UNIVERSITY OF ARIZONA

INSTITUTE OF ATMOSPHERIC PHYSICS

Technical Reports on the Meteorology and Climatology of Arid Regions

No. 1

VARIABILITY OF PRECIPITATION IN AN ARID REGION: A SURVEY OF CHARACTERISTICS FOR ARIZONA

James E. McDonald

This is the first in a series of reports on studies corollary to the University of Arizona - U. S. Weather Bureau Cooperative Punchcard Climatology Program. The analysis reported herein has been carried out as one part of an investigation supported by the National Science Foundation under Contract NSF-G1101.

December 31, 1956

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd.)

ABSTRACT

A number of statistical and meteorological aspects of the temporal and spatial variability of precipitation in Arizona have been examined in terms of their bearing on the water resources of the arid southwestern United States. Most of the work summerized has been of the nature of initial exploratory investigations made in order to lay the foundation for the much more extensive studies that will shortly be begun as part of the University of Arizona-U. S. Weather Bureau Cooperative Punchcard Climatological program.

A selected sample of long-record Weather Bureau precipitation stations in Arizona were analyzed for their historic variability properties, a number of statistical and calculational techniques were tested, and a general plan has been developed for the next phases of the Institute's variability program.

It is believed that these findings will be of interest not only to investigators in arid regions themselves but also to investigators chiefly concerned with more humid areas; for, in many respects, the statistical characteristics of arid-lands precipitation pose the most stringent of all requirements on statistical methodology. In that sense, the quantitative results of the present report may serve as useful indicators of upper bounds on the effects of non-normality, skewness, and heteroscedasticity of precipitation frequency distributions for North America in general.

VARIABILITY OF ARIZONA PRECIPITATION:

A SURVEY OF GENERAL CHARACTERISTICS

I. INTRODUCTION AND GENERAL DISCUSSION OF VARIABILITY PROGRAM.

Precipitation is regarded by the inhabitants of any given area of the world as notoriously variable in that locale. However, it is those who live and work in the arid regions, especially those living in the low-latitude deserts of the Saharan or Sonoran type, who may with most justification decry the unreliability of their local rainfall. It is a general principle, familiar to climatologists, that nowhere in the world is the relative variability of precipitation so extreme as in the arid zones. Working under this double penalty of low absolute amounts of rainfall plus large relative variability, those who inhabit the deserts and the desert fringes must always be keenly interested in the nature and causes of both temporal and spatial variability of rainfall. Viewing the clouds and precipitation of an arid region as the most important of that region's natural resources, one must regard the extreme variability of these critical resources as a regrettably dominant characteristic.

Arizona, lying as it does on the northern edge of the Sonoran desert, falls into this category of areas of excessive variability of precipitation. Consequently fundamental investigations into the climatological nature and the meteorological causes of time- and space-variations of precipitation have constituted one of the principal objectives of the research program of the University of Arizona's Institute of Atmospheric Physics since its establishment early in 1954. It is the hope of all those who have contributed to the establishment of the Institute that its investigations of climatic problems, hydrometeorological problems, and cloud physical problem of Arizona and the Southwest will shed useful light on the much broader problems of arid regions throughout the world. Hence, future Institute studies will, whenever possible, seek to be comparative in nature, examining characteristics and processes studied in the Southwest as these illuminate and are illuminated by corresponding feature of other arid regions of the earth. The present report is only the first of what will be an extended series of Institute reports during coming years treating many aspects of arid lands climatology and meteorology. As such the present study is essentially a pilot study.

In the present pilot study, the precipitation records of a limited number of Arizona stations have been subjected to a variety of statistical analyses in order to accomplish three principal objectives:

(1) To explore quantitively certain very general characteristics of Arizona's temporal and spatial precipitation variability as a means of blocking out those main problem areas wherein intensive effort can most profitably be made in the near future.

(2) To secure variability data for use in other portions of the survey of cloud and precipitation resources that the Institute is carrying out with the support of the National Science Foundation.

(3) To test and compare certain analytical techniques prior to their extensive application in the punchcard climatological investigations to be begun shortly.

Very useful findings in each of these three directions have come out of this pilot study and these have been summarized in the form of the present report. In addition, the report will include brief summaries of a number of miscellaneous studies which have been carried out in the course of the past two years of Institute work as individual surveys of minor precipitation characteristics about which design information had to be secured incidental to other Institute researches.

In the end, the kind of questions for which we badly need answers in Arizona, in the Southwest, and in arid regions in general, involve such points as the following:

(a) What are the long-term yearly, seasonal, and even monthly variabilities (i. e., standard deviations and coefficients of variation) of precipitation in the Southwest and in other arid regions? Are there significant differences between one portion of the region and another, and are there significant differences between one season and another? Elementary as such initial questions are from a statistical viewpoint, one finds in the literature virtually no quantitative information on these points for many areas of the world, including the Southwest. One variability study on a global basis was made about fifteen years ago by Conrad (1941) and a recent study for Mexico has been published by Wallén (1955), but the writer knows of no sources to which to turn for detailed information on relative variability of precipitation in the Southwest. Information concerning geographical and seasonal differences in the degree of precipitation in a given area provides basic clues to the nature of precipitation mechanisms and recent development of experimental cloud-modification techniques has underscored the need for all possible progress in delineating and explaining precipitation variability, for it is variability above all else that has stood in the way of every effort to secure an immediate answer to the question of the efficacy of cloud seeding techniques.

(b) The overall variability of a given station's precipitation series may be regarded, in first approximation, as composed of two parts--the variability contributed by secular trends or oscillation, about the long-term mean, and the variability contributed by shorter period (as, yearly or seasonal) "random" fluctuations about the secular trend. For many reasons one needs to know the comparative magnitude of these two components of total variability, so a preliminary study of secular trends at seven Arizona stations has been carried out and has been followed by analysis of the relative importance of secular oscillations and short-period random fluctuations for two selected stations. In the first of these studies results of such interest were found that steps have already been taken to extend considerably this type of investigation, and forthcoming reports will summarize this work.

(c) Full knowledge of the nature of tempora variability of precipitation on a short-term basis cannot be said to be achieved until quantitative measures of the degree of persistence of anomalies form one season to the next or from one year ro the next are evaluated. In general, such studies of persistence as have been made elsewhere in the world (unfortunately few in number) indicate that the degree of persistence is usually low when one considers time periods much greater than a few days or at most a few weeks. In the present report, representative values for Arizona precipitation are presented for interyearly and interseasonal persistence as measured by the autocorrelation coefficient. The values are found to be low-- so low that it is believed that little further work need be done on this problem except, perhaps, to examine somewhat more closely at some future time a faint suggestion of geographical variation from one part of the state to another and to check more fully the representativeness of three or four statistically significant, though small, autocorrelation coefficients found during the study.

(d) In addition to investigating temporal variability of precipitation from many viewpoints, the Institute will in the near future be studying spatial variability of Southwestern precipitation. Such studies must be conducted because the results are of basic interest in themselves for light they shed on geographical differences in intensity and reliability of precipitation mechanisms, for the information they provide concerning the characteristic scale of precipitation-releasing disturbances and related motion fields. Exploratory correlation studies to discern spatially homogeneous subregions are also prerequisite to a number of our future studies. Furthermore, it may be noted that the subject of cloud-seeding cannot be really intelligently attacked in a given region, at least by those frequently-tried methods of comparing target and control areas, until a large amount of investigation of the details of spatial variability has been completed. In many areas empirical seeding trials have been begun and even completed with only the most superficial analyses of these matters. Inasmuch as cloud-modification research represents one of the areas of strong potential interest to the Institute, it is a doubly valuable investment to begin at this early date a thorough analysis of spatial variability problems.

The preceding introductory comments should be regarded as a general statement of objectives of a continuing study of the broad problem of precipitation variability. The Institute program will embrace attacks on many fronts: Very detailed statistical analyses for Arizona, synoptic-climatological studies on a scale of the whole Southwest, general circulation studies of ocean-continent influences in the aridity and variability of the Southwest and similar climatic regions, and a variety of researches on the cloud physics of Southwestern precipitation.

The preliminary results reported below are those which have proved useful in planning the first two of these several approaches, namely a detailed study of Arizona precipitation characteristics, and further study on a geographically broader basis for the whole Southwest.

II. TEMPORAL VARIABILITY.

Any given locality experiences a seemingly chaotic sequence of "wet" and "dry" seasons and years. There are many features of the variability of such time series that need elucidation in the Southwest. In this section of the present report, variations of precipitation for selected Arizona stations will be examined in terms of the coefficients of variation of the annual and seasonal totals, the decadal-moving-average seasonal precipitation amounts, the autocorrelation coefficients of seasonal and annual totals, and several other statistics.

Most of the analyses in this pilot study dealt with the records at seven Weather Bureau cooperative stations chosen for length of record and geographical distribution over the state. The locations of these seven stations are shown in Fig. 1, where the seven principal stations' names are printed in bold letters and the names of other Arizona stations mentioned occasionally in the report are printed in small letters.[1]

Yuma, with an 83-year mean annual precipitation of 3.2 inches, lies in the driest area of the United States at an elevation of 190 ft. above sea level. Natural Bridge, with a 63-year mean annual precipitation of 24.0 inches, is at an elevation of about 4600 ft. on the south slopes of the Mogollon Plateau in central Arizona. These two stations have the smallest and largest mean precipitations, respectively, of the stations whose records are examined here. Flagstaff at 6900 ft. msl lies north of the Mogollon escarpment on the south edge of the Colorado plateau and has had a 48-year average precipitation of 20.5 inches. Prescott, also south of the Colorado plateau, lies near the north edge of a range of mountains that are outliers of the plateau, topographically speaking. Prescott's elevation is 5400 ft. msl (average of a number of past station locations) and its 78-year average precipitation has been 18.7 inches. Phoenix lies near no major mountain ranges, at an elevation of about 1100 ft. msl, in the confluence of the Salt and Gila valleys. Its low elevation and absence of nearby major orographic influence precludes its receiving large precipitation amounts since, in the Southwest, these two factors are of dominant importance. The 75-year record yields an average of 7.6 inches. Tucson and Bowie in southeastern Arizona have altitudes of about 2500 ft. and 3800 ft. msl respectively. Proximity to major mountain masses contributes to the prcipitation of both of these stations so that Tucson's 87-year average has been 11.2 inches while Bowie's 67-year average has been 12.0 inches.

Although it is not yet known whether these records can truly be said to be representative of the several regions which they were selected to "represent", this group of stations does include examples from what are probably the chief precipitation provinces of the state. It should be noted that almost all of the analyses reported here (exception: computation of one portion of Bowie's decadal moving average) have been performed on the data as reported in the several published sources from which the data were extracted. No adjustments for non-homogeneity of reported data were made. This was necessary simply because no previous studies exist to shed light on the degree of internal homogeneity of any of the southwestern precipitation records. Indeed, one of the several purposes of this study was to secure background for that very type of homogeneity study. Later, in this report, the results of a double-mass analysis of these seven stations will be described and the apparent degree of homogeneity of each of these records

---

1 All of the data analysed in this study are drawn from published summaries of the U. S. Weather Bureau. In order to secure data covering a long time-period in each locality of interest, it was found necessary to combine certain records. Statistical examination of the validity of this step is discussed later in this report.

will be discussed. In anticipation of that discussion, it may be noted that several of the records showed irregularities but only Bowie's record contained so obvious a break (at time of transferring observation from old Fort Bowie on the foothills of the Chiricahua Mountains to the town of Bowie fourteen miles north-northwest) as to demand adjustment in the study of secular trends. Hence, although the numerical quantities computed below must not be regarded as final values, it can at least be said that the adjustments for non-homogeneity that are soon to be made as part of a systematic Institute study using the double-mass method will probably impose only rather small changes on most of the statistics computed in this preliminary study (with the possible exception of Bowie). Prescott's double-mass curve is suspect, but no really major breaks appeared; so for the present purposes no adjustment was attempted.

Precipitation record-lengths for these stations are not identical, as noted above, but range from 48 years at Flagstaff to 87 years at Tucson. No attempt has been made in this study either to reduce all records to the shortest period available for all (i.e., that of the Flagstaff record), or to extend any of the shorter records by extrapolation methods which could at best be only very approximate in the present state of our knowledge of the finer details of Arizona precipitation characteristics. In the case of the secular trend studies, of course, it was highly desirable to use every bit of available data, since there are many ecological and hydrological problems as well as meteorological problems in this area which demand study with respect to the oldest obtainable rainfall data. In all, this nonuniformity in record length must be expected to bias slightly some of the comparisons made below, but it certainly does not vitiate the kind of results sought here -- namely results that can orient the next steps forward in the series of successive approximation that must characterize any careful investigation of historic precipitation (statistics in a region where little previous work has been done.)

One of the first steps in the present scrutiny of the temporal variability of selected Arizona stations was the computation of standard deviations[1] and associated coefficients of variation for the annual totals, winter totals, and summer totals. The results will next be summarized and then other temporal variability statistics discussed.

A. Coefficients of variation. The most useful single statistic expressing the degree of variability of a station's precipitation is the coefficient of variation, that is, the ratio of the standard deviation of the precipitation amounts to the mean amount. Whereas a two-inch standard deviation of, say, annual precipitation at a mountain station such as Natural Bridge would imply only small relative variability, that same two-inch standard deviation of annual rainfall at Yuma (and this is the approximate standard deviation at Yuma) implies a very extreme degree of relative variability; hence the standard deviations themselves are less revealing than are the dimensionless quotients of standard deviations divided by corresponding means, i.e., co-efficients of variations.

In Table 1 the results of the computations of means, standard deviations, and coefficients of variation are presented for the seven principal stations plus Oracle, which lies about 30 airline miles north-northeast of Tucson on the northern foothills of the Santa Catalina Mountains.

[1] Readers who object to the use of the term "standard deviation" when applied to non-normal frequency distributions may substitute "root-mean-square-deviation". For brevity here, the former term is used.

Table 1

Means, standard deviations, and coefficients of variation of seasonal and annual precipitation at eight Arizona Weather Bureau stations.

Station	Means (inches)			Record Length (yrs.)	Standard Deviations (inches)			Coefficients of Variation and standard errors		
	$\bar{w}$	$\bar{s}$	$\bar{y}$	N_y	$\bar{w}$	$\bar{s}$	$\bar{y}$	C_w	C_s	C_y
Bowie	4.98	6.98	12.01	67	3.08	2.97	4.41	.62±.05	.42±04	.37±.03
Flagstaff	10.14	10.37	20.53	48	3.72	2.91	4.81	.37±.04	.28±03	.23±.02
Natural Bridge	13.46	10.62	23.99	63	5.75	3.67	7.16	.43±.04	.35±.03	.30±.03
Oracle	9.98	9.15	19.17	51	4.70	2.72	5.54	.47±.05	.30±.03	.29±.03
Phoenix	4.31	3.28	7.62	75	2.45	1.83	3.05	.57±.05	.56±.05	.40±.03
Prescott	9.34	9.32	18.74	78	4.66	2.91	5.49	.50±.04	.31±.02	.29±.02
Tucson	4.51	6.69	11.20	87	2.43	2.59	3.28	.54±.04	.40±.03	.30±.02
Yuma	1.91	1.35	3.25	83	1.43	1.27	2.01	.75±.06	.94±.07	.62±.05

Note: The letters w, s, and y denote, respectively, winter, summer, and yearly quantities. By the definition used throughout this report, "winter" is the period from 1 November to the following 30 April, and summer is the remainder of the year.

In the first three columns of Table 1 are given the computed winter, summer, and yearly mean precipitations for each station and in the fourth column is the total number of full years of precipitation data entering into the yearly mean for each station listed. (For reasons of spottily missing data, the number of full years is, in every case but that of Tucson, a few years smaller than the number for the summers and winters. In the next three columns are the calculated standard deviations, and in the following three columns are given the seasonal and annual coefficients of variation. Following each tabulated coefficient of variation is the associated standard error given approximately by the quantity $c/\sqrt{2n}$.

1. Comparison with other areas. The most obvious feature of the Arizona coefficients of variation listed in Table 1 is the generally large value of these coefficients for all Arizona stations. There are not, in the literature, many published data on coefficients of variation for other areas. Foster (1948) has published some for midwestern U. S. stations for annual precipitation and Soucek and Howe (1938) have computed some for Iowa stations and for six other stations around the U. S., also for annual totals only. Landsberg (1951) has presented values for Oahu and Wallen (1955) has published a fairly extensive tabulation for Mexico. The data from the first two sources just mentioned are used here to provide some comparative data for regions outside of Arizona. In Table 2 these values are listed, along with the length of record used in their computation. In all cases, the record length seems adequate to insure stability of the statistic to about ten per cent of its computed value, if one accepts normal-distribution sampling theory as a basis for a rough estimate of reliability.

Comparison of the values of C_y of Table 1 and Table 2 reveals that with the exception of Flagstaff, Arizona C_y's exceed all non-Arizona values given in Table 2 except for the values for Pueblo, Colo. and San Diego, Calif. In a general way, this is just what is to be expected on the principle that relative variability of precipitation tends always to vary inversely with mean total precipitation (actually the relation is of roughly hyperbolic form).

Arizona is a semiarid state and the relative degree of variability should be quite large, as verified by Tables 1 and 2. The case of Pueblo, with a mean of 11.51 inches seems exceptional, for Tucson (with about the same mean annual precipitation) has an annual coefficient of variation of only 0.30. This raises the important question of which of these two cases constitutes a climatological anomaly. Conrad (1940) has examined this kind of question (using, however, the ratio of mean deviation to mean precipitation, sometimes called the "relative variability") and has shown that the use of "variability anomalies" can reveal significant climatic features of a region. Until completion of further studies, it will not be possible to say whether Arizona or eastern Colorado is the anomalous area, though Conrad's findings suggest that the Pueblo data are the exceptional ones.[1]

[1] About forty years ago, Hazen (1916) published a map of the general distribution of coefficients of variation for U.S. annual precipitation, but gave no tabulation of individual values. There are numerous points of contradiction between Hazen's map and one constructed by the writer from the data given by Conrad (1940). Although some of the differences would follow from skewness effects (Conrad used the ratio of mean deviation to mean precipitation rather than the coefficient of variation) the disagreement of the patterns of variability is large enough to indicate need for the extensive computation of U.S. coefficients of variation that we are now preparing to undertake.

Table 2

Comparative values of coefficients of variation of annual precipitation for non-Arizona stations. Data taken from Foster (1948) and from Soucek and Howe (1938)

Station	Years of Record	Mean annual precipitation (inches)	Coefficient of variation C_v
Bismarck, N. Dak.	66	16.26	.25
Boston, Mass.	60	41.02	.16
Cheyenne, Wyo.	70	14.61	.25
Cleveland, O.	60	34.02	.16
Iowa City, Ia.	70	35.17	.14
Kansas City, Kans.	63	36.10	.18
Mexico, Mo.	63	38.16	.18
Ogden, Utah	60	16.22	.25
Pierre, S. Dak.	49	15.85	.25
Portland, Ore.	60	42.25	.19
Pueblo, Colo.	52	11.51	.46
San Diego, Calif.	60	9.90	.44
Sheridan, Wyo.	47	15.00	.27
Sioux City, Ia.	83	26.41	.26
Wichita, Kans.	52	29.22	.26

2. Differences in winter and summer variability. In the preceding section it has been noted that the annual Arizona coefficients of variation are large compared to most of the values available for non-Arizona coefficients, but it has been pointed out that this is not climatologically very surprising. However, there is another quite unexpected feature of Table 1 associated with seasonal differences in variability. The fact that all but one of the eight stations exhibits a larger coefficient of variation for winter precipitation than for summer precipitation contradicts what had been the writer's preconceptions about winter-summer differences in Arizona precipitation characteristics, and also contradicts what appears to be the popular notions of the long-time residents in this area. Summer convective rains are so evidently spotty in a spatial sense that it seems reasonable to attribute a high degree of randomicity in general to the summer rains of Arizona, while winter precipitation from migratory Pacific cyclones with widespread cloud cover is rather naturally expected (incorrectly) to exhibit a generally lower measure of temporal randomicity. That one cannot so equate spatial variability and temporal variability in Arizona is clearly shown by the winter-summer differences of seasonal coefficients of variation listed in Table 1.

It will be shown later (in discussions of interstation correlations of seasonal precipitation at these stations) that whereas Table 1 shows the winter precipitation to be more variable from year to year at any given station than is the summer precipitation, the winter precipitation in any given year is less variable from station to station than is the summer precipitation. That is, the visually evident spottiness of summer thunderstorm precipitation in Arizona on any given day does show up even in seasonal totals in the form of low summer interstation correlation coefficients, yet this statistical characteristic goes hand in hand with a greater degree of long-term temporal reliability of the summer precipitation than of winter precipitation at individual stations.

Put in still more general terms, these initial results on seasonal comparisons of coefficients of variation and of coefficients of correlation (the latter to be described below) show that the general-circulation factors which govern the relative amount of cyclonic winter precipitation in Arizona are intrinsically more variable from year to year than are those large-scale factors governing arrival of summer moisture and its precipitation; but these winter precipitation paramenters tend to bring either a generally wet or a generally dry season to all parts of Arizona at one time whereas the summer shows much less tendency in the latter direction.

The aspect of this that is difficult to understand in terms of our presently available knowledge about the meteorology of this region is why the winter cyclonic rainfall is subject to significantly greater interyearly fluctuations than is the summer convective rainfall. Since each of the seasonal precipitation regimes of Arizona are almost certainly related to the vagaries of the Atlantic and Pacific Ocean anticyclones, it will be a problem deserving much intensive Institute research to determine the details of the effect of the oceanic highs on this difference in seasonal reliability of precipitation. That the differences are generally significant, in the statistical sense, may be seen by inspection of the standard errors of the coefficients of variation as given in Table 1. For all stations except Phoenix and Yuma the winter coefficient of variation differs from the corresponding summer value by an amount equal to or greater than twice the larger of the two standard errors of the two respective seasonal coefficients of variation.[1]

The case of Yuma is distinct from that of all others shown in Table 1 in that its summer variability is relatively larger than its winter variability. This difference is chiefly a reflection of the substantially lower summer mean rainfall at Yuma and cannot be regarded as a very meaningful point, though the tabulated coefficients are unambiguously descriptive of the numerical properties of the historic Yuma precipitation.

The unexpectedly large magnitude of the winter coefficients of variation constitutes the point of greatest interest in the table of these coefficients for the eight Arizona stations, but one other implication is to be noted. The fact that just within the area of the state of Arizona the computed coefficients range from as low as 0.23 for the annual precipitation at Flagstaff to as high as 0.62 for Yuma's annual precipitation emphasizes the need for a much more detailed study of the patterns of variability for the state.

3. Related frequency distribution characteristics. The variability of precipitation that is summarized in a single standard deviation or a single coefficient of variation can also be exhibited, less concisely, but in some ways more informatively, through histograms of the frequency distributions of precipitation amounts.

In Figure 2 we are shown such frequency histograms for five Arizona stations for winter, summer, and annual totals over the full period of record for each. (See Table 12 for record lengths.) The ordinate function in every one of the histograms represents the per cent of all cases that fall within class intervals of width specified on the axis of ordinates. Means are labelled on each histogram for reference purposes.

Little discussion of these histograms is needed, for these depict graphically the large variability of Arizona precipitation in self-evident manner. That so many tall narrow spikes appear in histograms based on such long records, above all in cases such as the Natural Bridge winter and annual diagrams, suggests that one would have to have something like several centuries of observations before a smooth frequency curve would

---

[1] These "standard errors" of Table 1 are based on formulas drawn from normal-distribution theory and hence are, strictly speaking, to be regarded as only approximations to standard errors. A study of non-normality effects in Arizona precipitation distributions that will be reported shortly by the writer has revealed that, despite the skewness in these distributions, statistics computed with the raw data, assuming normality, are remarkably close to those obtained by each of three (square root, cube root, and logarithmic) normalizing transformations tested. For the present purposes, in particular, the tabulated "Standard errors" can quite safely be interpreted in conventional sense.

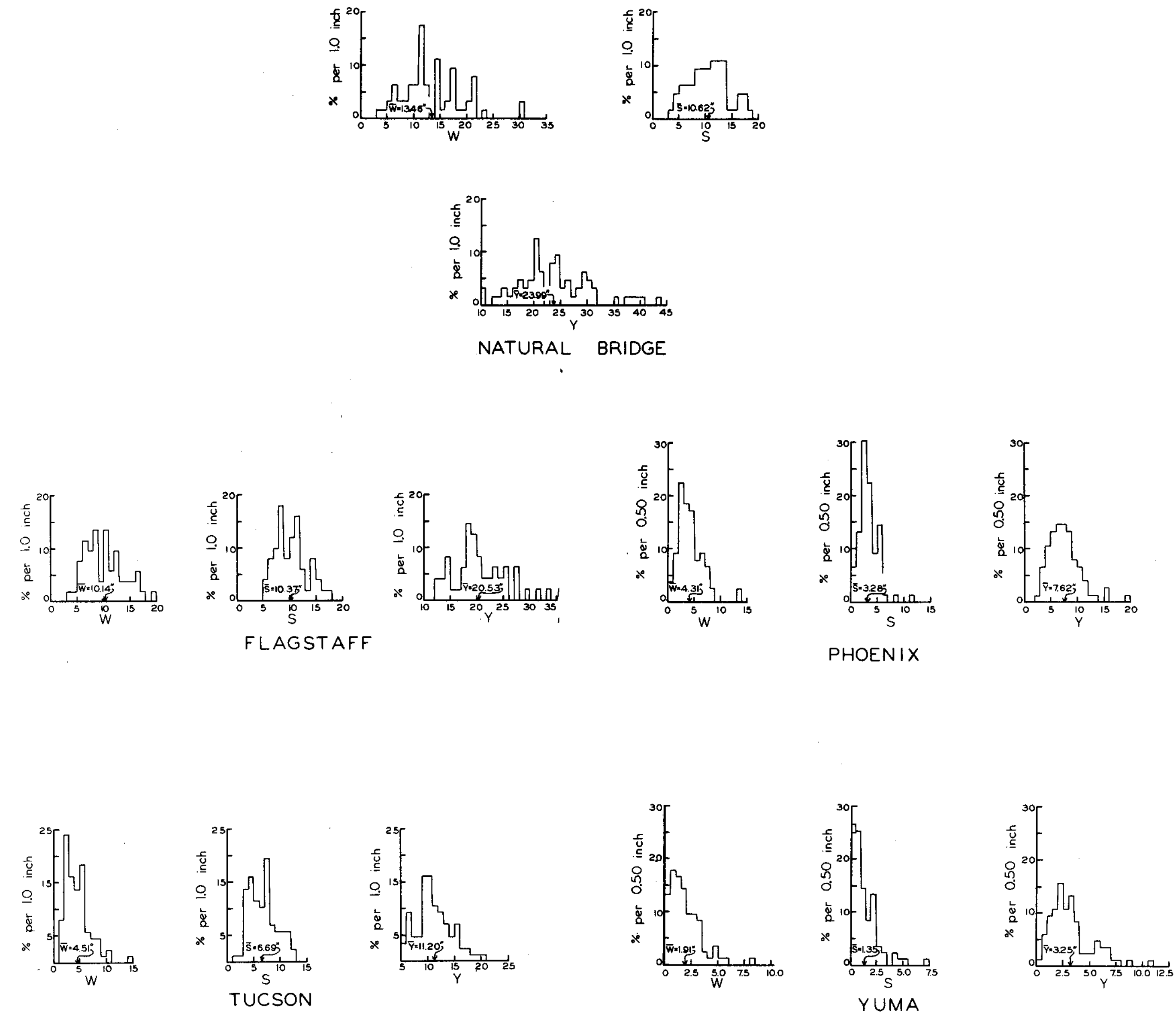


Figure 2. Frequency-distribution histograms for precipitation at five Arizona long-record stations. Letters, "w", "s", and "y", refer, respectively, to winter, summer, and yearly total amounts. Winter is the half-year from Nov. 1 to the following Apr. 30, summer is the remainder of the year. Ordinate function is the per cent of all cases falling in the specified class-interval widths on the abscissa. For number of cases used, see Tables 1 and 1a. The mean for each distribution is plotted within each histogram.

emerge.[1] Note that differences in variability between winter and summer are not uniformly evident in these histograms, though Natural Bridge is outstanding in this respect.

The skewness typically associated with arid-zone precipitation is readily discernible in these histograms, particularly in the case of Yuma. Analysis of the frequency distributions revealed that, though it is not uniformly apparent from the histograms, all of these distributions are skewed, and all in the same sense as that of the Yuma skewness.

A summary of the numerical measures of skewness, and of a number of other frequency-distribution characteristics for the winter and summer precipitation at the five stations represented in Figure 2 will be found in Table 1a. For each station and season, the record-length is given, then the mean and the median. Following that, the 10th, 25th, 75th and 90th percentile values of precipitation are listed (all amounts in inches). Next the percent of all cases, in each station-season category, lying beyond one standard deviation from the mean are listed, the first of the two values referring to the cases above the mean plus the standard deviation, the second referring to the cases lying below the mean minus the standard deviation. Finally, two different measures of skewness are evaluated; First, the quotient of three times the mean-median difference divided by the standard deviation, and second the measure used by Landsberg (1951) in his study of Hawaiian rainfall, namely one hundred times the quotient of the absolute value of the difference, mean minus median, divided by mean.

A few brief comments on the contents of Table 1a will suffice here. The medians may be thought of as part of the tabulation of percentiles, being equal, by definition, to the 50th percentiles. Note that as a rough rule of thumb, the 90th percentiles are very nearly three times the 10th percentiles at all of these stations in the winter. This constitutes a climatologically very large ratio, and is merely one more measure of the high winter variability of Arizona precipitation.

In an ideal normal distribution, approximately 16 per cent of all cases will lie, respectively above and below one standard deviation from the mean of the distribution. Inspection of Table 1a shows that for most of the distributions, this relationship holds to a rather unexpected degree, the worst exceptions being the Phoenix winter and the Yuma summer distributions. Briefly, one can say that, by and large, no extremely serious errors of inference will be made in simple deductions based on

[1] I am indebted to Dr. R. A. Bryson for calling my attention to the fact such a statement may be unwarranted. If further investigation should reveal that, say, the summer precipitation in this area results from varying frequencies of occurrence of several meteorologically distinct synoptic regimes having quite different individual frequency distributions of precipitation amounts, no length of record would obliterate polymodality. This viewpoint has been recently discussed by Schneider-Carius and Essenwanger (1955).

Table 1a

Some frequency-distribution characteristics for winter and precipitation at five Arizona stations. See text for explanation of table entries. All precipitation amounts are in inches.

