AN ANALYSIS OF PRECIPITATION DATA FOR TESTING THE VALIDITY OF POOLING RAINFALL OCCURRENCES WITHIN PERIODS OF SIMILAR RAINFALL PROBABILITIES

By

ROBERT BECK DUFFIN

Bachelor of Science

Utah State University

Logan, Utah

1950

Submitted to the faculty of the Graduate School
of the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1963

AN ANALYSIS OF PRECIPITATION DATA FOR TESTING THE VALIDITY OF POOLING RAINFALL OCCURRENCES WITHIN PERIODS OF SIMILAR RAINFALL PROBABILITIES

Thesis Approved:

Thesis Adviser

Dean of the Graduate School

ACKNOWLEDGEMENTS

Acknowledgement is gratefully extended to the Agronomy Department of Oklahoma State University and the Federal Government for the National Defense Fellowship made available to pursue the advanced degree. Gratitude is due the Agronomy, Meteorology, and Statistics Departments of Oklahoma State University for use of the necessary facilities and equipment.

The author appreciates the assistance of his major adviser, Dr. John F. Stone, for his guidance during this study and thesis preparation. Valuable assistance in techniques for obtaining the data from the punched cards was received from Robert L. Peace, of the Meteorology Department. Acknowledgement is due Dr. Carl E. Marshall for advice concerning statistical problems of the study.

TABLE OF CONTENTS

LIST OF APPENDIX TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Year by year, the competition for Oklahoma's available water supplies steadily increases. Since it is necessary to obtain most water either directly from the atmosphere through surface supplies or indirectly from ground water sources, the future key to a more beneficial water use obviously lies in a careful water management of these supplies. Although this study is more particularly concerned with the validity of dividing the year into periods of similar rainfall probabilities, a few considerations regarding general application of rainfall probability data to solving agronomic problems will follow.

Daily precipitation data is now available in punch card form. Effective use of this information should provide some assistance toward a more beneficial use of our available water supplies. This rainfall data can yield probabilities of both rainfall and dry periods that should be valuable in agricultural planning and decision making. Questions such as how many acres of certain crops to plant, or whether to convert to irrigation farming, or whether to build surface water storage can certainly be answered in part by these past rainfall and dry-season probabilities.

Analyses of precipitation data would seem to be justified from several different agronomic standpoints. Soil moisture is considered to be the first limiting production factor in western Oklahoma. The question as to the frequency that a farmer might expect serious crop losses because of insufficient soil moisture is of concern in all sections of Oklahoma.

Conversely, the problem might be one of too much moisture at one particular time, such as during wheat harvest. The probability of lowered cotton, peanut, and alfalfa quality because of excess moisture at harvest time is a definite concern.

The objective of this study is to determine if the year can be divided into portions in which (1) the probability of receiving rain (P_1), and (2) the probability of receiving specified amounts of rain provided rainfall occurs (P_2), are characteristic within different periods of the year. It is proposed that if valid periods of characteristic P_1 and P_2 values can be established with these two rainfall probabilities, then all rainfall occurrences within each period can be combined for purposes of further analyses. If it is possible to establish periods of characteristic rainfall probabilities of sufficient length, then the pooling of rainfall occurrences within each period might serve as a technique to obtain rainfall probability data with less class bias than occurs from short periods.

It is believed that a natural smoothing of rainfall frequency distributions will result when rainfall occurrences are pooled with maximum-length periods of similar rainfall probabilities. This study was designed to test the hypotheses (1) that there is a valid basis for these periods to be established and (2) that once established they are of sufficient length to result in smoothed rainfall frequency curves. If tenable, it would then follow that rainfall probabilities and other data could be directly obtained from tabulated frequencies. For example, probabilities of receiving specified amounts of rainfall on any particular day (P) could readily be obtained.

An affirmative result for these two hypotheses would indicate that this method of obtaining smooth rainfall frequency distributions might produce more representative results than those obtained by attempting to fit a mathematical function to tabulated frequencies of arbitrary one- to three-week periods.

CHAPTER II

REVIEW OF LITERATURE

The published material relating most directly to rainfall probabilities, and hence, to this study, will be considered in this review of the literature.

The 1941 Yearbook of Agriculture, "Climate and Man," affords a good general treatise on precipitation and other climatological data. Kincer (5) emphasizes that long-time precipitation data has more economic significance than any other aspect of the climate.

Precipitation Types

It is of interest that large-scale readjustments of atmospheric circulation have been shown to be associated with seasonal precipitation types (2,7,8). This condition would seem to suggest that the year might have well-defined periods of characteristic rainfall probabilities.

Weekly Rainfall Probability Studies

The following references relate to studies which report weekly rainfall probabilities of receiving specified rainfall amounts during the climatological year. A rather widely adopted mathematical function (incomplete gamma distribution) is used to artifically smooth rainfall distribution curves. This approach in smoothing rainfall distribution data is in contrast to that tested in this thesis. This thesis is concerned with testing the hypothesis that a natural smoothing might result within relatively long-time periods, each of which has constant or well-defined rainfall probabilities within the period. All results of this thesis are based on daily rainfall occurrences whereas in the weekly rainfall probability studies, weekly rainfall summaries are used as the basis for obtaining various data.

Thom (10) reported on the history, use and limitations of the incomplete gamma distribution in the analysis of rainfall probabilities. He developed the "approximate solutions" for the maximum likelihood equation for the incomplete gamma distribution. In analyzing rainfall data of the Ames, Iowa station, Barger and Thom (1) illustrate the decreasingly skewed curves for plotted frequency distributions as n increases in analyzing rainfall totals for n weeks. The incomplete gamma distribution used in smoothing the data was checked against the frequency histograms by the chi-square test, and a similar fit was reported in most cases.

In a preliminary study by Friedman and Janes (4), emphasis was placed upon the need for caution in interpreting empirical probability curves of rainfall amounts. Sampling fluctuations were the cause for irregularities in the shapes of plotted rainfall frequency curves. The incomplete gamma distribution was used as a method of smoothing out cumulative probability curves. These authors believe that an objective and consistent basis for smoothing out the irregularities can be reached by assuming that the population probability distribution from which the sample was attained has the characteristic of a known mathematical model. They also emphasized the convenience of the mathematical model in smoothing the plotted values for analytically investigating the space and time changes related to rainfall probability periods.

A North Central Regional Project reported by Barger, Shaw and Dale (2) makes available in tabular form two- and three-week precipitation distribution parameters for selected stations of the north central region of the United States. These probabilities were computed by fitting the incomplete gamma distribution to weather bureau records of precipitation totals. In a recent 1960 report, these authors (9) published another precipitation probability report for the north central states, which presents probabilities by use of graphs and maps of the more bulky tabulated information in the previous report.

CHAPTER III

METHODS AND MATERIALS

Historical Background

Since this study involves the obtaining and analyzing of data from climatological punch cards, a brief history of the punch card program at Oklahoma State University will follow.

In 1948, the Weather Bureau started cooperative programs with state universities and other state agencies to assist in processing weather records up to that time. The project in Oklahoma started in 1955 and was cooperative with the Agricultural Experiment Station of Oklahoma State University. The Weather Bureau provided all blank cards and original data. The University agreed to punch all cards from the beginning of the record through 1946 and furnish duplicate copies for the Weather Bureau. The Weather Bureau has agreed to provide to the University data on punched cards representing daily climatological events, starting with 1947. These punch cards are available for any Oklahoma station in an equal number to those which the University has processed and made available to the Weather Bureau.

Up to the present time, cards have been prepared for 24 stations. Twenty of the 24 are long-record stations of the central one-third of Oklahoma. Four short-record stations (Lookeba, Frederick, Jefferson, and Tipton) have been punched.

