A MODIFIED MONTE CARLO TECHNIQUE TO SOLVE THIESSEN COEFFICIENTS

by

Sun-Fu Shih & Robert L. Hamrick
Central and Southern Florida Flood Control District
West Palm Beach, Florida 33402

ABSTRACT

A modified application of Monte Carlo methods to solve the Thiessen Coefficients was developed and also extended to compute the new coefficients for the precipitation with missing record. A general principle is to choose the boundary segments of the entire watershed which represents the planform of the watershed by a polygon with as few sides as possible without changing the basic shape of the boundaries. A new technique was demonstrated to determine where a random point falls within an arbitrary shaped boundary. The convergence of the computed weights based on the number of random trials and the relative area ratio of the watershed boundaries to the enclosing rectangle was discussed. If the number of random points falling outside the boundary of the watershed differs greatly from that of the enclosing rectangle, then the three techniques of equal rectangles; unequal rectangles; and single rectangle can be used to obtain the computed weights.

INTRODUCTION

The average depth of precipitation over a study area, either on storm, monthly, seasonal, or annual basis, is required in many types of hydraulic problems. In general, Linsley, et.al (1958) indicated that there are three common methods that are available to use. First, the simplest one is called the arithmetic mean method, and this method gives good estimates in flat area under the condition of gages that are uniformly distributed and the individual gage catches do not vary widely from the mean. However, this method does not take into account the stations outside, but near the boundaries of the area. Second, the Thiessen Method (1911) attempts to allow for nonuniform distribution of gages by providing a weighting factor for each gage. Horton (1923) found that the results obtained by the Thiessen Method are usually more accurate than those obtained by simple arithematical averaging. But, Linsley et.al. (1958) said that the greatest limitation of the Thiessen Method is its inflexibility, because a new Thiessen diagram is required every time there is a change in the gage network. Third, the isohyetal method is highly dependent upon the skill of the analyst to construct the isohyetal map. If linear interpolation between stations is used, the results will be essentially the same as those obtained with the Thiessen Method. Moreover, an improper analysis may lead to serious error. This method also involved a degree of subjectivity because a different user could draw different isohyetal lines from which a different area rainfall can be obtained under the same rainfall event. Thus, the Thiessen Method is less dependent upon the skill of the analyst and can be easily performed by Monte Carlo techniques as reported by Diskin (1969). A more detailed description of Monte Carlo application in the fields of science and engineering is summarized in books by Hammersley and Handscombe (1964) and Shreider (1967). However, the Monte Carlo technique as reported by Diskin have some limitations to use in practical applications. For example, limitations of application to a watershed with arbitrary shaped form, perform the convergence of computed weight, and to divide a subarea for a hypothetical L shaped watershed, etc. are needed to be modified to that the Monte Carlo application can be more widely adopted.

The purposes of this study are:

(1) to demonstrate a new technique used for determining the random point position;

(2) to show a new concept for performing convergence of weights;

(3) to introduce a general rule for choosing the boundary segiments in the entire watershed;

(4) to devise a new technique for estimating a hypothetical L shaped watershed; and

(5) to extend this Monte Carlo technique to estimate the new weighting factors for some rainfall stations with missing data.

METHODOLOGY DESCRIPTION

The following five terms concerning the methodology of Monte Carlo applications are described as follows:

A. New Technique to Determine a Random Point Position:

Diskin (1969) presented a method which can be used to check a random point falling within the boundary. However, the watershed boundaries in Figure 1 not only show that the boundary has an irregular shaped form, but also a shaded area exists within the boundary. The computer program as developed by Diskin is used to find the area of Figure 1 and comparing the results with those obtained by the graphical procedure. The results are 43.9 and 36.4 in Diskin's program and graphical method, respectively. As can be seen from this result, the deviation between these two methods is about 20 percent. In other words, the Diskin method does not appear to be suited to estimating areas in watershed with complicated boundaries. Therefore, Shih and Hamrick (1973) developed a new technique to determine whether a random point falls within an arbitrary shaped boundary.

This method is based on the concept that, given any completely bounded region, a radial line constructed in any direction from a given point must cross the boundary an odd number of times if the point is located within the boundary or make an even number of intersections if it is located outside the bounded entity (assuming zero to be an even number). This is true with any degree of deformation of the boundary. It will also account for bounded, excluded subregions completely surrounded within the region. In general, the result for any stated point is ambiguous only in the case where the point is co-incident with the boundary. Some arbitrary ruling must be made in these cases.

The specific application here is to the two-dimensional case. The boundary is defined by a series of linear segments between node points, such that the accuracy of representation is suitable for the particular use at hand. Rule 1 below gives the methodology for determining whether the randomly generated point falls within the boundary or not. Rule 2 solves the ambiguity that exists when the radial line penetrates the boundary at the node point. Passing through a node point may possibly be scored as even number of intersections (0 or 2) or an odd number of intersections (1) as illustrated.

A computer program relying on these two rules and a detailed technique of application introduced by Shih and Hamrick (1973) were used to estimate the area of Figure 1 as given. The result is 36.8 which is very close to that obtained by a normal graphical method. The deviation is almost negligible.

B. Convergence of Weights:

Diskin's paper indicated that the computations are organized on cumulative sets of randomly generated points. Each set of computations contains as assigned arbitrary large number of points. At the end of each set, the weights of the stations are computed and compared to the weights obtained at the end of the previous set. The computations are terminated when the differences between each of the new values of the weights and the previous values are less than an arbitrarily set small quantity. A later example will show this concept of convergence is not quite applicable in some cases.