Station	Season	No. Years	M Mean	m Median	Percentiles 10	25	75	90	Percent of all cases: Above $M+\sigma$	Below $M-\sigma$	Measures of Skewness: $3(M-m)/\sigma$	$100\|M-m\|/M$
Flagstaff	w	57	10.14	9.18	5.95	7.31	12.45	15.75	16	16	.77	10
	s	57	10.37	9.75	6.78	8.16	11.66	14.46	16	16	.64	6
Natural Bridge	w	64	13.46	11.95	6.80	9.62	17.34	20.16	17	16	.78	11
	s	64	10.62	10.42	5.42	7.96	13.35	16.41	12	17	.16	2
Phoenix	w	78	4.31	3.83	1.94	2.63	5.55	7.25	8	13	.58	11
	s	78	3.28	3.06	1.40	2.75	4.69	5.48	15	12	.36	7
Tucson	w	87	4.51	4.10	2.00	2.84	5.75	5.90	10	13	.51	9
	s	87	6.69	6.52	3.42	4.47	8.01	10.95	17	19	.20	3
Yuma	w	84	1.91	1.56	0.43	0.94	2.72	3.59	13	13	.73	18
	s	84	1.35	0.95	0.11	0.45	2.00	2.95	5	11	.94	30

the observed standard deviations for Arizona stations lying outside of the driest (southwestern) sections of the state.

The first listed measure of skewness will not be discussed here. It is given only for reference purposes for those who may wish to see numerical values of this particular skewness measure. (The writer is not certain who these readers will be, for the literature with which he is familiar contains no tabulations of this skewness measure for any other areas of the world. This first measure is, nevertheless, a standard measure suggested in works on statistical methods and has therefore been computed as an incidental part of these calculations.)

The second measure of skewness was computed to permit comparison with the only other numerical values of skewness with which the writer happens to be familiar, namely those which Landsberg (1951) has presented for a number of stations on Oahu in the Hawaiian Islands. On comparing the two sets it appears that the Hawaiian precipitation distributions are substantially more skewed (for comparable mean amounts) than are the Arizona distributions. Landsberg tabulates only annual skewness figures, but since annual total rainfall in Hawaii is chiefly the winter wet-season total, and since Pacific frontal-cyclonic activity controls both the Hawaiian and the Arizona winter amounts, it seemed not unreasonable to compare Landsberg's annual skewnesses with the Arizona winter skewnesses. Plotting the five winter values of $100|M - m|/m$ for Arizona stations on Landsberg's Figure 4 (a scattergram of his annual skewness measure vs. mean precipitation amounts), one observes a striking disparity between the location of the closely clustered Arizona points and the hyperbolic trend of the Hawaiian annual points. Whereas 10% is a rough average of the Arizona winter skewnesses, the values for corresponding means in Hawaii as read from Landsberg's Figure 4 would be so large as to lie well above the limit to which Landsberg's plot is carried, perhaps 60-70% if one extrapolates his fitted curve. Indeed, Landsberg's fitted curve does not steady off at the low-skewness limit of about 10% until annual means of the order of 100 inches are reached.

Landsberg's Figure 4 is based on an annual skewness measure which was computed as an average of twelve monthly skewnesses. To check whether there might be some unrecognized influence of the summer months, the twelve monthly values for Honolulu were computed and were found to show no undue influence of this sort, though the May and the August values of about 40 percent do tend to increase the average a bit. (The Honolulu 12-month average skewness is 32 percent for a mean annual rainfall of 25.2 inches). The clue to the seemingly large discrepancy between the Arizona winter skewness and the Hawaiian annual skewness was finally recognized from closer inspection of the differences between annual means and annual medians in Landsberg's Table 4; for Honolulu's difference is only about 0.3 inches compared to an annual mean of 25.2 inches (i.e., skewness of annual totals of 100x0.3/25.2 or only 1 percent) and most of the other stations studied by Landsberg have differences between annual means and medians that yield skewnesses as small, or smaller than the Arizona skewnesses. Thus, the conclusion reached here is that no obvious discrepancy exists after all.

To test this conclusion in a single readily accessible case, the monthly precipitation data given by Landsberg for Honolulu (his Table 2) were used to compute winter (Nov.-Apr.) totals for every year from 1874 to 1945. From the resultant frequency distribution, the median value of 16.0 inches was found for Honolulu winters, and since the corresponding mean is 17.6 inches (Landsberg's Table 4 gives 17.5 for the 1917-44 period, which is surprisingly close to this 72-year Honolulu mean), the skewness of Honolulu winter amounts is found to be 100x1.6/17.6 = 9 percent. Natural Bridge, Arizona, with a winter mean of 13.46 inches has, by way of comparison, a skewness of 11 percent, so it is quite likely that (as one might expect) the skewness of Arizona precipitation is at least as large as, and probably even a bit larger than, that of Hawaii, when the two areas are compared on the same basis.

Use of the phrase "skewness of annual rainfall frequencies" to describe an average of twelve monthly skewnesses seems to the writer to be undesirably misleading. Some such descriptive identification as "mean monthly skewness" for the quantity employed by Landsberg would obviate the misconception which the writer fell into on first comparing the Hawaiian with the Arizona data. That these two climatological statistics must be quite clearly distinguished is amply demonstrated by the Honolulu data, for the difference between Landsberg's figure of 32 percent based on the average of twelve monthly skewnesses and the value of 1 percent obtained above by using the annual mean and median for this same station is very great. The skewness of individual monthly frequency distributions will almost certainly be much greater for any climatic station than the skewness of the annual amounts, though it is an open question whether Honolulu's thirtyfold ratio would be found representative. (It might also be noted that if there is appreciable difference in mean precipitation from month to month at a station, simple averaging of monthly skewnesses suffers the same kind of weakness as does averaging percentages).

B. Confidence intervals for mean precipitation amounts. Because computation of an arithmetic mean is a simple procedure, and because a mean so nicely summarizes in a single statistic the general magnitude of the series of variates averaged, it is one of the most commonly used of all descriptive statistics. It is tempting, however, to quietly slip it from the category of descriptive statistics over into the category of predictive or inferential statistics, and tacitly to draw conclusions without giving due attention to the sampling variations in means of finite samples. Hence, in beginning a statistical study of any new problem it is a wise precaution to make at least a few quantitative estimates of the reliability of the available means of the items under study. In Table 3 are given such estimates for the seasonal and annual mean precipitation amounts that are tabulated in the first three columns of Table 1.

The quantities used to display the reliability (or rather the unreliability) of the precipitation means are the half-widths of the 95% confidence intervals, expressed in per cent of the respective means. The meaning of these half-widths is best explained in terms of a specific example, say that for winter precipitation at Bowie: The available record for Bowie covers a total of 69 winters, two more than the number of full years of record shown in Table 1. The computed mean winter precipitation based on this period is seen (Table 1) to be 4.98 inches, with an historic coefficient of variation of 0.62. The latter variability is large--so large that one must be cautious about avoiding simple statements such as "The mean winter prcipitation at Bowie based on its long record is known

Table 3

Half-widths d of the 95% confidence interval for mean seasonal and annual precipitation, expressed in terms of percent of the respective means, for eight Arizona stations*

Station	N_w	d_w	N_s	d_s	N_y	d_y
Bowie	69	15	71	10	67	9
Flagstaff	52	11	50	8	48	7
Natural Bridge	63	11	64	9	63	8
Oracle	53	13	55	8	51	8
Phoenix	76	13	76	13	75	9
Prescott	79	11	80	7	78	7
Tucson	87	12	87	9	87	6
Yuma	84	17	83	21	83	14

*N is the number of years of record; subscripts w, s, y refer to winter, summer, and yearly precipitation totals respectively, winter is taken to cover precipitation from 1 November to the following 30 April, and summer is the remainder of the year.

to be 4.98 inches." In fact, the only statement that we can justifiably make in the face of such variability is that the Bowie winter mean is somewhere near 5 inches.[1] For an estimate of how near (if we momentarily preserve the fiction that there exists a "true sample mean Bowie winter precipitation"), we may use the 95% confidence half-width given (see for example Ostle, 1954, pp. 80-82) by the approximate relation

$$d = 200C/\sqrt{N} \tag{1}$$

where d is expressed in terms of percent of the mean of N items having a simple coefficient of variation C. Use of (1) here is approximate on two grounds: the factor 200=2x100 contains the merely approximate value of 2 for the 0.05 probability level of Student's t, and the basic relation applies strictly only to samples drawn from normally distributed populations. Since in Table 3, N is seen to exceed 48 in every instance and since the 0.05 value of t is very nearly 2 for samples of 50 or over, the first approximation is acceptable here. Since a separate study of the effects of non-normality in the data of five of the stations tabulated in Table 3 has shown that these effects are not climatologically very disturbing, the second difficulty may be put aside and the tabulated d-values used as helpful, even if not exact, measures of unreliability of the means. In the sample case of Bowie winter precipitation, the tabulated d_W of 15 per cent implies that even with our 69-year record <u>we cannot fix the "true" winter mean to better than about 15 percent if we wish to be 95% confident that 69 years of sampling fluctuations are</u> not giving us an "atypical" Bowie winter mean. Since 15 percent of 4.98 inches is about 0.75 inches, we are, <u>at most</u>, able to say that we are 95% confident that the "true mean" lies somewhere between the limits 4.23 and 5.73 inches, this large interval of uncertainty being produced by the vagaries of "random sampling" from a population exhibiting (or rather yielding by sampling) the historically large relative variability of winter precipitation at Bowie.

The last statement, plus several others made above must now be rather strongly qualified. The nature of secular trends in all meteorological quantities is such that the notion of a "true winter mean precipitation amount" at Bowie, i.e., the notion of a parent population of winter precipitation amounts having distinct characteristics and some unique (even if unknown) distribution is an assumption that is at least debatable. Furthermore, the physical processes leading to the ultimate determination of a given winter's precipitation at Bowie may bear some, but only slight relation to the notion of random sampling from a parent population. Hence, it must be stressed that the above-stated definition of the tabulated d-values cannot be taken quite literally. The magnitudes of the d-values should be taken chiefly as a warning against making loose statements about rather small differences between means of various kinds in this region.

---

[1] The point here is, of course, entirely one of the distinction between inferential and descriptive statistics: <u>The</u> mean winter precipitation at Bowie during the past 69 years <u>was</u> 4.98 inches as far as actual readings go. The difficulty enters only <u>when</u> one infers what may be expected in the future.

In general, all of the d-values are about 10 percent, so if we want to aim at significant statements and confidence statements at the 95% probability level, we must be sure we are dealing with difference between means or departures from Arizona precipitation means that are in the neighborhood of 10 percent or more. This is useful, if inexact, knowledge.

It might be suggested that if one only knows Bowie's winter mean to be 4.98 ± 0.75 inches, he should drop the convention of listing the "mean" to three significant digits! However, as a purely descriptive statistic one can (just barely) defend computations of means to hundredths of an inch so the usual practice is followed here and in the following.

A single practical application of the knowledge of d-values is relevant. If cloud seeding were to be done in Arizona and were evaluated in terms of departures from mean precipitation by a rancher reading his own rainguage near, say Oracle, then even if that rancher were to be in the unlikely position of having a 50-year record, he could not be sure of his "true" mean annual precipitation value to much closer than 10 percent. Hence, he would be well advised to look with caution upon a rainmaker who sold services on a claim of being able to achieve positive departures from "normal" of, say, 10 or 15 per cent on that ranch. And since, many such rainmaking claims have been made in terms of "normals" based on much shorter than 50-year records, the corresponding confidence half-widths spread out to more than even 10 per cent. Wider lay appreciation of this kind of statistical implication of precipitation variability would have obviated many of the unprincipled uses of statistics in cloud seeding "evaluations" during the past decade.

A useful rule of thumb can be derived from the Arizona d-values of Table 3. Noting that roughly a 10% half-width has been found characteristic for record-lengths such that $\sqrt{N}$ is between about 7 and 9, it follows that the much more numerous stations in Arizona having precipitation record lengths of about 15 to 20 years can be expected to be characterized by 95% confidence half-widths of about 20 percent, individual values being somewhat higher for winter precipitation than for summer, and values being higher in the southwestern corner than in other parts of Arizona.

A final conclusion of rather general climatological interest may be drawn from a joint implication of Table 1a and Table 3. The skewness measure tabulated in the last column of Table 1a represents simply the difference between mean and median expressed in terms of percent of the mean. It will be seen that this difference is in the neighborhood of 10 percent for most of the seasonal values of Table 1a. Now, Table 3 reveals that most of the 95 percent confidence half-widths for Arizona seasonal means may also be expected to equal about 10 per cent of the respective means. That is, the small sample of Arizona data examined here suggests that the percentage uncertainty in the present values of historic means is of about the same magnitude as the percentage difference in historic means and medians. It is sometimes suggested that the use of the arithmetic mean should be abandoned in favor of the median in climatology, especially in the case of precipitation (see discussion in Foster, 1949, pp. 106-108, for example, or Landsberg, 1951), but that the advantages would be dubious in arid regions such as Arizona would seem to follow from the near-equality of the two quantities just discussed. Briefly, we can scarcely say that we know the "true" Arizona means or medians to such a degree of reliability that any decisive choice between these two measures of central tendency may be made at the present time.

C. Secular trends in seasonal precipitation. One of the many aspects of Southwestern U. S. precipitation which will be thoroughly examined in the long-term program of the Institute is that of the secular trends and secular oscillations in precipitation in this region. The economy of the Southwest has alternately suffered and flourished because of swings from prevailingly wet to prevailingly dry periods since the settlement of the region. At present, the Southwest is experiencing a drouth that is generally regarded as having begun in 1942 and that, according to Schulman (1956), is revealed by dendroclimatic evidence as one of the most severe drouths in the past 350 years, possibly even exceeding in severity all drouths in the Southwest since the Great Drouth of the late thirteenth century. It seems a wholly sound meteorological inference to conclude that a prolonged drouth such as the present one must result from anomalies of some sort in the general planetary circulation. It could scarcely be local (i.e. southwestern) in origin and still last so long and affect such wide geographic areas as the present drouth has. Elucidation of the causal factors involved in such drouths must be a major goal of any program of climatological and meteorological research in this region. This is indeed a major goal of the Institute program and constitutes one very important part of the overall variability investigation.

There is a dearth of published material in the literature of recent years to shed light on the mere nature of the recent precipitation oscillations of the Southwest, so in the present pilot study it was decided to examine, on an exploratory basis, secular variations of precipitation for the seven key stations. In view of the marked winter-to-summer differences in synoptic patterns associated with Arizona precipitation, the data were broken down into seasonal totals for every year, and the winter half-year and summer half-year analysed separately to search for evidence of similarities and differences. As before, "winter" is the period from 1 November through 30 April, and "summer" the balance of the year.

1. Choice of analysis method. Secular variations of precipitation may be studied in a variety of ways that yield correspondingly various insights. Of the several standard graphical approaches, that of plotting moving average precipitation amounts was chosen here in preference to the mass method or the residual-mass method, for, though an experienced person can discern the significant trends in any of these graphical presentations, the moving-average method has the strong advantage of showing directly the smoothed mean precipitation amounts that have been characteristic of any given period of record while the other two show this kind of information only through slope-changes. (There are, of course, other ways in which the mass type of plotting holds advantage over the moving average type, but these were of less interest here).

2. Choice of averaging period. In using a moving-average method, the choice of the period of averaging is in the end arbitrary, but it is to be chosen on the basis of how far one wishes to go in smoothing out year-to-year fluctuations in order to more readily perceive long-term movements. A 30-year averaging period will almost completely conceal fluctuations of periods of the order of 10 years but leaves only slightly distorted those oscillations having periods of the order of 30 years and longer. However, from, say, a 60-year record one can only get thirty-one 30-year moving average values; so in any event, one suffers such a seeming loss of information by averaging over anything like 30 years that shorter periods seem more attractive. Finally the suspicion that in some areas, including the Southwest, there may be meteorological oscillations related to sunspot variations, leads rather naturally to use of a period of around the 11-years. Computational advantage is then attained by dropping to 10 years as the averaging period, for in continuous records the averaging can then be done by simple shifting of the decimal point in moving sums. In all of the present work, as well as in current extensions of the secular studies reported here, as 10-year moving average is used.

In the moving-average graphs presented here it will be apparent that some individually very wet or dry years exert a sustained influence on all of the 10-year means that include them. Since the seasonal coefficients of variation of Table 1 are mostly of the order of 0.3 to 0.4, it follows that a single season in which the seasonal departure from the long-term mean has value equal to two standard deviations will produce by itself about a six to seven per cent departure in the decadal mean that contains it. From this fact, and from the relevant probability considerations, it follows that narrow (under 10 year) peaks and troughs with amplitudes much below ten percent of the mean cannot be safely inferred to reflect truly persistent anomalies in decadal moving mean precipitation for stations in this region. These, for the present purposes, are of the nature of "noise", so one may say, very roughly that the noise amplitude in the moving average curves is about ten per cent in this study.

3. Double-mass analysis. Individual station records such as those of the several Arizona stations studied here are actually composites, in almost all cases, of two, three, or more segments associated with slightly different instrument locations or at least associated with slightly different instrument exposures. In the case of precipitation observations, slow growth of trees near a rainguage in a fixed location can impose slow changes in catch. All of these non-meteorological factors serve to make the given long-term record temporally inhomogeneous in the climatological sense (Conrad and Pollak, 1950, pp. 222). In any investigations of secular trends, it is clearly dangerous to work with records that may be markedly inhomogeneous and it is especially dangerous to work with just one or two records from a given area if these are not thoroughly checked for homogeneity. In a single record, in particular, it is simply impossible to distinguish true secular trends and true climatic discontinuities on the one hand from trends in exposure or discontinuities in location and exposure on the other.

One of the most generally useful tests for homogeneity is the method of double-mass analysis, developed by Merriam (1937). A recent discussion of this method has been given by Kohler (1949). Rational use of this technique requires that one have initial knowledge of the spatially homogeneous subregions within a region whose precipitation records are to be tested for temporal homogeneity. In any such preliminary study as the present one, this initial knowledge is lacking, being one of the very goals of preliminary study. Consequently only a poor job can be done at best. For this reason, not only a poor job but also a rough job has been done in the way of double-mass analysis here. Briefly, the seven station-records whose homogeneity (internal consistency) was in question have been used to form the comparison series by adding, for each season of each calendar year the seasonal totals for that year, and the time series of these seven-station totals forms the abscissa function in the double-mass plots for each of the seven individual records. So crude a first approximation can reveal little more than really major breaks in any one station's record, but this seemed acceptable in this pilot study. A very much more thorough double-mass analysis of Arizona data has been begun quite recently, making use of information gained from the present study. For the present purposes reliance has to be placed on only the first-approximation analysis represented graphically in Figure 3.

Cases of missing data were handled by inserting the 10-year moving mean value corresponding to the missing year. This is sufficiently accurate for the present purposes but is being replaced by a technique of estimating missing amounts using nearby reported amounts in the analyses now being carried out by the Institute.

Because of dissimilar record-lengths, it was profitable to run double-mass analyses on two overlapping periods, 1898-1954 and 1877-1902, with all seven stations used in the former, but only five available for the latter. Some overlap was necessary to avoid failure to detect a major break in or near 1898, the earliest year for which all seven records existed. Separate analyses for the two seasons are shown, and the alignment of points corresponding to the beginning of each calendar decade is indicated by dashed vertical lines labeled with the year.

It is apparent, upon close examination of the individual sets of points depicting the relation of individual station cumulatives to the cumulative totals for all seven (or five) stations, that a large amount of small-scale waviness is present, particularly in the plots for the stations with heavy mean precipitation. This undesirable feature results from use of so small a total number of stations in the comparison series and from several of these being much drier stations than others and hence contributing only a small weight to the comparison series (abscissa function). In other words, the plots of Figure 3 are documentary evidence of the strong advisability of using a fairly large number of records (a dozen is the suggestion of Kohler) and using records for stations of nearly the same mean precipitation to equalize weighting as much as possible, though the latter becomes less important as the total number of stations within a spatially homogeneous region increases in the comparison series.

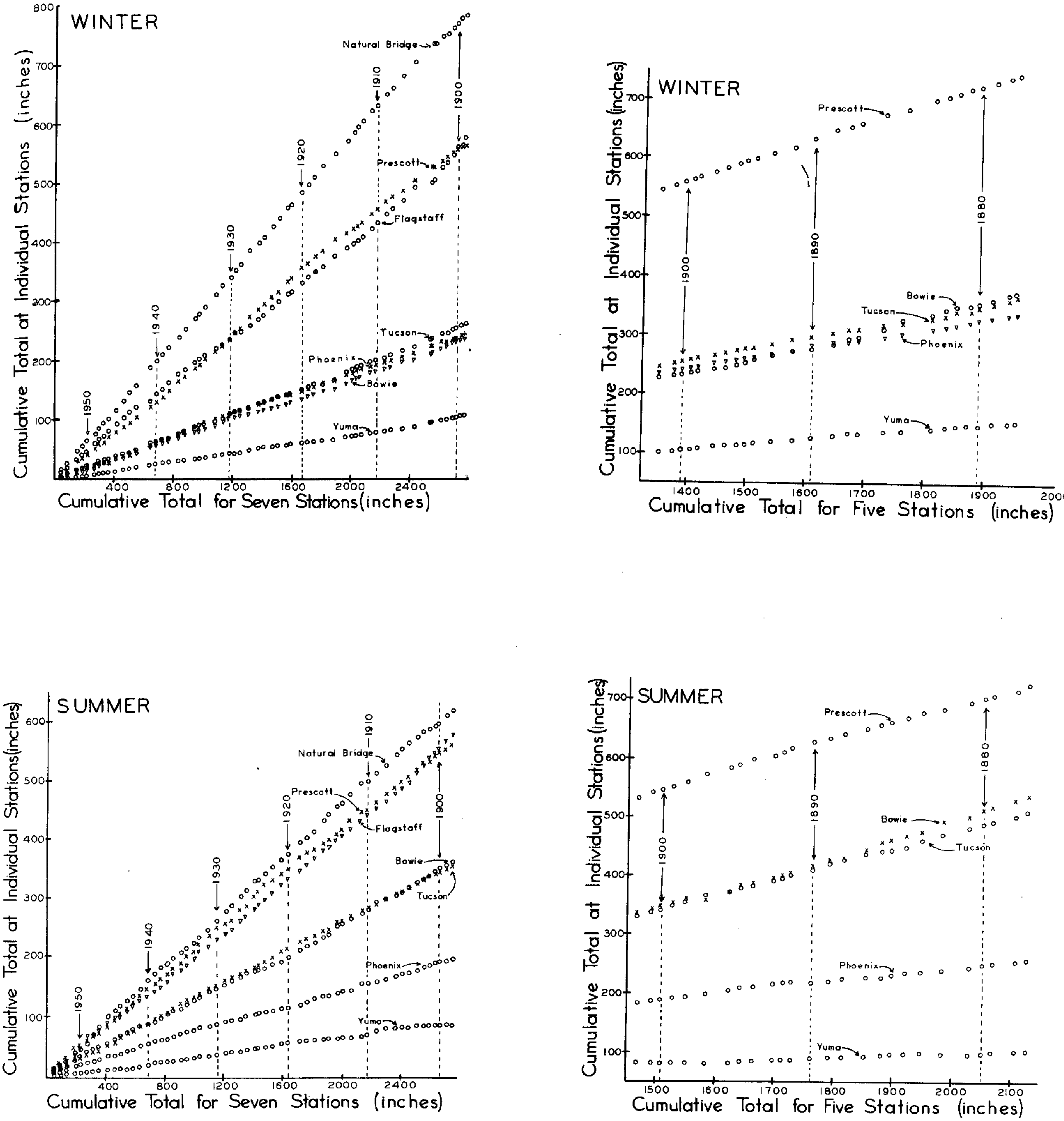


Figure 3. Double-mass plots for selected long-record Arizona seasonal precipitation records. Flagstaff and Natural Bridge records extend back only to 1898, so there is one pair of plots running back to that date and a second pair for cumulatives of just the remaining five stations from 1902 back to 1877. In general, total numbers of stations used here are too low to provide a stable base-function.

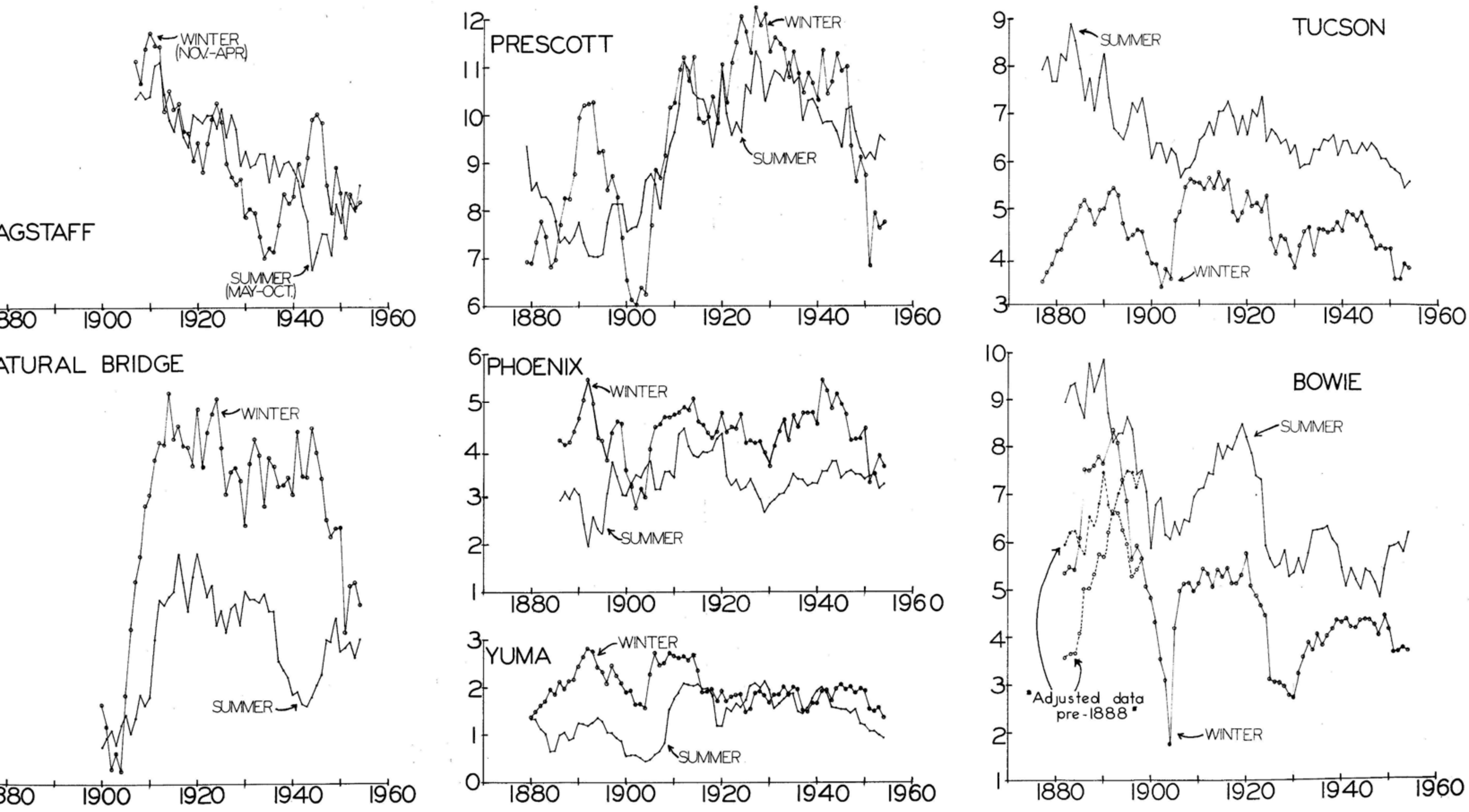
Because of these difficulties, which proved even more serious than had been anticipated, only one major record-break was clearly established and adjusted for -- a break in Bowie's data in 1888. Other breaks (e.g., Bowie near 1925, Natural Bridge near 1905, Yuma near 1909, etc.) seemingly indicated by the plots did not correspond to known station-changes and hence were not corrected. The Bowie 1888 break, however comes at a date when the raingauge was transferred from old Fort Bowie on the foothills of the Chiricahua Mountains to the town of Bowie, 14 miles north-northwest, out in the valley north of the mountains. The double-mass plots for bowie showed about 33 per cent higher precipitation rates prior to 1888 than after, which was compatible with this known station relocation, so the data before 1888 were reduced by the factor obtained from the slope-ratio, namely by a factor of 0.66.

It is highly likely that other adjustments are needed, but on the general principle that undercorrection is advisable when in doubt, all other corrections must await completion of an extensive double-mass analysis of Arizona seasonal precipitation totals that is now in progress using some 70 stations in five subregions of the state. In the meantime the plots of Figure 3 may at least be used to detect points of suspicion in the secular trend curves described below.

The raw data for these stations contained occasional breaks in the continuity of the record. When one or more (up to as many, but no more than five) amounts were missing in a given decade, the total for all amounts reported was divided by the number of reported amounts and the quotient used as the "decadal moving mean". This procedure is less accurate than estimating missing years from nearby stations, and the latter approach is being used in the follow-up work currently in progress.

4. Decadal moving average seasonal precipitation. Using the previously computed time series of seasonal totals for the seven Arizona stations, 10-year running means were computed. These are plotted in Figure 4 using the convention that the 10-year mean for the decade ending in, say, 1906, is plotted above the 1906 point of the abscissa time-scale. Note that, in general, the origin of the precipitation scale on the ordinate is not zero inches.

For Bowie, both the raw data and the adjusted data for the period before station relocation in 1888 are plotted to illustrate graphically the seriousness of failing to carry through homogeniety checks in precipitation data used to discern secular trends. In the recent literature one might cite several papers wherein moving-average precipitation curves were plotted for isolated stations (hence precluding even visual comparison by the reader) for which no consistency checks were made, and one must be suspicious that effects as serious as those for Bowie may be involved in any inferences drawn from such plots. Indeed, one must view each one of the of the other six sets of graphs in Figure 4 with that same suspicion inasmuch as their double-mass plots of Figure 3 do show breaks that were left uncorrected. However, the entire group of seven plots of Figure 4 afford an opportunity to sort out broad similarities from small differences that may represent still-uncorrected heterogeneities in the individual records.



IO-YEAR MOVING AVERAGE PRECIPITATION (INCHES)
(Decadal means plotted for years ending IO-year periods)

The salient features of the fourteen decadal moving average plots of Figure 4 will next be discussed briefly in order to call attention to points that demand further investigation that was attempted in this pilot study.

a. The recent drouth. Definition of the term, "drouth", has plagued climatologists and meteorologists for a long time, and no new effort will be made to define it here. By the term, "recent drouth" is to be understood here the period of generally deficient precipitation beginning early in the 1940's (say, 1942 to exclude the wet winter of 1941) and extending either to the present data or to the early fifties, depending on which part of Arizona is considered and which season is considered.