This thesis study is devoted exclusively to the examination of precipitation data from punch cards of the Stillwater station. Stillwater is the first long-record station in Oklahoma to be studied in this manner. It is hoped that this study might serve, in part, as a guide in obtaining and analyzing precipitation data from other Oklahoma stations.

Rainfall observations for the Stillwater station, from January, 1893, through December, 1948, were made at a University location approximately 75 yards north of the present library. From 1948 to 1955, observations have been made at a campus location west of the home economics building and, since then, at either the Stillwater airport or the agronomy farm adjoining Stillwater to the west.

Rainfall Probabilities

Rainfall was characterized in this study by evaluating variations of two rainfall probabilities within various sized groupings of days throughout the climatological year. All data were obtained from Weather Bureau IBM cards, upon which were recorded daily occurrences of either no rain, trace, or a particular amount of rainfall in hundredths of an inch for each day. All data are based on the continuous 65-year record of the Stillwater station from 1893 through 1957. The week numbers of the climatological year with their corresponding beginning and ending calendar dates in Appendix Table I are a necessary conversion throughout this study. The first week in March is climatological week number one. Week number 52 is the concluding week in February.

The first probability (P_1) can be described as the average probability of receiving rainfall on any randomly selected day within the particular grouping of days. The second probability (P_2) is the average probability that provided rainfall occurs on any randomly selected day within a particular period of days, a specified amount of rainfall will result.

These definitions of P_1 and P_2 should become more evident after considering the manner in which the values were obtained.

Grouped Daily Rainfall Occurrences

The different size groupings of days in which P_1 and P_2 were obtained include 1, 3, 5, and 7-consecutive-day periods throughout the climatological year. From the consecutive seven-day groupings, larger consecutive groupings of 14 and 28 days were made by pooling daily rainfall occurrences in two and four consecutive seven-day periods, respectively. It was desired to detect significant unbiased variation in rainfall patterns within the shortest period possible. In order to evaluate variations in P_1 and P_2 in this manner, the different sized groupings of days were selected. The degree of bias of rainfall distribution data due to insufficient rainfall occurrences on record had not been previously established for any Oklahoma station.

Only the IBM Type 083 and 101 sorters were used in this study. This selection of equipment was thought advisable because of the ready availability of these machines and also the exploratory nature of this investigation.

The IBM Type 083 was used to first arrange the WB-1009 daily weather punch cards from a chronological order by years into a grouping in which all the March 1, all March 2, etc., on record were grouped together.

The IBM Type 101 sorter and recorder was used to separate and record the number of cards with trace, missing data, and no rain into separate divisions. Rain-day cards were then separated into various incremental amounts. This division of rainfall increments was obtained up to 4.99

inches, by making a separate pass of rain-day cards through the machine to obtain data from .01 through .99 inch, from 1.00 through 1.99 inches, etc., up to the 4.00 through 4.99 inch division. The rainfall increments greater than trace to .99 inch were .01 - .09, .10 - .19, .20 - .29, .30 - .39, .40 - .49, .50 - .59, .60 - .69, .70 - .79, .80 - .89, and .90 - .99. The same incremental divisions were made within each one-inch range up to the 4.99-inch rainfall amount.

The P_1 values for one-day groupings throughout the year were then obtained from the number of rain days, missing data, trace, and no-rain cards by the following relationship:

$$P_1 = \frac{\text{number of days with rain}}{\text{days on record}} = \frac{\text{total - (missing data + trace + no rains)}}{\text{total - missing data}}$$

The P_2 values were calculated as follows from the previously described data obtained and recorded by the IBM Type 101 sorter:

$$P_2 = \frac{\text{number of rain days within a specified rainfall increment}}{\text{total rain days of the particular period}}$$

The P_1 and P_2 data for each of the remaining 3, 5, and 7, 14, and 28-day consecutive groupings throughout the year was obtained in a similar manner as that previously described for the single day and described as follows. A sorting of all cards with respect to the particular grouping was necessary. For example, with the 3-day grouping, a sorting with the IBM Type 083 sorter was made to separate all the March 1, 2, and 3 day cards for the 65 years on record into one group, all the March 4, 5, and 6 cards into another group, etc., throughout the climatological year. With this separation made, each 3-day grouping of cards could be further separated according to no rain, trace, and missing data groups with the rain days separated into incremental amounts and all numbers of the various piles recorded by the IBM Type 101 sorter and recorder. Values for P_1

and P_2 were then calculated as previously described for the single-day groupings, except that all data applied to the particular grouping of days. Note that the rainfall data obtained from the various groupings were still on a single-day-event basis and not the total for the particular grouping of days.

At this point, it might be well to again consider the definitions of P_1 and P_2. Both P_1 and P_2 are average probabilities since for a particular grouping of days the data obtained ignore variations from year to year and from day to day that occur within the grouping. Hence, if P_1 and P_2 are reported with significance on a daily basis, they must be defined as average and they must be defined with respect to any randomly selected day within the period.

Determination of Periods of Characteristic P_1 and P_2 Values

It was previously stated that for purposes of this study, rainfall patterns throughout the year would be characterized by variation in P_1 and P_2.

In order to evaluate significance of P_2 variation, a procedure of semi-log graphing of P_2 values (percentages of total rainfall occurrences in a given rainfall increment) against the X values (incremental rainfall amounts) was employed. A straight line was plotted on each graph by visually best fitting the line to the plotted data (see Figure 2 in results for example). Variation of P_2 can now be studied throughout the year.

An understanding of the physical significance of these slopes and their variations is important for the proper evaluation of results. Upon examination of the Figure 2 example, it becomes evident that the semi-log plotting of data points is approximately linear. This holds true quite well throughout the year. Because of the heavy grouping of small rains, this is not entirely true, however. This comparison of slopes of straight lines of consecutive periods of days throughout the year enables a comparison of variations in the relative proportions of the different incremental rainfall amounts in relation to the total rainfall occurrences. These slope values provide a measure of whether the percentages of larger and smaller rains are increasing or decreasing from one time of the year to another.

For example, consider two consecutive 14-day periods. If, in the first period, a greater percentage of the total rainfall occurrences consists of larger incremental rainfall amounts than is true for the second period, this difference could be detected by comparing the slopes of the plotted data. All slope values are negative since the straight line slopes downward to the right, which makes ΔP_2 negative when ΔX is positive. In the first period, with its greater comparative percentage of larger rains, the line would have a smaller absolute slope value than in the second period.

The equation of the straight line, as seen on the plot, is $\log_{10} P_2 = mX + b$. For convenience let $Z = \log_{10} P_2$. Then the expression $Z = mX + b$ represents the equation for each plotted straight line. Solving for the slope m yields the following equation: $m = (Z - b)/(X)$. If the slope is evaluated where $Z = 0$, then $m = -b/X_i$ where X_i = X intercept. Note that b = Z- intercept. Transforming back to P_2 notation, m = ($-\log_{10} P_2$ intercept)/(X intercept). For purpose of comparison, absolute values of slopes were used, so absolute slope = ($\log_{10} P_2$ intercept)/(X intercept).

Results of variation of slope values for different sized consecutive groupings of days throughout the year were used in this study as the criteria for dividing the year into periods of characteristic P_2 values.

Variations with P_1 values were evaluated in a more direct manner. A plotting of percent rain days within each size period against the time of the year in which the period occurs provides a separate comparison for P_1 values. With both P_1 and P_2 data on hand for the different periods, variations in both probabilities can be evaluated separately. This evaluation was used to establish periods within which both P_1 and P_2 can be considered essentially constant.

