D L Cheng 206

Precipitation Analysis for Hydrologic Modeling

National Symposium sponsored by
the Precipitation Committee
of the AGU Section of Hydrology

June 26–28, 1975 Davis, California

AN OVERVIEW OF THE PRECIPITATION PROCESSING SYSTEM AT THE SOUTHWEST WATERSHED RESEARCH CENTER [1/]

D. L. Chery, Jr. and R. S. Kagan [2/]

SWWRC PRECIPITATION DATA FILE

Introduction

The Southwest Watershed Research Center (SWWRC), USDA, ARS, has amassed an extensive file of precipitation data for several semiarid areas of the Southwestern United States. This Center maintains hydrologic records begun in 1939 by the Soil Conservation Service on small watersheds near Safford, Arizona and Albuquerque, New Mexico. Two major research watersheds were established by the Agricultural Research Service in 1954: one of 67.0 square miles on Alamogordo Creek near Santa Rosa, New Mexico and one of 57.6 square miles on Walnut Gulch that has the town of Tombstone, Arizona within its boundary. On both watersheds the rain gage network has been expanded until now the Walnut Gulch Watershed has 98 gages and the Alamogorodo Creek Watershed has 69 gages. Another minor location of watersheds was established in 1966 near Fort Stanton, New Mexico. With the 30 rain gages on three minor watershed locations, a total of 197 gages contribute records to the SWWRC data file.

The digital reduction of the accumulated analog recordings from these rain gages has produced over 757,000 records of breakpoint and informative rainfall data for use by the ARS researcher. Data presently being processed will add an estimated 250,000 records to the file, exceeding a total of one million points. The data are increasing at a rate of 85,000 records annually. These breakpoint records are augmented by identifications, codes, and an extensive set of documentation describing storm characteristics, instrumentation and processing procedures that cannot be numerically coded.

The gathering and processing of precipitation data is part of a comprehensive research program to (1) study water yield of semiarid rangeland watersheds in the Southwest in relation to conservation measures and forage production, (2) determine optimum utilization of water yield for local and downstream uses, and (3) obtain information needed for planning and designing measures to control flash flood and sediment damage.

Understandably, more data than just precipitation is collected to accomplish the general research program. Records of stream flow are the major complementary data file. In addition, vegetation surveys, soil surveys, infiltration tests, and meterological, geologic, geomorphic, and soil moisture data have been collected for each of the

---

1/ Contribution of USDA, ARS, Western Region.

2/ Research Hydraulic Engineer and Mathematician, USDA, ARS, Southwest Watershed Research Center, 442 E. 7th St., Tucson, Arizona, 85705.

watershed locations. These data files complement the precipitation data, giving the researchers a total data bank that not only quantitatively describes hydrologic processes, but also their effects on the environment.

The SWWRC Precipitation Processing Scheme

The rainfall charts collected at the Arizona and New Mexico locations are sent to the SWWRC for coding, digitizing and final processing. Precision of the instruments and accuracy of the methods employed are assessed periodically to determine the accuracies of the data. Care is taken to maintain a data bank that is free from error and consistent in form and content.

While maintaining such a file, the Center realizes the importance of a data processing system that is not only efficient but also flexible and able to implement up-to-date computer methods and technology. Many times, such flexibility is restricted by the volume and continuity of the data.

The flow of data in the current SWWRC processing system is shown in Figure 1. The operation is dynamic; modifications are constantly being implemented to increase the basic efficiency and respond to new processing demands.

Field Instrumentation

Networks of rain gages are used to measure precipitation; however, it is the operation and precision of each gage that determines the quality of the data. On the SWWRC networks, rainfall is sampled by a weighing rain gage with an unshielded orifice 8 inches in diameter located about 36 inches above the ground surface. Each rain gage produces an analog (ink line on a paper chart) record of accumulated rainfall versus time. Generally, rain gage charts are changed weekly. The clock time, as read from a wristwatch, is noted on the charts and a tick is made with the recording pen. Any additional information which might be useful in processing and interpreting the record is usually noted.