Comparing the winter and the summer curves of Figure 4 reveals that for most of the state the recent drouth has been primarily a winter drouth. It is almost entirely the winter cyclonic precipitation that supplies runoff for the irrigation districts of central Arizona and it has been the marked decline in winter precipitation of the past ten to fifteen years that has become cause for so much concern for water resources in this state. The statewide decline in winter precipitation in recent years is even more emphatically brought out in Table 4, which shows for both winter and summer the 10-year means for the decades ending 1942 and 1954 and also gives the 1954 means expressed in terms of percent of the 1942 means for each season and station. Except for Bowie and Flagstaff (both of whose decadal moving mean winter precipitation curves of Figure 4 will be seen to have attained their most recent winter maxima a bit later than 1942) all stations exhibit only about 70 to 80 percent as much winter precipitation for the decade 1945-54 as for the decade 1933-42. (Note that the picture with respect to summer trends is more confusing -- no simple geographic pattern is evident from this sample of only seven stations, so more data will be needed to pin down this point).

b. The general decline since the 1920's and concurrent runoff trends. In a recent study of rainfall-runoff relations for the Salt River Basin, of central Arizona, Anderson (1956, p. 53) has reached the conclusion that winter-season precipitation accounts for about 85 percent of total annual stream flow in the basin even though the winter precipitation comprises only about 60 percent of the annual estimated total for the basin. Anderson states (ibid.,p. 57), "The results of the study made support the conclusion that the basic cause for the prevailing below average stream flow lies in the generally deficient precipitation within the basin since the year 1932." His analysis indicated (ibid.,p. 53) that of the total water precipitated on the Salt watershed in summer (comprising about 40 percent of the annual total), some 95 percent is "lost" in evapotranspiration (ignoring storage in the long-term mean), and that of the total winter precipitation (comprising about 60 percent of the annual total), some 85 percent is so "lost". For the year as a whole the loss above the stream-gauging points averages about 90 percent.

In Figure 5 are plotted 10-year moving average runoff values for the Salt and Verde basins (data from unpublished records of the Salt River Valley Water Users Assn.). Comparing this curve with the 10-year moving averages plotted for Natural Bridge and Prescott (stations in this basin) in Figure 4 reveals a degree of similarity so great that it inevitably

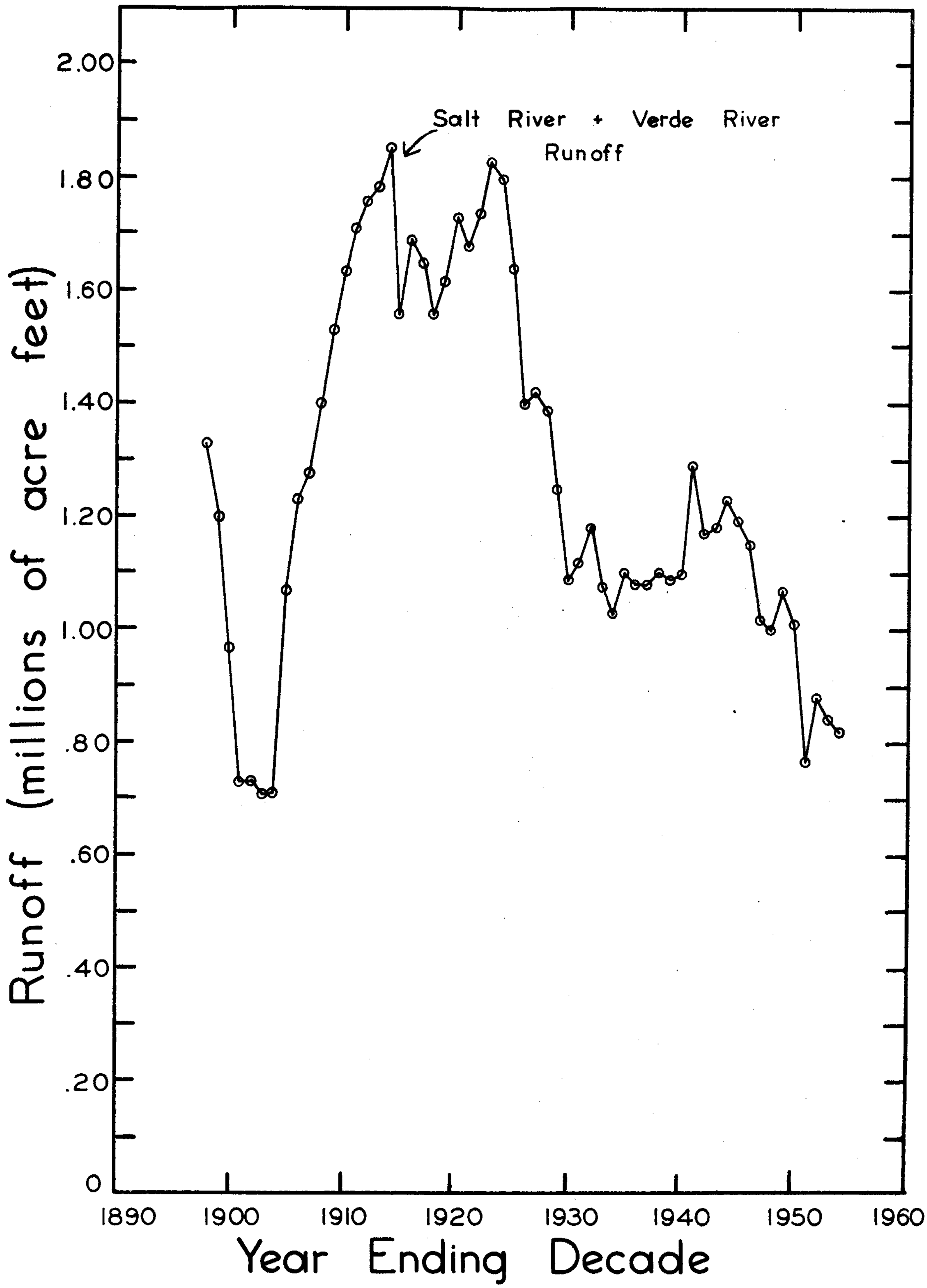


Figure 5. Secular trends in runoff in the Salt River basin as revealed by 10-year moving averages. Runoff values are water-year totals for the combined flow of the Salt near Roosevelt, the Tonto near Gun Creek, and the Verde below Bartlett Dam. Decadal means are plotted above the year ending the decade.

Table 4

Comparison of 10-year average seasonal precipitation at seven Arizona stations for the two decades ending 1942 and 1954.

Station	Winter: Decade ending 1942 (in.)	Winter: Decade ending 1954 (in.)	Winter: Ratio 54/42 (%)	Summer: Decade ending 1942 (in.)	Summer: Decade ending 1954 (in.)	Summer: Ratio 54/42 (%)
Bowie	4.23	3.71	88	5.46	6.21	113
Flagstaff	9.53	9.18	96	9.11	9.55	105
Natural Bridge	14.43	11.74	81	9.68	11.01	114
Phoenix	5.23	3.63	69	3.53	3.26	92
Prescott	10.42	7.77	74	9.86	9.48	96
Tucson	4.89	3.78	77	6.16	5.57	90
Yuma	1.94	1.36	70	1.87	0.94	50

forces one to the conclusion (as reached in Anderson's study and in similar studies in other basins in the West) that water yield is controlled almost entirely by meteorological factors of which precipitation is the obvious prime factor. But that it is prime only in a certain chronological sense and not in a numerical sense, follows from existing general knowledge as to the typical magnitudes of the "loss" factors for basins in the West. It is, indeed, to be regarded as quite remarkable that runoff curves such as that in Figure 5 parallel so very closely the concurrent basin-station precipitation curves such as those of Natural Bridge and Prescott, since most (e.g., some 90 percent for the Salt basin) of the measured precipitation reevaporates before the precipitated water can reach a gauging point. Thus, the numerically dominant factor in water yield is the loss factor and this is, like the initial precipitation, governed to large extent by meteorological factors such as temperatures, wind speeds, humidities, and so on, that are themselves only indirectly related to precipitation. Hence one cannot but marvel at the fact (typical of most basins) that runoff, the small residual difference between precipitation and evapotranspiration, correlates so well with precipitation. A complete understanding of the meteorological and ecological factors responsible for this typical yet always amazing degree of correlation, which often involves correlation coefficients of the order of 0.90-0.95, must be a goal of all workers concerned with the water resources of the semi-arid regions.

Here it is only to be noted that the declining trend in runoff in the Salt basin has gone hand in hand with general downward trends in Arizona precipitation (especially in the Salt watershed as shown by the Flagstaff, Prescott, and Natural Bridge curves of Figure 4) since the mid-twenties. General circulation changes have in some way been responsible for a major maximum in Arizona winter precipitation in or near the 1916-1925 decade and for a fluctuatory decline since then. Further study is clearly needed.

c. The drouth of the early 1900's. No single decade in the recorded history of Arizona rainfall exhibits wider oscillations of precipitation than the decade centering on about 1904. In every curve of Figure 4 for which data for that period was available one notes an astonishing decline in the late nineties to decadal means lower, in general, than any since experienced, and then an equally rapid recovery to the level of the prevailingly wet period in or near the early 1920's. Other studies now underway show clearly that this pattern is found throughout much of the Southwest -- i. e., it must have resulted from a large-scale disturbance in the general circulation over the Pacific and North America. Perhaps no more emphatic argument for the pressing need for research on the relation of precipitation variability to large-scale meteorological phenomena could be given than to point out that, at present, there is no reliable scientific basis for predicting such anomalous developments, nor even for explaining them *post mortem*.

Some speculative comments on the nature of the turn-of-the-century oscillation will be given below in a general and very preliminary discussion of large-scale controls of Arizona precipitation. Here the oscillation is only noted as an outstanding example of the amplitude through which Nature can swing in short time periods. It is a matter of historical record that the drouth ending with the outstandingly wet year of 1905 in Arizona was one that had severe effects on Arizona's economy, just as the recent drouth threatens to injure the present Arizona economy. It is almost superfluous to point out that intelligent planning to cushion the shock of such oscillations can never be realized until the physical causes of these great swings are clearly identified.

d. Differences in secular trends within Arizona. In several of the preceding comments on Figure 4, similarities of behavior in the moving average curves have been emphasized, but differences have also had to be cited. Turning briefly to these differences it must be stressed that they demand very careful investigation. Why, for example do Bowie, Tucson, and Phoenix exhibit a summer-drouth condition near 1930 while Prescott, Natural Bridge and Flagstaff do not? And why did Natural Bridge and Flagstaff experience summer-drouth conditions in the early 1940's while Tucson and Phoenix experienced no minima at that time? Are these actually due to uncorrected effects of station changes or instrument exposure changes? No answers can be given until the extensive homogeneity check now in progress is completed.

However, it is pertinent to call attention to the reasons for wishing to settle such questions decisively. First, if there are true meteorological trends persisting in given portions of even so small an area as Arizona and differing significantly from concurrent trends in immediately adjacent areas, then this is a climatological feature of fundamental interest that demands physical explanation. Secondly, if such conflicting secular trends can actually occur in regions only a hundred or so miles apart, then this fact and its cause must be ascertained as part of the Institute's fundamental research on the broad problem of cloud modification. The latter point follows from the fact that one of the standard techniques for assessing the efficacy of cloud seeding consists in comparing precipitation amounts received in target areas and in nearby unseeded "control" areas. A regression analysis of historic relationships between precipitation in these target and "control" areas forms the basis for an indispensable statistical estimate of the amount of rain that might have been expected in the seeded area during the seeding period had no seeding been done. Now, if within a decade or so, secular precipitation trends in regions as close together as those represented by the stations used in Figure 4 can diverge as much as seems to be indicated by the curves of Figure 4, then there is indeed serious danger of either over or under-estimating the efficacy of cloud seeding by use of the standard type of target-control regression analysis. Thus, if a seeder had been seeding in southern Arizona near 1930 and had used the precipitation observed in the mountainous area of the Mogollom Rim as his "control", Figure 4 shows that the results would have seemed highly favorable, whereas in fact no cloud seeding was going on at all during that early period.

5. Contribution of secular variation to total variance. A number of questions that are implicit in the discussions of station coefficients of variation and of secular trends should now be asked. Joint presence of high variability (as measured by the coefficient of variation) and large-amplitude secular trends in the Arizona precipitation data have been found to be characteristic. One must ask, then, whether the presence of secular trends may not make a substantial contribution to the total variability (total variance, say). If this is true, then the coefficients of variation computed from the historic records do not actually reflect randomicity; and inferences made from them, such as deductions of confidence half-widths for the means, will be misleading. It is necessary, in other words, to break the total variance into components which in some way reveal the relative contribution of secular trends and of year-to-year

randomicity to the total historic variance.

To accomplish this, the following breakdown can be used. Let the seasonal (either w or s) precipitation for the ith year be p_i, the mean over the full historic record be $\bar{p}$, the decadal mean centered on the ith year be $\bar{p}_{di}$, the departure of the ith decadal mean from the historic mean be

$$\Delta p_i = \bar{p}_{di} - \bar{p},$$

and the departure of the ith seasonal total from the ith decadal mean be

$$\delta p_i = p_i - \bar{p}_{di}$$

Then, by definition,

$$p_i = \bar{p} + \Delta p_i + \delta p_i, \; p_i = \bar{p} + \Delta p_i + \delta p_i,$$

and, consequently, the total sum of squared deviations of all of the individual seasonal totals from the long term seasonal mean is

$$\sum(p_i - \bar{p})^2 = \sum(\Delta p_i + \delta p_i)^2$$

$$= \underset{(A)}{\sum(\Delta p_i)^2} + 2\underset{(B)}{\sum(\Delta p_i \cdot \delta p_i)} + \underset{(C)}{\sum(\delta p_i)^2} \quad (2)$$

The letters, A, B, C, entered under the three terms of the last member of (2) will be useful in identifying the components of total sum of squares (total variance). Term A measures the contribution to total variance resulting from departure of decadal trend from historic mean, term C measures the contribution resulting from departure of seasonal totals for the individual years from the concurrent decadal trend, and term B reflects the covariance of the factors whose contributions are identified with A and B.

If term C is not large compared to A, then the usual interpretation placed upon the coefficient of variation is inappropriate, but if C is substantially larger than A, it will be safe to ignore effects of trend on station standard deviations and coefficients of variation.

To test this important point, the records for winter and summer at two stations in two areas of markedly different topography were broken down into the sums identified in (2) as A, B, and C. Tucson was used as a test station for low elevations in southern Arizona, and Natural Bridge was used as a test station for the mountainous areas south of the Mogollon Rim in central Arizona. Table 5 presents the results of the computations.

Table 5

Analysis of total sum of squares into components due to secular oscillations (based on decadal moving means) and to year-to-year randomicity. For meaning of symbols A, B, C, see Equation (2).

	No. yrs. used	A	B	C	$\frac{A}{A+B+C}$
Tucson					
Winter ppn.	77	30.0	-0.3	450.9	0.06
Summer ppn.	77	43.0	-35.2	510.9	0.08
Natural Bridge					
Winter ppn.	54	224.5	134.0	1564.0	0.12
Summer ppn.	54	72.7	1.2	662.8	0.10

Inspection of the last column of Table 5 shows that for both Tucson and Natural Bridge and for both winter and summer, secular trend and secular variations contribute to total variance of precipitation amounts only to a minor degree. Nor is any appreciable seasonal difference suggested by Table 5. That Natural Bridge has so much larger a secular contribution to variance than Tucson is hard to understand. The difference may not be significant; but fortunately the writer does not know how to make a significance test of this difference of ratios of sums of squares involving moving averages. In any event, it is chiefly to be emphasized that even for Natural Bridge, the contribution of secular variations to total variance is small enough (about one-eighth) that it seems likely that no seriously misleading inferences will be made in treating Arizona precipitation coefficients of variation as measures of "random" variation about historic means free from appreciable influence due to long-term trends and oscillations. This is a result that will prove useful in a number of future studies.

D. Autocorrelations of seasonal and annual precipitation. Up to this point, three topics under the general heading of temporal variability have been discussed -- coefficients of variation, confidence intervals for means, and secular trends. The next topic to be considered haere is that of the extent to which the variations of precipitation from season to season and year to year exhibit persistence of anomalies; i.e., we next consider the extent to which the Arizona precipitation time series exhibits autocorrelation.

Put qualitatively, the questions considered here are these: Is there any appreciable tendency, in Arizona, for wetter than average winters of one year to be followed by wetter than average (or, conceivably, by drier than average winters) of the next year? Are there any such interyearly persistence or opposition tendencies for annual total amounts? And finally, is there a significant tendency for persistence or opposition in anomalies between one winter season and the immediately succeeding summer season of the same year?

It is generally acepted by climatologists that most areas of the world exhibit very little autocorrelation for time-lags of the order of six months or a year, so the expectation was that this would also hold true for Arizona. However, there appears to be no published quantitative check of this point for any area in the Southwest, so it was felt that a small number of autocorrelation analyses should be run to provide quantitative answers to the questions put above.[1]

One additional reason for performing some trial autocorrelations is found in the fact that, in any given region there are often some persons who have developed strong impressions to the effect that there are local tendencies toward persistence (or opposition) in seasonal or yearly rainfall, so to be able to offer definite evidence confirming or disproving such notions is to be able to settle questions that are sometimes of real lay concern. For instance, within recent months, there has been a petition circulated in the Arizona area urging discontinuation of atomic bomb tests on the ground of supposed interference with normal rainfall patterns in this region. The chief reason for the supposition of interference seemed to be based upon the suggestion that in certain recent years wherein the first portion of the year (winter period) seemed to show promise of a good rain year, the ensuing portion of the year, following bomb tests, exhibited a disappointing deficiency of rainfall. If it is, in fact, possible to thus anticipate summer rainfall on the basis of preceding winter rainfall, then this is first of all a point of great theoretical interest, and secondly the supposition of bomb effects on precipitation might then deserve further scrutiny, notwithstanding its meteorological improbability. On the other hand, if only negligible interseasonal autocorrelation is found by statistical analysis of historic rainfall data, then this point needs to be clarified in the minds of persons who may have developed misconceptions thereon.

To provide quantitative answers to the questions posed above, four different types of autocorrelation analyses were carried out for the seven key stations of earlier discussions, and for Oracle. The results of the analyses are presented in Table 6.

The first type of autocorrelation coefficient is termed, for brevity, "Winter", and represents the autocorrelation of successive pairs of winter precipitation totals made up of the total for the ith year and the total for the (i + 1)th year, where i ranges from 1 to N-1, and N is the record length of the station when there are no breaks in the record (but some smaller number if breaks exist). Thus, the "Winter" autocorrelation coefficients are measures of the extent to which winter precipitation, considered alone, tends to exhibit persistent anomalies (or opposing anomalies, should the autocorrelation coefficient be negative) from one winter to the next.

The second type of autocorrelation is termed, "Summer", and is defined in the same way as the "Winter" coefficient, except that the "Summer" autocorrelation coefficient measures the persistence (or opposition) from one summer to the next summer. Similarly, the "Yearly" coefficient measures autocorrelation form one year to the next for annual totals, thus lumping together both summer and winter precipitation characteristics. Note that for all three of these types, it is a one-year-lag autocorrelation that is involved.

[1] That product-moment autocorrelation techniques were used here, implies that only linear persistence tendencies could be discerned. Other persistence tests might well be tried in the future.

Table 6

Autocorrelation coefficients for Arizona seasonal and annual precipitation. See text for explanation of terms and discussion of standard errors. N_y is the number of years of record used in the "Yearly" case and is in general slightly smaller than the number used for the "Winter" and "Summer" correlations but is equal to the number used in the "Interseasonal" case. Standard errors are in parentheses following the autocorrelation coefficients to which they apply.

Station	N_y (Yrs.)	Winter	Summer	Yearly	Interseasonal
Bowie	59	0.29(.12)	0.01(.13)	0.31(.12)	0.29(.11)
Flagstaff	46	-0.16(.14)	-0.21(.14)	-0.08(.15)	0.01(.14)
Natural Bridge	61	0.11(.12)	0.12(.12)	0.33(.11)	0.12(.12)
Oracle	45	0.05(.15)	-0.16(.14)	0.07(.15)	0.12(.14)
Phoenix	73	-0.11(.12)	0.09(.12)	-0.17(.11)	-0.01(.12)
Prescott	73	0.03(.12)	0.05(.12)	-0.14(.14)	-0.01(.11)
Tucson	86	0.06(.11)	-0.21(.10)	-0.04(.11)	-0.15(.11)
Yuma	81	0.03(.11)	0.21(.11)	0.13(.11)	0.10(.11)

The fourth type is termed, "Interseasonal", and measures the degree to which the winter precipitation of the ~~ø~~ ith "rain year" (i.e., the period from 1 November of the (i-1)th calendar year to 30 April of the ith calendar year) is correlated with the summer precipitation of that same ith rain year. That is, the magnitudes of this fourth autocorrelation coefficient will reveal whether it really is possible to predict, to any reliable degree, the summer rainfall in Arizona on the basis of knowledge of the preceding winter season's precipitation. The above-mentioned concern over effects of atomic bomb tests was based on the premise that such predictions can reliably be made. This fourth case, then, represents an autocorrelation with only six-month lag.

In Table 6 are given, after each autocorrelation coefficient, the "standard error of the coefficient" as found from the (normal-distribution) formula

$$\sigma_r = \frac{1-r^2}{\sqrt{N-1}} .$$

Although non-normality and existence of serial correlation do preclude strict interpretation of σ_r in the usual fashion, it is believed that these difficulties are quite unimportant here -- partly because non-normality has been found (in a separate study soon to be reported) to have negligible effects on interstation correlation coefficients, and partly because the very magnitudes of the autocorrelation coefficients of Table 6 remove suspicion of serious consequences of serial correlation in the time series concerned. The fact that all of the r's are so small removes the boundedness difficulty so it seems unnecessary to test significance by use of the Fisher z' transformation here. Even for the largest r of Table 6, there was found to be only a slight difference between the implications of the z'-test applied (at the 5% level) to the null hypothesis that the "true" autocorrelation coefficient is zero and the implications of the rule of thumb (cited below) that r must exceed $3\sigma_r$ for significance of r.

Inspection of Table 6 shows that, as most climatologists would have anticipated, there is almost no persistence from year to year or even from one season to the next in Arizona precipitation amounts. Certainly there is no persistence strong enough to be of any practical use in forecasting seasonal precipitation amounts, and certainly popular notions to the effect that in Arizona "a wet winter means a wet summer" (or vice versa) are completely contradicted by the autocorrelation coefficients found in this study. For example, on the usual argument that the fraction of variance predictable on the basis of correlation is equal to r^2, one sees that even in the case of the highest of all thirty-two autocorrelation coefficients in Table 6, that for the Natural Bridge yearly totals, only a mere 10 per cent of variance is related to autocorrelation in the time series. And for over half of the cases of Table 6, only about one per cent of the variance is predictable on the basis of autocorrelation!

As a graphic means of emphasizing this principal implication of Table 6, the thirty-two autocorrelation coefficients of Table 6 have been plotted against their respective standard errors in Figure 6. Lines representing loci of the relations $r = \pm \sigma_r$ and $r = \pm 3\sigma_r$ have been plotted in Figure 6 to

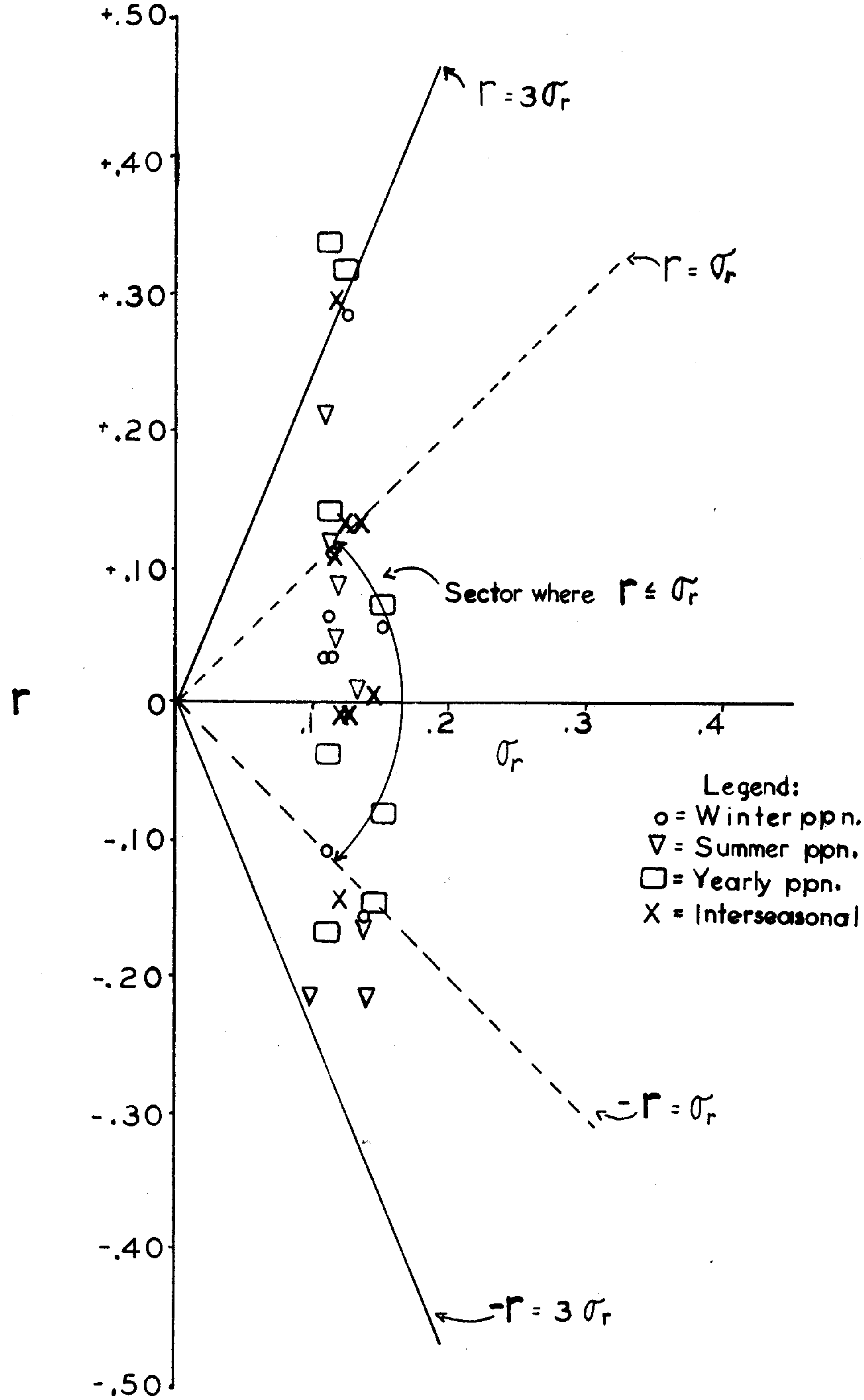


Figure 6. Graphical representation of the insignificance of Arizona precipitation autocorrelations. The solid lines bound the sector within which autocorrelation coefficients (ordinate functions) are of smaller than three times the absolute magnitude of their respective standard errors (abscissa). Dashed lines, similarly, bound the sector within which the correlation coefficients' absolute magnitudes do not even exceed their standard errors.

bring out the insignificance of the Arizona autocorrelations. In general, correlations coefficients must be regarded as of no significance unless they are about four or five times greater than their standard errors; and meteorologists, even though they relax the stringency of this criterion, commonly set the limit at no less than three times the standard error. (This rough rule of thumb becomes inapplicable for large correlation coefficients where σ_r itself is to be put aside and Fisher's z' test used, but it will suffice to reveal the lack of significance in the very low coefficients obtained here.) Hence, all of those points in Figure 6 falling inside the sector bounded by the lines $r=3\sigma_r$ and $r=-3\sigma_r$ are well below the level of acceptable significance. Those falling within even the sector bounded by the $r=\pm\sigma_r$ lines have standard errors larger than the autocorrelation coefficients themselves. Seventeen of the thirty-two values fall in this sector of utter insignificance!

Only three of the thirty-two autocorrelation coefficients fall outside the $r=\pm3\sigma_r$ sector of Figure 6. Two are "Yearly" values and one is an "Interseasonal" value, but all three fall so close to the $r=3\sigma_r$ line that no great emphasis can justifiably be placed upon these three coefficients exceeding three times their standard errors.

One point was examined a bit further: Bowie's "Yearly" and "Interseasonal" coefficients comprise two of the three values falling just barely above the threshold of significance, and it can also be seen in Table 6 that this same station's "Winter" coefficient is noticeably higher than all other "Winter" coefficients. Thus, this one station in southeastern Arizona seems to exhibit an interesting tendency toward persistence that was felt to warrant further consideration here.

The principal question to be answered is that of whether Bowie's record is or is not typical of its surrounding area. As a crude test of this, tetrachoric autocorrelation coefficients were computed for several stations in this area as well as for five of the key stations whose product-moment autocorrelation coefficients were given in Table 6. These latter five tetrachoric correlation analyses were computed to test (empirically) the concordance between correlations computed in these two very different ways (the tetrachoric being much the cruder and harder to interpret quantitatively, but being very much the easier to compute, of course). The results for these "Yearly" tetrachoric autocorrelation values are presented in Table 7. (See Figure 1 for locations of the stations not previously identified.)

The station in Table 7 closest geographically to Bowie is Willcox and the second closest is Benson. These yield conflicting implications for the representativeness of Bowie's product-moment autocorrelation coefficient. Bisbee and Douglas are next in order of proximity, and this pair also yields conflicting evidence. There is thus raised at this point a question concerning the reliability of the tetrachoric coefficient, especially when used with somewhat skewed data such as we are here dealing with. It was to obtain some crude indication of this reliability that the five stations already analysed by product-moment methods were then subjected to tetrachoric auto-correlation. It will be seen from Table 7 that acceptable agreement exists for the case of Bowie itself, and for Natural Bridge, Tucson, and marginally for Yuma, while the results for Phoenix are quite different when the two methods of correlation are used. In the writer's opinion, Table 7

Table 7

Tetrachoric autocorrelation coefficients (one-year lag) for annual precipitation totals at nine Arizona stations. Where product-moment autocorrelations have been previously computed, these coefficients are shown in parentheses following the tetrachoric value.

Station	N (Yrs.)	Tetrachoric Autocorrelation Coefficient
Benson	49	0.10
Bisbee	40	0.09
Bowie	59	0.25 (0.31)
Douglas	53	0.49
Natural Bridge	61	0.28 (0.33)
Phoenix	73	0.37 (-0.17)
Tucson	86	0.03 (-0.04)
Willcox	30	0.58
Yuma	81	0.02 (0.13)

simply does not settle the question of the representativeness of the Bowie product-moment autocorrelation coefficients. Rather, it yields a confusing mixture of agreement and disagreement that requires further study at some future time. It must be emphasized that this further study is desirable, not because there is appreciable indication that any prediction-value may exist in autocorrelation coefficients for stations in the Bowie area, but rather because there is a climatological point of academic interest if this corner of Arizona exhibits even a systematic tendency towards barely significant coefficients. (The Bowie, as well as the Natural Bridge "Yearly" coefficients, it might be pointed out differ from zero by an amount that was found to be significant at the 5% level using the Fisher z' transformation test, a test somewhat more appropriate than use of the "standard errors" employed above.)