Because the Monte Carlo method relies on the laws of probability, a large number of random walks should be taken if the result is to closely approximate actual value. The common method used to estimate the sample size and accuracy is the large sample normal approximation. Using the Central Limit Theorem the binomial distribution can be approximated by a normal distribution for a large N where N is the total number of trials. It also shows that a sample error is proportional to $1/\sqrt{N}$. The convergence is a statistical convergence, i.e., the probable error is proportional to $1/\sqrt{N}$. For example, a watershed as shown in Figure 2 with eight rainfall stations was used to discuss this concept of convergence of weights. The results of the computed weights and relative area ratio formulated by Diskin's computation are given in Table 1. As can be seen from Table 1, Diskin said the computations are terminated when the differences between set 8 and set 9 are less than an arbitrarily small quantity such as 0.001. However, according to the Central Limit Theorem the more sample sizes taken, the more accuracy of estimation can be obtained. For instance, combining the number of sets shown in Table 1 into a form with more number of random points such as 4000, 6000,..., 18000 random points gives a new relative area ratio and computed weights as shown in Table 2. The results of Table 2 indicate that the 4000 points will give a more reliable estimation because the statistical error of $1/\sqrt{N}$ is 1.581% in 4000 points, and 2.236% in 2000 points. It is obvious that 4000 points is better than 2000 points. In a similar case, the statistical errors are 1.291%, 1%, and 0.745% in 6000 points, 10000 points, and 18000 points respectively. In general, if 2% of the statistical error is admitted, then a number of trails from 2000 to 4000 is recommended. It should be noted that one must increase the random point by a factor of 4 in order to halve the error. Therefore, the difference of the relative area ratio between set 8 and set 9 in Table 1 is less than 0.001, from which a convergence of weights was concluded by Diskin. The difference between two consecutive sets within a tolerance is a case of chance only and not a case of consistent phenomenon. In other words, if a different group of random numbers is used, the different relative area ratio within a tolerance 0.001 may be non existant between set 8 and set 9. Another example of watershed as shown in Figure 3, will show that the Diskin's concept of convergence of weights is not applicable. The results of Figure 3 indicate that 113 sets were run and the differences between two consecutive sets were never less than the tolerance 0.001. The computing time was more than an hour in this test, and a different number of random points were used to test the sensitivity of convergence. The results of computed weights and relative area ratio difference between different random points admitted are shown in Table 3. As shown by Table 3, the convergence is very slow after the random point is over 20000 points. Therefore, a computing process with too many sets or too many random points, is not recommended in practical application because a longer computing time is spent and little accuracy of estimation is improved. In general, 2000 to 6000 random points can give a quite good results. A detailed discussion of the convergence concept in practical application will be explained in a later section.

B. Selecting Boundary Segments:

Diskin (1969) defined the boundaries of the watershed by giving the coordinates of successive points along the boundary (in a clockwise direction) and considering the boundary between each pair of successive points to be a straight linear segment. The actual boundaries could be approximated as closely as desired by increasing the number of such points. But, the user should note that the more segments chosen, the more computing time and user's time is required in setting up the more segments. A later example will show the effects of the many irregularities of natural boundaries such as lake shorelines, and the watershed may be averaged by reducing the boundary planform to a simplified polygon. The general principle of this procedure is to represent the planform by a polygon with as few sides as possible without changing the basic shape of the boundary. An example of this is the Upper Kissimmee River Basin as shown in Figure 3 which has two different cases of selecting boundary points, i.e., 103 segments and 10 segments. Comparision of the computed weights, relative area ratio, and computing time differences between 2000 and 6000 random trials are given in Table 4. As can be seen from Table 4, although the boundary segments are reduced from 103 segments to 10 segments, the results of computed weights and relative area ratio do not change appreciable, but the computing time can be reduced about 60-70%.

D. Dividing a Large Irregular Shaped Watershed:

An expected accuracy of computation by this Monte Carlo method depends upon not only the number of random points, but also the shape of the watershed. The number of random points that affects the sample error has been discussed in a previous section. However, the efficiency of this Monte Carlo method affected by the shape formed by the watershed should also be discussed because a large number of random points fall off the outside boundaries of the watershed which differs greatly from that of the enclosing rectangle. This difficulty can be overcome by following three techniques:

(1) Equal Rectangles: Diskin (1969) suggested that the watershed is enclosed by a number of smaller rectangles of equal area that have a common edge which cuts the watershed. In order for this method to be used more widely, the following relationship should be introduced. Let A_1, ..., A_n be the relative area ratiofalling within the boundaries of sub-watershed 1, ..., n. Then the new relative area ratio, R_1, ..., R_n, of sub-watershed 1, ..., n are equal to

$$R_1 = A_1 / \sum_{i=1}^{n} A_i \qquad (1)$$

$$\vdots$$

$$R_n = A_n / \sum_{i=1}^{n} A_i$$

The final computed weights, W_1, ..., W_m, of raingall stations 1, ..., m are equal to

$$W_1 = \sum_{i=1}^{n} R_i E_{i1}$$
$$\vdots \qquad \vdots$$
$$W_m = \sum_{i=1}^{n} R_i E_{im} \qquad (2)$$

where E_{ij} is the computed weights of rainfall station j in subrectangle i; j includes the rainfall station from 1 to m and i includes the subrectangle from 1 to n. For example, m is the total number of rainfall stations and n is the total number of subrectangles. A hypothetical L shaped watershed with five rainfall stations as shown in Figure 4 was divided into two equal rectangles. Such as efgh and opik; i.e., n=2 and m=5. The results of the computed weights of 2000 random trials, 4000 trials, and 6000 random trials are shown in Table 5. In 2000 random trials, the values of R_1, R_2, W_1, W_2, W_3, W_4 and W_5 are calculated by equations 1 and 2, i.e.

$$R_1 = 0.545/(0.545 + 0.573) = 0.487477$$
$$R_2 = 0.573/(0.545 + 0.573) = 0.512522 \qquad (3)$$

$$W_1 = 0.314 \times 0.487477 + 0 \times 0.512522 = 0.153$$
$$W_2 = 0.686 \times 0.487477 + 0.087 \times 0.512522 = 0.379$$
$$W_3 = 0 \times 0.487477 + 0.068 \times 0.512522 = 0.035 \qquad (4)$$
$$W_4 = 0 \times 0.487477 + 0.609 \times 0.512522 = 0.312$$
$$W_5 = 0 \times 0.487477 + 0.236 \times 0.512522 = 0.121$$

These results compare quite favorably with graphical data. However, the procedure of selecting equal rectangles to include a similar weight in each rectangle is relatively difficult and time consuming especially for a watershed including more than two equal rectangles.

(2) Unequal Rectangles: The technique of unequal rectangles is similar to the equal rectangles method except that the watershed is enclosed by a number of smaller unequal rectangles. Let S_1, ..., S_n represent the area of enclosing rectangles 1, ..., n; $A_1, \ldots A_n$ and R_1, ..., R_n are defined in the case of equal rectangles. The value of R_i is

$$\begin{aligned} R_1 &= A_1 S_1 / \sum_{i=1}^{n} A_i S_i \\ &\vdots \\ R_n &= A_n S_n / \sum_{i=1}^{n} A_i S_i \end{aligned} \qquad (5)$$