To assist in determining the time periods in which the year might be divided, slope data were prepared for the 1, 3, 5, 7, 14 and 28-day groupings throughout the year. These various sized groupings enabled a gradual increasing of the total rainfall occurrences on record for which a single slope function could be prepared. It seems advisable that selection of the smallest possible grouping which can still be considered reasonably reliable would enable more closely defined division dates of the periods.

Analyses of Data Within Periods

With the year divided into periods within which both P_1 and P_2 can be considered characteristic, the pooled rainfall occurrences within each of these periods can be analyzed to obtain a wide variety of data. However, the primary hypothesis to be tested in this study with regard to these periods was that there exists on record a sufficient number of rainfall occurrences within these periods of the year to result in a natural smoothing of the plotted points of the rainfall distribution of incremental rainfall amounts.

Pooled rainfall occurrences within periods in which variations of either P_1 or P_2 are considered insignificant were analyzed to determine average probabilities (P values) that on any randomly selected day the indicated amounts of rainfall would occur. This new probability (P) was determined by taking the products (P_1 x P_2) of the two probabilities concerned for the respective periods in which they were considered constant or characteristic. To obtain this new probability (P), the (P_1) values used in obtaining these data were obtained from the 28-day grouping of daily rainfall occurrences. If consecutive P_1 values within each of the (P_2) periods of the year varied more than 10 percent from one 28-day grouping to another, the separate P_1 values were used in determining the product of the probabilities. When several P_1 values were very similar (less than 10 percent variation) during several consecutive 28-day groupings within a particular (P_2) period, they were averaged for purpose of obtaining the products of P_1 and P_2 within a period.

CHAPTER IV

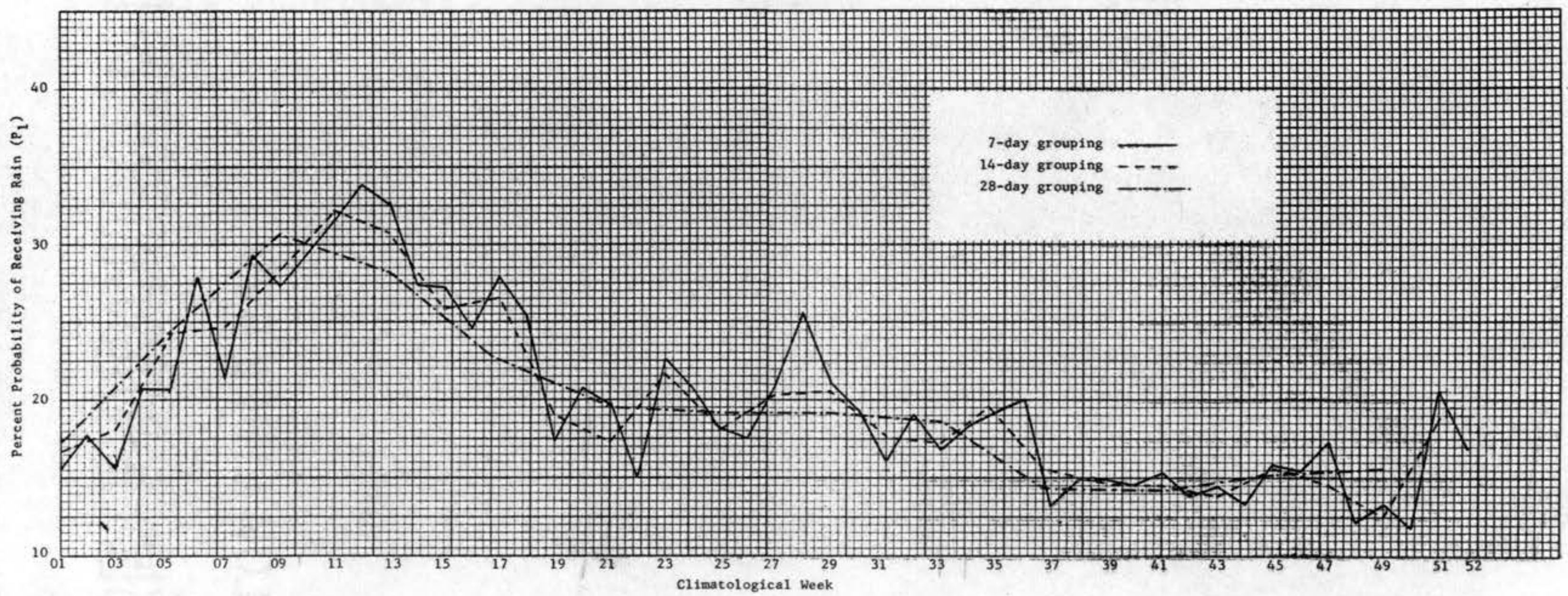
RESULTS AND DISCUSSION

Probability of Rain (P_1)

Results of the initial investigation concerning P_1 values are reported with 7-day, 14-day, and 28-day groupings in Figure 1. In Figures 1, 3, and 4 each plotted point is located at the beginning day of the particular grouping. Rainfall probabilities using 5-day, 3-day, and 1-day groupings of individual-day rainfall data were also determined. The results of these shorter periods were extremely variable because of smaller number of total rainfall occurrences on record within the particular period. As a result, the probability of rain on any given date is not considered as useful as the average probability of rain for a date selected at random within a grouping of several days. It is observed that this latter probability is progressively less variable, as the grouping increases from 7-day to 14-day and to 28-day periods of time. It is evident that probabilities obtained from longer groupings of days are more significant because of the removal of class bias.

An examination of the plotted rainfall probability data of the various groupings shows a similar range of probabilities during the time period between the 19th and 36th climatological weeks. Another separate range of similar rainfall probabilities seems to exist from the 37th through the 50th climatological weeks. This range constitutes the lowest range of P_1 values during the year. It is observed that the

Figure 1. Average percent probability of obtaining rain on any randomly selected day within 7, 14, and 28-day groupings throughout the climatological year.