In sumary, it may be said that there appears to be need for a bit more examination of several barely significant autocorrelation coefficients found in the present exploratory study. No further analysis seems necessary to settle the practical question of whether "a wet summer follows a wet winter" -- it definitely does not, in terms of any commonsense viewpoint. Statistical analysis has shown that, for a representative sample of Arizona stations, there exists no usable degree of year-to-year or season-to-season persistence of precipitation excesses and deficits.[1]

E. Persistence tendencies for monthly precipitation amounts. One may ask this question: Even though we have just found that interyearly and interseasonal precipitation autocorrelations are generally very low, might it not still be possible that, significant persistence of precipitation anomalies between successive months within a season exists? This question was examined through the use of a very simple techniques of counting numbers of occurrences of pairs of successive months exhibiting the same or opposite algebraic signs of departures of individual monthly amounts from the long-term mean monthly precipitation amounts for each station. For brevity, this scheme will be called an analysis of opposition frequencies.

A total of 32 Arizona stations was examined, using monthly precipitation amounts for just the 22-year period 1931-52, since data for this period are very conveniently tabulated in the Weather Bureau's Supplement to Bulletin W. The opposition frequencies for four pairs of successive months were determined. Three pairs were from the winter half-year (Jan.-Feb., Feb.-Mar., Mar.-Apr.) and the fourth pair was taken from the summer rainy season (Jul.-Aug.). In the absence of any tendency toward persistence or opposition (i.e., in the limit of zero autocorrelation), the a priori probability that, say, both the January, 1913 and the February, 1913 monthly

---

[1] This finding does not contradict the findings of the secular-trend study discussed earlier in this report. Ten-year running means smoooth out the year-to-year variability that lowers autocorrelation, as shown in Part 5 of Section C above. There is definite persistence in the Arizona decadal moving averages, as is qualitatively apparent in Figure 4.

precipitation amounts should exhibit departures of the same algebraic sign from, respectively, the January and the February 22-year means for 1931-52, is one-half; similarly the a priori probability that they should exhibit opposite signs is one-half.[1] Hence, out of all twenty-two cases of January-February pairs for a given station during the 1931-52 period, one expects eleven cases of opposition of sign and eleven cases of parity of sign if no autocorrelation between successive months exists.

All 4x32=128 values of the opposition frequencies are presented in Table 8. Averages over 32 stations for each of the four month-pairs are entered at the bottom of the table, and below these are their respective deviations from the a priori (zero-autocorrelation) expectation values of 11.0 oppositions in twenty-two years per station.

Since the 128 individual values shown in the body of the table exhibit a noticeably large variability, it is clear that the deviations from 11.0 oppositions in the 32-station average values at the bottom may be non-significant. To examine this important question, the standard deviation of the 32 opposition frequencies for each pair of months was determined. (These, it should be noted, were computed as root mean square deviations not from 11.0 but from the observed mean values found in each sample of 32 month-pairs.) The results are as listed in Table 9.

Table 9

Standard deviations and standard errors of opposition frequencies shown in Table 8. (See text for discussion).

Month-pair	Jan.-Feb.	Feb.-Mar.	Mar.-Apr.	Jul.-Aug.
Std. Devn.	1.8	2.6	1.9	2.1
Std. error of means(N=32)	0.3	0.5	0.3	0.4

If we use these standard deviations of Table 9 to estimate the significance of the deviations of the column-means from 11.0 we may use as an estimate of the standard error of the 32-station means the values of sample standard deviations divided by the square root of one less than the sample size, i.e., $\sqrt{32-1} = 5.6$. In the last row of Table 9 are listed these estimated standard errors, rounded off to tenths in keeping with their very approximate nature. Very roughly, one says that the 32-station means listed at the bottom of Table 8 must depart from 11.0 by twice the magnitude of these standard errors in order that we may regard those departures

[1]More precisely, these a priori probabilities apply to departures taken with respect to the medians, but despite Arizona precipitation skewness, it is believed that median and mean differ by little enough to be overlooked here.

Table 8

Opposition frequencies for 32 Arizona precipitation stations and for four pairs of months, for the 22-year period 1931-52. In the absence of autocorrelation, each entry would have an expectation value of (approximately) eleven. See text for explanation.

Station	Jan.-Feb.	Feb.-Mar.	Mar.-Apr.	Jul.-Aug.
Ajo	14	8	10	12
Apache Powder Co.	13	14	10	11
Ashfork	12	8	12	14
Bisbee	7	12	5	13
Canelo Ranger Sta.	10	9	10	10
Childs	12	5	11	12
Clifton	9	13	7	13
Cordes	13	7	11	11
Eagle Creek	13	9	9	13
Florence	12	9	8	8
Globe	11	7	12	13
Grand Canyon HQ	13	8	10	13
Holbrook	15	10	10	12
Jerome	10	9	11	10
Leslie Canyon	8	13	7	14
Litchfield Park	14	10	7	15
Miami	11	6	8	15
Mesa Exp. Farm	12	10	9	9
Mormon Flat	11	8	9	10
Natural Bridge	13	7	13	12
Patagonia	14	9	7	10
Payson Ranger Sta.	11	6	12	14
Roosevelt	13	7	10	12
Rucker Canyon	10	14	10	12
Salome 6 SE	11	5	8	13
San Carlos Res.	15	12	10	12
San Rafael Ranch	11	10	7	11
Superior	15	6	10	15
Tempe	13	7	9	14
Tucson U of A	14	10	11	15
Walnut Grove	12	5	11	9
Williams	14	8	10	10
Average for 32 stations	12.1	8.8	9.5	11.3
Deviations from 11.0	1.1	-2.2	-1.5	0.3

as significant at (very approximately) the 5% significance level.

We note, in fact, that the Jan.-Feb. mean of Table 8 departs from 11.0 by almost four standard errors, the Feb.-Mar. by over four standard errors, the Mar.-Apr. by five standard errors, and the Jul.-Aug. by slightly less than one standard error. Thus, notwithstanding the crudity of the underlying tests employed here, it appears reasonably safe to say that the three month-pairs taken from the winter half-year exhibit non-random opposition frequencies while the summer month-piar exhibits only random frequencies. More specifically, Tables 8 and 9 imply that a statistically significant tendency seems to exist for successive January and February precipitation amounts to depart from their respective long-term means in opposite sense (i.e., in the sense of Jan.-Feb. negative autocorrelation), while February-March and March-April exhibit significant persistence tendencies (positive autocorrelation). Note that the difference in the 32-station mean opposition frequencies for the Jan.-Feb. versus the Feb.-Mar. case is 3.3, or over six times the larger of the two relevant standard errors, implying a highly significant difference in the persistence characteristics of these two successive month-pairs in the middle of the winter wet season.

The writer sees no way in which these several results concerning intermonthly autocorrelation of precipitation might have been anticipated. The one seemingly reasonable presupposition that had been made prior to doing the above analysis was the following: Since the summer rainy season tends to be only a little greater than a month in total length, and sometimes begins early in July, but sometimes rather late in July, a tendency towards opposition was expected to be found in the July-August pair. This was, contradictorily, the one non-significant case out of the four month-pairs considered, so the presupposition was quite unsupported by the analysis.

The odd pattern in winter opposition and persistence tendencies revealed by the above analysis is simply not now understood by the writer. Inasmuch as the analysis method employed was a relatively crude one for which only quite rough significance tests could be made, it is recommended that more refined methods be used in the near future to check the conclusion reached here, for they concern features of the winter precipitation regime that may be made to yield telltale evidence of the large-scale factors controlling variability of successive portions of the winter precipitation season. It was regarded as beyond the scope of the present study to search further for confirmation or for synoptic causal factors -- all that can be done here is call attention to what appears to be a point worth careful scrutiny in the Institute's more intensive precipitation analyses by punchcard techniques.

As one last interesting feature of Table 8 to be noted here, it should be pointed out that a listing of individual station-cases where the 22-year mean station opposition frequencies (body of Table 8) departed by more than two standard deviations from 11.0 revealed 17 to be in the winter half-year and only 4 in the summer. Of the 17 in winter, 13 were cases where the 22-year means were more than two standard deviations under 11.0 in Feb.-Mar. or the Mar.-April month-pairs. And, it is to be emphasized, these 13 cases were quite widely distributed geographically over the state, suggesting no subregional tendencies toward greater or lesser persistence.

In qualitative conclusion of the above examination of inter-monthly precipitation persistence characteristics, the following four statements may be tentatively made.

i) A wet January in Arizona tends (to a slight but statistically significant extent) to be followed by a dry February, and vice versa.

ii) A wet February in Arizona tends (to a slight but statistically significant extent) to be followed by a wet March, and vice versa, and the same relation holds, in turn, for March and April

iii) A wet July implies neither a wet nor a dry August, i.e., neither opposition nor persistence is evident in the July-August data.

It bears repreating that each of the above conclusions warrants further testing, using more refined statistical techniques than the simple one here employed. It seems quite likely, however, that the one finding that will not be contradicted by further analysis is this:

iv) A marked difference in persistence characteristics appears to exist between the pairs Jan.-Feb. and Feb.-Mar.

The synoptic-climatological explanation for this last result must be found, for the explanation will almost certainly shed very useful light on the still poorly delineated picture of how general-circulation features that control precipitation in the in the Southwest. It is recommended that further studies, if they employ the "opposition frequency" approach, be carried out in terms of departures from median rather than mean, for use of the mean introduces (for skewness reasons) a slight bias towards non-real persistence, the quantitative effects of which have not been determined in the present pilot study.

F. Seasonal distribution of precipitation. Arizona lies in a region of east-west transition between (1) the Mediterranean type climate of southern California with its strong concentration of precipitation in the winter period, and (2) the southern Plains type of climate of eastern New Mexico and northwest Texas with its concentration of most of the annual rain into the summer period. In addition, Arizona is transitional in a meridional sense too, in that there is rapid variation of seasonal distribution from central Arizona southward into the Mexican portions of the Sonoran and Chihuahuan deserts, the latter receiving the bulk of their annual precipitation in the summer months and getting only slight amounts of Pacific cyclonic-type precipitation. Arizona, therefore, is a state whose patterns of seasonal precipitation distribution are more than ordinary interest.

Over and above the general points of interest just cited, there are two particular reasons for examining seasonal distribution statistics. At this initial phase of the Institute program we are not yet well-informed on the nature and location of the main climatic divides within Arizona, and one of a number of useful indicators of climatic divides and gradients is to be found in the geographic pattern of mean seasonal distribution. Secondly, the critical dependence of evapotranspirational loss on temperatures is such that secular trends in seasonal distribution within various subregions of Arizona and the Southwest are of fundamental concern in hydrologic studies. There is basic need for much more detailed information on long-term swings, all over the Southwest, between one and the other type of seasonal distribution. These hydrologically important swings have unusually high probability of appearing in Arizona precipitation records just because of Arizona's unique location, climatologically speaking.

For these several reasons, some preliminary analysis of the seasonal distribution problem was done as part of the present study in order to see where further, more intensive, investigation can best be directed.

1. Statewide patterns of mean seasonal distribution. In order to delineate the principal features of the differences in winter-summer precipitation amounts from one part of Arizona to another it is necessary to use a much larger number of station-records than the number that has been utilized in most of the pilot studies reported above. The great importance of not having secular-trend effects distort this pattern led to the decision to confine the analysis to data from a uniform period of time; and in the interest of getting the largest possible number of stations, a recent period was preferable. Recent publication by the Weather Bureau of the Arizona Supplement to Bulletin W filled the need adequately, for it was possible to obtain seasonal averages based on the 22-year period 1931-52 for 97 Arizona stations from that publication. In a very few cases out of this sample of 97, records slightly under this 22-year length were accepted if they fell in regions of the state having relatively few Weather Bureau substations. The figures are given, for reference use, in Table 10.

Of the several ratios that might be employed to dipict the seasonal distribution, the percent of annual total that falls in summer was chosen for use here. This ratio, expressed as a percentage, was computed for the 97 stations, plotted geographically, and isopleths of percentage drawn. The result is presented in Figure 7. Individual station values were also plotted in Figure 7 in order to reveal the goodness of fit of the isopleths and to reveal where in the state paucity of data calls for only tentative acceptance of the isopleths (e.g., much of the southwestern and northeastern corners of the state).

A first point of interest in Figure 7 is seen in the fact that it was possible to draw a smooth and reasonable pattern of isopleths that faithfully match the individual station percentages. Hence the data used are likely to be quite representative of their surrounding areas. The only significant exception to this statement is found in the case of Bright Angel Ranger Station just across the canyon from Grand Canyon Headquarters. Bright Angel's 36% value does not fit the general pattern in Coconino

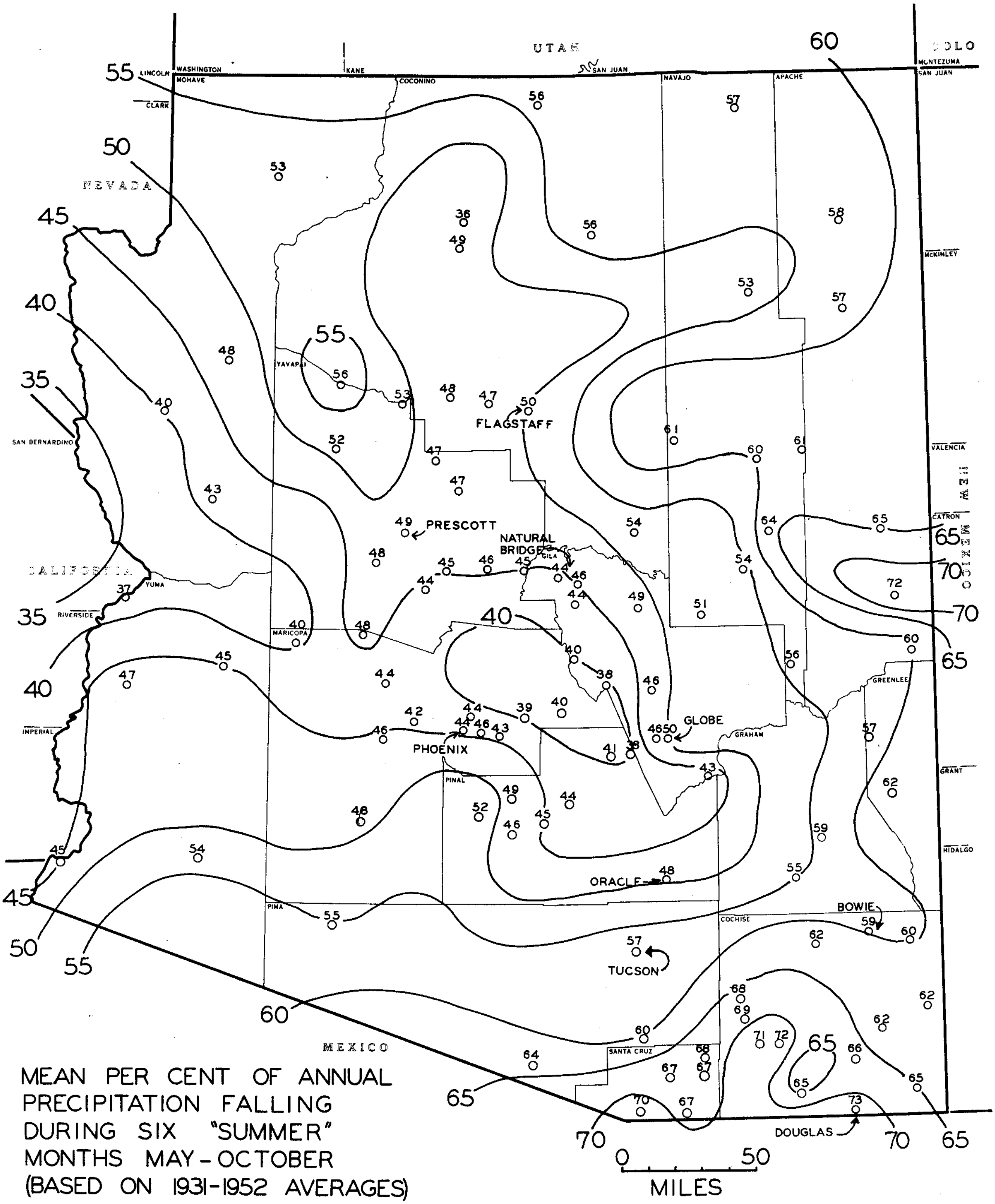


Figure 7. Seasonal distribution of Arizona precipitation. Isolines depict summer mean precipitation amounts expressed as a percentage of total annual mean precipitation. Data for 97 stations for the period 1931-52 have been used in preparing this chart. Isolines depict only the pattern implied in the distribution of the 97 cases plotted, with no attempt at interpolation for effects of topography.

Table 10

Seasonal mean precipitation amounts for Arizona stations. Data taken from Supplement to Bulletin W for 1931-52 only. In a few cases the tabulated means are for slightly less than 22 years. Winter is the period from Nov. 1 to April 30, summer the remaining six months. All means are in inches.

Station	Winter Mean	Summer Mean	Annual Mean	Percent in Summer
Aguila	5.94	4.01	9.95	40
Ajo	4.23	5.18	9.41	55
Alpine	8.00	11.98	19.98	60
Apache Powder Co.	3.74	8.50	12.24	69
Ashfork	6.35	7.10	13.45	53
Benson	3.79	8.12	11.91	68
Bisbee	6.06	11.43	17.49	65
Bowie	4.03	5.78	9.81	59
Bright Angel R.S.	16.33	9.06	25.39	36
Buckeye	4.19	3.56	7.75	46
Canelo R. S.	6.05	12.21	18.26	67
Casa Grande	4.37	4.00	8.37	46
Casa Grande Ruins	4.64	3.82	8.46	45
Cedar Glade	7.34	6.58	13.92	47
Childs	9.83	7.97	17.80	45
Chinle	4.13	5.77	9.90	58
Cibecue	9.44	10.02	19.46	51
Clifton	4.64	7.43	12.07	62
Cordes	7.10	5.82	12.92	45
Crown King	14.76	11.77	26.53	44
Douglas Smelter	3.22	8.57	11.79	73
Eagle Creek	6.56	8.87	15.43	57
Elgin	4.71	9.88	14.59	68
Fairbank	3.42	8.28	11.70	71
Flagstaff	9.16	9.13	18.29	50
Florence	5.48	4.25	9.73	44
Fort Grant	5.97	7.44	13.41	55
Fort Valley	12.03	10.81	22.84	47
Ganado	4.98	6.73	11.71	57

Table 10 (Cont'd.)

Station	Winter Mean	Summer Mean	Annual Mean	Percent in Summer
Gila Bend	2.98	2.76	5.74	48
Gisela	9.46	7.42	16.88	44
Globe	7.79	7.77	15.56	50
Gould's Ranch	4.42	3.41	7.83	44
Grand Canyon H. Q.	7.80	7.35	15.15	49
Granite Reef Dam	5.47	3.46	8.93	39
Holbrook	3.07	4.65	7.72	60
Intake	6.95	5.91	12.86	46
Jeddito	5.35	6.09	11.44	53
Jerome	9.04	7.88	16.92	47
Kayenta	3.48	4.66	8.14	57
Kingman	6.90	4.54	11.44	40
Lee's Ferry	2.59	3.28	5.87	56
Leslie Canyon	4.26	8.45	12.71	66
Litchfield Park	4.71	3.42	8.13	42
Maricopa 8 SSE	3.31	3.58	6.89	52
Mesa Exp. Farm	4.44	3.31	7.75	43
Miami	10.40	8.75	19.15	46
Mohawk	2.05	2.42	4.47	54
Mormon Flat	7.91	5.34	13.25	40
Mount Trumbull	6.07	6.90	12.97	53
Natural Bridge	13.96	10.37	24.33	44
Nogales	4.97	11.36	16.33	70
Oracle	9.62	8.95	18.57	48
Parker	3.22	1.90	5.12	37
Patagonia	5.63	11.41	17.04	67
Payson R. S.	11.22	9.71	20.93	46
Petrified Forest N. Mon.	3.56	5.52	9.08	61
Phoenix	4.32	3.36	7.68	44
Pinal Ranch	15.55	9.70	25.25	38
Pinedale	8.61	10.08	18.69	54
Portal	6.10	9.77	15.87	62
Prescott	9.64	9.30	18.94	49
Quartzite (only 10 yr. mean)	3.49	3.08	6.57	47
Reno R. S.	11.50	7.64	19.14	40
Roosevelt	9.82	6.03	15.85	38
Rucker Canyon	7.29	11.93	19.22	62
Sacaton	4.70	4.51	9.21	49
Safford	3.68	5.33	9.01	59
Saint Johns	4.06	7.47	11.53	65

Table 10 (Cont'd.)

Station	Winter Mean	Summer Mean	Annual Mean	Percent in Summer
Salome 6 SE	4.45	3.59	8.04	45
San Carlos Reservoir	7.83	5.81	13.64	43
San Rafael Ranch	5.74	11.61	17.35	67
San Simon	3.43	5.09	8.52	60
Santa Margarita	5.63	10.09	15.72	64
Santa Rita Exp. Range	8.10	12.00	20.10	60
Seligman	4.40	5.49	9.89	56
Snowflake	4.22	7.66	11.88	64
Springerville	3.36	8.79	12.15	72
Stephens Ranch	4.48	8.19	12.67	65
Superior	10.64	7.40	18.04	41
Sycamore R. S.	9.03	7.79	16.82	46
Tempe Vegetable Farm	4.26	3.57	7.83	46
Tombstone	3.92	9.87	13.79	72
Truxton	6.19	5.82	12.01	48
Tuba City	3.01	3.82	6.83	56
Tucson U. of A.	4.50	6.07	10.57	57
Wallace R. S.	8.89	10.44	19.33	54
Walnut Creek	7.88	8.55	16.43	52
Walnut Grove	8.92	8.17	17.09	48
Whiteriver	8.37	10.52	18.89	56
Wickenberg	6.07	5.60	11.67	48
Wikieup	6.16	4.63	10.79	43
Willcox	4.53	7.53	12.06	62
Williams	11.09	10.38	21.47	48
Winslow	2.79	4.43	7.22	61
Wittman	5.61	4.38	9.99	44
Young	10.47	10.10	20.57	49
Yuma Citrus Station	1.81	1.51	3.32	45

County, while the value at Grand Canyon Headquarters fits it rather well. This is probably an orographic peculiarity or a local wind effect since Bright Angel is at an elevation that is about 1500 ft. above Grand Canyon Headquarters on the leeward side of the Canyon, but for the present purposes this question had to be passed over and the Bright Angel value ignored in drawing the isopleths.[1]

The second general point to be made in reference to Figure 7 is that it reveals a very complex statewide pattern of seasonal distribution in Arizona. Large differences occur because of Arizona's aforementioned peculiar location in a region of strong longitudinal and latitudinal gradients of cyclonic precipitation; and orography superimposes subregional irregularities on these gradients, with the result that the pattern is quite intricate. It is to be noted that this degree of complexity is not indicated in the only other regional examination of this problem that the writer found in the literature, namely the study of Dorroh (1946). It seems likely that Dorroh used a very much smaller number of stations in his analysis of seasonal distribution patterns over the Southwest than the number used in Figure 7. This inference suggests that caution should be used in drawing conclusions from even as large a sample as underlies Figure 7, for doubling or tripling the sample may well add still further significant complexities to the pattern.

A number of general features of the pattern of Figure 7 merit comment. The low percentages of summer precipitation in the west and southwest do not imply large absolute amounts in winter, but rather result from the unusually low summer precipitation in these parts of Arizona. In fact, winter precipitation amounts increase eastward and northeastward from the Yuma corner for orographic reasons, and the low summer amounts result, presumably, from the low frequency with which the moist current of maritime tropical air advancing anticyclonically from the Gulf of Mexico reaches the Yuma corner.

The relatively large amounts of summer precipitation in the southeastern corner of the state, on the other hand, do also represent fairly large absolute amounts (for Arizona). The closed 40 percent isopleth northeast of Phoenix results from the particularly heavy (for Arizona) winter precipitation released, presumably, on the eastern sides of migratory cyclonic storms as southwesterly flow strikes the Mogollon escarpment.

---

[1]Much more careful analysis of the important question of altitudinal differences in seasonal distribution needs to be carried out by the Institute in the near future, using appropriate adjustment methods to utilize those mountain stations having short or incomplete records. A number of related problems have been examined by Hiatt (1953), using northern Arizona data. It is evident from his work that paucity of high-altitude stations will be a serious obstacle to attainment of firm conclusions in those very areas where answers are most pressingly needed -- in the mountainous regions of high runoff yield.

The relatively high percentage of summer precipitation in the northeastern corner of the state is not clearly understood in its relation to the synoptic climatology of the Southwest. Frequency of frontal passages and cyclone passages increases steadily northward across Arizona, judging from preliminary Institute analyses, so one might anticipate a rather regular decline in per cent of annual precipitation falling during summer. That the relative importance of summer precipitation instead rises again as one goes northward from the Mogollon Rim is surprising. Whether this is a rainshadow effect of the Rim itself is not clear, for it is certainly not apparent that this should be a markedly more important phenomenon in winter than in summer. Further study seems required before meaningful speculations can be made.

The general implications of Figure 7 for the problem of spatial homogeneity of Arizona precipitation characteristics are the following: The southwestern corner and the northeastern corner seem to be the most homogeneous as judged by gradients of percentage isopleths.[1] The area containing the closest packing of isopleths is that just to the south and to the north and east of Phoenix. A particularly tight gradient shows up just west of Globe, due almost certainly to topographic factors. A significant, though not intense gradient in a north-northwest to south-southeast line shows up in the southeastern corner of the state. Thus, in going from Oracle to Douglas one finds a difference of twenty-five percentage units in the relative importance of summer precipitation, the only irregularity appearing near Bisbee where the summer percentage falls locally to sixty-five.

The climatic and cloud physical explanations for these several features of Figure 7 will have to await further Institute investigation. The objective of this preliminary examination has only been that of depicting these features and calling attention to some of the associated indications of climatic gradients.

2. Secular trends in seasonal distribution. All of the data for Figure 7 refer to the period 1931-52. It is not necessarily true that the pattern of Figure 7 has held over the past seventy or eighty years during which precipitation observations have been made. In order to gain some notion of the probable extent to which seasonal distribution characteristics have altered, if at all, the curves of Figure 8 were constructed.

In Figure 8 are given 10-year moving average values of the per cent of annual total precipitation falling in summer at four Arizona stations. Each value is the quotient of a decadal moving average summer precipitation divided by the corresponding decadal moving average annual precipitation (i. e., the plotted values are not averages of ten percentages, but rather are percentages obtained from a ratio of two separate averages). Straight horizontal lines are drawn from 1931 to 1952 at the values corresponding to each station's 1931-52 mean per cent summer precipitation (as plotted in Figure 7) to indicate the sense of individual departures from these 22-year means.

---

[1] The fact, noted earlier, that these very corners are the areas of lowest density of available records argues caution in drawing firm conclusions as to their spatial homogeneity. The situation may resemble that of the typical simplicity of isobaric patterns over the oceans.

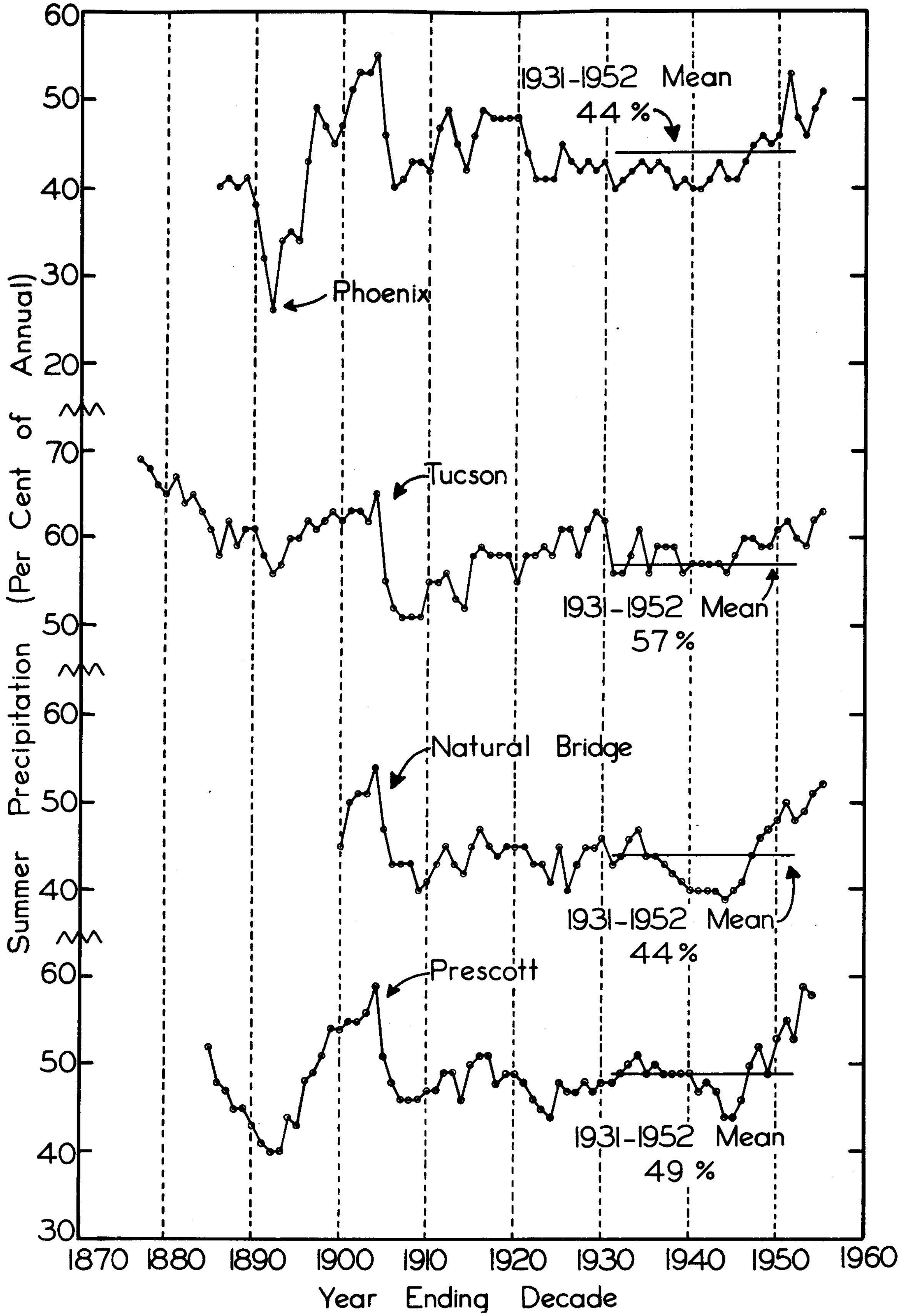


Figure 8. Secular trends in seasonal precipitation distribution at four Arizona long-record stations. Ordinates are 10-year moving mean summer precipitation amounts expressed as percentages of decadal mean annual amounts, plotted above the year ending the decade. The horizontal lines spanning the 1931-52 period show the mean per cent summer precipitation during that period for comparison with historic oscillations. Ordinate scales are non-overlapping but the time-scale is common to all four plots.