The final computed weights, W_1, ..., W_m of rainfall stations, 1, ..., m, are similar to equation 2, except that the R_i values are replaced by equation 5. For example, the watershed as given in Figure 4 was divided into two unequal rectangles, i.e., efgh and mnik. If a scale as shown in Figure 4 is used, the values of S_1 = 45.36 square miles and S_2 = 30.8125 square miles are obtained. The values of R_1, R_2, W_1, W_2, W_3, W_4, and W_5 in 2000 random trials case are calculated as follows:

$$R_1 = \frac{0.495 \times 45.36}{0.495 \times 45.36 + 0.565 \times 30.8125} = 0.523269$$

$$R_2 = \frac{0.565 \times 30.8125}{0.495 \times 45.36 + 0.565 \times 30.8125} = 0.436730$$

$$W_1 = 0.274 \times 0.52369 + 0 \times 0.436730 = 0.154$$

$$W_2 = 0.655 \times 0.523269 + 0 \times 0.436730 = 0.369$$

$$W_3 = 0.006 \times 0.523269 + 0.074 \times 0.436730 = 0.036$$

$$W_4 = 0.065 \times 0.523269 + 0.665 \times 0.436730 = 0.327$$

$$W_5 = 0 \times 0.523269 + 0.271 \times 0.436730 = 0.118$$

The values of equation 5 are also shown in Table 5. Comparing these results of unequal two rectangles with the graphical method gives a good agreement. The random trials with 4000 and 6000 are performed in a similar way. The results are also shown in Table 5. These unequal small rectangles are more convenient to use as desired by changing the area of each subrectangle. However, additional time is needed in this method for dividing the entire watershed to obtain the computed weight of each station. In practical application, this method is only recommended in the case of a watershed such as new moon shape which is really needed to divide into several subrectangles. The reason for choosing unequal rectangles in a special shaped watershed will be discussed in the following section of single rectangle.

(3) Single Rectangle: Based on the Central Limit theorem, the more random points chosen, the greater the accuracy of the estimate obtained. As previously discussed, a number of random points between 2000 and 6000 is recommended in a general shaped watershed. This is true only if the relative area ratio of the watershed to the enclosing rectangle is not less than a certain percent. For example, the watershed shown in Figure 2 has about 68% as shown in Tables 1 and 2. The percentage implied that 68% of the total random point will fall within the watershed. If a watershed with relative area ratio is less than 0.3 and only 2000 random trials are used, the random points falling within the watershed will be less than 600 points. This may cause a bias estimation because the sample error $1/\sqrt{N}$, used in estimating the portion of the watershed area may be over 4 percent. If the relative area ratio is only 0.1, the error may be over 10 percent. Therefore, a watershed which has a lower relative area ratio, more random trials should be taken. For example, a watershed as given in Figure 4, the relative area ratio to the rectangle abcd is about 0.4, and the random trials with 2000, 4000, and 6000 are used in this study. The results of computed weights and relative area ratio in those

different number of trials are shown in Table 5. As can be seen from Table 5, comparing this single rectangle, abcd, technique with other techniques indicates that the single rectangle is as accurate as other methods. But, it should be realized that the single rectangle technique is relatively easy to use. Another example is the Kissimmee River Basin, Florida as shown in Figure 5, with 16 rainfall stations used to exemply the technique of a single rectangle. It should be noted that the coordinates chosen for enclosing the rectangle is an important factor. The principle of choosing coordinates for the enclosing rectangle can be performed as small as possible from which a more relative area ratio of watershed can be obtained. As can be seen from Figure 5, the x and y coordinates are chosen from the enclosing rectangle performed as small as possible. The results of these estimates with random trials of 2000, 4000, and 6000 are shown in Table 6. Comparing the single rectangle with the graphical technique as shown in Table 6 indicates that the single rectangle is applicable to a watershed with a long and narrow shape.

As mentioned in the section of unequal rectangles, a new moon shaped watershed from which the relative area ratio to the enclosing single rectangle may be less than 0.1. In that case, a user desiring to decrease the statistical error can be overcome by two ways. First, the user can increase the random points to 10,000 or more from which a 1000 random trial may be falling within the new moon shaped watershed. Second, dividing the new moon shaped watershed into two smaller subrectangles, and the relative subarea ratio to enclosing subrectangles may be assumed equal to 0.5. Therefore, the user should be able to justify in an economical way the accuracy of computation for special shaped watershed by either increasing the random points or dividing them into a set of smaller rectangles.

E. New Thiessen Coefficients for Missing Data:

As Linsley et.al. (1958) indicated the greatest limitation of the Thiessen method is its inflexibility, because a new Thiessen polygon is required every time there is a change in the gage network. This modified Monte Carlo method can be used to overcome this limitation. In general, there are two cases of missing data. Case 1: The missing data of each rainfall station are priorly known, and any missing period of record are assigned as a new station set. The distance of a random point from all rain measuring stations are calculated simultaneously in each station set, and the random point is assigned to the nearest rain measuring station in each set. Case 2: the missing data of each rainfall station is posteriorly known, i.e. how many stations with missing data are unknown. In this case, if a station with a missing record is found, then that station is omitted and a new gage network is considered. Based on this new gage network, a computed weights is performed by a repeating procedure. The detail description of these procedures will be discussed in the section of computer program.

A watershed of the upper Kissimmee River Basin is used in this study to test the applicability of this missing data case. The results of new computed weights for the missing data of each station in these eight stations are shown in Table 7. The results indicated that the modified Monte Carlo techniques are applicable to overcome the limitation of changing the gage network.

COMPUTER PROGRAM

As described in previous sections of new coefficients for missing data, the case 2 is more common in practical application. Based on this case 2 problem, the following procedures are developed to compute the weights of rainfall station:

(1) Enclose the watershed boundaries with a rectangle whose coordinates are also recorded.

(2) Read the x and y coordinates of the boundary segments.

(3) Compute the weighting factor of each boundary node according to Rule 2 introduced.

(4) Generate random points with uniform probability over the enclosing rectangle.

(5) Draw an imaging line from the random point and parallel to the x-axis.

(6) Count the number of intersections of this line with the boundary in either the left or the right hand side of the random point.

(7) Test whether the random point is falling within the boundary according to the Rule 1 indicated.

(8) If the above test fails, increase the counter of rejected points by one, otherwise, increase the counter of accepted points by one.

(9) If the random point is accepted, then the random point is assigned to the nearest station.

(10) Repeat the processes 5, 6, 7, 8, and 9 until a large number assigned is reached.

(11) Compute the relative area ratio of the watershed boundaries to the enclosing rectangle by dividing the accepted points by total number of random points.

(12) Calculate the computed weights of each station by dividing the assigned points to each station by the total accepted points.

(13) Check whether the record is missing data or not.

(14) If the check is yes, the processes from 4 to 13 are repeated.

The above procedure is simulated to a flow chart figure as shown in Appendix A. The flow chart was also converted to a computer program for CDC 3100 computer with Fortran IV language. The computer program is available from the authors.

SUMMARY AND CONCLUSIONS

A modified application of Monte Carlo methods to solve the Thiessen coefficients has been successfully developed. After applying this newly developed technique to several practical problems, the following conclusions were drawn:

(1) A new technique used for determining the random point falling within the watershed boundaries was demonstrated.