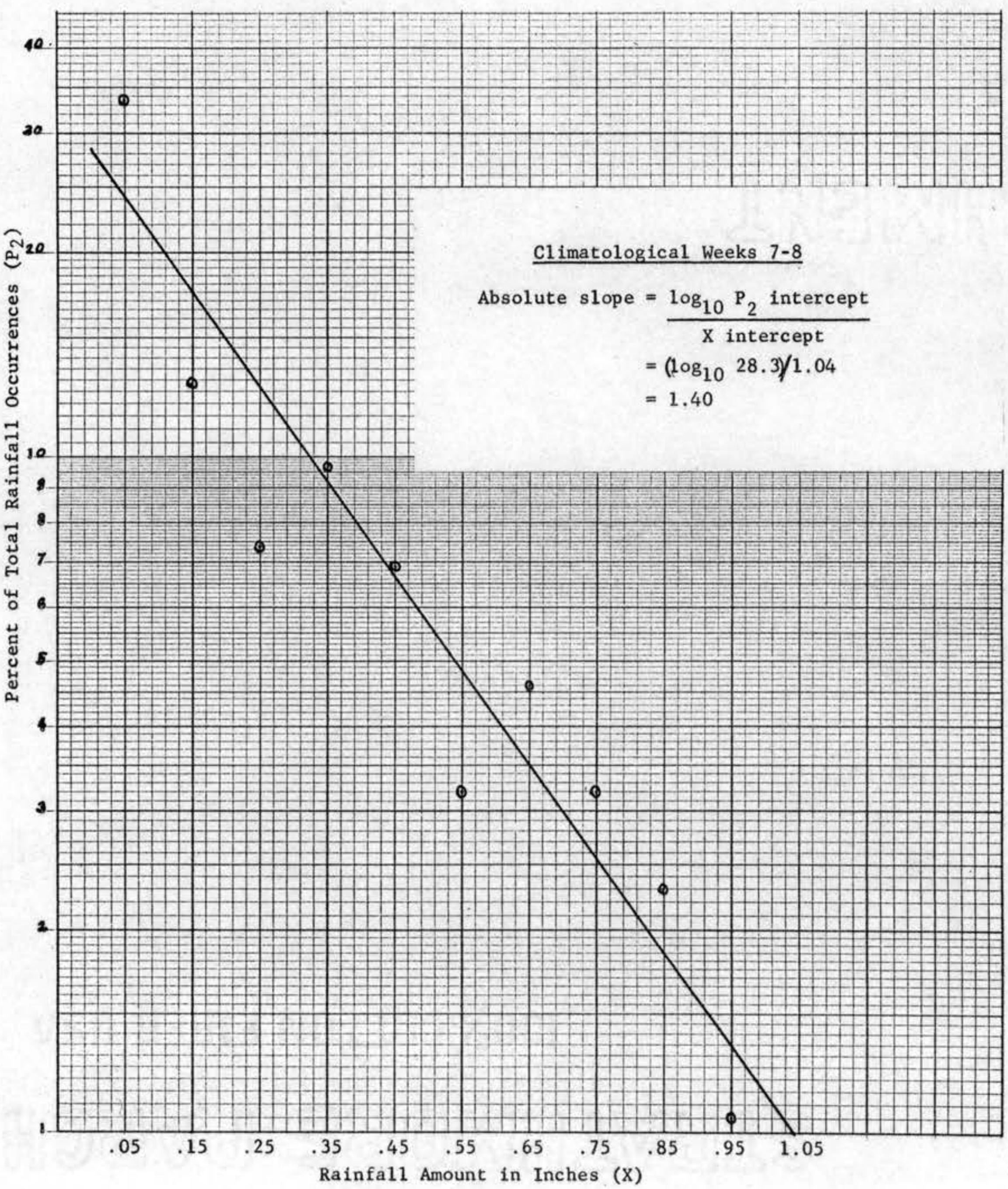
highest rainfall probabilities exist during a period starting with the 4th climatological week and ending with the 18th climatological week. Unlike the previously described periods, however, this time interval is characterized by a general increase in rainfall probability from approximately 20 to 30 percent, followed by a general decrease in rainfall probability back to 20 percent. From these observations it would seem that three periods of the year can be considered characteristic with regard to P_1 values. Of these three periods only the period from the 51st week through the 18th week contains P_1 variations greater than 10 percent within the period. This latter period must be divided into several smaller periods if less than 10 percent variation is desired within the different periods.

Probability of Receiving Various Amounts of Rain (P_2)

Results of the study concerning the variation of P_2 throughout the year are reported in Figures 2 through 5. It should be re-emphasized here that all data were obtained by different groupings of individual daily rainfall occurrences throughout the year. A given probability, then, would represent an average probability that if it does rain, the designated amount of rainfall will be obtained on any randomly selected day within the particular size grouping of days.

An example of the procedure and the results obtained in determining absolute slope values is reported in Figure 2. A 14-day grouping was used in this example. The relative degree of scattering of this semi-log plotting of points is somewhat intermediate as compared to the other 14-day plottings throughout the year. The first plotted point represents the mid-point of the lowest rainfall incremental range .01 through .09 inch. This point was observed to be always above the

Figure 2. Semi-log graphing of P_2 values against X values for climatological weeks 7 and 8. Also shown is the visually best fitting the straight line to the plotted data. Rainfall amounts as shown represent mid-points of rainfall ranges.

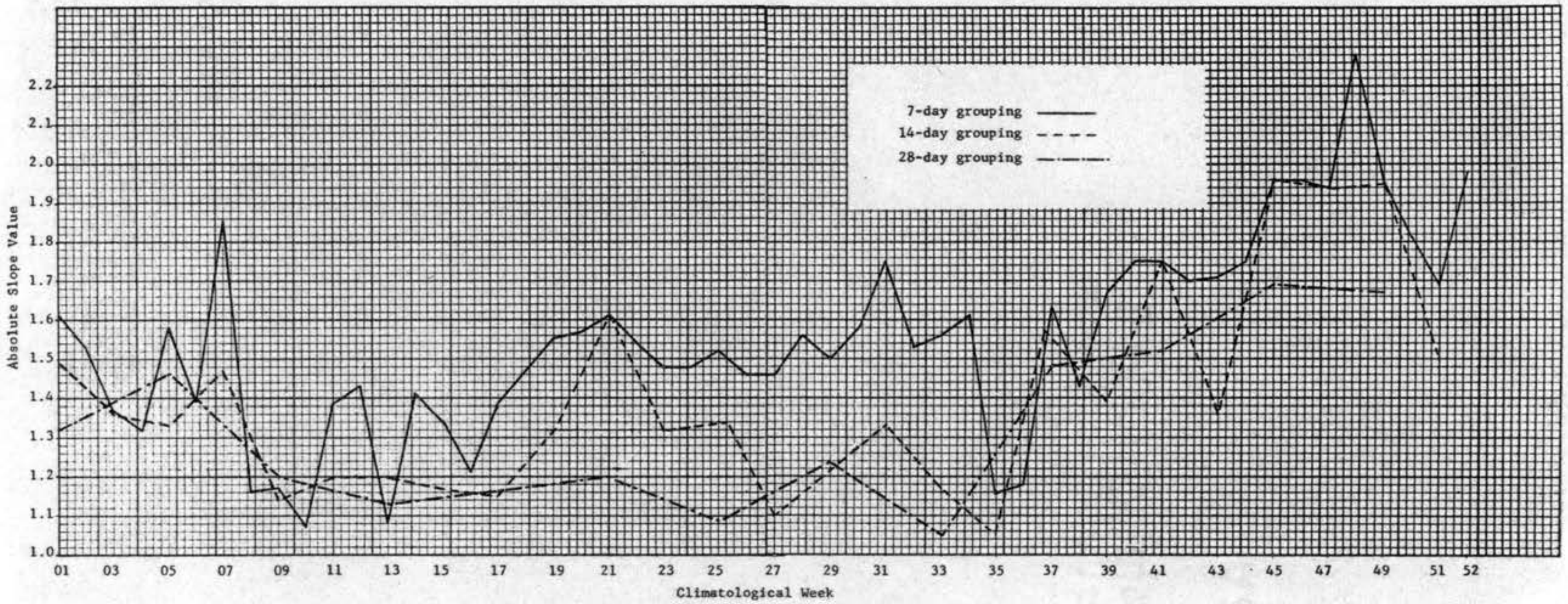


visually fitted straight line. A very high percentage of total rainfall occurrences during any period of the year is apparently of small incremental amount.

The slope functions within each 7, 14, and 28-day grouping throughout the climatological year are plotted against time (Figure 3). This graph presents a comparison of these slopes which individually designate the manner in which percent of the total rainfall occurrences within each period varies with respect to the various rainfall increment amounts. This plotting of the 28-day grouped data in particular shows that slopes are quite similar during the time period between weeks number 9 and 36. This 17-week period is characterized by having a higher percentage of large rains than any other time periods of the year. Although greater variation in 28-day absolute slope values occurs between weeks number 37 through 6, the slope values can be considered to be in a definitely higher slope range than the time period defined by weeks number 9 through 36. These higher absolute slope values broadly define a time period during the climatological year when a relatively higher percentage of the total rainfall occurrences is smaller rains.