Gages having two different depth measuring scales (one inch on chart equals 1 or 0.333 inch of rain) and four different time scales (1 inch on chart equals 31.3, 62.6, 125.2 or 1001.7 minutes) contribute recordings to the data file. Each gage/clock combination has a different resolution of the depth and time measurement.

In addition to the performance of each gage, the entire network of gages must also be considered as a group. There is no common time base among the gages. The time of each clock is set approximately once a week with reference to a technician's wristwatch. Because of the time resolution of the 125.2 min./in. charts and the use of the wristwatch, time at one gage at any instance will be, at best, within $\pm$ 5 minutes with any other gage in the network. Gages can be in phase or up to 10 minutes out of phase.

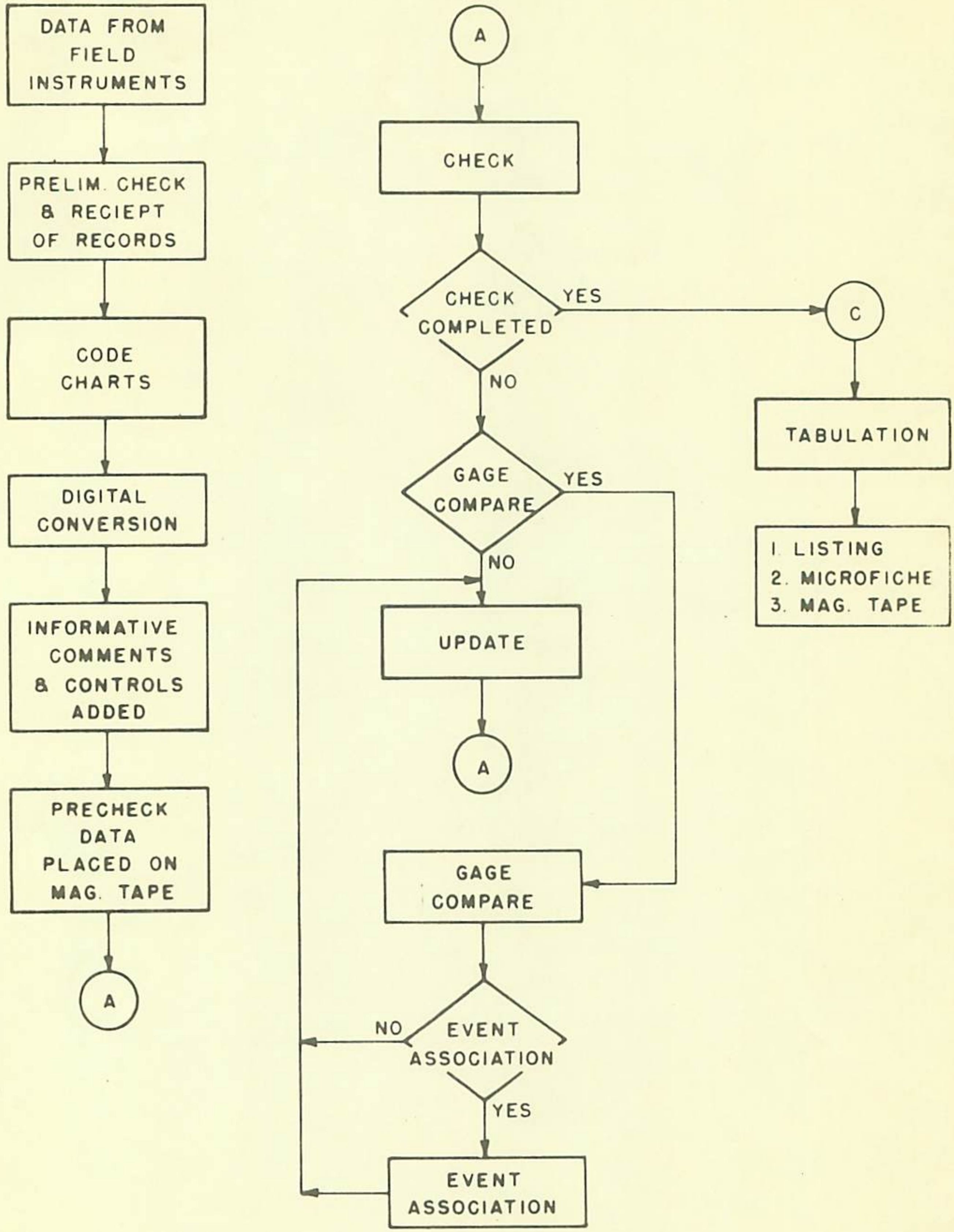

Figure 1. Processing Scheme for SWWRC Precipitation Data.