(2) A new concept to perform the convergence of computed weights was introduced. The convergence of computed weights depends upon not only the number of random trials but also the relative area ratio of the watershed boundaries to the enclosing rectangle. i.e., if the watershed boundary has less relative area ratio, the more number of random points should be run.

(3) A general principle to choose the boundary segments of the entire watershed is to represent the planform of the watershed by a polygon with as few sides as possible without changing the basic shape of the watershed boundaries.

(4) A large number of random points falling outside the boundary of the watershed which differs greatly from that of the enclosing rectangle can be overcome by three techniques - equal rectangles, unequal rectangles, and single rectangle.

(5) This modified Monte Carlo method has been extended to compute the new weighing factors for the missing data case.

ACKNOWLEDGEMENTS

Authors would like to express our appreciation to Mr. William V. Storch, Director of the Resource Planning Department, Central and Southern Florida Flood Control District, for his supervision and guidance.

REFERENCES

Diskin, M. H., Thiessen coefficients by a Monte Carlo procedure. Journal of Hydrology, Vol. 8 (3): 323-335, 1969.

Hammersley, J. M., and D. C. Handscombe, Monte Carlo methods, Methuen, London, 1964.

Horton, R. E., Accuracy of areal rainfall estimates, Monthly Weather Review Vol. 51 (7): 348-353, 1923.

Linsley, R. K. Jr. M.A. Kohler, and J. L. H. Paulhus, Hydrology for Engineers, McGraw - Hill Book Co., Inc., New York, 1958

Shih, S. F. and R. L. Hamrick, A technique used to determine random point position: I Theory. Submitted to Water Resources Bulletin, 1973.

Shreider, Y. A. (ed.), The Monte Carlo method, Second Impression, pergamon, New York, 1967.

Thiessen, A. H., Precipitation averages for large areas. Monthly Weather Review, Vol. 39: 1082-1084, 1911.

Table I. Computed weights (in percent) and relative areas for watershed shown in Fig. I, (Diskin, 1969)

Set No.	Station No. 1	2	3	4	5	6	7	8	9	Rel. Area
1	0.88	4.66	24.78	5.38	15.27	16.67	0.22	13.72	13.42	0.678
2	0.81	10.02	23.47	5.49	15.92	16.76	0.26	13.85	13.41	0.679
3	0.70	10.09	22.87	5.49	16.43	17.21	0.24	13.66	13.31	0.686
4	0.72	9.95	22.58	5.75	16.05	17.60	0.34	13.73	13.26	0.691
5	0.78	9.59	23.24	5.41	16.06	17.59	0.35	13.46	13.52	0.687
6	0.79	9.52	23.36	5.49	16.32	17.41	0.38	13.14	13.59	0.686
7	0.89	9.42	23.24	5.50	16.48	17.56	0.34	13.43	13.43	0.687
8	0.86	9.36	23.34	5.44	16.56	17.59	0.35	13.09	13.42	0.687
9	0.86	9.22	23.42	5.42	16.73	17.54	0.38	13.00	13.43	0.687

Table II. Modifiedly computed weights (in percent) and relative areas for watershed shown in Fig. I.

No. of Random Points	Station No. 1	2	3	4	5	6	7	8	9	Rel. Area
2000	0.880	9.660	24.780	5.380	15.270	16.670	0.220	13.720	13.420	0.6780
4000	0.845	9.840	24.125	5.435	15.595	16.715	0.240	13.785	13.415	0.6785
6000	0.797	9.923	23.707	5.453	15.873	16.880	0.240	13.743	13.380	0.6810
8000	0.778	9.930	23.425	5.528	15.518	17.060	0.265	13.765	13.350	0.6835
10000	0.778	9.862	23.388	5.504	15.946	17.166	0.282	13.684	13.384	0.6842
12000	0.780	9.805	23.383	5.502	16.008	17.207	0.298	13.593	13.418	0.6845
14000	0.796	9.750	23.363	5.501	16.076	17.257	0.304	13.570	13.420	0.6849
16000	0.804	9.701	23.360	5.494	16.136	17.299	0.310	13.510	13.420	0.6851
18000	0.810	9.648	23.367	5.486	16.202	17.326	0.318	13.453	13.421	0.6853

Table III. Comparison of the computed weights and relative areas obtained with different numbers of random points.

Method	No. of random points 2000	Rainfall Station								Ratio of area
		Lake Alfred	Kiss. II	Isle-Worth	Orlando	Bithlo	Lake Hart	Indian Lake State	Mountain Lake State	
Monte Carlo method	1	0.0360	0.2980	0.2910	0.0570	0.0010	0.1510	0.2500	0.1160	0.5830
	10	0.0324	0.3163	0.0879	0.0522	0.0010	0.1558	0.2410	0.1133	0.5897
	20	0.0348	0.3112	0.0905	0.0533	0.0008	0.1589	0.2386	0.1120	0.5907
	30	0.0341	0.3118	0.0897	0.0528	0.0007	0.1609	0.2381	0.1121	0.5910
	40	0.0349	0.3120	0.0892	0.0520	0.0008	0.1609	0.2378	0.1126	0.5898
	50	0.0343	0.3136	0.0895	0.0520	0.0009	0.1597	0.2376	0.1126	0.5893
	60	0.0343	0.3133	0.0897	0.0520	0.0009	0.1594	0.2384	0.1122	0.5896
	70	0.0342	0.3124	0.0890	0.0527	0.0009	0.1607	0.2383	0.1118	0.5907
	80	0.0345	0.3122	0.0894	0.0529	0.0008	0.1609	0.2379	0.1116	0.5903
	90	0.0346	0.3120	0.0895	0.0528	0.0009	0.1610	0.2379	0.1115	0.5899
	100	0.0346	0.3121	0.0890	0.0531	0.0009	0.1610	0.2381	0.1112	0.5899
	110	0.0347	0.3123	0.0888	0.0531	0.0008	0.1606	0.2383	0.1114	0.5896
	113	0.0347	0.3123	0.0888	0.0530	0.0008	0.1605	0.2383	0.1113	0.5898
Graphical method		0.034	0.288	0.091	0.052	0.001	0.160	0.259	0.115	

Table IV. Comparison of computed weights, relative areas and computing time with different numbers of boundary segments.

		Number of boundary segments			
		103 segments		10 segments	
Rainfall	Thessien	Random points		Random points	
stations	Polygon	2000	6000	2000	6000
Lake Alfred	0.034	0.036	0.030	0.034	0.030
Kiss. II	0.288	0.298	0.321	0.300	0.321
Isleworth	0.091	0.091	0.087	0.090	0.085
Orlando	0.052	0.057	0.054	0.059	0.055
Bithlo	0.001	0.001	0.001	0.001	0.001
Lake Hart	0.160	0.151	0.150	0.155	0.153
Ind. Lake ST.	0.259	0.250	0.248	0.258	0.252
Mountain Lake	0.115	0.116	0.109	0.103	0.102
Ratio of area		0.583	0.586	0.580	0.586
Computing * time,sec.		146	247	91	147

* CDC 3100 Computer system with Fortran IV Language was used in this comparison

Table V. Comparison of Thiessen coefficients determined by different techniques used on the hypothetical L shaped watershed shown in Fig. 4.