It should be noted that one type of difficulty encountered in the interpretation of the secular trend curves of Figure 4 should be almost nonexistent in the problem of interpreting the secular trends in summer-to-annual ratios of Figure 8. Reference is made to the difficulty arising out of temporal inhomogeneities that, in turn, result from station relocations, slow or sudden changes in raingauge exposure, etc. It seems so probable (though exceptions can be imagined) that the latter type of nonmeteorological influences must serve to increase or decrease both winter and summer raingauge catch in about the same proportion, that one may be fairly confident that the secular trends in percent summer precipitation that are shown in Figure 8 are meteorological in origin and are not artificial effects such as must always be at least suspected in the curves of Figure 4, based as these are on unadjusted records (exception, Bowie pre-1888).

A number of extremely interesting features of Figure 8 deserve brief comment:

First, it is immediately apparent that secular trends in seasonal distribution of Arizona precipitation have at times involved ranges of variation that are by no means insignificant as judged against, for example, the differences in 1931-52 means from one area to another as shown in Figure 7. Thus, at Prescott, between the decades ending in 1893 and 1904, an increase amounting to 19 percentage units occurred, while Phoenix, during almost the same period rose by 29 percentage units. Viewed in terms of Arizona synoptic climatology, such significant changes reflect major changes in the general circulation parameters affecting one season but not another. (Reference to Figure 4 shows that the particular cases just cited resulted from the sharp decline of winter prcipitation just before the 1905 termination of the turn-of-the-century drouth.)

Second, it is equally apparent from Figure 8 and Figure 4 that, notwithstanding these occasional oscillations of seasonal distribution, there is much less variability of the summer/annual ratio that of the summer or annual precipitation amounts themselves. Indeed, it seems that one must regard as quite remarkable the general monotony of the curves of Figure 8, for the summer and the winter precipitation regimes are associated with such seemingly different synoptic patterns in the special case of Arizona. However, keeping in mind that both summer and winter regimes are alike in being strongly subject to the control of oceanic anticyclones as these influence, respectively, the intensity of the summer Caribbean tradewind importations of moisture and the winter invasions by Pacific cyclones, the monotony becomes at least understandable in very general terms, even if not less remarkable. Viewed in this light, Figure 7 is interpretable as graphic evidence of the extent to which really large-scale meteorological factors act to control the precipitation received in the Southwest.

Third, the strong similarity in most of the oscillations (or sometime lack thereof) for all four of the curves of Figure 8 bespeaks the tendency for all of Arizona (at least south of the Mogollon Rim) to be similarly affected by the major changes in the general circulation that have influenced Arizona's historic precipitation variations. Slight differences in the several curves might be cited and discussed, but these seem less important than the quite evident similarities. Recalling here the remarks made above concerning the relative insensitivity of the curves of Figure 8 to raingauge-exposure changes, it would seem reasonable to draw the following two very useful conclusions:

i) A substantial part of the dissimilarities of the several curves of Figure 4 are probably due to exposure changes.

ii) Secular trends in precipitation in Arizona during the past half-century have probably been rather similar throughout the state (at least south of the Mogollon Rim).

The latter point could not be deduced from Figure 4 but is a point whose verification is badly needed in order to make intelligent decisions as to how to proceed in many climatic studies of interest to the Institute program. Here it must be regarded as a tentative conclusion deserving further scrutiny in the near future.

Fourth and last in the series of features of Figure 8 that warrant special comment is a point that refers to the portion of the curves corresponding to the last ten to fifteen years. It will be noted that each of the four curves, but especially those for Prescott and Natural Bridge, shows a fairly steady upward climb in relative importance of summer precipitation during almost exactly the period usually identified with the present southwestern drouth. Reexamination of Figure 4 shows that this upward climb cannot be simply ascribed to statewide decline of winter precipitation (though such statewide winter decline does appear to exist) nor to statewide increase in summer precipitation amounts. At both Tucson and Prescott, the absolute summer amounts have been generally falling off, but not as rapidly as have winter amounts, while at Natural Bridge and Phoenix, absolute amounts of summer precipitation have been climbing simultaneously with a general winter decline. Until more extensive study has been completed one cannot safely conclude that these differences in detailed nature of the recent increase in relative importance of summer precipitation contradict conclusion ii) of the preceding paragraph, but it is clear that one must argue caution in too quick acceptance of the latter conclusion.

The hydrologic implications of this fourth and last feature of Figure 8 are extremely interesting, but since these will be discussed more completely in a report, now in preparation, dealing just with rainfall-runoff relations in the Salt River basin, these implications will not be elaborated here.

G. Diurnal variation of January and February precipitation. The next topic to be considered in this report under the general heading of temporal variability concerns the question of whether there is any appreciable diurnal variation in precipitation amounts and/or in time of onset of precipitation during the two winter months, January and February. This is only one of many important questions that must be investigated in relation to diurnal variability before this matter is even properly described, let alone explained. The case of the January and February diurnal tendencies had to be examined during the winter of 1955-56 in order to plan intelligently a program of flight observations of winter cyclonic storms carried out with the support of the National Science Foundation and done with the cooperation of the University of Chicago cloud physics group. Briefly, the practical question asked was this: Is there a tendency for most of the winter-season precipitation to fall during the night? Personal observations over just one winter by the writer had led him to suspect a trend in the direction of a majority of winter precipitation coming between sunset and sunrise, and if this were true, arrangements for the nighttime flights as well as daytime flights had to be made.

A quick analysis of Weather Bureau hourly precipitation data for a number of southern Arizona stations for just five years' records was immediately made. It indicated that, although the preceding year had actually exhibited a marked tendency towards nocturnal winter precipitation, this was a departure from five-year normal that could scarcely be expected to occur during any given year, so no special plans for predominantly night flights were made.

Since the same type of question could arise in other future Institute operations, (e. g., winter radar observations, cloud seeding operations, etc.), and since information of any kind of diurnal characteristics of precipitation processes is of basic cloud physics interest, the initial analysis was extended to cover the 16-year period 1940-55 for which hourly precipitation data were available in the Institute reference files. The results are summarized here, even though they are fragmentary compared with what still needs to be done in this general problem area, chiefly to get them into form useful to others in the Institute and in the region, pending initiation of a more thorough study in the future.

Twelve southern Arizona stations reporting hourly precipitation were used for the 1945-55 period, and a decreasing number for years earlier than 1945, since there were not then as many stations reporting hourly data. The 1945-55 stations will be listed, but no more detailed station-information given here. The twelve were Ajo, Bisbee, Bowie Jct R-15 (or W-5), Casa Grande Ruins, Cochise P.H., Douglas, Nogales, Oracle, Phoenix WBAP, Phoenix PO Bldg (or WB City), Tucson Nursery, Tucson WBAP.

The day was divided into eight 3-hour intervals for tabulation purposes. The time notation used in the Weather Bureau Hourly Data publications will be used here: Thus the period denoted "1-3" begins at 0000 and ends at 0259, and "10-12" begins at 0900 and ends at 1159, etc. The first part of the study was extended only to occurrence vs. non-occurrence and did not concern diurnal variation of amounts, though the latter is of great interest and must be investigated at an early time. The number of occurrences of any amount of rain of 0.01 inch or more per hour was tallied. Thus, a single station could, on a single day, and for a single 3-hour category, contribute from zero to three to the tally, and with twelve stations and thirty-one days (January), a single 3-hour category could show a maximum count of 3x12x31 or 1116. (Commentary on Arizona's freedom from long showery spells is found in the fact that the higest observed monthly count in any single 3-hour interval was 93 for midnight to 0259 during January of 1955!)

The results of the tabulation are presented in Tables 11 and 12. In each table the first two columns give, respectively, the year and the number of stations whose hourly data were used (uniformly 12 only after 1945) during that year. In the next eight columns are given the percentages of total observed precipitation occurrences (in excess of 0.01 inch) for the given month and year that fell in each of the 3-hour time intervals denoted by the

Table 11

Diurnal variation of occurrence of precipitation in January for the period 1940-55 in southeastern Arizona. Figures tabulated in the body of the table give the per cent of all hours of occurrence of each year's January precipitation (in excess of 0.01 inch per hour) that were reported in the given 3-hour interval indicated at the top. See text for discussion of remainder of table.

		Per Cent of Total Occurrences by 3-hour Intervals									
Year	No. Sta.	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	Total Hours Ppn. For Avg. Sta.	Total Jan. Ppn. for Avg. Sta. (inches)
1940	4	17	14	17	17	8	6	11	11	9	0.44
1941	9	21	15	11	7	8	9	13	16	20	1.32
1942	12	20	27	15	14	11	5	4	4	11	0.53
1943	11	8	15	14	17	16	12	12	6	20	1.10
1944	11	15	7	9	12	15	19	14	8	11	0.63
1945	12	19	18	16	14	9	7	7	9	26	0.96
1946	12	10	11	8	10	11	16	19	15	33	2.00
1947	12	14	20	16	20	17	6	4	4	6	0.28
1948	12	4	25	14	0	25	18	0	14	2	0.02
1949	12	13	16	13	14	15	13	10	8	48	2.20
1950	12	10	8	9	18	16	12	12	16	12	0.42
1951	12	16	13	12	15	16	13	8	7	19	1.50
1952	12	14	8	11	7	11	16	19	14	15	0.75
1953	12	15	10	15	0	0	20	30	10	2	0.12
1954	12	10	12	16	16	11	11	13	11	16	0.92
1955	12	17	13	10	9	9	10	15	17	46	2.10
16-yr. avg:		14	14	12	13	12	12	12	11	19.1	0.84

Table 12

Diurnal variation of occurrence of precipitation in February for the period 1940-55 in southeastern Arizona. Figures tabulated in the body of the table give the per cent of all hours of occurrence of each year's February precipitation (in excess of 0.01 inch per hour) that were reported in the given 3-hour interval indicated at the top. See text for discussion of remainder of table.

Year	No. Sta.	Per Cent of Total Occurrences by 3-hour Intervals								Total Hours Ppn. for Avg. Sta.	Total Jan. Ppn. for Avg. Sta. (inches)
		1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24		
1940	6	13	11	16	14	13	14	9	10	21	1.51
1941	9	15	15	15	11	11	11	12	11	31	1.52
1942	12	10	16	15	14	11	14	12	9	25	1.20
1943	11	22	37	2	6	4	4	4	22	5	0.28
1944	11	12	12	12	21	15	17	8	4	27	1.33
1945	12	21	21	21	18	11	3	3	3	6	0.23
1946	12	10	13	15	13	16	13	15	5	5	0.15
1947	12	28	24	19	9	0	2	3	16	5	0.12
1948	12	11	7	7	11	14	20	15	15	34	1.16
1949	12	10	14	20	16	12	10	9	11	16	0.24
1950	12	4	6	3	5	14	25	30	19	13	1.03
1951	12	9	16	13	11	14	20	11	7	10	0.31
1952	12	24	22	12	0	7	2	5	27	3	0.17
1953	12	5	14	17	34	34	27	20	13	16	1.07
1954	12	6	8	8	11	19	14	15	18	6	0.48
1955	12	17	9	12	7	3	9	19	24	6	0.28
16-yr. avg:		11	13	12	13	13	15	12	11	13.9	0.69

eight column-headings. Thus, in January, 1940, for a total of four recording-raingauge stations available in southeastern Arizona, 17 per cent of all station-hours showing reported precipitation fell in the "1-3" time-category. In the next to the last column of each table is given the actual number of hours for the given month and year during which precipitation was falling, expressed as monthly total hours for an "average southeastern Arizona station". Thus, in January of 1940, an average station (based, in this instance on a small sample of but four) had a total of only 9 hours during which precipitation fell in excess of 0.01 inch/hr. Finally, the last column of each table gives, for comparison purposes, the average total monthly precipitation for the given month and year as computed (in general) from the same dozen southern Arizona stations. Thus, in January, 1940, the average monthly total for a southern Arizona station was 0.44 inch. (Data for the period prior to 1945 include a few stations not entering into the hourly-precipitation tabulation in order to achieve a more stable average monthly figure.)

In the last row of each table are given the 16-year average percentages (in columns three through ten) that summarily reveal the 16-year average diurnal variation of occurrence of precipitation for these two winter months. In the same bottom row of each table, the next-to-last column entry gives the 16-year average number of hours of precipitation per month for an average southeastern Arizona station, and the corner entry is the 16-year average precipitation amount for the average southeastern station. (It is to be noted clearly that the term, "average southeastern Arizona station", is only used here for brevity to denote the sample of a dozen used here.)

The implications of Table 11 and Table 12 are several. First, the 16-year average percentages at the bottom of each table show very little diurnal variation from the a priori expectation value of 12.5 per cent per 3-hour interval. The greatest departure from this value occurs in February when 15 per cent of all occurrences in the 1940-55 period were found to fall in the "16-18" time interval. It should be pointed out that these grand-average percentages are not simple averages of the percentages tabulated in the body of the table, but rather are computed from weighted total number of hours of occurrence of precipitation in each interval. The latter method avoids the exaggerated influence of a few large or small percentages contributed by years of very low total number of hours of precipitation.

Individual years exhibit rather wide fluctuations in diurnal variation in both months, but this is seen to be largely associated with occurrence of very dry years, when the percentages become very unstable on a small-sample basis. The chief reason for tabulating these individual percentages to serve as very reliable predictors of the diurnal variations in any given year. In any field operation or observational program, it is exactly that kind of prediction of the diurnal characteristics of a given year's January or February precipitation that would be useful in planning activities, the body of Tables 11 and 12 reveal clearly that statistical prediction for planning purposes is going to be rather uncertain. Thus, if flight or radar equipment had been arranged, in February of 1950, only for the normal work day, about 62 per cent of the total hours of precipitation would have been missed, while in February, 1953, operations planned for just the normal work day would have covered about 95 per cent of all cases of precipitation.

The entries in the next-to-last column of each table are very informative. Examining first just the 16-year average monthly total number of hours of occurrence of over 0.01 inch of precipitation at an average southeastern Arizona station, we see that only 19.1 hours per month was the January average for an average station. There are 744 hours in January, so a mere 19.1/744 or only 2.6 per cent of all January hours during 1940-55 were rainy in southeastern Arizona. Similarly, there are about 678 hours in an average February, allowing for leap years, so during 1940-55 only 13.9/678 or only 2.1 per cent of all February hours were rainy. Briefly, there is only about one chance in forty or fifty that a given hour in January or February will have over 0.01 inch of precipitation in southeastern Arizona under present climatic conditions!

The rainiest January in the 16 years studied was in 1949, during which about 48 hours were rainy (i.e., had over 0.01 inch per hour), and the least rainy January came in 1953 with only 2 hours at an average station during the whole month.[1] For February the rainiest case came in 1948 with about 34 rainy hours at an average station for the whole month, while the least rainy case was 1952 with a mere 3 hours for the month. Recalling that the coefficients of variation were generally high for the winter precipitation at the Arizona stations considered earlier in this report, one is less surprised at this very broad range in total monthly number of hours with precipitation. From a field-research viewpoint, this variability implies the difficulty of being quite uncertain that, in any given January or February, one will have an adequate number of hours during which the synoptic situation yields good case material with which to work.

Casual comparison of the entries in the last two columns of both Table 11 and Table 12 suggests that there is a rather close correlation between average number of hours of precipitation in a given month and average precipitation amount for that same month. Any marked absence of such correlation would be of great interest with regard to the cloud physics involved, so to be quite certain that the correlation was strong, scattergrams were prepared. These confirmed the casual inspection -- the relationship is a strong one. As a quantitative measure of this relation, the Spearman rank-difference correlation coefficients were computed for the sixteen January cases and for the sixteen February cases. The correlation coefficients were, respectively, 0.95 and 0.84. Hence, it seems possible to say that, during the 1940-55 period at least, the intensity, or release-rate, of winter storms has not varied much, for, in general, the dry months have simply been those in which only a few hours of precipitation were reported while the wet months have, conversely, resulted from synoptic situations that gave longer intervals of passage overhead of precipitation-releasing cloud systems. This is a rough, but useful inference in the sense that it excludes the converse from immediate need of consideration.

---

[1] For brevity, the term "rainy" will be used here to denote either rain or snow. Southeastern Arizona gets most, though not all, of its winter precipitation as rain.

Intercomparison of January and February average monthly hours of precipitation for given years, reveals cases where a January characterized by a large number of rainy hours was followed in the same year by a February with very few rainy hours. The year 1948 is a case in point, for January of that year had a 16-year record number of rainy hours, yet the next month was next to the lowest in number of rainy hours for February! The rank-difference correlation coefficient of the sixteen pairs of average monthly (Jan. vs. Feb.) number of hours was found to be -0.21, a result which, in a general way, could have been predicted on the basis of the results cited earlier in this report concerning the tendency for opposition in sign of departures from normal for January and February precipitation in Arizona.

Tables 11 and 12 show that whereas the 1940-55 average number of rainy hours at a southeastern Arizona station in January was 19.1, that for February was 13.9. Does this imply a real difference in the storm characteristics of these two successive winter months? A more appropriate pair of figures for this kind of comparison is obtained by dividing the average number of rainy hours by the average precipitation amount. For January, the 16-year figures then yield 0.044 inches per hour of activity and the February figures yield 0.049 inches per hour. These two quantities are near enough to equality to indicate that no very marked difference in intensity of precipitation distinguishes January storms from February. Put in other words, during the past 16 years, January has been a bit wetter chiefly because its storms have been a bit longer-lived than those of February.

All that has been said up to this point concerns just the question of occurrence or non-occurrence of precipitation in given hours of the day without regard to time of onset or ending of the precipitation. In the initial stages of considering the problem of diurnal variations in winter precipitation, it was suggested by a colleague[1] that the January 1956 storms seemed to tend to start precipitating in the Tucson area in the evening hours. Although no statistics on this tendency were obtained for immediate use in planning the 1956 flight activities, this suggestion was felt to be worth further study, for if it were true that there is any prevailing tendency for winter storms to begin precipitating after sunset, this might well be evidence of some unrecognized cloud physical process associated with the initial phases of development of winter cyclonic precipitation. Consequently the point has been investigated and the results are summarized here.

Because winter storms frequently bring more than a single continuous period of precipitation, it is necessary to introduce a working definition of the "time of onset of precipitation", in order to avoid counting as a new beginning each successive renewal of precipitation in a period of intermittent precipitation. The working criterion selected was this: An "onset" was regarded as occurring at a given station when 0.01 inch or more was reported in an hour that was preceded by 12 or more hours of rainless (under 0.01 inch) observations. The 12-hour interval was based on the time a 250-mile wide storm area would require to pass a given point moving at about 20 mph, these figures being reasonable estimates for winter storms, though the overall criterion is admittedly a bit arbitrary.

---

[1] Milton Draginis of the University of Chicago flight group.

A frequency count of such "onsets" was made for the same set of southeastern Arizona stations previously used for the study of occurrence vs. non-occurrence, and the same 16-year period from 1940-55 was examined. Also, the same breakdown into eight 3-hour intervals per day was employed, and only the months of January and February data were studied.

Only the 16-year average figures will be reported here, since individual yearly values just reveal the same kind of rather wide variability that the occurrence data did with respect to individual years. The averages have been based on weighted values such that the early years, for which the numbers of stations available were less, have been adjusted to yield data equivalent to those obtained from the 12 stations used since 1945. In Table 13, the results are presented in the form of 16-year average percentages of total mean monthly number of "onsets" per 3-hour time interval per month per average southeastern Arizona station. The notation used for time intervals is identical with that used in Tables 11 and 12.

Table 13

Average diurnal variation of time of onset of precipitation in January and February for the period 1940-55 in southeastern Arizona. Based on data for 12 stations for 1945 and later, and for somewhat fewer before 1945. See text for further explanation.

	Per cent of total mean monthly onsets by 3-hour intervals								Average No. Onsets per Sta. per Mo.
Month	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	
January	14	12	11	13	11	12	13	14	3.1
February	12	14	15	15	14	12	11	7	2.4

The 16-year average percentages of Table 13 reveal that there is some slight basis for the suspicion which initially led to this tabulation--namely, a tendency for January storms to start precipitation in southeastern Arizona in the evening hours, but the table makes clear the fact that this is only a very weak tendency. In the half-day from 1800 to 0559 MST, during the past sixteen Januaries, an average of 53 per cent of all onsets occurred, with 47 per cent occurring in the remaining half-day. Inspection of the year-to-year variability in the original frequency counts (not reproduced here) shows that so slight a difference as six percentage units cannot be regarded as very significant.

Unlike January, February has had on the average during the past sixteen years, rather more onsets in midday than during the night-time hours. Specifically, 56 per cent of all February onsets were found in the half-day from 0600 to 1759, and 44 per cent in the remaining twelve hours. Year-to-year variability is again felt to be too great to warrant attributing any great significance to this difference in diurnal patterns in January and February, though no objective test of significance has been applied. The February data show a wider amplitude of diurnal variation of onsets than do the January data, with about twice as many February onsets in the forenoon hours as in the three-hour period ending at midnight. In a very general way, it appears that daytime field operations and flight operations might be a bit more successful in February than in January; though with such slight diurnal variations as have come to light here, it is evident that a larger sample of data would be needed to attribute much significance to this conclusion.

In the last column of Table 13 are listed the average total number of onsets (as here defined) per month per average southeastern Arizona station. These numerical values bespeak the preciousness of single cyclonic storms moving in from the Pacific with trailing cold fronts sweeping over Arizona; for southeastern Arizona, during the past sixteen years, has enjoyed so meagre a number of these per winter month that one more or less makes a substantial difference in monthly totals. It must be remembered that the data have been taken from a period which has been below normal in precipitation, so these values are subnormal too, presumably. Viewed from the standpoint of Institute research, the fact that, on the average, storms start precipitating (in the sense of the criterion here employed) only about three times in every January and only from two to three times in every February, shows how very few opportunities per year one has to conduct observational studies of the winter precipitation situations, a difficulty that is aggravated by the fact, implicit in Tables 11-13, that roughly half of the few cases available must be expected to occur at night.

One last useful bit of information may be gleaned from Tables 11-13: A rough estimate of the average duration of individual precipitation periods is obtainable by dividing the monthly mean total hours of precipitation (Tables 11 and 12) by the corresponding mean number of onsets per month (Table 13). In this way, one gets 19.1/3.1 or 6.1 hours of precipitation per storm (strictly per "onset") in January and 13.9/2.4 or 5.8 hours per storm in February.[1] Once more looking at the problem from the viewpoint of Institute field operations, the very shortness of winter precipitation periods, coupled as it is with 50 per cent probability of nocturnal onset, increases the chances of missing a significant number (because small at best) of the winter storm-precipitation situations if only a daytime operation is conducted. This last additional point stems, of course, from the fact that if storms are actively precipitating for only about six hours, then those that begin in the evening hours (say, 1800 to 0300, from a practical viewpoint) will have run their short course before daytime operations can begin.

---

[1] The fact that rains of duration as short as a minute are tallied indistinguishable from those lasting sixty minutes (so long as over a trace is recorded), introduces uncertainties that would have to be allowed for, in a thorough duration-study, by assuming each rain to begin on the half-hour. Here, such refinement is out of order.

In summary of the overall discussion of diurnal variations of January and February precipitation, it appears that three principal conclusions may be drawn:

i) Contrary to initial suspicions, there is only slight tendency in winter for any one part of the day to have more precipitation or to be a more likely time for onset of precipitation than any other part of the day. Hence, no special question concerning cloud physical processes are posed by diurnal variations.

ii) There appears to be a faint tendency for January storms to begin precipitating in the nighttime hours and for February storms to begin precipitating in the daytime hours, based on the data since 1940. This contrast between the two months has not been tested for statistical significance, but is probably not significant in view of the large year-to-year variability of most times.

iii) The total number of hours during which precipitation falls, the total number of times precipitation starts, and the consequent duration of percipitation periods are, per average winter month, all so small as to pose serious operational difficulties in a research program designed to secure a good sample of obervational data on winter storms in Arizona. A daytime-only schedule enjoys expectation of encountering only about 7-8 hours of natural precipitation per winter month. Even in the rainiest month of the 16-year period studied, January of 1949, only about 24 day-time hours of rain were reported by an average southeastern Arizona station in the whole month!

H. Diurnal variation of July and August precipitation. An analysis of diurnal variation of summer precipitation was carried out in the same manner as for the case of the winter precipitation. July and August are the two wettest months in Arizona's short summer rainy season, so attention was confined to these two months, and, as before, the 16-year period 1940-55 was examined. Again, only southeastern Arizona stations were studied, because these data are of immediate use in operational planning of Institute field activities. More extensive studies for the remainder of the state will be carried out later. It is expected, however, that the behaviour of the summer rain in the southeastern corner will not differ markedly from that throughout the rest of Arizona in summer.

Since all of the same criteria were used for July and August as for January and February above, they will not be restated here. A summary of results is presented in Table 14 and Table 15, whose content and interpretation are exactly analogous to those of Tables 11 and 12 above.

The entries in the bottom row of each table give the 16-year average diurnal variations in the form of percentages of total observed hours of precipitation that fell within each 3-hour interval. It is evident that, as general meteorological considerations would predict, the summer rainfall exhibits very much greater diurnal variability than does the winter. Thus, in July, the past sixteen years in southeastern Arizona have been characterized by only about 5 per cent of all reported precipitation falling in the

Table 14

Diurnal variation of occurrence of precipitation in July for the period 1940-55 in southeastern Arizona. Figures tabulated in the body of the table give the per cent of all hours of occurrence of each year's July precipitation (in excess of 0.01 inch per hour) that were reported in the given 3-hour interval indicated at the top. See text for discussion of remainder of table.

		Per Cent of Total Occurrences by 3-hour Intervals									
Year	No. Sta.	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	Total Hrs. Ppn. for Avg. Sta.	Total July Ppn. for Avg. Sta. (inches)
1940	8	13	8	3	6	20	21	21	8	10	0.82
1941	10	15	9	11	13	9	14	15	14	20	1.72
1942	12	8	4	2	4	10	26	31	15	15	1.53
1943	11	14	6	6	4	12	15	25	18	17	1.28
1944	10	19	12	4	4	7	11	20	23	20	1.70
1945	10	15	11	3	5	12	20	17	17	25	2.67
1946	12	8	8	4	5	12	22	25	16	28	3.08
1947	11	10	8	3	2	16	20	26	15	18	1.58
1948	12	13	7	3	3	8	20	26	20	19	1.90
1949	12	13	11	12	6	11	20	12	15	29	2.97
1950	12	7	5	6	13	21	16	18	15	37	4.07
1951	12	15	6	2	1	12	22	22	20	22	1.94
1952	12	15	8	3	0	10	16	24	24	19	1.65
1953	12	13	8	2	3	11	21	21	21	30	2.74
1954	12	5	5	5	5	15	25	25	15	31	3.09
1955	12	15	8	6	7	11	18	19	16	45	4.27
16-yr. avg:		12	8	5	5	13	19	21	17	24	2.31

Table 15

Diurnal variation of occurrence of precipitation in August for the period 1940-55 in southeastern Arizona. Figures tabulated in the body of the table give the per cent of all hours of occurrence of each year's August precipitation (in excess of 0.01 inch per hour) that were reported in the given 3-hour interval indicated at the top. See text for discussion of remainder of table.

Year	No. Sta.	Per Cent of Total Occurrences by 3-hour Intervals 1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	Total Hrs. Ppn. for Avg. Sta.	Total Aug. Ppn. for Avg. Sta. (inches)
1940	8	6	5	6	5	5	17	32	24	17	1.82
1941	10	12	3	3	3	10	25	27	17	24	2.16
1942	12	7	5	4	2	12	24	31	15	18	1.86
1943	11	13	14	12	9	14	15	14	9	32	3.84
1944	10	18	14	11	9	9	9	11	19	24	1.57
1945	10	14	14	4	3	9	18	18	20	25	2.98
1946	12	8	6	7	6	10	21	22	20	30	3.13
1947	11	2	4	5	8	17	29	22	13	28	2.22
1948	12	15	7	2	2	9	21	25	19	17	2.41
1949	12	6	4	9	8	15	24	24	10	12	1.33
1950	12	9	4	7	13	18	15	18	16	36	1.06
1951	12	14	13	13	10	10	13	13	13	41	3.51
1952	12	8	5	5	5	17	21	24	17	24	2.33
1953	12	5	4	4	2	18	24	27	16	9	0.94
1954	12	14	9	6	7	16	24	16	8	27	2.37
1955	12	14	8	4	3	12	17	23	19	31	3.57
16-yr. avg:		11	8	7	6	13	19	20	16	25	2.32

7-9 interval, the same fraction in the 10-12 interval, and, by contrast, 19 per cent in the 16-18 interval and 21 per cent in the 19-21 interval. Thus, the time of maximum precipitation activity in July is about four times rainier than the time of minimum activity. Comparison with Tables 11 and 12 shows that the winter precipitation maximum and minimum times in the month of greater diurnal variability, February, stood in the ratio of only 15/11, so the summer max.-min. ratio is some three times greater than the winter max.-min. ratio. Table 15 shows that August has only slightly less diurnal variability than does July.

For both July and August, the lowest frequencies of precipitation-occurrence come in the six-hour period ending at noon, while the highest frequencies of precipitation-occurrence are found in the six-hour period ending with 2100. These summer data were actually tallied by hours, and in view of the large variation in raininess throughout the day, it is of interest to examine the hour-by-hour changes for the 16-year averages. Table 16 displays these data in hourly-frequency form.

Table 16

Percentages of total monthly hours for which precipitation in excess of 0.01 inch per hour fell at an average southeastern Arizona station over the 16-year period 1940-55, by hours, for July and August.

Hour Ending	Per cent Occurrence July	Per cent Occurrence Aug.	Hour Ending	Per cent Occurrence July	Per cent Occurrence Aug.	Hour Ending	Per cent Occurrence July	Per cent Occurrence Aug.
01	4.2	3.5	09	1.4	2.1	17	6.9	6.4
02	4.1	3.8	10	1.5	1.9	18	6.8	6.7
03	4.0	3.6	11	1.7	2.0	19	7.2	6.6
04	3.2	2.8	12	2.0	2.5	20	6.9	7.3
05	2.5	2.9	13	2.9	3.4	21	7.0	6.6
06	1.9	2.5	14	4.2	4.1	22	6.3	6.3
07	1.8	2.3	15	5.5	5.2	23	5.9	5.5
08	1.7	2.5	16	5.7	5.9	24	4.9	4.0

The values of Table 16 indicate that the least precipitation activity of all hours of the day during the summer rainy season in southeastern Arizona comes at about 0800 to about 1000 MST. In July, this two-hour period accounts for only 2.9 per cent of all hours of raininess; and in

August, this period accounts for only 4.0 per cent of all August raininess according to the sample here considered. On the other hand, the greatest precipitation activity has come in the 1800-1900 period in July (7.2 per cent average) and in the 1900-2000 period in August (7.3 per cent average). Actually, these percentage values must not be regarded as accurate to tenths of a per cent, so about all that is really justifiably deducible from Table 16 is this: In general, the least rainy summer hours are those from about 1700 to 2100 MST. Note well, that all data used here refer to precipitation at stations that are chiefly in valley locations, so the above must only be applied to valley precipitation.