Results of the study of slopes obtained from 14-day and 7-day groupings indicate that considerably more variation in slope values is obtained from these smaller groupings. It is interesting to note, however, that even with this increased variability of results within the previously described period, the time of division of these periods corresponding to that of the 28-day grouping still seems to remain quite abrupt. Slope values obtained in the 14-day groupings still show well-defined periods similar to those of the 28-day grouping in which a division might be made. Results of the 14-day groupings are considered sufficiently free

Figure 3. A comparison of absolute slope values for 7, 14, and 28-day groupings throughout the climatological year.



from extreme variability to use in making divisions in the year. The 14-day groupings which suggest definite division times within the year are between weeks 07-08 and 09-10 and weeks 35-36 and 36-37. The study of slopes obtained from the 7-day groupings throughout the year suggests that insufficient rainfall occurrences are on record to sufficiently smooth the curve obtained from plotting slopes against time of year. It is interesting to observe, however, that between weeks 07-08 and 36-37, there are two quite definite dividing points in the year which coincide closely to those of the 14-day data.

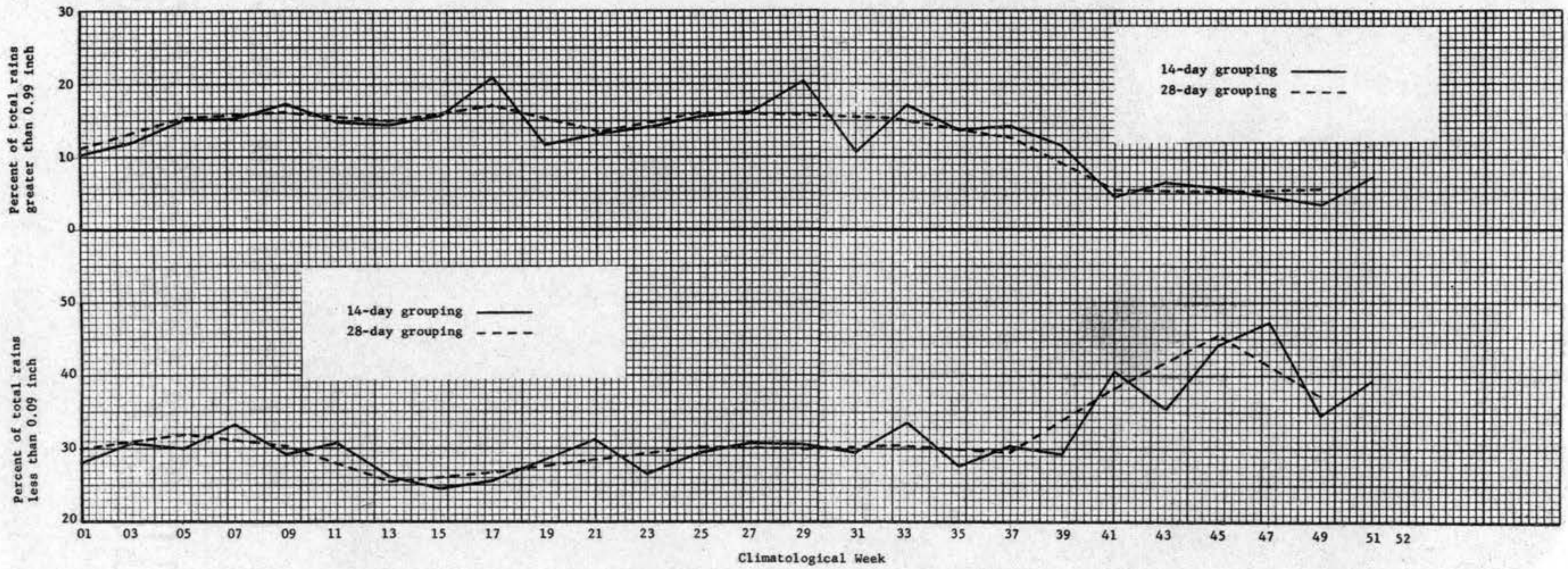
Seven-day slope values are quite consistently of higher absolute magnitude than the 28-day values. This characteristic is due to the nature of the semi-log plotting of the data. Heavy rainfall occurrences are infrequent compared to light occurrences. With the 7-day grouping the probability of the large rains being included in the total rainfall occurrences is less than with the 28-day grouping. It must be understood that in the semi-log plotting, a larger rainfall occurrence has a much greater influence on the slope value than a smaller occurrence. This possibly is desirable since the effect of larger rains is generally greater than those of smaller rains.

It is noteworthy that relative differences are actually quite small between absolute magnitudes of slope values for the 7-day as compared to the 28-day groupings.

In addition to determining slope values for the 7, 14, and 28-day groupings, they were also determined for samples of 3-day and 5-day groupings. The variability of these data exceeded that of the 7-day grouping. It was therefore considered of questionable value. Slope results from the 14-day groupings were used in dividing the year into periods of characteristic P_2 values. It was quite apparent that insufficient rainfall occurrences were on record within the 3, 5, and 7-day groupings to yield slope data which could be relied upon as being representative.

Results of plotting the percent of the total rains greater than 0.99 inch and the percent of total rains less than 0.09 inch during each week of the year (Figure 4), seem to strongly confirm the validity of using results from the slope data as a basis for dividing the year into periods of similar rainfall probabilities. It is noted that the times in the year which most noticeably define times of major change from predominately small rains to larger rains or vice versa coincide closely with the dates obtained from the slope values.

Figure 4. Percent of total rainfall occurrences which are relatively large (greater than 0.99 inch) and relatively small (less than 0.09 inch) as determined by 14 and 28-day groupings throughout the climatological year.



Pooled Data Within Periods

Plotted probabilities of receiving specified rainfall quantities are reported in Figure 5. These data were obtained from dividing the year into periods within which P_2 can be considered characteristic. Data from each of the rainfall increments employed, i.e., .01 - .09, .10 - .19, .20 - .29, etc., are plotted on the next highest hundredth-inch increment, i.e., .1, .2, .3, etc., for Figures 5, 6, 7, and 8. It is readily observed that very little scattering of points exists in any of the three plotted rainfall frequency distributions. The least degree of scattering occurred for Period 2 (starting with week 9 and ending with week 36). This would seem to support the argument that a smooth plotting of points of tabulated rainfall frequencies, representing various amounts, would result if sufficient rainfall occurrences are tabulated.

Now that results have been obtained for both P_1 and P_2, those for the third probability P can be reported. Figures 6, 7, and 8 report the products of the two probabilities ($P_1 \times P_2$) representing each of the three periods in which the year has been divided on the basis of similarity of P_2 values within each period. At this point it is essential to understand the significance of the probability P as explained in the procedure.

Since P_1 seems to vary quite independently of P_2, it was necessary to make divisions of the year within which variation of either P_1 or P_2 is considered insignificant. Each plotted point represents an average probability that on any randomly selected day within the specified periods in which both P_1 and P_2 are assumed constant, the indicated amounts of rainfall will occur.

In the study of the plotted $P_1 \times P_2$ of the divided periods, there resulted a smoothing of points similar to that of P_2 data.

Figure 5. Distribution of probabilities of receiving specified rainfall quantities during periods 1, 2, and 3.