No. of random points	Rainfall station	Graphical method	Monte Carlo method						
			one rect-angle, abcd	Two equal rectangles			Two unequal rectangles		
				efgh E_{1j}	oplk E_{2j}	Total W_j	efg'h' E_{1j}	mnlk E_{2j}	Total W_j
2000	1	0.155	0.143	0.314	0	0.153	0.274	0	0.154
	2	0.376	0.381	0.686	0.087	0.379	0.655	0	0.369
	3	0.040	0.037	0	0.068	0.035	0.006	0.074	0.036
	4	0.315	0.323	0	0.609	0.312	0.065	0.665	0.327
	5	0.114	0.116	0	0.236	0.121	0	0.271	0.118
	Ratio of area, A_i		0.406	0.545	0.573		0.495	0.565	
	New ratio factor, R_i			0.487477	0.512522		0.563269	0.436730	
4000	1	0.155	0.139	0.318	0	0.156	0.273	0	0.153
	2	0.376	0.390	0.682	0.087	0.379	0.665	0	0.374
	3	0.040	0.036	0	0.064	0.033	0.006	0.068	0.033
	4	0.315	0.320	0	0.629	0.320	0.056	0.676	0.327
	5	0.114	0.115	0	0.220	0.112	0	0.255	0.112
	Ratio of area, A_i		0.408	0.558	0.579		0.498	0.571	
	New ratio factor, R_i			0.490765	0.509234		0.562157	0.437842	
6000	1	0.155	0.139	0.311	0	0.153	0.264	0	0.148
	2	0.376	0.393	0.689	0.079	0.378	0.679	0	0.379
	3	0.040	0.038	0	0.066	0.034	0.006	0.073	0.036
	4	0.315	0.316	0	0.640	0.326	0.051	0.675	0.326
	5	0.114	0.114	0	0.215	0.110	0	0.252	0.111
	Ratio of area, A_i		0.417	0.560	0.581		0.497	0.578	
	New ratio factor, R_i			0.490797	0.509202		0.558660	0.441339	

Table VI. Computed weights and relative area ratios of Fig. 5.

Rainfall station	Graphical method	Monte Carlo method		
		No. of random points		
		2000	4000	6000
1	0.029	0.036	0.034	0.030
2	0.102	0.107	0.105	0.107
3	0.002	0.0	0.0	0.001
4	0.058	0.054	0.057	0.057
5	0.166	0.179	0.188	0.190
6	0.113	0.117	0.118	0.114
7	0.053	0.044	0.042	0.044
8	0.071	0.054	0.069	0.068
9	0.089	0.081	0.079	0.084
10	0.015	0.016	0.013	0.012
11	0.053	0.047	0.043	0.046
12	0.043	0.044	0.042	0.046
13	0.080	0.096	0.085	0.080
14	0.078	0.066	0.073	0.071
15	0.046	0.058	0.051	0.050
16	0.002	0.002	0.001	0.001
Ratio of area		0.621	0.624	0.630

Table VII. Computed new weights and relative area ratios with missing data for the upper Kissimmee River Basin as shown in Fig. 3.

Rainfall station	Station with missing data								
	0	1	2	3	4	5	6	7	8
Lake Alfred	0.034	-	0.109	0.034	0.034	0.034	0.034	0.034	0.066
Kiss. II	0.300	0.324	-	0.348	0.300	0.300	0.386	0.341	0.314
Isleworth	0.090	0.090	0.134	-	0.127	0.090	0.090	0.090	0.090
Orlando	0.059	0.059	0.063	0.100	-	0.059	0.088	0.059	0.059
Bithlo	0.001	0.001	0.001	0.001	0.001	-	0.040	0.001	0.001
Lake Hart	0.155	0.155	0.271	0.155	0.177	0.156	-	0.156	0.155
Indian Lake State	0.258	0.258	0.283	0.258	0.258	0.258	0.258	-	0.315
Mountain Lake	0.103	0.114	0.140	0.103	0.103	0.103	0.103	0.320	-
Ratio of area	0.580	0.580	0.580	0.580	0.580	0.580	0.580	0.580	0.580

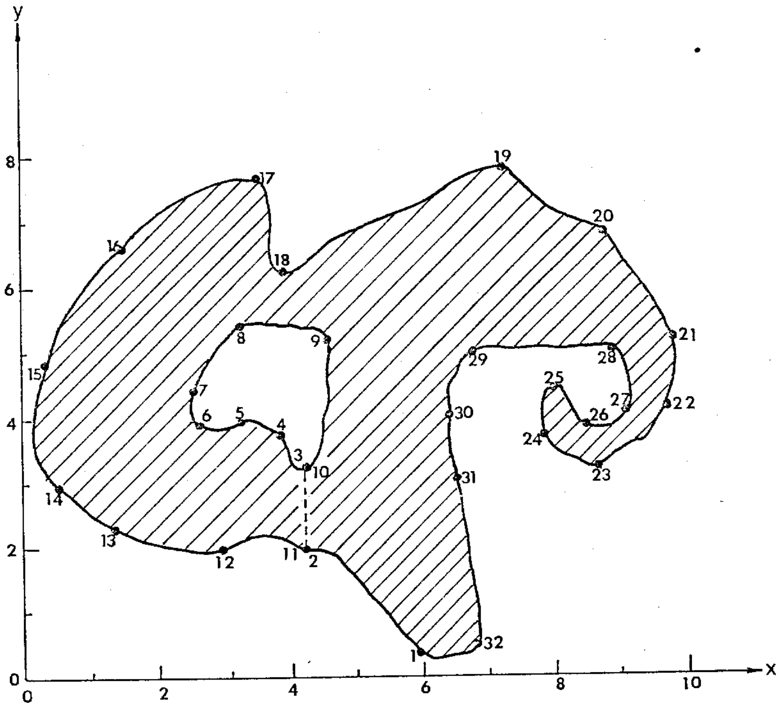


Fig. 1. Area under an irregularly shaped form.