The next-to-last columns of Tables 14 and 15 give the year-by-year average number of hours of precipitation (over 0.01 inch per hour) per month for an "average" southeastern Arizona station. The average of all of these sixteen monthly averages, as listed at the bottom of this column is seen to be 24 hours in July and 25 hours in August. In the last column are given corresponding precipitation amounts for comparison. It is interesting to note that, for the 16-year average for the sampled stations, the July and August means, 2.31 and 2.32 inches, respectively, are nearly the same, as are also the 16-year mean total monthly hours of raininess per station. A good deal of year-to-year variability in hours of raininess is evident, with July, 1955 exhibiting the highest monthly mean number of station-hours of rain, 45 hours.

Since there are 744 hours in each of the two summer months under consideration, we may infer from these tables that an average southeastern Arizona has, during the past decade and a half, received precipitation in only about 24/744 or 3.2 per cent of all hours. This incidence of precipitation is roughly one-fourth to one-half higher than that of the winter months considered in the previous section of this report.

The average rate of precipitation for the July and the August precipitation for the stations and period studied is about 2.31/24 or 0.096 inches per hour, an intensity almost exactly twice as great as that found typical of January and February precipitation during the same 16-year period and in the same geographic area.

In the preceding analysis of the winter precipitation, it was found that there was quite high correlation between monthly mean hours of occurrence of precipitation and monthly mean precipitation amount (0.95 for January and 0.84 for February for the past sixteen years). To examine the same point in the case of the summer precipitation, rank-difference correlations were computed, yielding values of 0.96 and 0.58 for July and August, respectively. This large difference in degree of duration-vs.-amount correlation for these two consecutive months of the summer rainy season is surprising. Unfortunately, the sampling properties of the Spearman rank-difference correlation coefficient are not well known, so it is not possible to make any clear-cut confidence statements here. However, Mills (1955, p. 315) gives as the standard error of this statistic $(N-1)^{-\frac{1}{2}}$ for large samples drawn from uncorrelated populations, where "large" means something in excess of n = 25. Neither of these criteria of applicability of this formula for the standard error is satisfied in the present instance, but it is of interest to note that for N = 16, as here, the value of $(N - 1)^{-\frac{1}{2}}$ is 0.26. The difference between the July and the

August correlations amounts to 0.38 or only half again higher than the above-computed (but really inapplicable) standard error. One suspects that the July-August correlation difference is not statistically significant.

Tentatively, it may be concluded that certainly in July, and probably also in August, the variations in observed amounts of precipitation in southeastern Arizona have resulted principally from correspondingly varying durations of precipitation-producing situations -- that is, the present data do not in any definite way suggest that large year-to-year variations in release-rate of thunderstorms have occurred, though the low sample-value for August renders this conclusion less firm than in the case of the similar conclusions drawn for the winter months of January and February. A bit of reflection on the cloud physical implications of the above conclusion, combined with recognition of the sampling uncertainties inherent in the present analysis, is sufficient to indicate that one must take these conclusions as only useful gross indications of the actual state of affairs in so complex a question as the one involved here. About all that can be said here is that it is probably unsafe to conclude, on the basis of the 16-year sample here examined, that a real difference between July and August exists, and that this point should be examined more thoroughly by more adequate techniques than the rank-difference method. It is worth special note here, that over half of the sum of squared differences of rank for the August case was due to a single case, that of the year 1950, so one must be quite suspicious that the low value of the August correlation constitutes an effect of sampling fluctuations. In fact, the 1950 plus the 1948 cases together, comprise slightly over 80 per cent of the total of squared rank-differences for August, so if one had drawn a 14-year sample excluding these two years, a very much higher correlation would have been obtained. It would be of interest to make a special study of August precipitation characteristics for 1948 and 1950 in an effort to explain meteorologically the causes of the unusual features of these two periods, but this is beyond the scope of the present study.

The degree of persistence between station-average monthly hours of occurrence of precipitation in July and August was examined by computing the rank-difference correlation coefficient of July-August pairs, and the sample of 16 cases yielded a coefficient of 0.25. In the analysis of persistence between July and August, summarized earlier in this report, the July-August persistence was concluded to be non-significant. Since the rank-difference standard error discussed in the preceding paragraph does apply more closely here than for the case discussed in the preceding paragraph, it is worth noting that the value of $(N-1)^{-\frac{1}{2}}$ is here 0.26 and that this is essentially equal to the July-August persistence correlation coefficient of 0.25, so that one must conclude that this value may have resulted from sampling fluctuations alone. That is, the result here is not in disagreement with results of the earlier persistence analysis for these two summer months.

It is of interest to examine Tables 14 and 15 for their implications for scheduling of Institute field operations or cloud-census observational work in southeastern Arizona. From the 16-year mean percentage frequencies of the bottom line of each table one finds that operations carried out during only the normal workday period would embrace only 32 per cent of all hours of July precipitation and 33 per cent of all hours of August precipitation, if we base our estimates on this 16-year sample. Extension of observational activities from a termination hour of 1700 to one of 2400 adds to the observed sample a very significant portion of all precipitation occurrences, for this seven-hour period is found from the tables to contain 45 per cent of all hours of July precipitation and 43 per cent of all hours of August precipitation. By maintaining operations from 0800 to 2400, it then follows that one can expect to have under observation about 77 per cent (say, three-fourths) of all cases. During the past two summers, the Institute's radar and cloud-census observational activities have, in fact, been carried out during just this sixteen-hour portion of each day, so we may infer, after the fact, that our decision to carry out operations during those hours was a reasonable sound decision; and by the same token, it follows that future operations of similar nature should cover about this same period of time. For special purposes, other time-periods may be indicated, but it is useful to know that "morning-to-midnight" operations may be expected to cover almost three-fourths of all summer precipitation situations.

In summary of the overall discussion of diurnal variations of July and August precipitation, it appears that six principal conclusions may be drawn:

i) Arizona summer precipitation differs very markedly from Arizona winter precipitation with respect to degree of diurnal variability: The ratio of frequencies of occurrence of precipitation during the period of maximum and of minimum activity is three times greater in summer than in winter.

ii) July and August exhibit very similar patterns of diurnal variability of precipitation.

iii) For both July and August, the least rainy hours come from about 0700 to about 1100, while the rainiest hours are those from about 1700 to 2100 MST. These, and all other conclusions drawn here, must be applied only to valley locations and the further, very important question of behaviour of mountain storms must be left open until radar data can be adequately exploited in studies of diurnal variability.

iv) The average rainfall intensity in the summer rainy season was estimated here to be about 0.10 inches per hour, a rate almost twice as great as the corresponding January-February rate.

v) An average southeastern Arizona station has enjoyed only about 24 total hours of precipitation in excess of 0.01 inch per hour per month during July and August in recent years. This total duration is well correlated with observed amounts in July, but less closely correlated in August, the latter difference probably resulting only from sampling fluctuations.

vi) Field operations and summer observations of precipitation phenomena would cover only about one-third of all cases if activities were restricted to only the normal workday, but if extended to midnight would cover approximately three-fourths of all cases, judging from the present analysis.

III. SPATIAL VARIABILITY.

Precipitation is not only variable with respect to time at a given locality, but is also variable with respect to space at a given time (or over a given period). Arid regions can make no claim to unusually high variability in the spatial sense (as they can in the temporal sense), but in all climatic areas, including the arid regions, questions of spacial variability are of very great interest. From the viewpoint of the climatologist, detailed knowledge of spatial variability of precipitation (and, indeed, of all other climatic elements) is extremely useful in planning research, for one repeatedly encounters such practical questions as these: How large a geographical area can one assume is well represented by climatic statistics derived for Phoenix? Can runoff in the Salt River watershed be reliably estimated using precipitation data for the nearby Verde basin, where there happen to be more long-record stations than are found in the Salt Basin? In cloud-seeding experiments, how could one map out target and control areas to take optimum advantage of spatial homogeneity of Arizona climate? Can one safely assume that secular trends in northern Arizona faithfully mirror concurrent trends in the Yuma corner of the state? Again and again, in climatological research one finds that he needs to have at least general answers to such questions of spatial variability before he can even intelligently lay plans for a particular piece of research. Hence, as part of the present pilot study, a very small sample of spatial variability data has been examined, primarily for the purpose of securing just some first-approximation information that would be of use in the next stages of Institute climatological research, and secondarily for the purpose of testing certain computing methods.

A. Interstation correlations. In this exploratory study it was decided to choose one long-record station in southeastern Arizona stations in order to obtain some rough notions of probable geographic patterns of correlation. Because of the very different character of summer and winter precipitation in Arizona, the analysis was carried through not only for annual totals but also for winter and summer totals separately, where "winter" is, as before, defined to begin on November 1 and to end on the following April 30, and "summer" is the balance of the year.

The conventional Pearson product-moment correlation coefficient was computed for a total of 33 series of data (11 station-pairs, and three time-periods, winter, summer, and yearly). The computations were begun prior to installation of the major share of the Institute's IBM equipment and hence a number of the computations were initially done by desk calculator methods. Later, these were redone by IBM methods as an incidental part of another study (concerning effects of non-normality on derived statistics) and this step disclosed two extremely useful items of information:

i) Even with as slow an IBM calculator as the IBM 602A, the overall time for punching original data, wiring boards, calculating, and listing on an IBM tabulator was only about 5 per cent as great as the time for manual computations!

ii) A simple pitfall, seemingly neglected in the literature of statistical methods, was found to yield quite intolerable errors in computing product-moment correlations by the means most suitable to IBM equipment of the type used.

A brief explanation of ii) may be helpful to other readers. Because of the undesirability of employing any negative quantities in calculations in the IBM 602A, the preferred method of computing correlation coefficients is that in which, to avoid dealing with deviations from the mean, one computes squares and cross-products of the raw variates, following which the usual corrections for an assumed zero mean are applied. The serious difficulty into which we unwittingly fell concerns the latter corrections. These corrections involve, in one way or another, subtracting an actual mean, a square of a mean, a cross-product of a mean, or some other function of the actual mean. Now, we had previously computed means for all seasonal and annual precipitation amounts for all stations entering into our correlation computations, but had computed these for the total sample of all available years of record for that station. In any given correlation of Station X versus Station Y, missing data and dissimilar dates of beginning of the total records for each station almost insures that the variates entering into the correlation constitute less than the total of all available observations, and therein lay the pitfall. We used the total-record means in applying the subtrative corrections for assumed zero means in our correlation computations via IBM, whereas we should have used means based on just the overlap years common to both the X and the Y records in each case, i.e., we should have used different means for Station X for each new correlation in which Station X was involved, based on overlap years peculiar to that particular case. This point escaped our attention even after the marked discrepancy between our early desk-calculator results and our subsequently rerun IBM results had appeared. The reason it was not immediately apparent is that it seemed quite unreasonable, at first glance, that the very slight differences between total-record means and overlap-record means could engender errors of correlation coefficients such as, for example, yielding 0.35 for a case wherein desk-calculator methods (involving direct calculations of deviations from means) had given 0.60. However, closer inspection of the arithmetic showed that these small differences in means were capable of throwing the results off by just such very large amounts simply because the correction for assumed zero mean is a small-difference-between-two-large-quantities situation, and squares or cross-products of raw variates, produced differences that might be off by almost 50 per cent! All of the IBM correlations were rerun to correct for the above errors, and the corrected values are presented here. Readers who may have been unaware, as were we, of this pitfall should profit from our mistake.

In Table 16 are displayed the several coefficients for six stations correlated with Tucson and for five stations correlated with Natural Bridge. Conventional standard errors have been computed to suggest the order of reliability of these coefficients, and the number of years of data underlying each correlation is shown for each of the 33 coefficients in the table. Airline interstation distances are given for each station-pair in the table to assist the reader in estimating intensities of climatic gradients.

Table 16

Interstation correlation coefficients for Arizona seasonal and annual precipitation amounts. Two-base-stations, Tucson and Natural Bridge, are used, and winter, (w), summer (s), and yearly (y) correlations given for a series of other Arizona stations. In the body of the table are given, for each station-pair and time-period, the correlation coefficient, the standard error of the correlation coefficient (in parentheses), and the total number of years of data used in the computation. Thus, the Tucson-Bowie correlation of winter precipitation for the past 69 years is seen to be 0.64 with standard error 0.07. The figure in parentheses after each station-name is the approximate airline distance, in miles, from that station to the base-station in question. See text for further discussion.

		Bowie(90)	Flagstaff(210)	Oracle (30)	Phoenix(110)	Prescott(190)	Yuma(220)
Tucson	w	0.64(.07)69	0.58(.09)52	0.93(.02)53	0.82(.04)76	0.60(.07)79	0.65(.06)84
	s	0.59(.08)71	0.44(.12)50	0.51(.10)55	0.38(.10)76	0.36(.10)80	0.19(.11)83
	y	0.50(.09)67	0.53(.10)48	0.82(.15)51	0.58(.09)75	0.39(.10)78	0.33(.10)83

		Flagstaff(60)	Phoenix(70)	Prescott(60)	Tucson(150)	Yuma(220)
Natural Bridge	w	0.57(.10)51	0.85(.03)63	0.87(.03)59	0.78(.05)63	0.68(.07)63
	s	0.63(.08)50	0.61(.08)64	0.60(.08)59	0.54(.09)64	0.22(.12)63
	y	0.53(.10)47	0.73(.06)63	0.82(.04)58	0.70(.06)63	0.52(.09)62

By far the most interesting feature of Table 16 is its clear-c indication that interstation correlation is substantially higher ir ter than in summer in Arizona. With the single, interesting excep of the Natural Bridge-Flagstaff correlations, every case shows stronger winter than summer correlation. Using the relationship that the square of the correlation coefficient measures the fraction of total variance accounted for by regression, one sees that in the Tucson-Phoenix case over four times as strong a correlation holds in winter as in summer, while in the Tucson-Yuma case over eleven times as strong a correlation exists in the winter as in summer.

At this point, the seasonal differences in coefficients of variation should be recalled (see Table 1), for these revealed the fact that relative temporal variability of precipitation is higher for Arizona winters than for Arizona summers. Hence, as has been pointed out earlier in this report, one has the peculiar combination of high temporal variability but low spatial variability in winter as compared with summer in Arizona's precipitation regime. The low spatial variability in winter is easily accepted, for one expects this of the kind of widespread cyclonic storms that bring the winter precipitation (though recent Institute radar studies of winter storms render this a less obvious point than it might seem to be).

Some miscellaneous questions raised by and implications of Table 16 may be noted briefly. Why should Tucson be more closely correlated with Flagstaff in summer than Tucson is with Flagstaff in summer than Tucson is with Phoenix and Prescott, both of which are closer? Perhaps the standard errors point to the answer, for the difference in the correlation coefficients does not exceed one standard error. On the other hand, a seemingly statistically significant difference of correlations appears to exist between Natural Bridge vs. Flagstaff on the one hand and Natural Bridge vs. Tucson and several other stations on the other hand, in winter. Flagstaff on the north side of the Mogollon Rim (but only about 60 airline miles away from Natural Bridge) has a correlation coefficient of 0.57 with respect to Natural Bridge in winter, while Tucson, 150 miles to the south, has a coefficient of 0.78 with respect to Natural Bridge in winter. Even Yuma, 220 miles southwest of Natural Bridge, is more strongly correlated with Natural Bridge than is Flagstaff; so in all, Table 14 gives strong indication that a climatic divide must lie, in winter, just a short distance north of Natural Bridge. This point demands considerably more analysis will be made in the near future.

Comparison of the correlations of Tucson's winter precipitation with that for Phoenix and Bowie, suggests that a marked correlation gradient may exist east of Tucson in winter. Additional correlations should be carried out to examine this point, too.

Table 16 indicates that the wintertime correlation gradient between Natural Bridge and Flagstaff, i.e., in the area of the Mogollon Rim, is replaced in summer by relative homogeneity, for of all the coefficients for summer precipitation shown in Table 16, none is higher than that for Natural Bridge vs. Flagstaff, although the coefficients for Natural Bridge vs. Phoenix and Prescott are not significantly lower. It would appear that the entire area from Phoenix northward at least to Flagstaff is, in summer, characterized by about the same rainfall regime in the restricted sense that interstation coefficients are all close of 0.6 (i.e., about one-third of each station's total variance is predictable in terms of regression on the other station). Since the Tucson-Natural Bridge correlation in summer is 0.54, one is tempted to extend this statement to the entire area from Flagstaff to Tucson, but that this is probably unwarranted is suggested by the 0.38 value for Tucson-Phoenix in summer.

An interesting summer-winter contrast in correlation is to be found in the Tucson-Oracle case. Oracle, lying a mere 30 miles north-northeast of Tucson, but on the north rather than the south side of the major orographic barrier of the Santa Catalina Mountains, is very strongly correlated with Tucson in summer. Reference to Figure 7 reveals that there is a substantial difference in per cent of annual precipitation falling in summer at these two stations, with Oracle getting relatively less summer precipitation. The exact cloud physical explanation of these interesting differences must await further Institute study. It may be noted that Table 10 shows that during the past 22 years, Oracle's winter precipitation has been almost double that at Tucson, a difference not obviously accounted for just in terms of the 200-ft. greater elevation of Oracle.

In summary of the above comments on interstation correlation coefficients, it must be admitted that the eleven sets of correlations here reported have not yielded very satisfactory initial indications of climatic gradients in Arizona. More questions have been raised than have been answered, and it appears clear that until a considerably greater number of correlations have been computed, there will be no clear picture of the dominant patterns for the state. This conclusion, in itself, is useful, however, and steps are being taken to secure a much larger sampling of correlation coefficients soon. As part of these steps, studies are being made of the acceptability of using simpler types of correlation such as the tetrachoric and the rank-difference correlation techniques.

It is to be repeated, finally, that the one fairly definite conclusion deducible from Table 16 is that one must expect to find significantly higher correlations between Arizona stations in winter than in summer, a conclusion that really only constitutes objective confirmation of what was fairly confidently anticipated on general meteorological grounds.

B. Interstation regressions. Inasmuch as the same sets of squares and cross-product sums derived for the above correlation computations can be utilized to calculate regression coefficients and associates standard errors of estimate, and inasmuch as plotted regression lines and scattergrams can exhibit, in a useful way, a number of important characteristics of the climatic relationships here under scrutiny, a number of regressions were computed. Only the winter and the summer seasonal total regressions were determined, since these are of greater interest than the annual total regressions, and only seven station-pairs were considered. In Figures 9, 10, and 11, all of the results are presented.

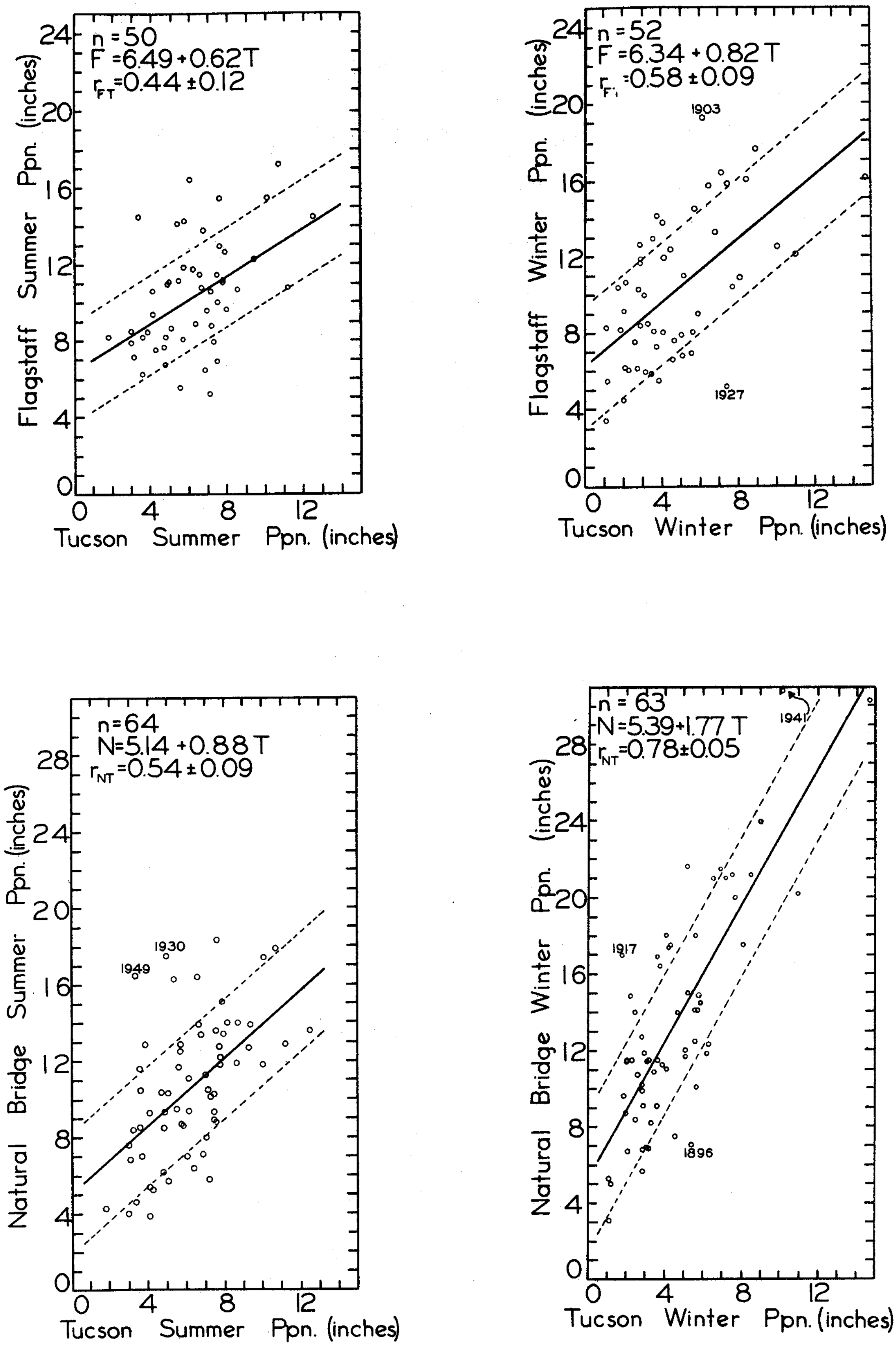


Figure 9. Regressions of Flagstaff and Natural Bridge seasonal precipitation on Tucson's. The bold line is the regression line, the dashed lines are one standard error of estimate above and below the regression line. The legend on each diagram gives the number of years' data employed, the computed regression equation, and the associated correlation coefficient and its standard error. A few outstanding departures from regression have been labelled with their year of occurrence.

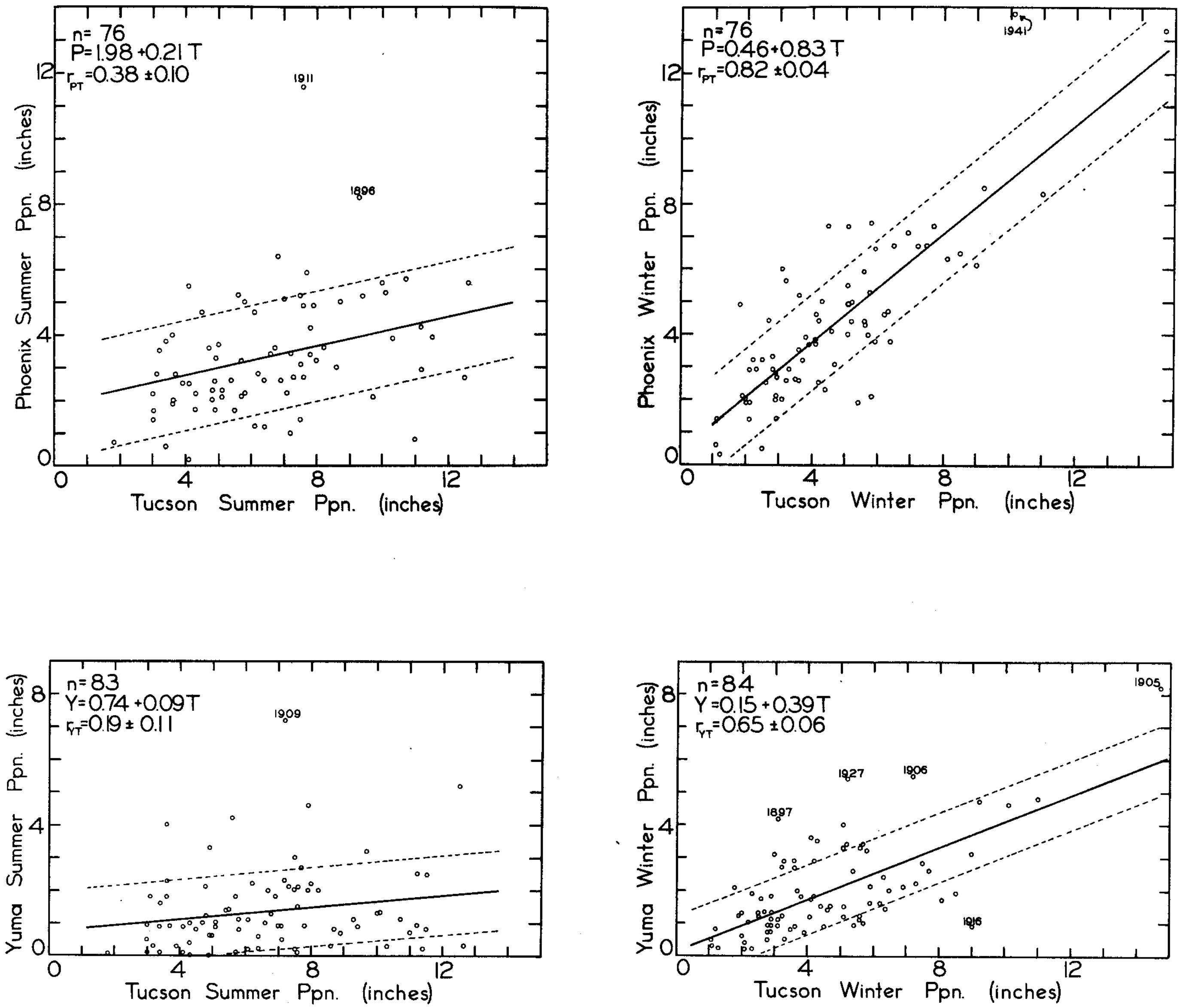


Figure 10. Regression of Phoenix and Yuma seasonal precipitation on Tucson's. For explanation of diagrams, see Figure 9.

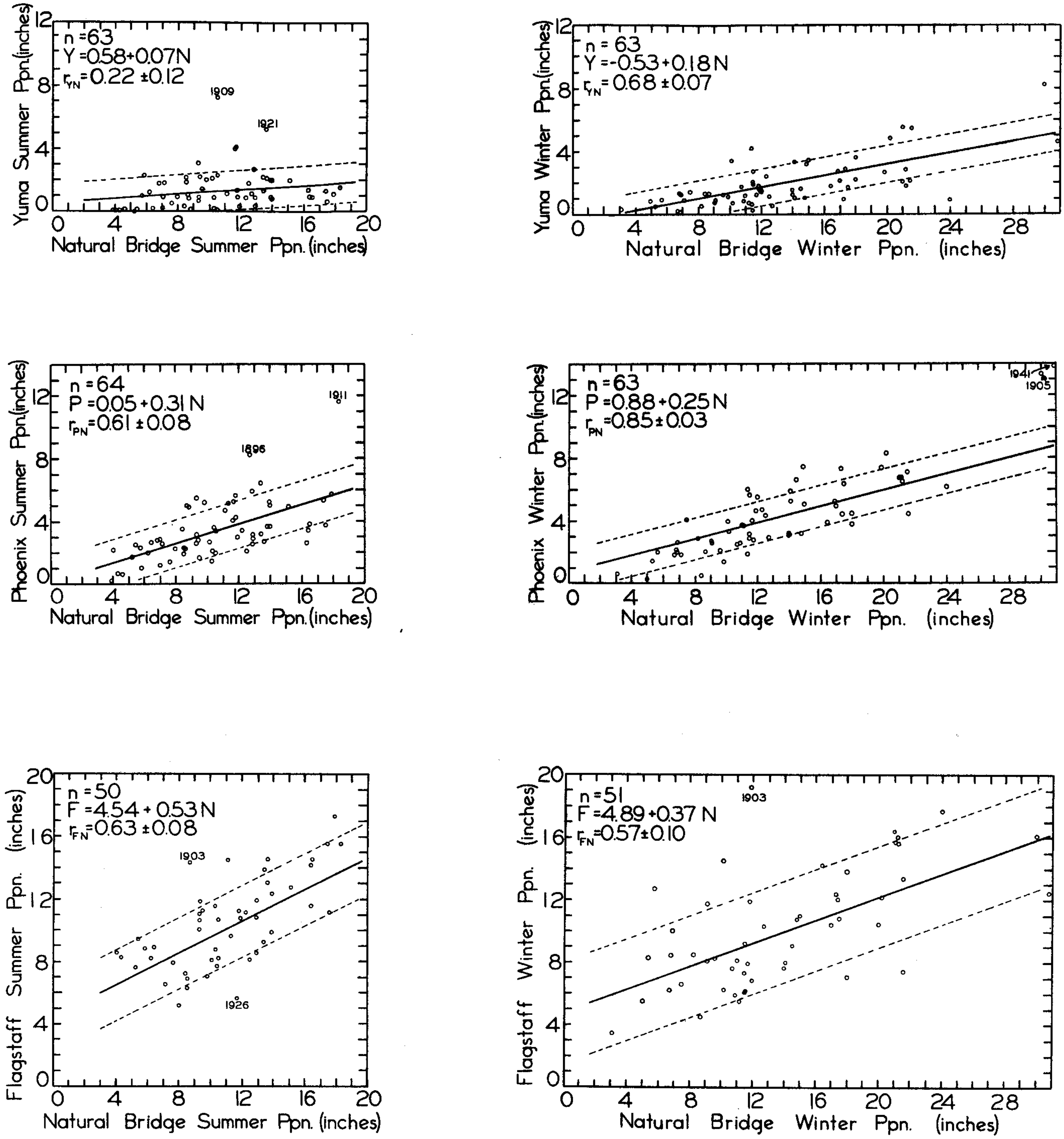


Figure 11. Regression of Yuma, Phoenix, and Flagstaff seasonal precipitation on Natural Bridge's. For explanation of diagrams see Figure 9.

In each panel of Figures 9-11 will be found the following items: A scattergram for the full period of record of overlap years for the station-pair involved, the linear regression line (heavy line), the upper and lower limit of one standard error of estimate departure from the regression line (the light dashed lines above and below the heavy regression line), a few points labeled with the years to which they apply when thses are cases of exceptional departure from regression, the number of years of record entering into the regression (given by the quantity n in the legend), the linear regression equation (with the initial letter of the station-name used to symbolize that station's seasonal precipitation amounts, and the correlation coefficient and associated standard error (with initial letters used as subscripts to denote the station pair involved in the correlation).