Before 1968, occasional checks of rain gage accuracy were made by volumetric measurments of the water remaining in the bucket of the gage. The measurement was written on the chart, and consistent differences between the reading on the chart and the measured amount were used to determine a correction factor. Since 1968, each gage has been checked and adjusted annually for weighing accuracy with a set of standard weights (823.00 $\pm$ 0.01 gram equals 1 inch of water).

Record Processing

The charts and records received by the Tucson office are checked for continuity and completeness, and a receipt is sent to the submitting field location. The precipitation charts are placed in a queue for coding. In coding the charts, a technician ascertains the date, beginning time and classification codes of each precipitation event. Precipitation events are classified as significant or nonsignificant. A significant event causes runoff through any of the flow-measuring structures or has total rainfall of 0.25 inch with intensities greater than 0.50 inch per hour at any gage. This determination requires that all the records of a single event be examined as a group and the runoff records be consulted. The chart ON and OFF times are checked, and necessary time corrections are distributed linearly throughout the recording period.

For missing records, the coder makes estimates (discussed further on). The charts are checked for skewed traces (a full vertical trace of the recorder pen that does not coincide with the time reference lines on the charts). These charts are separated for special processing that compensates for deviation of the pen trace.

The coded precipitation charts are digitized by an analog-to-digital converter coupled with a card punch. The operator must first enter the coded information (date, begin time of event, type of precipitation, significance of event) and then break the pen traces into appropriate line segments that accurately describe the event. This procedure is crucial, since it significantly affects intensities computed later in the data reduction process. The operator is also responsible for scaling the charts and encoding estimate codes which may fall under a series of five classifications ranging from Known Event (estimated record) to Gap-In-Record. Since a majority of these classifications are already coded on the chart, a coder is able to reduce digitizing time by minimizing the tasks the machine operator does.

At this point, the data begins a series of documentations and error scans that transform the raw data into a master file for final processing. Comments on gage and event characteristics are punched on cards and merged with the digitized data on magnetic tape. These comment records make note of time changes, gage starts and stops, and processes involved in recording a particular event.

As the master file tape is generated, records are checked for proper characters and format. An error listing is printed and data processing personnel correct the tape using updating programs.

The main CHECK program then scans the file for inter-record errors, such as mis-sequenced or logically redundant records. The generated listing flags possible errors, such as time intervals that may be less than the precision of the field instrumentation or possible contradiction of event codings. There are also provisions to double-check the logical organization of the precipitation data as defined by a set of header cards.

There are two updating programs. For quick record deletions and insertions, a simple correction program is used. For more complicated corrections, a larger, more flexible program is used.

When these two types of errors are corrected, a computer program is used to graphically represent each event by printing event summaries at the approximate coordinate of every gage. This printed map is used to spot inconsistancies that result from miscoded dates, begin-times, significance or type of precipitation. Apparent errors are checked against the original charts and corrected with the updating programs.

For small watershed locations, the data undergo one final error-scan program (Event Association) that correlates rainfall with its associated runoff. Errors involving time inconsistencies and estimates can be isolated and the appropriate correction made.

The formal error detection process is completed at this point, and the data is considered ready for tabulation. However, there is one last component of the correction process: detection of errors through use of the data file. Notes of these errors are recorded in a central precipitation error file and periodically these corrections are made and the basic data file is updated.