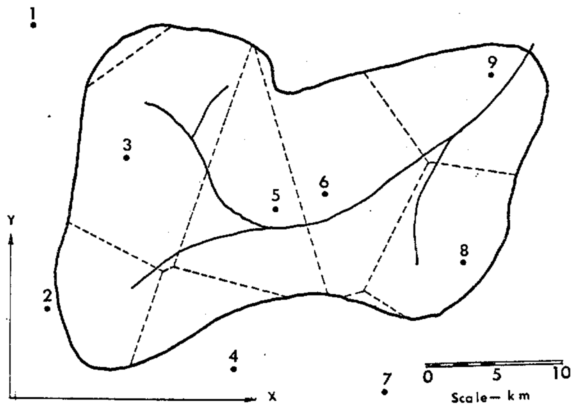


Fig. 2. Watershed Used In Example, (After Diskin, 1969).

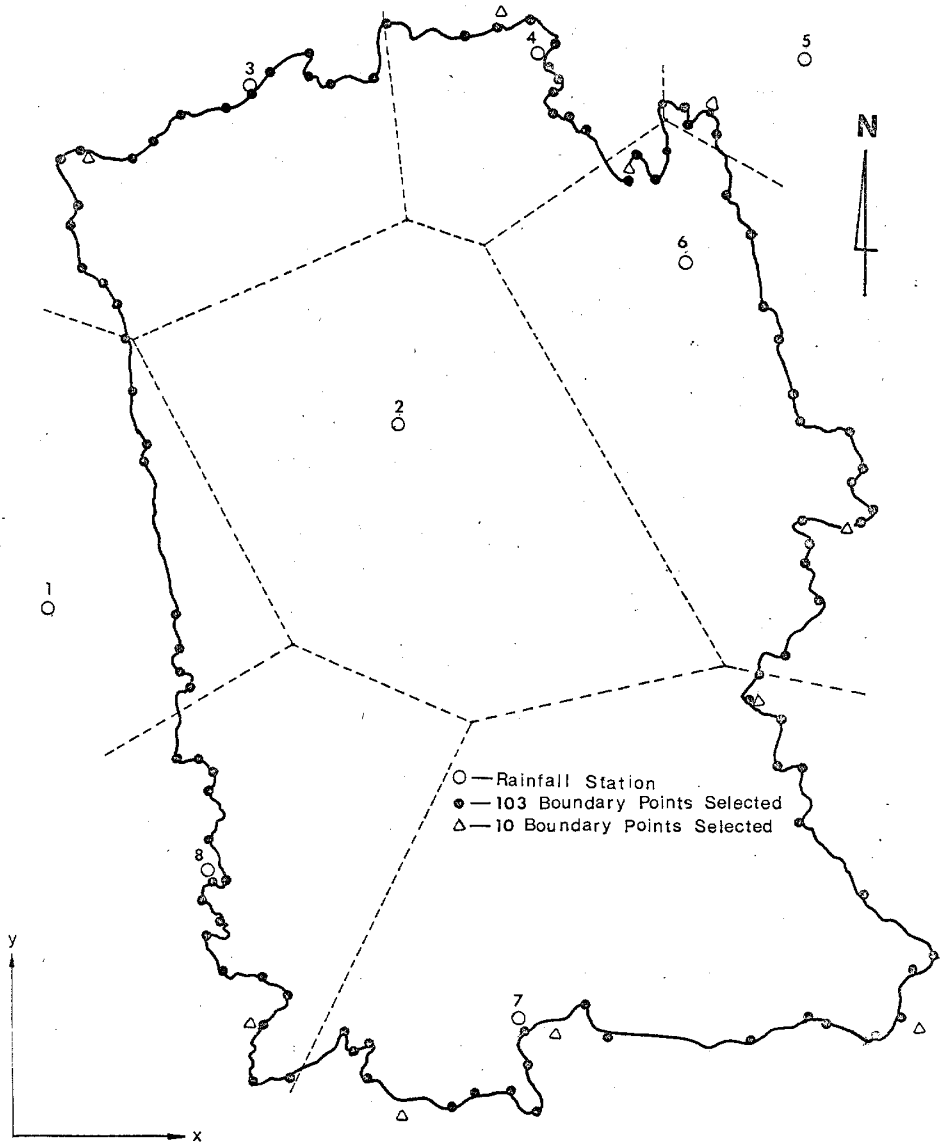


Fig. 3. Upper Kissimmee River Basin

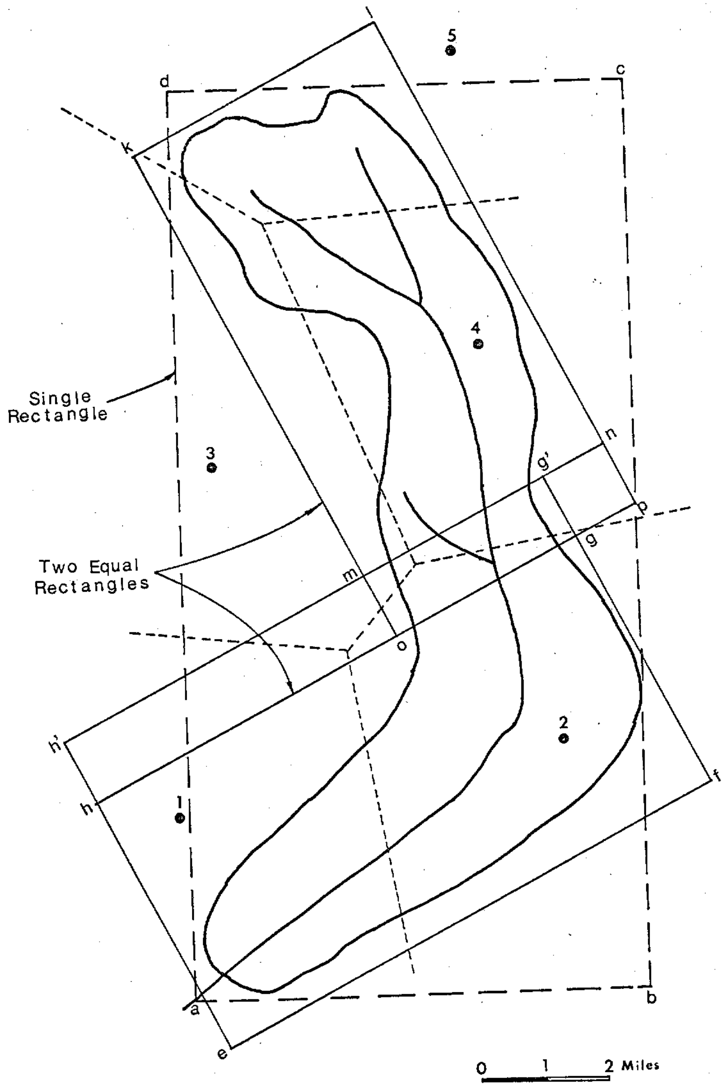


Fig. 4. Hypothetical L Shaped Watershed, (After Diskin, 1969).