Because these figures are almost self-explanatory and are really intended primarily for internal Institute reference use anyway, only a few comments will be appropriate.

1. Confidence limits on regression estimates. There are many popular and scientific reasons for wishing to know the degree of reliability with which one can expect to be able to estimate the precipitation received during a given season at one station when that at another station during the same season is given.

The popular reasons center around the not uncommon practice of describing as anomalous a given season where one part of Arizona gets somewhat more (or less) precipitation than another. Too few persons with an economic stake in precipitation characteristics seem to be aware of the highly variable nature of precipitation distribution, and partly because of this lack of awareness have often drawn incorrect inferences. In the particular instance of rainmaking activities, loose comparisons of a given seeded area's precipitation with concurrent values in some adjacent area have often been questionably used by commercial seeders in all parts of the world to argue efficacy of their seeding, when the observed differences were only of magnitudes that statistical analyses show might easily have resulted from chance variations of the sort revealed graphically in Figures 9-11.

There are many scientific reasons for wishing to have quantitative estimates of the degree of precision with which precipitation estimates can be made for one area using the reported precipitation in another area. One of them concerns the very evaluation problem in cloud seeding that was described above: One must have an objective yardstick of variability against which to measure observed departures from regression. Another concerns the practical problem of estimating missing data in using historic records containing the usual number of occasional breaks. Another deals with the search for climatically homogeneous provinces, and others concern problems of deciding when truly anomalous seasonal patterns have occurred in a given region.

For these reasons, the upper and lower dashed lines in Figures 9-11 are of more than casual interest, for they display the limits within which approximately two-thirds of all historic cases have fallen. Non-normality of the frequency distributions is such that this is only a rough statement, but actual counts of points in the scattergrams will reveal that it is close to the facts. These dashed limiting lines should actually be curves that flare out slightly at the lower and upper limits in accordance with the decreasing reliability of estimation for values of the independent variable that lie well off the mean, but the sample-sizes are so large that this effect is of small importance here.

The dashed lines have been laid off just one standard error above and below the regression line. A more commonly accepted standard of confidence is give by constructing lines lying at two standard errors above and below the regression line, for these have the property (in a truly normal-distribution case) that very nearly 95 per cent of all cases fall in the band they bound, leaving about 2.5 per cent of all cases above the upper and 2.5 per cent of all cases below the lower line. That is, such lines would bound the so-called 95 per cent confidence region. According to commonly used critieria in sampling theory, a given year whose point on the scattergram falls outside that region represents a "statistically significant" departure from regression in the sense that it has an empirical probability of occurrence of only one-twentieth, and this probability is rather widely accepted as "significantly" less than unity. Numerical values of the half-width of such bands are given (approximately) by twice the computed standard errors of estimate, and in Table 17 such half-widths are listed for convenient reference.

Inspection of Figures 9-11 will show that the values for at least one station represented in Table 17, Yuma, clearly cannot be interpreted in the manner just described, for the lower 95 per cent confidence interval is in the region of negative Yuma precipitation, a consequence of the high skewness and low mean precipitation at this station. For the other stations, with conditional exception of the Phoenix regressions, the half-widths of Table 17 can be interpreted, loosely, in the manner stated. Thus, one might say, for example, that if cloud seeding were being done near Natural Bridge, and Tucson were used as a single control station, the seeded winter total precipitation at Natural Bridge would have to exceed the Natural-Bridge-on Tucson regression value by about 7.2 inches before one could attach statistical significance to the result. Table 1 shows Natural Bridge to have a mean winter precipitation, during the past 60-odd years, of about 13.5 inches, so if during a seeded year Tucson received its mean amount, Natural Bridge would have to receive in excess of 13.5 plus 7.2 inches or over 20.7 inches before significance could be said to exist in this hypothetical and oversimplified application of Tabel 17. That is, an increase over "predicted" winter precipitation amounting to some 53 per cent at Natural Bridge would have to occur to be regarded as so large a departure from historic trends, in the above context, as to be more than could be laid reasonably to "change" meteorological fluctuations. It is of passing interest to note that just such an exceptional case has occurred naturally, namely, in the unusually wet winter of 1941, during which the Natural Bridge winter precipitation departed from the linear regression value by just slightly over two standard errors of estimate. This 1941 cases serves as a warning of the real difficulties of separating natural from man-made effects when cloud seeding is performed. Had 1941 actually been a year when the Natural Bridge area was seeded, a real effect would have appeared to be present with no more than the naturally occurring amounts that did fall in that very unusual year.

Other applications of Table 17 are obvious enough to deserve no further comment.

Table 17

Approximate half-widths of the 95 per cent confidence region symmetric with the regression lines of Figures 9-11. All values are in inches of precipitation, and are actually equal to twice the computed standard errors of estimate applicable at the mean value of the independent variable.

Regressions with respect to Tucson

Station	Winter	Summer
Natural Bridge	7.20	6.20
Phoenix	2.82	3.36
Flagstaff	6.08	5.20
Yuma	2.18	2.48

Regressions with respect to Natural Bridge

Station	Winter	Summer
Phoenix	2.70	2.98
Flagstaff	6.44	4.48
Yuma	2.18	2.52

2. Heteroscedasticity. One of the assumptions underlying most applications of regression analysis to problems of prediction, or to problems of making confidence statements, is the assumption that the variance of the dependent variate does not change systematically as one moves from low values of the independent variate to larger values of the independent variate. Bivariate distributions satisfying this condition of essentially uniform variance over the full range of the independent variate are said to be homoscedastic, while those failing to satisfy this condition are said to be heteroscedastic.

In the recent literature of cloud seeding evaluation one finds occasional references (e.g., Jeeves, et al, 1955) to the suspicion that ordinary precipitation data usually gives rise to heteroscedasticity that could prejudice the types of confidence statements normally made in evaluation analyses. Various normalizing transformations have been suggested to suppress this reported difficulty.

Visual inspection of the scattergrams of Figures 9-11 seemed to the writer to indicate only slight heteroscedasticity. There appears to be no standard quantitative test for heteroscedasticity, and, indeed, the statistical literature consulted by the writer gives no clue as to what degree of variation of variance is to be regarded as serious. In order to secure some objective measure of this characteristic for application to the Arizona precipitation data, the following test was devised: The numerical values of the independent variate corresponding to one standard deviation avove and one standard deviation below that variate's mean were computed and marked on the scattergram in question. Then the residuals for all points of the scattergram lying outside these bounds were computed and their squares summed separately for the upper and the lower tails of the scattergram. The corresponding mean-squares of residuals were next computed, and finally the ratio of the larger to the smaller mean-square was computed for each case, to permit use of the F-test for variance ratios.[1] The number of degrees of freedom is ambiguous in this situation, but since only an estimate of heteroscedasticity was being sought, it was decided simply to use the actual number of points of the scattergram entering into each mean-square as the corresponding number of degrees of freedom in making the F-test.

In Table 18, the fourteen mean-square ratios found from the fourteen scattergrams of Figures 9-11 are listed for inspection. Of the fourteen, ten were characterized by a larger mean-square for the upper tail than for the lower tail of the scattergram (as one would expect for most precipitation scattergrams). The remaining four cases are marked by an asterisk (example, regression of Natural Bridge on Tucson in winter). The number of degrees of freedom are not listed, but it is worth noting that there was no evident trend for these numbers to be systematically larger for either of the two tails. Values of F at the 5 per cent level are listed when the observed mean-square exceeds the 5 per cent level of F.

[1] A one-tailed F-test was employed, since the hypothesis to be tested was that the variance is greater for large than for small values of the indepent variable.

Table 18

Test for heteroscedasticity of bivariate distributions of Arizona seasonal precipitation. For each station-pair and season the table gives the ratio of the larger to the smaller means square of residuals from regression for cases lying above one standard deviation from the mean or below one standard deviation below the mean of the independent variable. The values of F, the variance-ratio, corresponding to the 5 per cent confidence level (and in certain cases, the 1 per cent level) are also listed for comparison. See text for further discussion.

A. Regressions with respect to Tucson

	Winter			Summer		
Station	Mean Square Ratio	F 5%	F 1%	Mean Square Ratio	F 5%	F 1%
Natural Bridge	1.9*	4.8	--	2.6*	3.1	--
Phoenix	2.7	2.8	--	3.6*	2.6	3.9
Flagstaff	1.6	4.8	--	1.0	4.1	--
Yuma	4.7	2.9	4.8	2.0	2.5	--

B. Regressions with respect to Natural Bridge

	Winter			Summer		
Station	Mean Square Ratio	F 5%	F 1%	Mean Square Ratio	F 5%	F 1%
Phoenix	7.7	3.0	4.8	12.7	3.3	5.7
Flagstaff	1.4	3.3	--	1.8	3.4	--
Yuma	20.2	2.9	4.8	9.0*	3.3	5.7

*An asterisk after a given mean-square ratio means that that particular case was one in which the larger mean square of residuals was found in the lower rather than in the upper tail of the corresponding scattergram.

Eight of the fourteen mean-square ratios are too small to be regarded as significantly different from unity at even the 5 per cent level, and a total of ten are too small to be significant at the 1 per cent level. All of the four that are significant at the 1 per cent level are regressions on Natural Bridge for the two stations, Phoenix and Yuma. Inspection of their scattergrams reveals that the seasonal precipitation amounts at these stations are so low in dry years, that zero-bound effects suppress the mean-square of the lower tail of the scattergrams in a way that almost inevitably yields heteroscedasticity, though the asterisk on the Yuma-Natural Bridge summer value of Table 18 reveals that in this instance dry years are not chiefly responsible for those particular heteroscedastic tendencies. Furthermore, in the extreme case of the Yuma-Natural Bridge winter distribution, inspection of Figure 11 reveals that the observed mean-square ratio of 20.2 is very heavily weighted by only two quite divergent points in the upper tail. In fact, these two contributed over half of the total sum of squares. And in the Phoenix-Natural Bridge summer case, a single Phoenix summer value (1911) contributed two-thirds of the total sum of squares of the upper tail. Thus, in the very cases where this sample indicates greatest heteroscedasticity, sampling fluctuations appear to have played an unfortunately large role.

Although the test employed here has been an arbitrary one, it seems to the writer to justify the following conclusion: With the qualified exception of the cases involving the very dry stations, Yuma and Phoenix, the present data indicate that heteroscedasticity is not a serious statistical difficulty in the Arizona precipitation data.

This conclusion seems to follow not noly from the majority of the results of the F-tests, but also from the fact that four of the fourteen mean-square ratios refer to cases where the variance of residuals was actually larger for points in the lower than in the upper tail of the scattergrams. In the few reports of precipitation heteroscedasticity which have come to the writer's attention, a suspected general tendency for variances to increase from low to high values of the independent variate was under consideration so the fact that four of fourteen cases in this sample exhibited inverse tendencies seems to argue real caution in accepting these suspicions without further quantitative evidence in support thereof, for the present data refer to samples containing rather long station records in a part of the continent where skewness and other non-normal distribution characteristics are probably better developed than in any other part of North America. It must, on the other hand, be kept in mind that six-month totals were here used, and such totals must show less serious fluctuations of sampling nature than would, say, monthly totals.

If the above conclusion as to the small importance of heteroscedasticity could be shown to be generally applicable in evaluation of seeding experiments and in other uses of precipitation data in climatology it would be comforting, for suppression of conditional variance effects is one of the principal reasons for going to the labor of performing transformations on raw precipitation data preparatory to using them in analyses of variance and in regression analyses. Elemination of the transformation step is certainly desirable if it can be justified, as seems to be the case for the Arizona data exclusive of the extremely arid southwest and west central part of the state. Using other bases stations than the two here employed, one must admit the possibility that the data of Table 16 may benon-representative for Arizona.

3. Estimation of missing data. The third topic to be considered under the general heading of spatial variability of precipitation concerns a recurrent question in climatological research -- that of the estimation of missing data in individual station-records. Completely unbroken climatological records are rare, for a variety of reasons needing no discussion here. In using such discontinuous records for climatological research, the investigator must either accept the occasional breaks or else undertake to estimate the missing data objectively. In some studies, the former course is satisfactory, but in most investigations it is much more desirable to fill the gaps by some systematic procedure and thereby secure solid records with which to deal. It happens to be particularly inconvenient to work with broken records when making use of punchcard analysis techniques, because of the complications in machine programming introduced by random breaks in record continuity, so considerable attention has had to be given to the estimation problem in the early stages of the Institute punchcard program. Here, the problem will be discussed as a problem in spatial variability and some exploratory tests of estimation methods reported briefly.

One finds surprisingly little in the climatological literature to serve as guide and reference in optimizing estimation schemes. Investigators have sometimes felt that it was profitable even to do no more than to insert into any break in a given station's record just that station's mean over a long period of time (a view-point which, by itself, documents the argument of experienced climatologists that almost any kind of an estimate is better than a blank in the time-series under study). Others have suggested using the average of two nearby stations' data for the time-period corresponding to the gap. In one early homogeneity analysis of seasonal precipitation data forming a part of the present study, missing seasonal data were replaced by the corresponding seasonal average over all reported data in the decade centered on the missing season, but it was clear that this was a rather poor approximation. None of the aforementioned schemes is conceptually very satisfactory, and simple checks reveal them to be numerically quite far off (when synthetic tests are run to "estimate" data for stations actually reporting the amounts regarded as "missing" for the purpose of the test).

A seemingly preferable estimation scheme that has been used in a number of Institute investigations already completed or currently underway is most concisely described in symbolic form: Denote by P_{ik} the precipitation reported for the ith time-period (e.g., the ith day of a month, ith month of a year, etc.) at the kth station of a group of climatological stations in a given area, and let $\overline{P}_{ik}$ be the mean value for that kth station and ith time-period over a specified long period of years. (Thus if one is dealing with monthly totals, i = 3 would specify March, and $\overline{P}_{3k}$ would be the long-term mean March precipitation at the kth station.) If the jth station has no report for the ith time period, then an estimate of this missing datum, denoted by (P_{ij}), is given by

$$(P_{ij}) = \frac{1}{n} \sum_{k=1}^{n} \frac{\bar{P}_{ij}}{\bar{P}_{ik}} P_{ik} \qquad (3)$$

where the sum over k from 1 to n and division of the right member by n implies that one uses as the final estimate an average of n individual estimates derived from actually reported values at n stations located near the jth station.

The following general comments about the scheme of (3) may be made:

i) This scheme would seem, on general principles, to be superior to any that gave (P_{ij}) as some function of only the concurrently reported values, P_{ik}; for (3) weights the estimate in proportion to the ration of the long-term-means at the jth and kth stations, and in actual practice one would usually have to deal with station-pairs involving unequal means. The scheme of (3) is, at the same time, clearly superior to mere insertion of $\bar{P}_{ij}$ as a crude estimate of (P_{ij}), for (3) allows any anomalous characteristics of the precipitation peculiar to the particular ith time-period in question to enter through the P_{ik}. Briefly, (3) gives weight both to long-term climatic relationships between the jth station and the predictor-stations and also to the departures from normal characterizing the ith period (as revealed by the departure from unity in the n ratios, P_{ik} $\bar{P}_{ik}$, that appear as factors in the n terms of the sum on the right of (3)).

ii) The scheme symbolized by (3) tacitly assumes that ratios of precipitations rather than sums or differences should be used in estimating missing data. This point is discussed in a general way by Conrad and Pollak (1950, pp. 235-237), who conclude that in the case of precipitation, it is the ratio and not the difference of two stations' receipts that are quasi-constant. Their conclusion rests, apparently, only on empirical evidence and this evidence is not elaborated by those authors. One might question the generality of the ratio-assumption on the grounds that interstation precipitation regression equations frequently contain non-zero intercepts whose magnitude is sometimes so great as to seemingly vitiate the ratio method. Thus, for example, note the regression of Natural Bridge on Tucson for summer precipitation in Figure 9. The N-intercept of 5.14 inches for that particular regression would seem to contradict the ratio assumption very seriously as far as its use in estimating Natural Bridge summer precipitation on the basis of concurrent Tucson reports. However, this difficulty must, in almost all cases, become unimportant when the n stations used in computing the right member of (3) are all chosen from the immediate geographic (climatographic) vicinity of the jth station, for then the several regressions of the jth on the kth stations can reasonably be expected to exhibit zero or nearly zero intercepts. Only in an extreme case would one attempt, that is, to estimate missing Natural Bridge data from reported Tucson data (although in the case of making estimates for early parts of long Arizona records, we have sometimes had to go that far to find a reporting station to use in an estimate).

iii) It is quite essential that all of the means, $\bar{P}_{ij}$ and $\bar{P}_{ik}$, be means computed for the same period of years, for otherwise secular trends in precipitation can very easily introduce systematic error into the estimation.

iv) If any information exists as to patterns of correlation between the jth station and surrounding stations, it is obvious that one should select the n comparison stations from regions of maximum correlation to optimize the estimate in (3), but lacking such information, proximity and topographic factors must serve as deciding considerations. In the Institute program, only the latter have been used thus far, for reasons of lack of detailed information concerning correlation patterns.

Now although the estimation scheme implied in (3) is not explicitly discussed in the literature with which the writer is familiar, he feels confident that this scheme has been used elsewhere.[1] It is, in fact, almost identical with an alternate method using double-mass slopes in place of means discussed in the ASCE Hydrology Handbook (1949), p. 14), and Conrad and Pollak (1950) suggests essentially this scheme. The unfortunate fact is, however, that there appears to be no published information on the degree of reliability of a scheme like (3). This reliability is a function of spatial variability of precipitation, i.e., of the intensity of correlation gradients in the region to which (3) is applied.

It was felt to be highly desirable to obtain some quantitative estimate of the error level to be expected for missing values interpolated by the method of (3) in the Southwest, so a synthetic run was carried out, "estimating" winter and summer seasonal totals for Tucson and Natural Bridge for all years from 1900 to 1929, using a variety of station combinations in sets up to n = 4 in (3). Inasmuch as the temporal and spatial variability of seasonal totals may reasonably be expected, for any region, to be less than that for monthly and more than that for annual totals, the present results constitute useful intermediate yardsticks for

[1] After completing the typescript of this report, the writer came upon a pertinent paper by Keeler (1944). Keeler used the scheme of equation (3), suggested taking n equal to three or more if feasible, stressed azimuthal uniformity in the distribution of these n comparison stations, felt that the means employed need only be over five years but did not stress identity of the averaging period (though he cites this consideration), and gave no figures of merit for the scheme. Keeler cites no previous literature references concerning such a scheme, it may be noted.

estimating precision for all three of these time-periods. On the other hand, until more is known about both Arizona and non-Arizona precipitation correlation gradients, the reader cannot entirely safely judge how applicable the present results are to other areas of the world. (It should be noted that it is only with respect to temporal variability, that the arid zones of the world are known to be unusual. As far as the writer knows, it is only an interesting, but unanswered question whether deserts are characterized by substantially greater spatial variability of precipitation than are humid areas.)

For the purposes of the test of (3), four stations were chosen for use in the series of synthetic estimates for Tucson, and four for the estimates for Natural Bridge.

These stations, and their airline distances from Tucson or Natural Bridge are as follows:

Tucson predictors:	Natural Bridge predictors:
Oracle, 30 miles north-northeast	Prescott, 60 miles west
Benson, 45 miles east-southeast	Flagstaff, 60 miles north-northwest
Phoenix, 110 miles northwest	Ft.Apache[1], 95 miles east-southeast
Bisbee, 85 miles southeast	Phoenix, 70 miles south-southwest

Map locations for these stations may be found in Figure 1, and comparative values of seasonal mean precipitation for the 1931-52 period may be ascertained from Table 10. It may be questioned why the test was applied to predictor-stations as far away as 110 miles when there are somewhat closer predictor-stations that would surely be preferred in actual applications of (3). The answer is twofold: First, it was desired to approximate the least favorable rather than the most favorable situations actually encountered, and second it was desired to apply the test using recent mean-ratios in combination with old reports, since this is the most frequent case encountered in practice. Specifically, in the Institute studies we have had to do the largest amount of estimation of missing data for periods before about 1920. But since the two means appearing in the right member of (3) must be means for exactly the same time period (or at least for periods which are so nearly identical as to rule out distorting effects of secular trends), it becomes almost indispensable to use in the right side of (3) means over some recent period, common to all record spans. Availability of the recently published Supplements to Weather Bureau Bulletin W has made it an obvious choice to use, throughout all of our estimation work, the 22-year means for 1931-52. These 1931-52 means were therefore used in the synthetic test, but were used to "estimate" Tucson and Natural Bridge data for the 30-year period 1901-29 as a fairly realistic test, and in the earlier portions of the period of record, one must often use predictor stations removed by many tens of miles from the station whose missing datum is to be estimated. For this reason, the present test was run with some fairly remote predictors.

---

[1] The 1931-52 mean used in the Ft. Apache estimations was that for Whiteriver, about 5 miles north of Ft. Apache. Homogeneity of these two records has not yet been checked.

Estimates were made for values of n equal to 1, 2, and 4 in (3), in order to gain some notion of how much increase in accuracy might be expected to result from variations in n. Because of gaps in the 1901-29 records for the four predictor-stations, it was not possible to obtain 30 estimates in each case, so for uniformity, the first twenty estimates available, beginning with 1901 were used to evaluate the average error for the present purposes.

The precision of the estimation scheme might be measured in any of a number of ways, of which that chosen for use here was the computation of the average of the absolute values of the twenty percentage errors in each predictor-season category. Use of the errors rather than their squares seems desirable since an occasional highly erratic prediction is weighted by a squaring process in an undue proportion for its seriousness. Use of averages of percentages is generally undesirable, and employment of logarithms of relative errors would have been somewhat better here; but since the individual percentage errors do not vary over an extremely large range, a simple average of percentages was felt to provide a fairly adequate measure of overall error, at least for the present purposes.

In Table 19 the results of the synthetic test are summarized. For each of the two stations (Tucson, Natural Bridge) and for each of the two seasons (defined as before in this report), a total of seven different series of estimates were made. The first four of each set of seven are single-station estimates, the next two are two-station estimates, and the last is a four-station estimate (that is the summation over k in (3) is taken in this last instance from 1 to 4). The percentage errors of which twenty are averaged to give each of the entries of Table 19 were obtained by dividing the difference between the individual estimate and the actually reported value by the actually reported value, or in the symbolism of (3) $(P_{ij}) - P_{ij}$ divided by P_{ij} where P_{ij} is the true Tucson or Natural Bridge seasonal total and (P^{ij}) the corresponding estimate.

Table 19 reveals nothing about the sense of the departure of estimated value from reported value, since only absolute values enter into that tabulation. Since secular trend effects can be different at two stations for two different periods of time, it has to be admitted that making estimates in an early period such as 1901-29 but employing means for a recent period such as 1931-52, introduces possibility of systematic bias, each one of the (fourteen) series of estimates was processed as follows: Using all available estimates for the entire period 1901-29 for each predictor-season combination, the per cent positive signs of departure of estimates from reported values were computed separately for each of the fourteen cases. The results are found in Table 20.

The implications of Tables 19 and 20 may be summarized as follows:

i) Very generally, all of the error-levels of Table 19 are in the neighborhood of 20-30 per cent. That is, it appears that the method of equation (3) must be expected to yield about a twenty-five per cent error for estimated seasonal totals in Arizona, when predictor-stations are taken at distances of several tens of miles.

Table 19

Errors of estimation of missing precipitation data based on a synthetic test for Tucson and Natural Bridge using Equation (3). Tabulated values are the averages of the absolute values of twenty percentage errors for each predictor or combination of predictors. All station-means used were 1931-52 means and estimates are for the first available twenty years out of the thirty-year period 1901-29. Estimates for the winter half-year are listed under the heading "w", and those for summer under "s".

Predictor Stations	Mean Error of Estimate for Tucson (Per cent)		Predictor Stations	Mean Error of Estimate for Natural Bridge (Per cent)	
	w	s		w	s
Oracle	23	45	Prescott	22	23
Benson	29	28	Flagstaff	43	35
Bisbee	21	16	Ft. Apache	25	27
Phoenix	26	36	Phoenix	18	34
Oracle, Benson	18	25	Prescott, Flagstaff	31	27
Phoenix, Bisbee	19	23	Ft. Apache, Phoenix	17	27
Oracle, Benson Phoenix, Bisbee	18	20	Prescott, Flagstaff Ft. Apache, Phoenix	20	21

Table 20

Errors of estimation of missing precipitation data based on a synthetic test for Tucson and Natural Bridge using Equation (3). Tabulated values are the percentages of all estimation errors within each predictor-season category exhibiting positive signs. Data same as for Table 19 except that all (rather than merely the first twenty) available cases in the thirty-year period 1901-29 were used in computing each percentage.

Predictor Stations	Percent of Tucson Errors with Positive Signs		Predictor Stations	Percent of Natural Bridge Errors with Positive Signs	
	w	s		w	s
Oracle	50	44	Prescott	63	59
Benson	32	36	Flagstaff	67	60
Bisbee	59	52	Ft. Apache	41	29
Phoenix	30	37	Phoenix	30	30
Oracle, Benson	32	35	Prescott, Flagstaff	66	62
Phoenix, Bisbee	37	59	Ft. Apache, Phoenix	30	29
Oracle, Benson Phoenix, Bisbee	35	45	Prescott, Flagstaff Ft. Apache, Phoenix	54	41

ii) An improvement results, in general, as one increases the number of stations used, i.e., as n is increased, but this trend is seen to be a bit erratic, and the gain is not extremely impressive.

iii) Distance between predictor-station and estimated station is not here correlated in any very strong way with error of estimation. Thus, for Tucson winter estimates, Bisbee is the best single-station estimator, though its distance from Tucson is twice that for Benson and three times that for Oracle. Similarly Phoenix is the best estimator of Natural Bridge winter data though it is slightly farther from Natural Bridge than either Prescott or Flagstaff. The latter case is additionally surprising inasmuch as Phoenix is well out of the mountainous area within which Natural Bridge, Prescott, and Flagstaff all lie. Noting the correlations for Natural Bridge in winter as listed in Table 14, one sees that Phoenix's 0.85 exceeds that for Flagstaff, 0.57, but is slightly smaller than 0.87 for Prescott. One can only say that distance and topographic similarity do not emerge as decisive clues to selection of optimal predictors in the present problem.

iv) In general, as would be expected from winter-summer correlation differences cited earlier in this report, one sees that estimation is a bit more accurate in winter than in summer. Thus, 10 of the 14 predictors of Table 19 show better estimates in winter than in summer. Again, however, the differences are not uniformly large even in the cases where winter is better.

v) The distribution of algebraic signs of estimation errors, as revealed by Table 20, presents a somewhat confusing picture. One can at most say that no uniform trends emerge, with the partial exception of the Tucson winter estimates, for which there is a noticeable tendency towards a deficiency of positive errors. It would seem that no very serious bias has been introduced into the present sample of estimates as a consequence of treating the practically important situation wherein one estimates <u>old</u> missing data using <u>recent</u> means; but beyond that observation, it seems difficult to generalize from Table 20.

On the basis of these results of the empirical test of the estimation scheme of equation (3), it is concluded that <u>two stations should be used as predictors</u>, but that the added labor of increasing n in (3) is scarcely justified. It is further concluded that, if these two can be selected from within a circle of radius under 100 miles centered on the station with a missing datum to be estimated, the error of estimate may be expected to be <u>in the neighborhood of 25 per cent</u> on the average.

In most practical applications in studies of precipitation in the Southwest (except for the very earliest periods of record when stations were few and far between), the above requirements can be met, and on the other hand, a 25 per cent error in a single missing seasonal total is sufficiently small (especially as compared with typical station coefficients of variation), that decided investigational advantage is gained by carrying out this type of estimate. A two-station application of (3) has already been extensively used in Institute studies, so the abovefindings constitute a posteriori justification for the previous applications and serve to warrant continuation of use of this method in future studies.

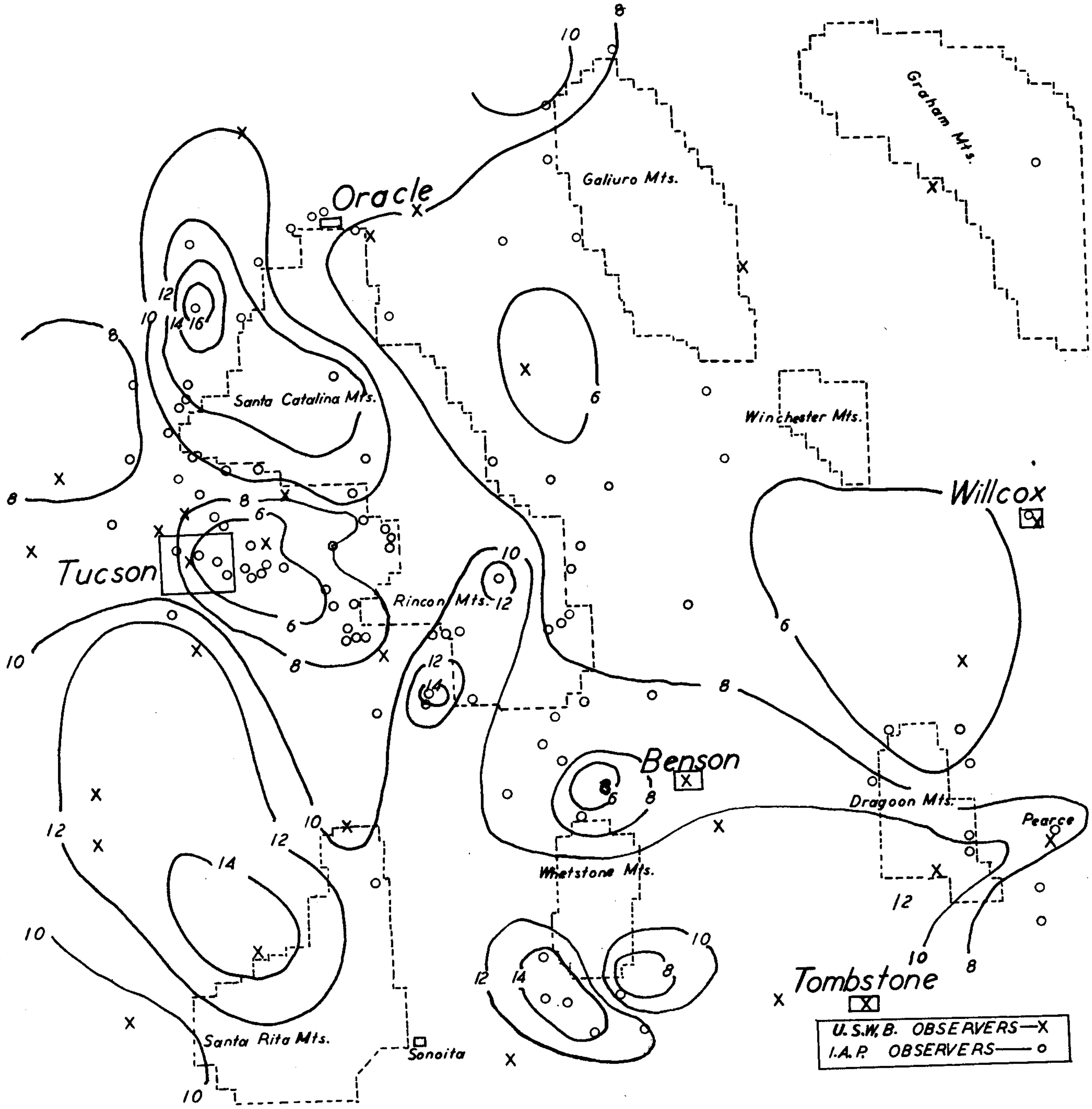
Finally, it may be suggested that one can deduce a reasonable estimate of percentage errors of estimate of monthly and of annual totals by increasing the 25 per cent value in proportion to the ratio of the monthly (or annual) to seasonal coefficients of variation (i.e., quotients of standard deviations divided by means), when independent information concerning these statistics is at hand. Up to the present no study of monthly coefficients of variation for Arizona precipitation has been made, but annual coefficients for a sample of stations have been reported earlier in this report. For eight long-record (over 50-year) Arizona stations, the average coefficient of variation of annual precipitation was only three-fourths as large as that for the same stations' seasonal coefficients of variation. This ratio must be recognized to be a very rough figure, for the winter coefficients are systematically higher than summer, and coefficients of variation are not quantities which should be averaged in so crude a fashion. Nevertheless, that information may be taken as an indication that the present estimation scheme, using two predictor stations in the manner here employed, should give relative errors of estimation of only about three-fourths of twenty-five per cent or just under twenty per cent.

