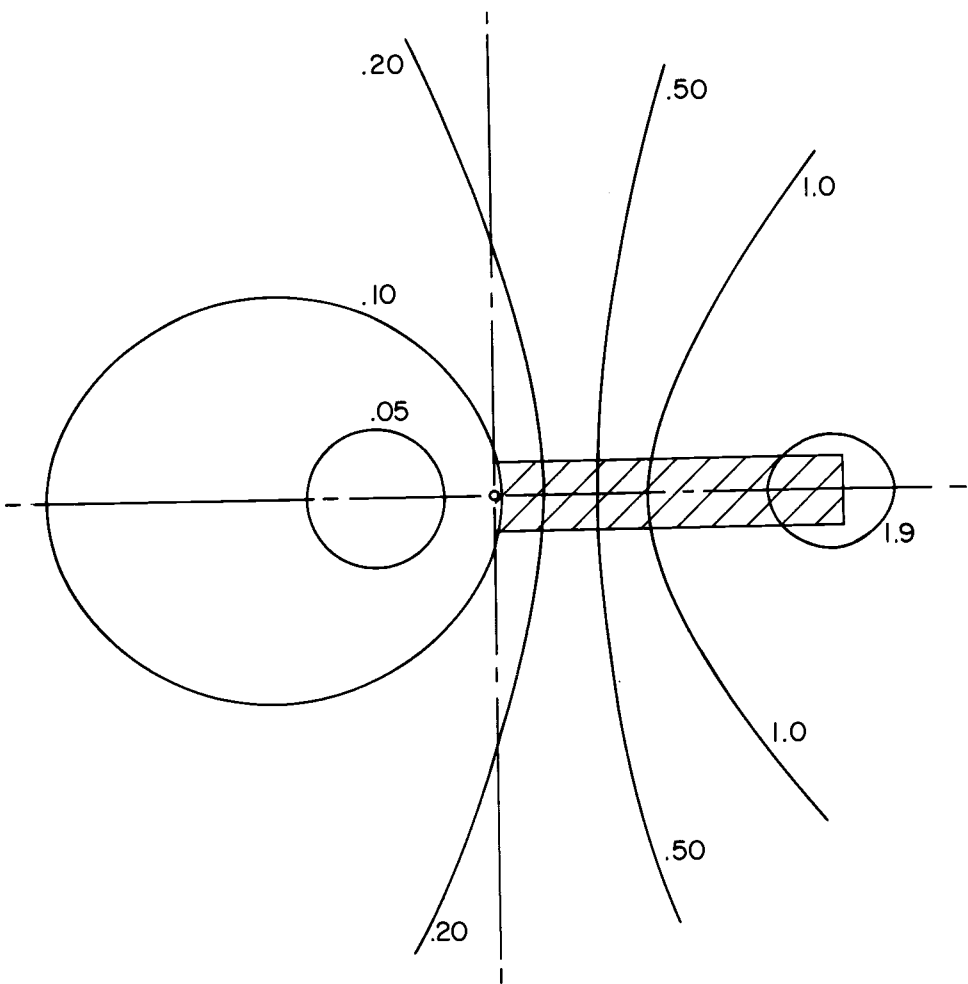
Figure 3 shows a topographical map of the areal rainfall of the watershed for this example. The countour lines represent lines of equal average rainfall over the watershed. The large circular shape to the left represents an average rainfall of 0.1 inches. Although the storms along this line grow larger as they move away from the raingage, they have a decreasing effect on the watershed. The converse effect occurs over the watershed. Near the right boundary, a storm must be quite large to produce a raingage of 0.4 inches and the effect of such a storm on the watershed is severe.

SEQUENCE OF PROGRAM CALCULATIONS



VARIABLE IS FIXED FOR STEP TO THE RIGHT

FIGURE 2



CONTOUR LINES OF EQUAL AREAL RAINFALL ON
WATERSHED

FIGURE 3

A plot of the likelihoods is more complex. A "crater" effect is seen where the most likely values are located at approximately 1.30 miles from the raingage. The likelihood falls off in both directions away from this radius. This is slightly biased over the watershed due to the adjustment for elevation.

Areal rainfall, given a raingage reading, is a conditional probability distribution function. Figure 4 presents the p.d.f. for the average rain on the watershed in the example.

EXTENSIONS

The empirical p.d.f. we have obtained may be used in conjunction with other techniques, such as watershed models, to evaluate the results of the rainfall (Parmele et al, 1972, for a closely related example). It is desired to locate the raingage so as to minimize the uncertainty in this evaluation (uncertain rainfall implies uncertain runoff). Since the variance of the p.d.f. is a crude but easily calculated indicator of the uncertainty we have calculated the variance of the p.d.f. obtained at various locations on the watershed assuming 9.4 inches in the raingage. The results are in Table 1.

We have found the optimum location for the raingage assuming the raingage reading. In reality we don't know what the raingage will read. However we usually have enough information to infer a probability distribution for the raingage reading. We have used an exponential with mean of 0.25 inches, with mean increasing for altitude in the same manner as the likelihood of a storm. Using this distribution we calculate the expected variance at each location. The expectation is calculated by Gaus-LeGuerre integration.