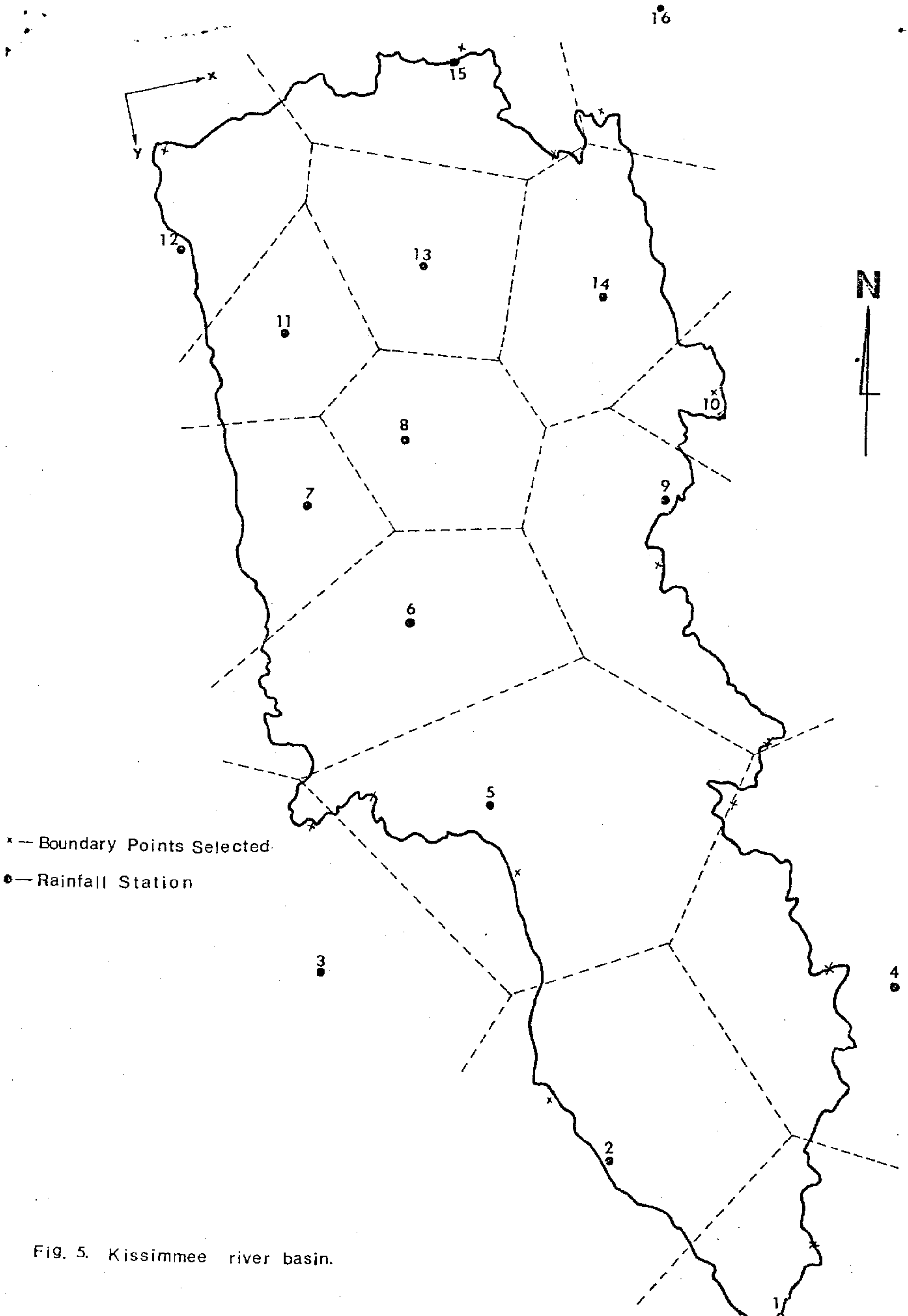
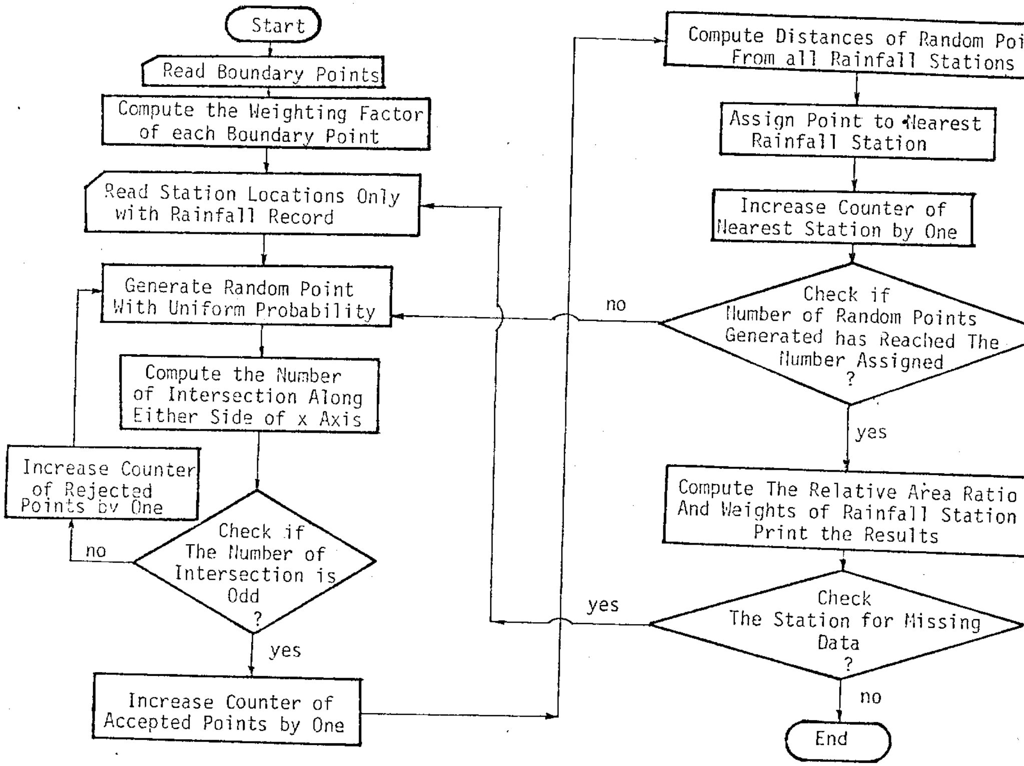


Fig. 5. Kissimmee river basin.

APPENDIX A: Flowchart for the modified Monte Carlo technique of computation.