C. Miscellaneous spatial variability data. In Figures 12 through 16 are presented miscellaneous isohyetal charts depicting spatial variability observed in certain specific time periods in Arizona. They are included primarily for convenience of internal Institute reference and so will not be discussed in any detail.

Figures 12, 13, and 14 show precipitation totals for the region south and east of Tucson for three different periods. Part of the data are taken from Weather Bureau substations, but a larger portion are taken from a network of cooperative stations set up by the Institute for detailed studies in the area of radar coverage in southeastern Arizona.[1] It may be noted that the summer of 1955 was an outstandingly wet summer in this area, that January and February of 1956 were just a bit above normal, and that the summer of 1956 was near normal to slightly under normal.

Figures 15 and 16 are copies of only rough-draft isohyetal patterns for two storm situations of January, 1955. Hourly precipitation data were totalled at each Weather Bureau hourly recording station in Arizona for the time-periods indicated, (chosen to fully span single storm passages across Arizona) the totals plotted, and storm-total isohyets drawn. The very complex pattern exhibited by individual storm totals is familiar to meteorologists, but lay appreciation of this point and its strong bearing on cloud seeding evaluation has been very inadequate.

[1] These figures were prepared by C. H. Reitan from his compilations of Weather Bureau and Institute raingauge data.



PRECIPITATION TOTALS: JULY, AUG., 1955

Figure 12. Isohyets of total precipitation for July and August, 1955 in southeastern Arizona based on combined observations of Weather Bureau substations and of cooperative observers in the Institute's special raingauge network.

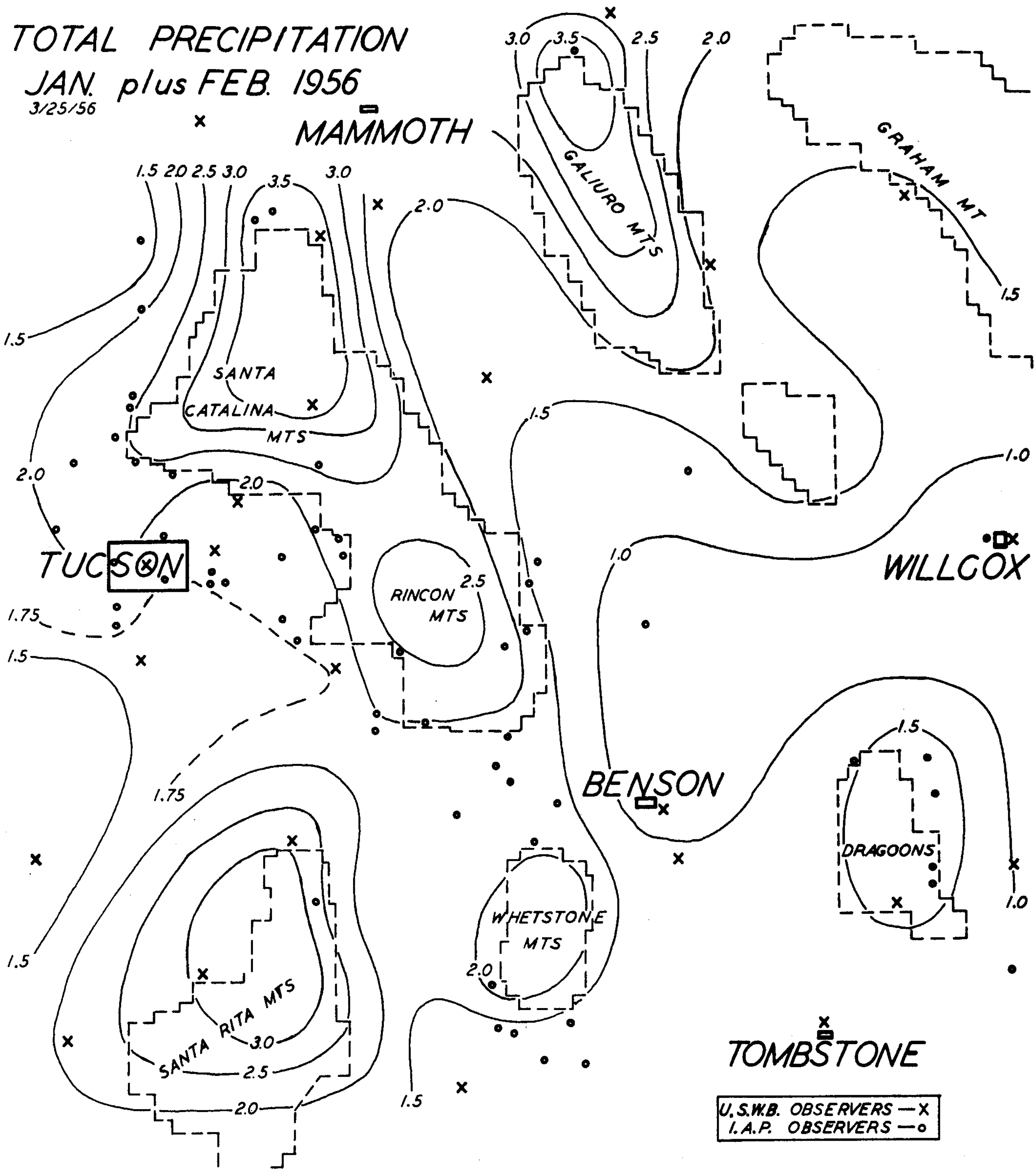


Figure 13. Isohyets of total precipitation for January and February, 1956. Data sources as for Figure 12.

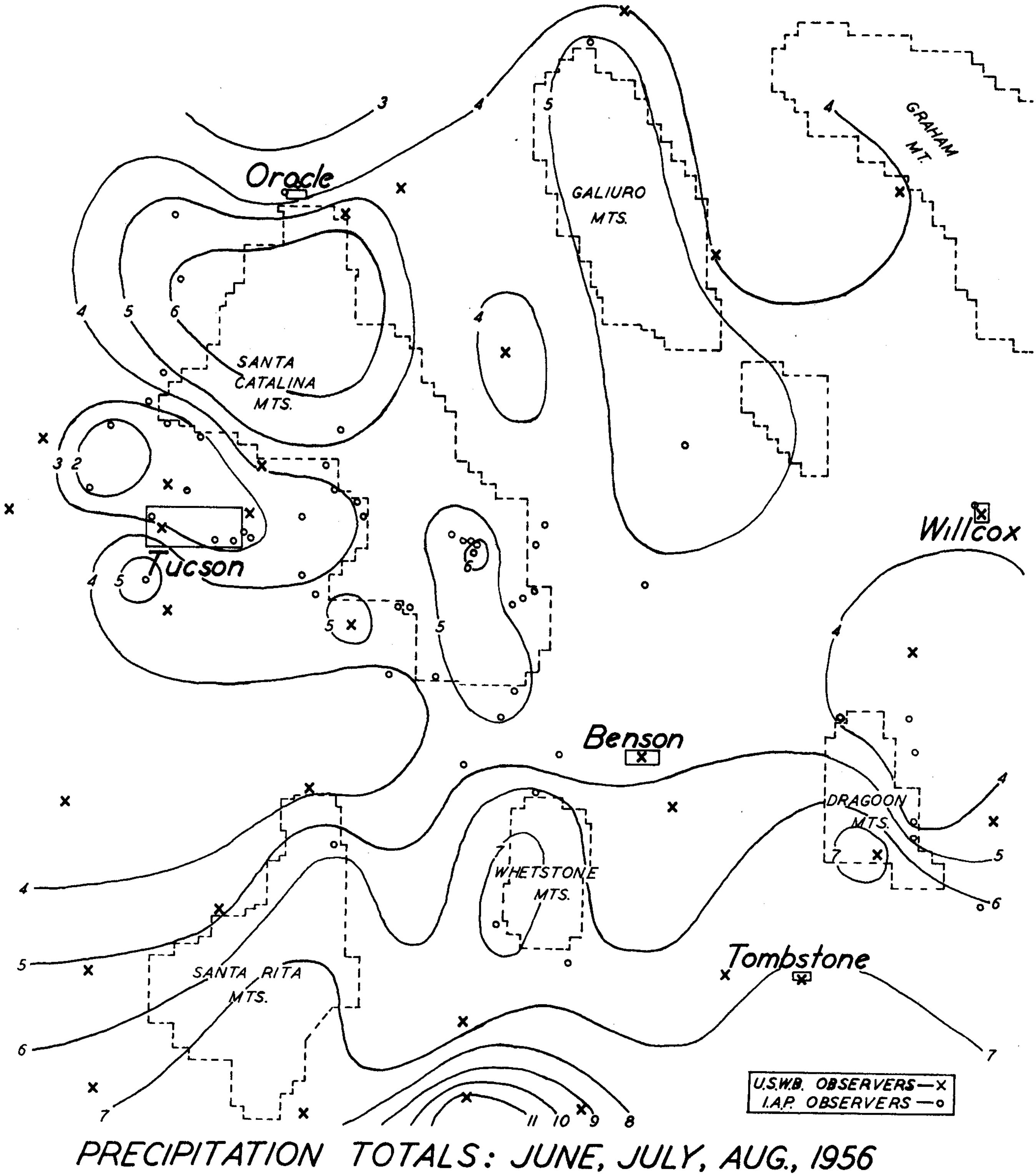


Figure 14. Isohyets of total precipitation for June, July, and August, 1956. Data sources as for Figure 12.

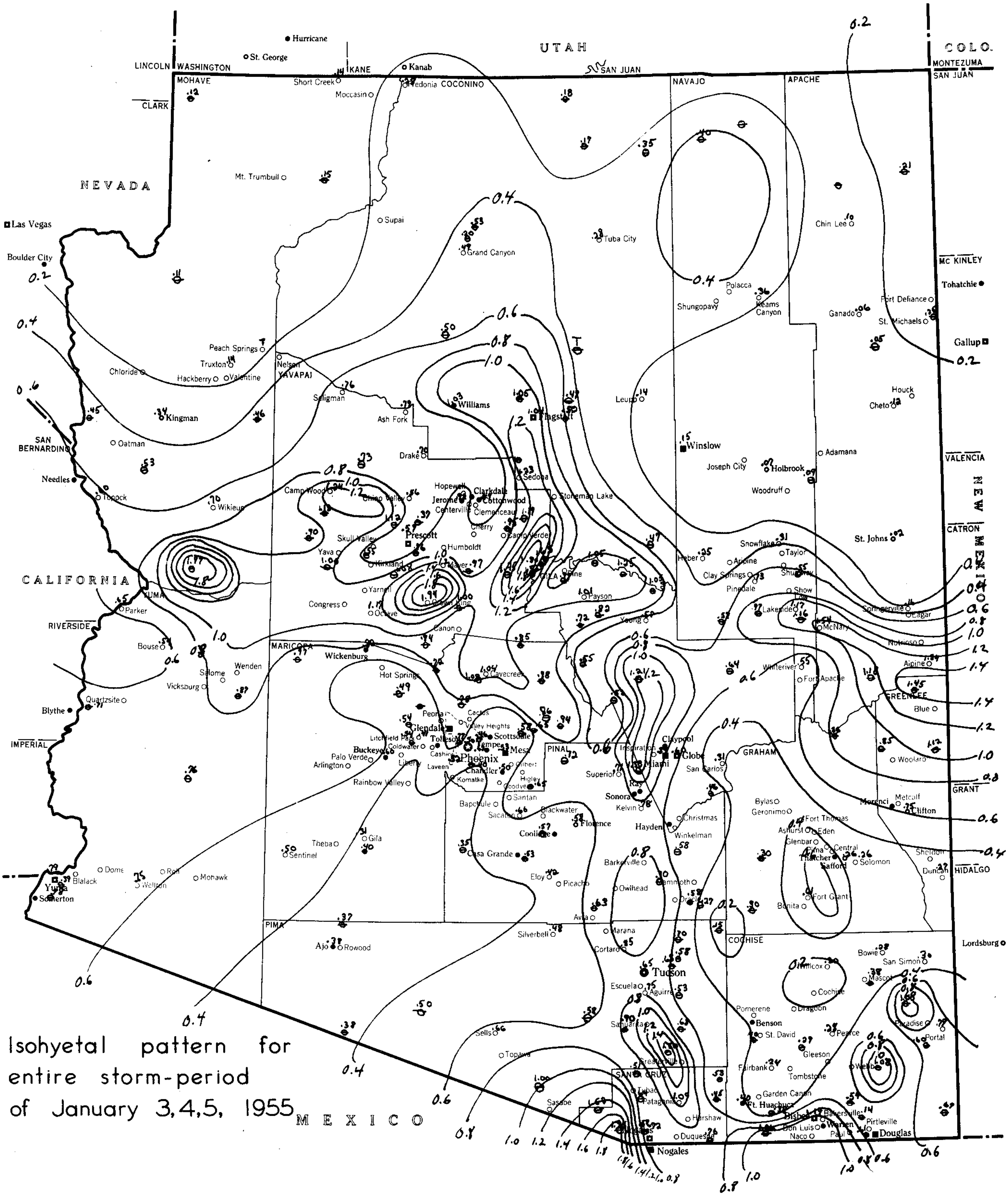


Figure 15. Storm total isohyets for a winter storm (January 3-5, 1955). Rough-draft copy for reference use.

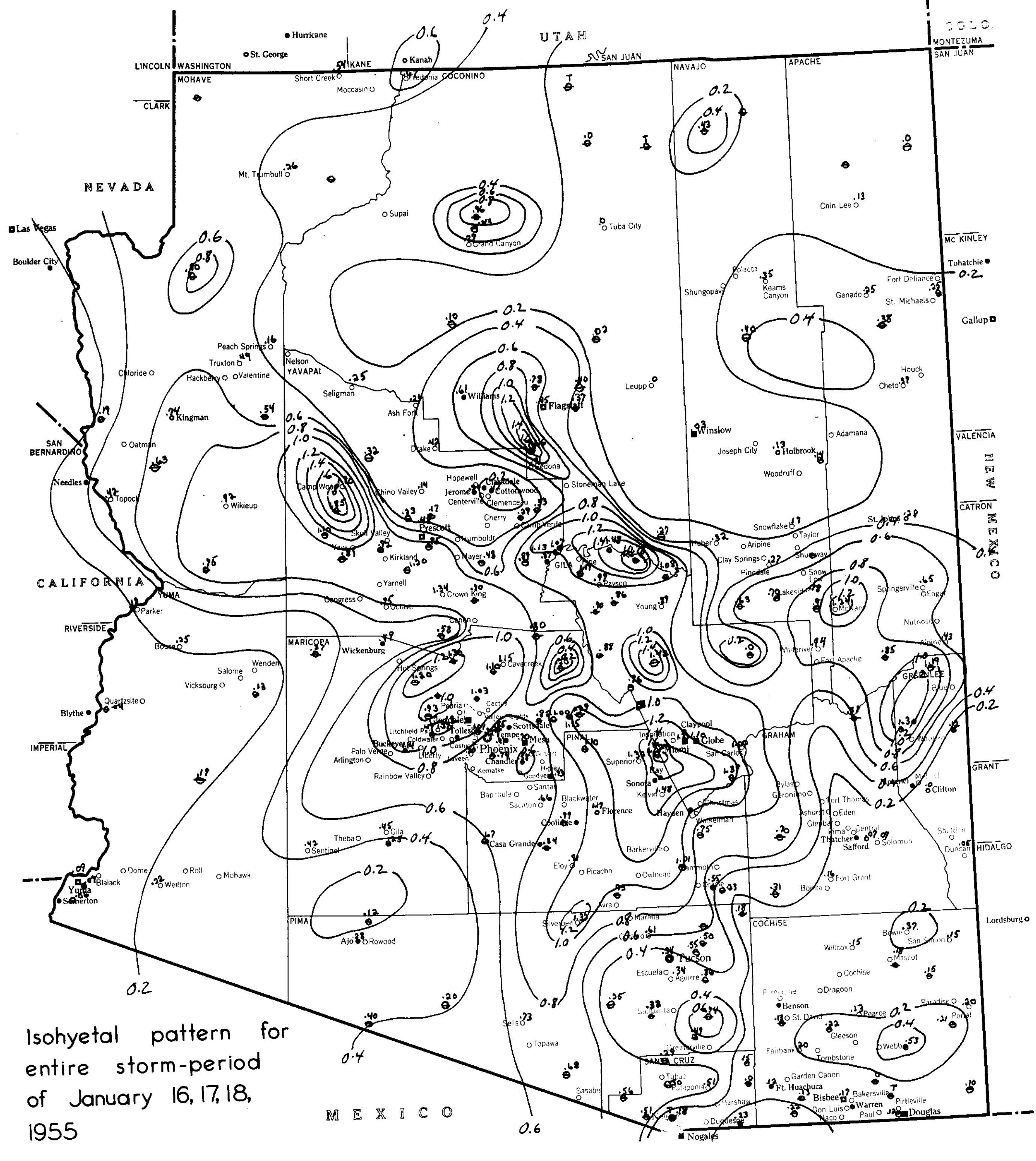


Figure 16. Storm total isohyets for a winter storm (January 16-18, 1955). Rough-draft copy for reference use.

IV. SUMMARY.

This report has summarized a number of pilot studies carried out as one preliminary phase of an extended Institute program of investigation of variability of precipitation for Arizona and for arid region in general.

Necessity for these pilot studies arose as a part of the planning of the Institute's punchcard climatological program, a program established by the University of Arizona in cooperation with the U. S. Weather Bureau. Optimum use of the punchcard data and automatic analysis resources of this program demanded that an initial exploration of certain Arizona climatological problems and testing of certain statistical and computational techniques be conducted prior to the start of the main analysis phase of the Institute punchcard program. Because a number of the findings of the survey are of general interest they are presented in the form of this report.

Only Arizona data were considered in detail here, and in fact for most studies reported here consideration was given only to a relatively small sample of Weather Bureau stations selected on the basis of unusual length of record and representativeness for distinct geographic areas of Arizona. In a few cases, much larger samples have been examined for special purposes. The series of problems considered fell naturally into the two categories of temporal variability and spatial variability of precipitation.

From an examination of seasonal and annual coefficients of variation it was concluded that there is a surprising difference in temporal variability of winter and summer Arizona precipitation, the wintercyclonic precipitation coefficients being higher (by statistically significant amounts) than the coefficients for summer thunderstorm precipitation for all but the driest portions of the state. Despite greater temporal winter variability, it was found from correlation analyses that the winters are characterized by significantly lower spatial variability in Arizona, an unexpected combination of conditions.

Computation of the ninety-five per cent confidence half-widths of seasonal and annual precipitation means has made clear how large an uncertainty must be recognized as existing even in the means of stations that have many decades of past records. Roughly, these half-widths amount to fully ten per cent of the mean themselves for the sampled long-record stations. The difficulties which these large half-widths imply for evaluation of any cloud-modification experiments in Arizona and for certain research studies are noted.

Secular trends in Arizona precipitation, studied through the use of 10-year moving-average plots, revealed a widespread pattern of general decline of precipitation since about the 1920's, but also revealed differences between stations that demand further careful study. Resolution of total variance into components associated with secular trend and with year-to-year flucuations indicated that the latter is definitely the dominating component, yet the decadal means have clearly varied through amplitudes of very great ecologic and economic significance.

Tendencies for wet winters to be followed by wet summers, and other similar persistence tendencies, were examined by use of the autocorrelation technique. It was concluded that it is quite out of the question to anticipate one season's (or one year's) precipitation anomalies from those of the preceding season (or year). On the other hand, some small but statistically significant relationships between precipitation of successive months in winter were found (strongest of which was a tendency towards opposition of anomalies between January and February), but no significant relation was found for the summer months of July and August, contrary to expectations.

About a hundred Weather Bureau records for the 22-year period 1931-52 were used to study patterns of seasonal distribution of Arizona precipitation, a study that shed some light on the important question of climatic provinces within the state. Long-term stability of these seasonal patterns was examined for a small sample of records, and the existence of substantial historic oscillations in seasonal distribution was established.

Diurnal variability of both winter and summer precipitation in southeastern Arizona were investigated, partly for purposes of planning field operations and partly for basic research purposes. No significant diurnal variation was found for January and February, but the anticipated marked variation of summer precipitation was found. A striking limitation on wintertime precipitation research is implied in the finding that, on the average, there are only about eight station-hours of active precipitation available for daytime study at a given locality during a typical January or February in southeastern Arizona. It was found that precipitation intensities of the July-August rainy season have averaged almost exactly twice those of the January-February period. For both winter and summer, correlation analysis suggests that amounts of precipitation measured at the ground have been rather closely proportional to corresponding number of hours of occurrence of precipitation, seemingly implying a relatively stable intensity of precipitation over the years.

Interstation correlations of seasonal and annual precipitation amounts not only revealed that winter precipitation is much more spatially homogeneous in Arizona than is summer precipitation, but also shed some light on climatic gradients within the state. However, it became clear that much more analysis of this problem is clearly needed. Fourteen sample interstation regressions were examined, and associated standard errors of estimate were determined for use as indicators of predictability in considering problems such as cloud modification and in estimating missing data. A test of the heteroscedasticity of these bivariate distributions was carried out and the conclusion was reached that, with the exception of the data for the driest parts of Arizona, heteroscedasticity is not a serious methodological difficulty. This conclusion is somewhat weakened by lack of any standard statistical criterion for seriousness of a given degree of heteroscedasticity, but a seemingly reasonable criterion was proposed and used in this study.

Necessity of estimating missing precipitation data in historic records prior to beginning certain analyses led to an objective test of one estimation scheme. Average errors of estimate of about twenty-five per cent for seasonal totals were found characteristic in situations typical of those actually encountered in practice.

The several findings of this pilot study are being used in planning the next stages of Institute investigation of variability of precipitation. A brief summary of specific problems that will be examined in the near future is appended.

V. SUGGESTIONS AND PLANS FOR FURTHER RESEARCH.

A. Effects of non-normality. Because non-normality of frequency distributions of precipitation introduces uncertainties into most of the standard statistical manipulations of precipitation data, a number of studies of non-normality effects is needed. Some are now underway and will be reported in the near future. The practical question is: Does non-normality really distort the sampling characteristics of the data in a way that is significant compared to the many other inherent weaknesses of the data and/or in a way that is significant compared to the kind of precision one is ultimately trying to attain in a given investigation. Even empirical answers to these questions would be helpful to climatologists, so work is now in progress in the Institute to examine effects of the conventional normalizing transformations on statistics such as the correlation coefficient and the variance, and on heteroscedasticity.

B. Additional variability data. The sample of Arizona seasonal coefficients of variation examined in the preceding report was large enough to reveal certain important trends in Arizona variability, but many questions were raised that can only now be answered by working with a very much larger sample. Consequently, cards are already being punched to permit study, by IBM methods, of the variability of some 70 Arizona stations. Furthermore, inorder to see these data against the background of similar data for more humid areas as well as for other portions of the arid West, monthly and annual total values for the full period of record are being placed on cards for about 75 non-Arizona stations in the Southwest, and for about 150 other U. S. stations for just the period 1905-55. From these cards, seasonal variability data will be computed and charts of these statistics for the entire U. S. prepared. In addition, detailed studies for the whole Southwest will become possible.

C. Skewness. Since skewness of frequency distributions is one of the non-normality characteristics of interest in the general area of item A above, it should be so treated in future studies. It is, however, a point of general climatological interest to gave available data on the geographic distribution of skewness for the entire U. S., so the decks of cards being prepared for the variability studies will be utilized to investigate skewness on an extensive basis. As for coefficients of variation, so also for skewness, charts of the normal skewness values as a function of mean precipitation will be prepared for ready comparison of individual area with norms, in order the more readily to discern anomalies.

D. Secular trends. Differences in trends from one part of Arizona to another, and even between fairly nearby stations, have come to light in the present pilot study. Inadequate homogeneity checks were made. Both of these difficulties must be removed by extensive analysis. Work is in progress on both of these matters. In addition, comparative secular trend data for other parts of the West are being examined in order to understand the relation of Arizona fluctuations to those of other parts of the continent.

E. Homogeneity. Secular trend studies are inseparable from homogeneity studies. A technique to replace graphical double-mass analysis by nearly equivalent IBM analysis is under study. If effective, the Arizona and other Southwest data will be checked for homogeneity by this method. If not, other means must be found.

F. Variability as related to runoff. The winter precipitation of the Southwest is the chief source of runoff in irrigation projects, yet the present study has shown the winter precipitation to be substantially more variable from year to year than is the summer precipitation. The detailed relationships involved need far more study than they have been given. Some first efforts in this direction are now being made with reference to the Salt River Basin. Secular trends in rainfall-runoff relations deserve much more thorough statistical study and such study is also in progress.

G. Persistence tendencies. The question of whether autocorrelations between successive seasons or years justifies any lay beliefs in the predictability of precipitation on such basis has been unequivocally and negatively settled. However, several weak but climatologically interesting persistence trends have come to light here that deserve further scrutiny in the future. No effort is being directed towards this goal at present, but the problem should be taken up again.

H. Interstation correlation analysis. The least adequate portion of the pilot study reported here has been the much too small sample of correlation data. It was not possible to see clearly any geographic patterns of correlation from just the data for eight stations considered here. Once the 70 Arizona seasonal-total decks are punched, extensive correlation analysis should be begun, for the question of what areas of the state are climatically homogeneous is a recurrent one in all of the present Institute precipitation studies. Some corollary studies now in progress may reveal that the product-moment technique can be replaced by the rank-difference approach, which is much faster when punchcard methods are used.

ACKNOWLEDGMENTS.

The support of the National Science Foundation is gratefully acknowledged. The aid of the U. S. Weather Bureau in innumerable ways has been indispensable; for not only has all of the present study utilized Weather Bureau data, but also individual Bureau personnel have repeatedly lent their assistance most generously.

A number of Institute staff members and students have assisted in various parts of the calculations underlying this report. They include Patricia Morgan, J. J. Riley, and R. M. Mitchell. Mr. R. B. Des Jardins set up and carried out almost all of the punchcard correlation calculations used here. Patricia Morgan has drawn almost all of the figures; Marilyn Mauldin provided several. Ruth Young, Erdean McVay, and Anna Seiler have assisted in various stages of preparation of the typescript. I am indebted to Dr. Reid A. Bryson for many helpful comments. Mr. C. H. Reitan has made available for inclusion in this report several isohyetal maps for southeastern Arizona.

REFERENCES.

Anderson, C. A., 1956: The water supply of the Salt River Basin. In Recovering rainfall, Arizona Watershed Program, 1-66.

A.S.C.E. hydrology handbook, 1949: American Society of Civil Engineers, Manuals of Engineering Practice, No. 28. New York, Am. Soc. Civ. Eng., 184 pp.

Bryson, R. A., and W. P. Lowry, 1955: Synoptic climatology of the Arizona summer precipitation singularity. Bull. Am. Met. Soc., 36, 329-339.

Conrad, V., and L. W. Pollak, 1950: Methods in climatology (2nd ed.), Cambridge, Harvard Univ. Press, 459 pp.

Dorroh, J. H., Jr., 1946: Certain hydrologic and climatic characteristics of the Southwest. Albuquerque, Univ. of New Mexico Publications in Engineering, No. 1, U. of New Mexico Press, 64 pp.

Foster, E. E., 1948: Rainfall and runoff. New York, Macmillan, 487 pp.

Hazen, A., 1916: Variation in annual rainfall, Eng. News, 75, 4-5.

Hiatt, W. E., 1955: The analysis of precipitation data. In The physical and economic foundation of natural resources, Vol. IV, Subsurface facilities of water management and patterns of supply -- Type area studies, Interior and Insular Affairs Committee. Washington, House of Representatives U. S. Congress, 186-206.

Jeeves, T. A., L. LeCam, J. Heyman, and E. L. Scott, 1955: On the methodology of evaluating cloud seeding operations. In Weather modification operations in California, Bull. No. 16, Sacramento, California, State Water Resources Board, 89-90.

Keeler, K. F., 1944: Some uses and deficiencies of climatological data. Trans. Am. Geophys. Un., 25, 420-438.

Kohler, M. A., 1949: Double-mass analysis for testing the consistency of records and making required adjustments. Bull. Am. Met. Soc., 30, 188-189.

Kraus, E. B., 1955: Secular changes of tropical rainfall regimes. Quart. J. Roy. Met. Soc., 81, 198-210.

Landsberg, H., 1951: Statistical investigations into the climatology of rainfall in Oahu (T. H.). In On the rainfall of Hawaii, Met. Monogr., 1, No. 3, 7-23.

Ostle, B., 1954: Statistics in research. Ames, Iowa, The Iowa State Coll. Press, 487 pp.

Merriam, C. F., 1937: A comprehensive study of the rainfall on the Susquehannah Valley. Trans. Am. Geophys. Un., 18, 471-476.

Mills, F. C., 1955: Statistical methods. New York, Henry Holt, 842 pp.

Schulman, E., 1956: Dendroclimatic changes in semiarid America. Univ. of Arizona Press, 142 pp.

Schneider-Carius, K., and O. Essenwanger, 1955: Analyse von Niederschlagswahrscheinlichkeiten ausgewaehlter Stationen Vorderindiens. Geofis. Pura Appl., 30, 205-221.

Souceck, E., and J. W. Howe, 1938: A study of the variability of precipitation. Trans. Am. Geophys. Un., 19, 520-525.

Wallen, C. C., 1955: Some characteristics of precipitation in Mexico. Geograf. Ann., 37, 51-85.