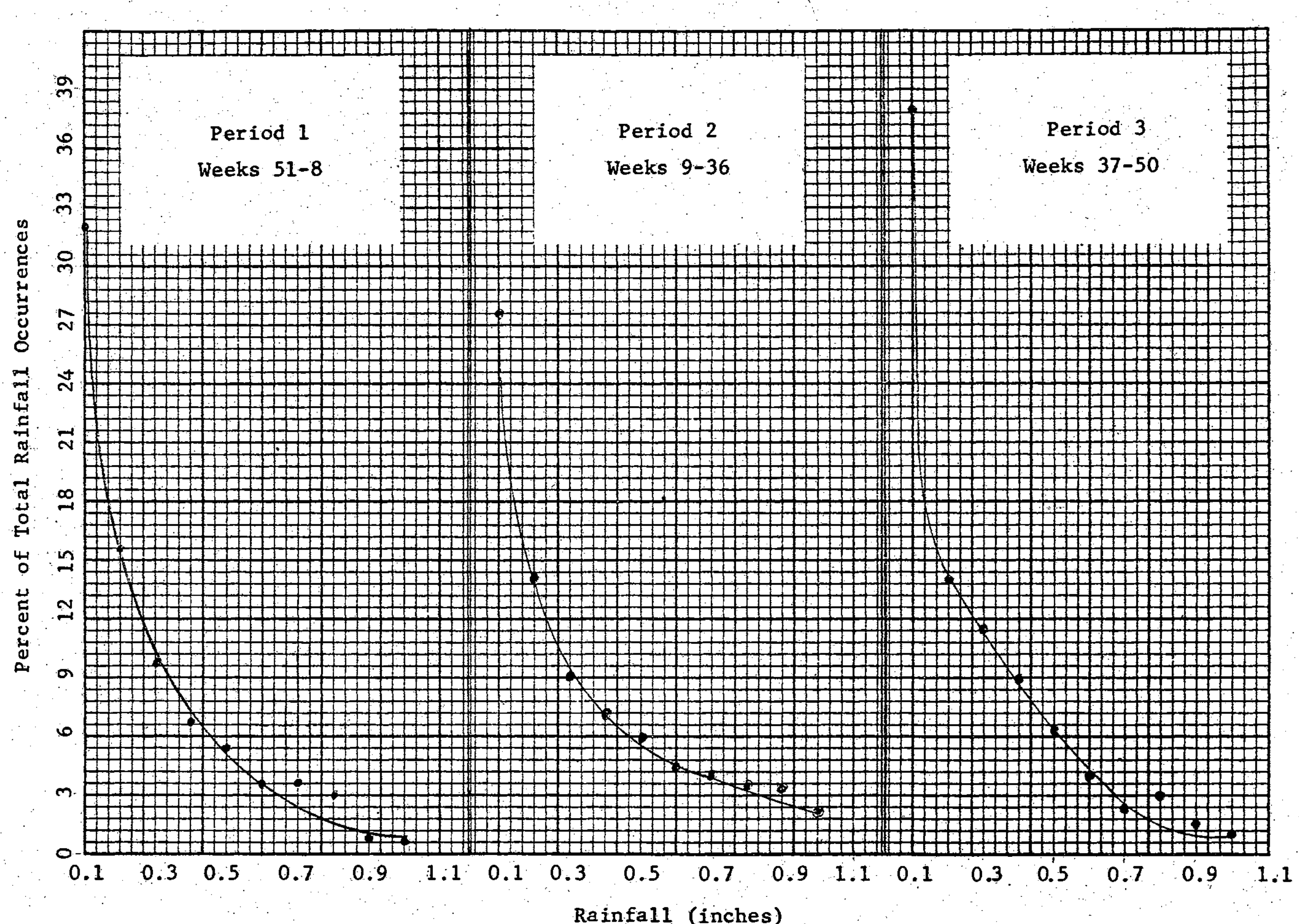


Figure 6. Average daily probability of receiving various amounts of rainfall within each of three periods from climatological weeks 51 through 8 in which both P_1 and P_2 are assumed constant.

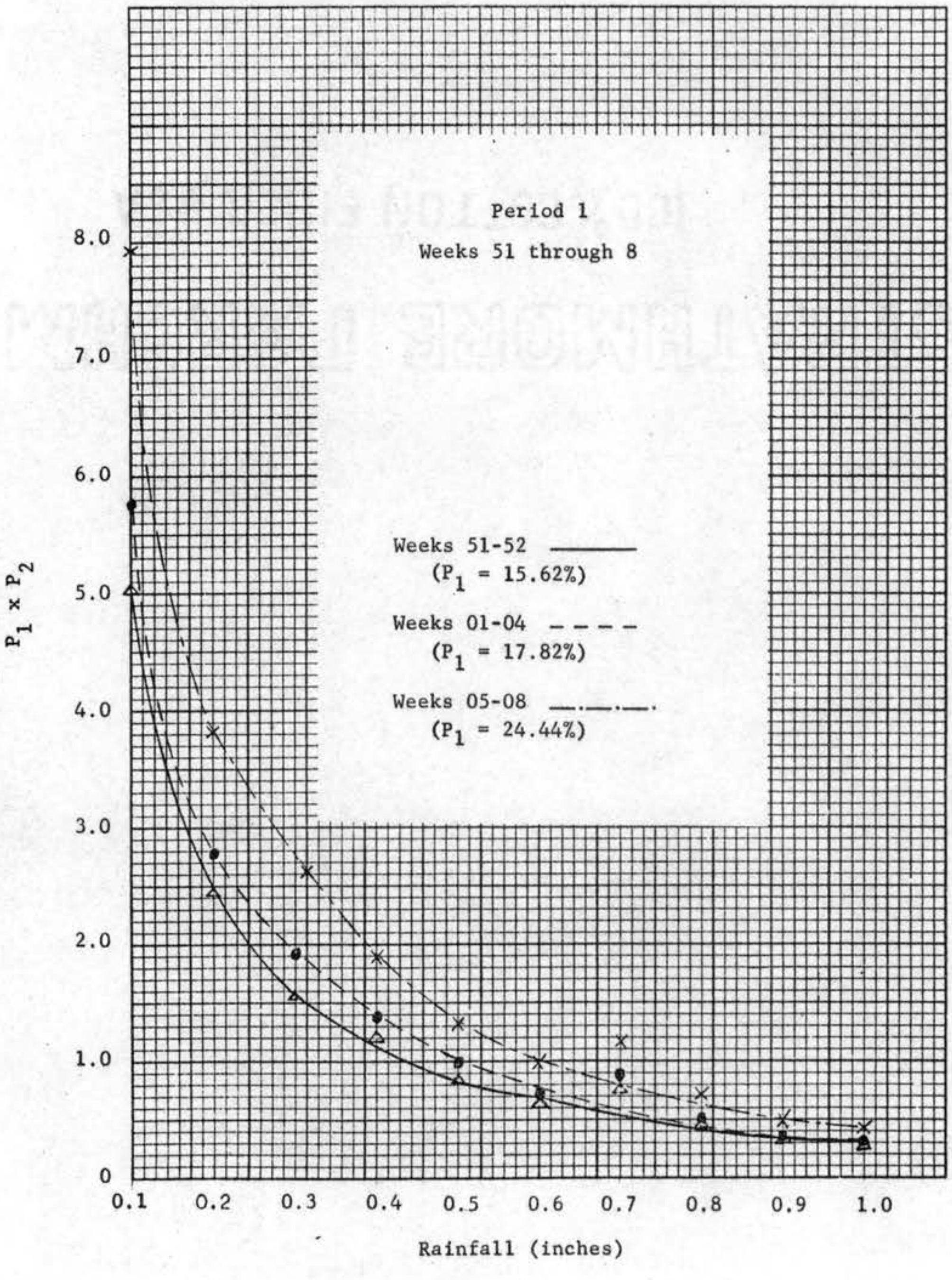


Figure 7. Average daily probability of receiving various amounts of rainfall within each of four periods from climatological weeks 9 through 36 in which both P_1 and P_2 are assumed constant.