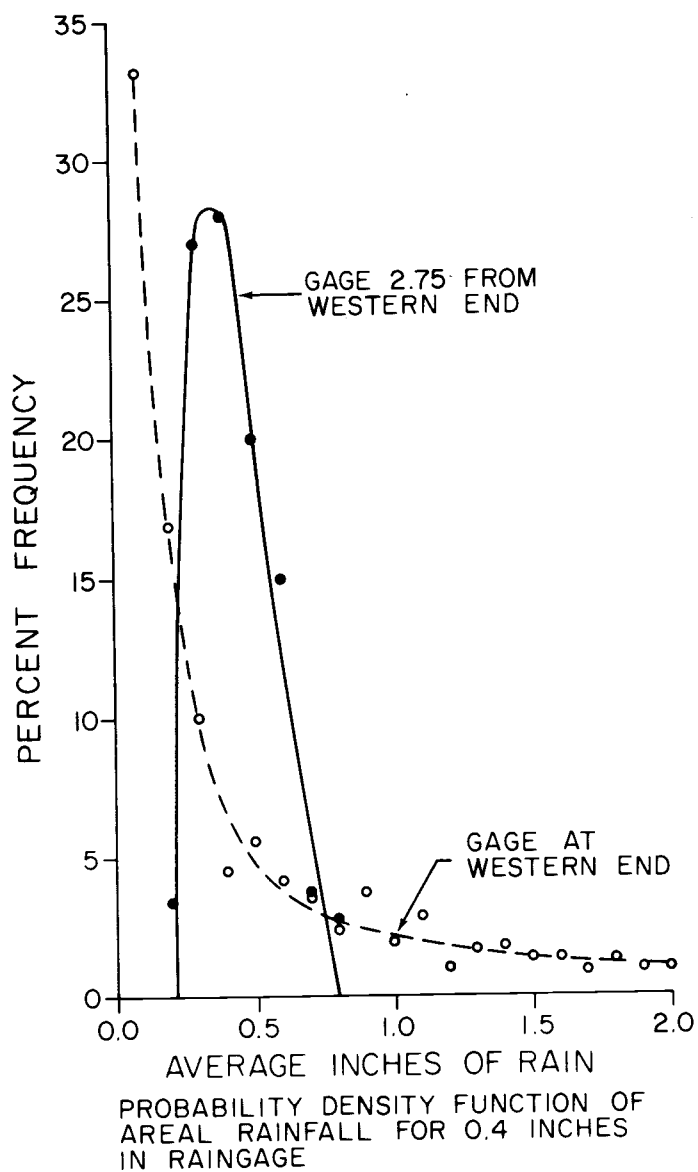


FIGURE 4

When this integration is performed, we have a single value for some location of the raingage. Now we move the raingage to a new location and repeat the simulation in its entirety to obtain a new value. Using a Fibonacci search, we can find the location of the logitudal axis where the variance is a minimum. This is shown in the scheme presented in Figure 2.

For the watershed that has been previously described, this location of minimum average variance is 2.66 miles from the western boundary. This result is hardly surprising because of the altitude correction factor. But our primary purpose was not to find this location, but to develop the single gage watershed model. With very little effort, we can define watersheds of different and possibly irregular shapes, different elevation factors including a contour way if necessary, and any storm model that we chose. Somewhat more difficult modifications would include multiple gaging and allowing the storm to move over some trajectory. Various objective functions might also be expected.

Distance from Western Edge	TABLE 1 Mean	Variance
0	.43	.230
1.25	.44	.093
2.50	.40	.018
2.75	.39	.017
3.00	.37	.020
3.75	.33	.050
5.00	.27	.120

Mean and Variance of Simulated p.d.f.'s with raingage reading of 0.4.

CONCLUSIONS

We have shown how a rainfall model may be broken into segments and implemented on a computer. In doing so we developed probability density functions for areal rainfall given the raingage reading and storm location. The criteria of minimum expected variance was then used to find the optimum raingage location.

This simulation was designed to provide information needed for a decision theoretic analysis. From a sample, a raingage reading, we developed a p.d.f. representing the possible amounts of areal rainfall on a watershed. We then found the optimum location to minimize the expected variance of this p.d.f. The raingage location could have been chosen to minimize other functions of the p.d.f.

Simulation is an extremely evaluable tool, especially when other methods of analysis fail. The process must be understood to insure that a model is defined and implemented correctly.

Acknowledgment. The results of this research were partially supported by a Title II grant on "Use of Models in Validation of Hydrologic and Related Data" from the Office of Water Resources Research, U.S. Department of the Interior, under grant No. 14-31-0001-3708.

REFERENCES

- Davis, D.R., C.C. Kisiel and L. Duckstein, "Bayesian Decision Theory Applied to Design in Hydrology", Water Resources Research, No. 1. 8, pp. 33-41, 1972.
- Dawdy, D.R. and J.M. Bergmann, "Effect of Rainfall Variability on Streamflow Simulation", Water Resources Research, No. 5, 5, pp. 958-966, 1969.
- Enshoff, J.R. and R.L. Sisson, Design and Use of Computer Simulation Models, McMillian C., New York, 1970.
- Fogel, M. and L. Duckstein, "Point Rainfall Frequencies in Convective Storms", Water Resources Research, No. 6, 5, pp. 1229-1237, 1969.
- James, L.D. "An Evaluation of Relationships Between Streamflow Patterns and Watershed Characteristics"; Through the Use of OPSET (A self-calibrating version of the Stanford Watershed Model), Research Report No. 36, Water Resources Institute, University of Kentucky, Lexington, Kentucky, 117 pp. 1970.
- Parmelee, L.H., E.T. Euyman and R.L. Hendrick, Proceedings, Second Symposium of Meteor Observations and Instrumentation, American Meteor. Society, San Diego, March 1972.