APPENDIX B: COMPUTER PROGRAM FOR DETERMINING THE THIESSEN COEFFICIENTS BY MONTE CARLO METHODS

```
      PROGRAM MCMTC
C     X IS THE X COORDINATE OF THE BOUNDARY
C     Y IS THE Y COORDINATE OF THE BOUNDARY
C     NN IS A WEIGHTING FACTOR FOR NODE POINT OF BOUNDARY
C     IA AND IX ARE INITIAL ODD INTEGER FOR USING IN  RANDU  SUBROUTINE
C     N IS THE NUMBER OF BOUNDARY POINTS CHOSEN
C     M IS THE NUMBER OF RAIN MEASURING STATION
C     NSET IS THE NUMBER OF RANDOM POINTS EXPECTED
C     XMIN AND XMAX ARE THE MINIMUM AND MAXIMUM RANGE IN X AXIS
C     YMIN AND YMAX ARE THE MINIMUM AND MAXIMUM RANGE IN Y AXIS
C     AX IS THE X COORDINATE OF THE RAIN MEASURING STATION
C     AY IS THE Y COORDINATE OF THE RAIN MEASURING STATION
      REAL L(500)
      DIMENSION X(400),Y(400),AX(500),AY(500),WF(500),NS(500),YY(400),
     1NN(400),TITLE(9)
C     READ INPUT DATA
   83 READ(60,3) IW,TITLE
    3 FORMAT(I5,5X,9A8)
      IA=5
      IX=7
C     CHECK WHETHER THE END OF STATION SET
      IF(IW.EQ.0) GO TO 85
      WRITE(61,5) TITLE
    5 FORMAT(1H1,//5X,9A8//)
      READ(60,8) N,M,NSET,XMIN,XMAX,YMIN,YMAX
    8 FORMAT(3I6,21X,4F8.4)
      WRITE(61,108) N,M,NSET,XMIN,XMAX,YMIN,YMAX
  108 FORMAT(//10X,29HNUMBER OF BOUNDARY SEGMENTS =,I4/10X,29HNUMBER OF
     1RAINFALL STATIONS =,I4/10X,25HNUMBER OF RANDOM POINTS =,I6/10X,16H
     2X-AXIS MINIMUM =,F8.4/10X,16HX-AXIS MAXIMUM =,F8.4/10X,16HY-AXIS M
     3INIMUM =,F8.4/10X,16HY-AXIS MAXIMUM =,F8.4//)
      READ(60,9) (X(I),Y(I),I=1,N)
    9 FORMAT(16F5.1)
      READ(60,4) (AX(I),AY(I),I=1,M)
    4 FORMAT(16F5.1)
C     ADD A SMALL VALUE TO EACH NODE FOR OBTAING THE NN FACTOR
      DO 106 I=1,N
  106 YY(I)=Y(I)+0.00001
C     COMPUTE THE WEIGHTING FACTOR NN FOR EACH BOUNDARY NODE
      DO 101 I=1,N
      NN(I)=0
      IF(Y(I).EQ.Y(I+1)) GO TO 102
      IF(YY(I).LT.Y(I+1).AND.YY(I).GT.Y(I)) NN(I)=NN(I)+1
      IF(I.EQ.1) GO TO 103
      IF(YY(I).LT.Y(I-1).AND.YY(I).GT.Y(I)) NN(I)=NN(I)+1
      GO TO 101
  103 IF(YY(I).LT.Y(N).AND.YY(I).GT.Y(I)) NN(I)=NN(I)+1
```

```
      GO TO 101
  102 NN(I)=1
  101 CONTINUE
      DO 107 I=1,M
      NS(I) = 0
  107 CONTINUE
      NA=0
      NR=0
      XMN=XMAX-XMIN
      YMN=YMAX-YMIN
      DO 500 IK=1,NSET
      IR=0
      IL=0
C   GENERATE THE RANDOM NUMBER
      CALL RANDU(IX,IY,RANDM)
      IX=IY
      XT=XMIN + RANDM*XMN
      CALL RANDU(IA,IB,RANDN)
      IA=IB
      YT=YMIN + RANDN*YMN
      X(N+1)=X(1)
      Y(N+1)=Y(1)
C   CALCULATE THE NUMBER OF INTERSECTION ALONG THE X AXIS IN EITHER SIDE
      DO 300 K=1,N
      IF(YT.EQ.Y(K).AND.XT.EQ.X(K)) GO TO 310
      IF(YT.EQ.Y(K).AND.YT.EQ.Y(K+1)) GO TO 10
      IF(YT.EQ.Y(K)) GO TO 20
      IF(Y(K).GT.YT.AND.YT.GT.Y(K+1)) GO TO 40
      IF(YT.GT.Y(K).AND.Y(K+1).GT.YT) GO TO 40
      GO TO 300
   10 IF(X(K).LT.XT.AND.X(K+1).GT.XT) GO TO 310
      IF(X(K).GT.XT.AND.X(K+1).LT.XT) GO TO 310
      IF(XT-X(K)) 11,310,12
   11 IR=IR+NN(K)
      GO TO 300
   12 IL=IL+NN(K)
      GO TO 300
   20 IF(X(K)-XT) 12,310,11
   40 XX=X(K)+ (YT-Y(K))*(X(K+1)-X(K))/(Y(K+1)-Y(K))
      IF(XX-XT) 42,310,41
   41 IR=IR+1
      GO TO 300
   42 IL=IL+1
  300 CONTINUE
C   CHECK WHETHER THE RANDOM POINT IS FALLING WITHIN THE BOUNDARY
      IF((IR-IR/2*2).EQ.0.OR.(IL-IL/2*2).EQ.0) GO TO 302
  310 NA=NA+1
C   ASSIGN THE FALLING WHITHIN BOUNDARY POINT TO THE NEAREST STATION
      L(1)=(XT-AX(1))**2 +(YT-AY(1))**2
      SAVE=L(1)
      ISUB =1
```

```
      DO 91 I =2,M
      L(I)=(XT-AX(I))**2 +(YT-AY(I))**2
      IF(L(I)-SAVE)13,91,91
 13   SAVE=L(I)
      ISUB=I
   91 CONTINUE
      NS(ISUB)  =NS(ISUB)+1
      GO TO 500
 302  NR=NR+1
  500 CONTINUE
C    COMPUTE RELATIVE AREA RATIO AND WEIGHTING FACTOR OF EACH STATION
      AF= NA/(FLOAT(NA+NR))
C    PRINT THE RESULTS
      WRITE(61,7) AF
    7 FORMAT(10X,21HRELATIVE AREA RATIO =,F9.6//)
      DO  21 I=1,M
      WF(I) =NS(I)/FLOAT(NA)
   21 WRITE(61,6) I,WF(I)
    6 FORMAT(10X,35HCOMPUTED WEIGHT OF RAINFALL STATION, I4,3H  =,F9.6/)
      GO TO 83
   85 CALL EXIT
      END
```