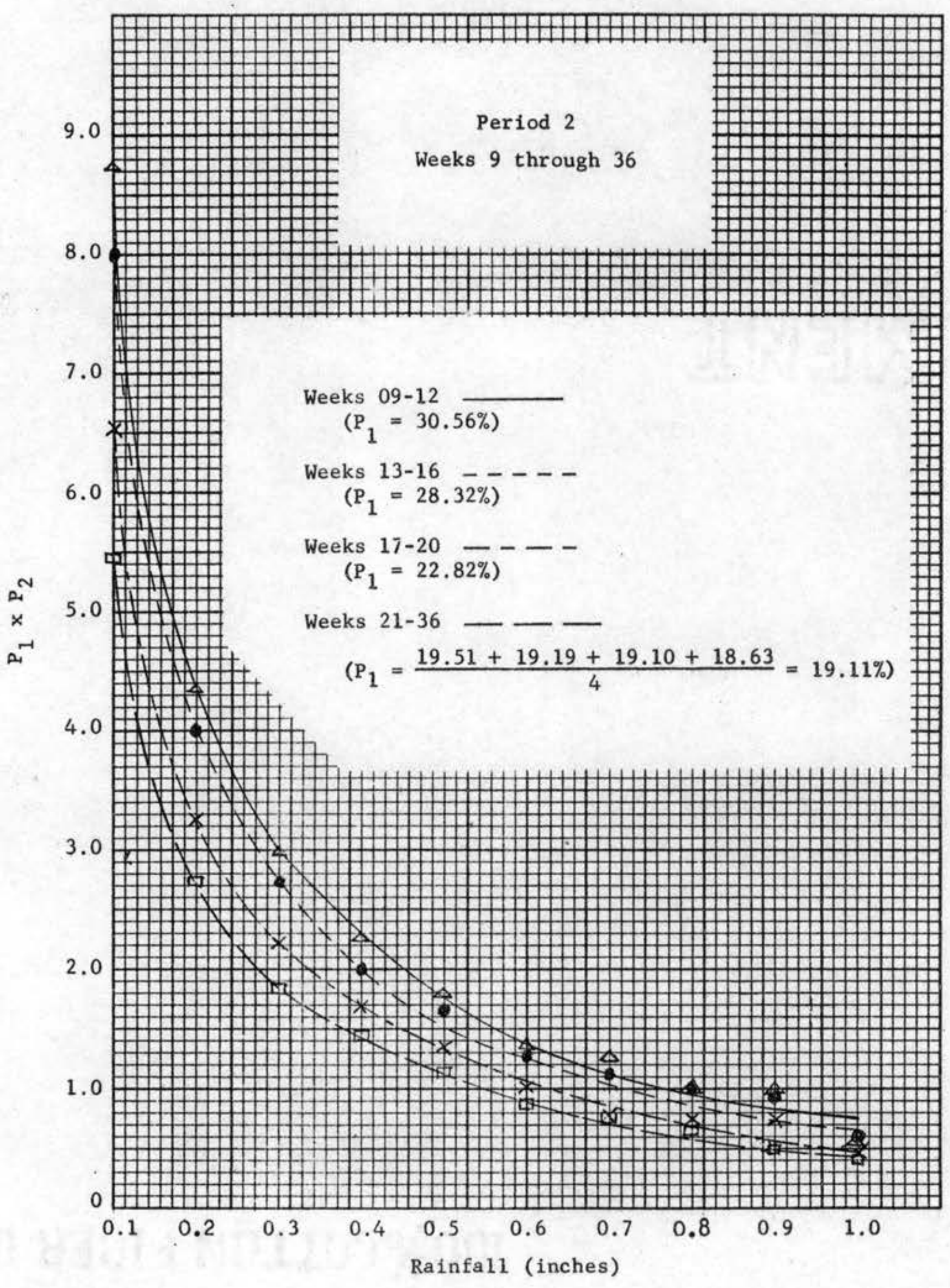
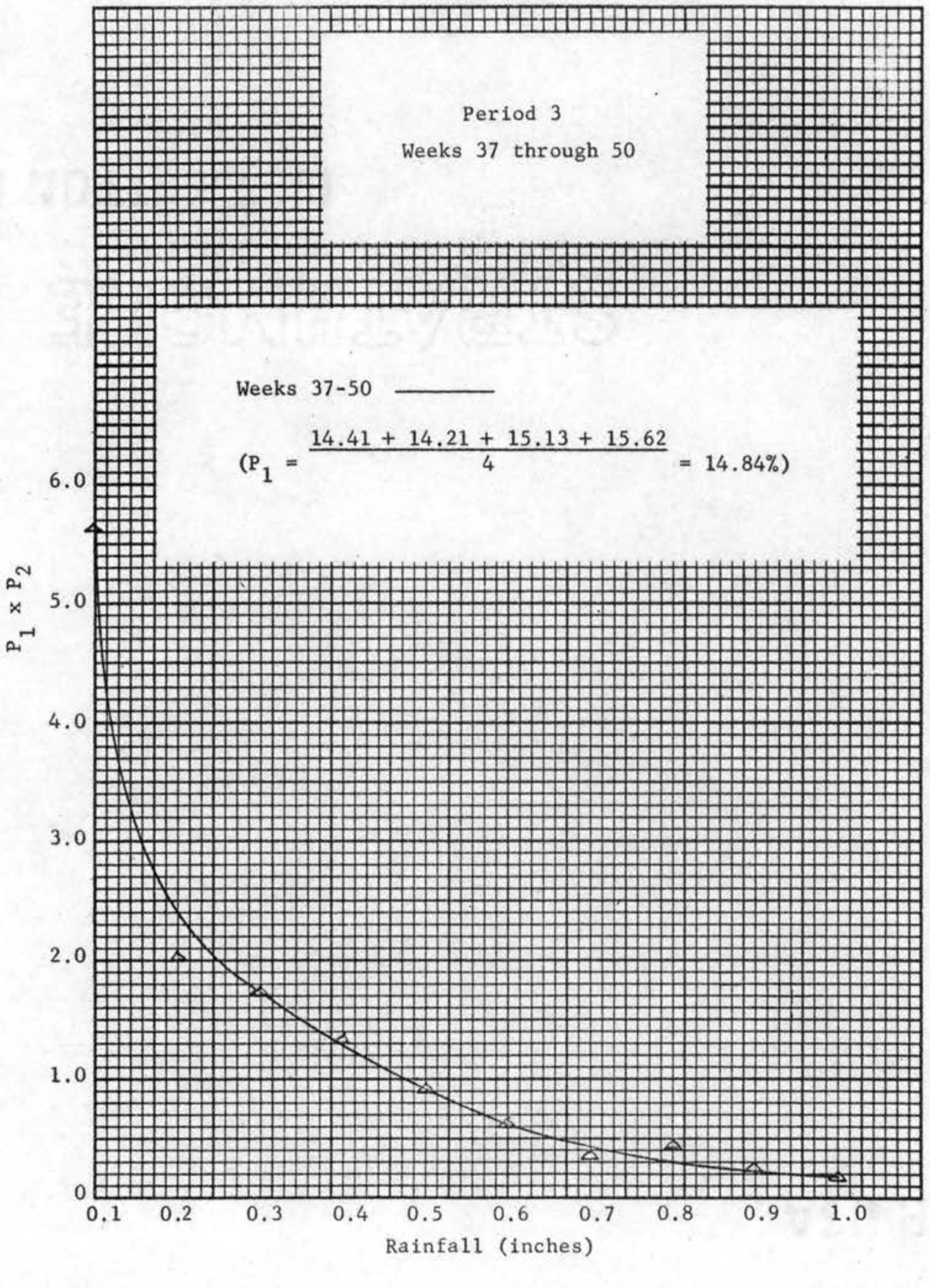


Figure 8. Average daily probability of receiving various amounts of rainfall from climatological weeks 37 through 50 in which both P_1 and P_2 are assumed constant.



CHAPTER V

SUMMARY AND CONCLUSIONS

A study involving analyses of 65 years of punched card daily precipitation data of the Stillwater station was conducted to determine whether the climatological year could be divided into portions where rainfall probabilities are representative. With the establishment of periods in which rains are similar, the total rainfall occurrences on record within each period can be pooled together for the purpose of further rainfall probability analyses.

The IBM Type 083 and 101 machines were used for sorting and tabulating frequency distributions of precipitation data. Results of individual climatological one-day rainfall distributions were very erratic. Groupings of days, including 3, 5, 7, and 14 and 28 consecutive days, were analyzed to include larger numbers of rainfall occurrences with each plotted distribution. The average probability of obtaining rain (P_1) on any randomly selected day within the grouping throughout the year was calculated. This probability was found to be as low as 14 percent in the 28-day grouping starting with week 41, and as high as 30 percent for the 28-day period starting with week 9.

The average probabilities of receiving a given amount of rainfall, provided rainfall does occur (P_2), were studied for variation throughout the year by comparing slopes of semi-log plottings of percentages of total rainfall (P_2), versus the rainfall increment amount. The slope data from the 14-day groupings of daily rainfall data throughout the year was used as the criteria for dividing the year into three portions in which P_2 values are sufficiently similar to be considered constant within each period. Slopes obtained from the 14-day groupings suggest that the following periods of similar P_2 values within each period can be made:

Period 1 includes weeks 51 through 8.

Period 2 includes weeks 9 through 36.

Period 3 includes weeks 37 through 50.

Next, all rainfall occurrences were pooled and the P_2 values assumed constant within each of the three periods. The average probability (P) of receiving various amounts of rainfall on any given day during a particular period of the year was expressed as the product of the two probabilities ($P_1 \times P_2$). In obtaining these P values, however, a further division of periods 1, 2, and 3 was necessary since P_1 seemed to vary independently of P_2. The P values, then, were obtained from pooled rainfall occurrences within a period of the year in which variations of both P_1 and P_2 are considered insignificant.

The probability of receiving 0.1 inch of rainfall on any day varied from 5.07 to 7.93 percent during Period 1, from 8.76 to 5.48 percent in Period 2, and was 5.62 percent in Period 3. The probability of receiving 1.0 inch of rainfall on any day varied from 1.77 to 0.43 percent during Period 1, from 0.66 to 0.41 percent during Period 2, and was 0.18 percent during Period 3. It is concluded from this experiment that rainfall can be characterized by the probability of receiving rain (P_1) and the probability of receiving rain in a given amount provided rainfall occurs (P_2). In this study the variations in these probabilities can be used as the criteria for dividing the year into relatively long periods.

It is proposed that a wide variety of rainfall data might be obtained from the frequency distributions of pooled rainfall occurrences within these periods.

LITERATURE CITED

1. Barger, Gerald L., and H. C. S. Thom. "Evaluation of Drought Hazard." Agronomy Journal, 41 (1949), 519-526.

2. _______, Robert H. Shaw, and Robert F. Dale. "Chances of Receiving Selected Amounts of Precipitation in the North Central Region of the United States." First Report to North Central Regional Technical Committee on Weather Information for Agriculture, Agricultural and Home Economics Experiment Station, Iowa State University, July, 1959.

3. _______, and William P. Lowry. "Synoptic Climatology of the Arizona Monsoon." University of Wisconsin Department of Meteorology Science Report No. 1, 1955.

4. Friedman, D. G., and B. E. Janes. "Estimation of Rainfall Probabilities." Connecticut Agricultural Experiment Station Bulletin 332, 1957.

5. Kincer, J. B. "Climate and Weather Data for the United States." Climate and Man, 1941 Yearbook of Agriculture. Washington: United States Government Printing Office, 685-699.

6. McDonald, James E. "Annual and Seasonal Persistence of Precipitation in Arizona." American Meteorological Society Bulletin, 41 (1960).

7. Reitan, Clayton H. "Distribution of Precipitable Water Vapor Over the Continental United States." American Meteorological Society Bulletin, 41 (1960), 79-87.

8. _______. "The Role of Precipitable Water Vapor in Arizona's Summer Rains." Technical Report on the Meteorology and Climatology of Arid Regions, No. 2. University of Arizona Institute of Atmospheric Physics, 1957.

9. Shaw, Robert H., Gerald L. Barger, and Robert F. Dale. "Precipitation Probabilities in the North Central States." University of Missouri Agricultural Experiment Station Bulletin 753, 1960.

10. Thom, H. C. S. "A Note on the Gamma Distribution." Monthly Weather Review, 86 (1958), 117-122.

APPENDIX

APPENDIX TABLE I

THE CLIMATOLOGICAL YEAR

Week Numbers with Inclusive Dates

Week Number	Beginning Date	Ending Date	Week Number	Beginning Date	Ending Date
01	Mar. 1	Mar. 7	27	Aug. 30	Sept. 5
02	Mar. 8	Mar. 14	28	Sept. 6	Sept. 12
03	Mar. 15	Mar. 21	29	Sept. 13	Sept. 19
04	Mar. 22	Mar. 28	30	Sept. 20	Sept. 26
05	Mar. 29	Apr. 4	31	Sept. 27	Oct. 3
06	Apr. 5	Apr. 11	32	Oct. 4	Oct. 10
07	Apr. 12	Apr. 18	33	Oct. 11	Oct. 17
08	Apr. 19	Apr. 25	34	Oct. 18	Oct. 24
09	Apr. 26	May 2	35	Oct. 25	Oct. 31
10	May 3	May 9	36	Nov. 1	Nov. 7
11	May 10	May 16	37	Nov. 8	Nov. 14
12	May 17	May 23	38	Nov. 15	Nov. 21
13	May 24	May 30	39	Nov. 22	Nov. 28
14	May 31	June 6	40	Nov. 29	Dec. 5
15	June 7	June 13	41	Dec. 6	Dec. 12
16	June 14	June 20	42	Dec. 13	Dec. 19
17	June 21	June 27	43	Dec. 20	Dec. 26
18	June 28	July 4	44	Dec. 27	Jan. 2
19	July 5	July 11	45	Jan. 3	Jan. 9
20	July 12	July 18	46	Jan. 10	Jan. 16
21	July 19	July 25	47	Jan. 17	Jan. 23
22	July 26	Aug. 1	48	Jan. 24	Jan. 30
23	Aug. 2	Aug. 8	49	Jan. 31	Feb. 6
24	Aug. 9	Aug. 15	50	Feb. 7	Feb. 13
25	Aug. 16	Aug. 22	51	Feb. 14	Feb. 20
26	Aug. 23	Aug. 29	52	Feb. 21	Feb. 27
			53	Feb. 28	Feb. 29

VITA

Robert B. Duffin

Candidate for the Degree of

Master of Science

Thesis: AN ANALYSIS OF PRECIPITATION DATA FOR TESTING THE VALIDITY OF POOLING RAINFALL OCCURRENCES WITHIN PERIODS OF SIMILAR RAINFALL PROBABILITIES

Major: Soils

Biographical:

Personal Data: Born August 11, 1926, at Aberdeen, Idaho, the son of Wallace and Thelma Duffin.

Education: Graduated from Aberdeen High School, Aberdeen, Idaho, 1944; attended Utah State University, Logan, Utah, 1948-1950, Bachelor of Science degree in Agronomy, 1950; graduate study at Oklahoma State University, 1959-1963.

Experience: Reared on irrigated farm at Aberdeen, Idaho; Soil Scientist for Bureau of Reclamation, summer, 1950; Assistant, Associate, and Extension Irrigation Specialist, Oklahoma Agricultural Extension Service, 1951-1959.

Member of: Sigma Xi, Epsilon Sigma Phi, and American Society of Agricultural Engineers.