The final phase of the data reduction is the tabulation of the master file by the program TABULAT. This tabulation performs all the necessary depth corrections and calculates summary data as well as breakpoint quantities such as intensity and incremental volumes. For significant events, first moments and maximum depths for selected time intervals are also calculated. This information is stored on magnetic tape in an format suitable for both visual and computer retrieval. Formal headings and labels are added by this program to event and gage commentary. To facilitate computer retrieval of the data, special line codes are generated. The organization of these tabulation tapes is similar to the basic data tapes, with each year's data sequenced by gage number, each file containing only one year. The tape is copied directly to computer print-out and stored at the Southwest Watershed Research Center, with back-up files stored at other ARS locations.

With the creation of the tabulation tapes and their listing on computer paper or microfiche, the data reduction process is complete.

Data Quality

With the great amount of automatic processing involved in reducing the field measurements, accuracies of the data are sometimes overlooked. However, for any data use, whether computerized or not, the quality of the initial measurements and final tabulations are extremely important. For this reason, the SWWRC processing system has been designed so that such descriptions of quality can be associated with the data. The quality of the precipitation data is discussed in the following paragraphs.

<u>Resolution</u>. Present rain gage, chart scales and ideal reading resolution for Location 63 are presented in Table 1. Theoretically, the analog-to-digital processing equipment should read 1 count in 1000 plus or minus 1 count (as is indicated in Table 1), but the actual resolution may be plus or minus 2 or 3 counts. Further, such physical limitations as ink line width, chart clip bulges, parallax and operator positioning of charts with respect to the reference lines, positioning of the reader cross-hairs and the confounding of a curved time line on charts all contribute to reducing the practical resolution by one-half or one-third. These remarks reveal nothing about the sampling properties of the gages nor their precision.

One consequence of the daily rain gage accuracy and processing resolution is the nonmeasurement of many small events (less than depths of 0.05 inch). W. D. Sellers [1/] called this situation to our attention in a comparison of frequency-depth curves for records from a daily recording gage and records from weekly gages (see Figure 2). Also, an investigation of an inductance transducer mounted on a standard weighing rain gage "showed a definite tendency of the electronic system to record more sensitively the beginning of an event. For these events, the electrical system measures small amounts of rain occurring up to ten minutes before the major amount fell. These amounts could not be detected on the mechanical recording." [2/]

<u>Amount of Estimation</u>. Because of a multitude of mechanical and human failings, portions of rainfall events are never measured. These missing measurements are estimated in an attempt to provide uninterrupted sequences of data.

Estimations are made from an abbreviated isohyetal plot of the total rainfall measured by gages surrounding the gage with the missing record. Notes or marks on the chart are also used to assist the estimation. Only the beginning time, storm duration and total depth are estimated. The extent of these estimates can be judged by the tally (Table 2) of percent of total storm duration and total storm depth that has been estimated. Those gages with estimation quantities of 10 percent or more were counted and noted in the last column of Table 2.

Table 2 has been separated into two parts. The first part summarizes the years 1955 through 1966, in which only a fraction of the

---

1/ Research Professor, Institute of Atmospheric Physics, University of Arizona, Tucson, Arizona, Personal Communication.

2/ Payne, E. L. and D. L. Chery, Jr., "Inductance rain gage transducer mounting, recording system and evaluation of records obtained in 1969," File Report Location 63.000, USDA-ARS-SWWRC, January 197

TABLE 1. Processing resolution of rain gage charts from Walnut Gulch, Location 63.

No. of Gages	Depth Scale 1 inch of chart = XX in. of rain	Ideal Processing Resolution (in.)	Practical Processing Resolution (in.)	Time Scale 1 inch of chart = XX minutes	Ideal Processing Resolution (min.)	Practical Processing Resolution (min.)
38	0.333	0.003 ± .003	0.006 ± .012	125.22	1.25 ± 1.25	2.5 ± 5
4	0.333	0.003 ± .003	0.006 ± .012	31.30	0.31 ± 0.31	0.6 ± 1.2
4	1.000	0.01 ± 0.01	0.02 ± .04	1001.74	10.0 ± 10	20. ± 40
52	1.000	0.01 ± 0.01	0.02 ± .04	125.22	1.25 ± 1.25	2.5 ± 5

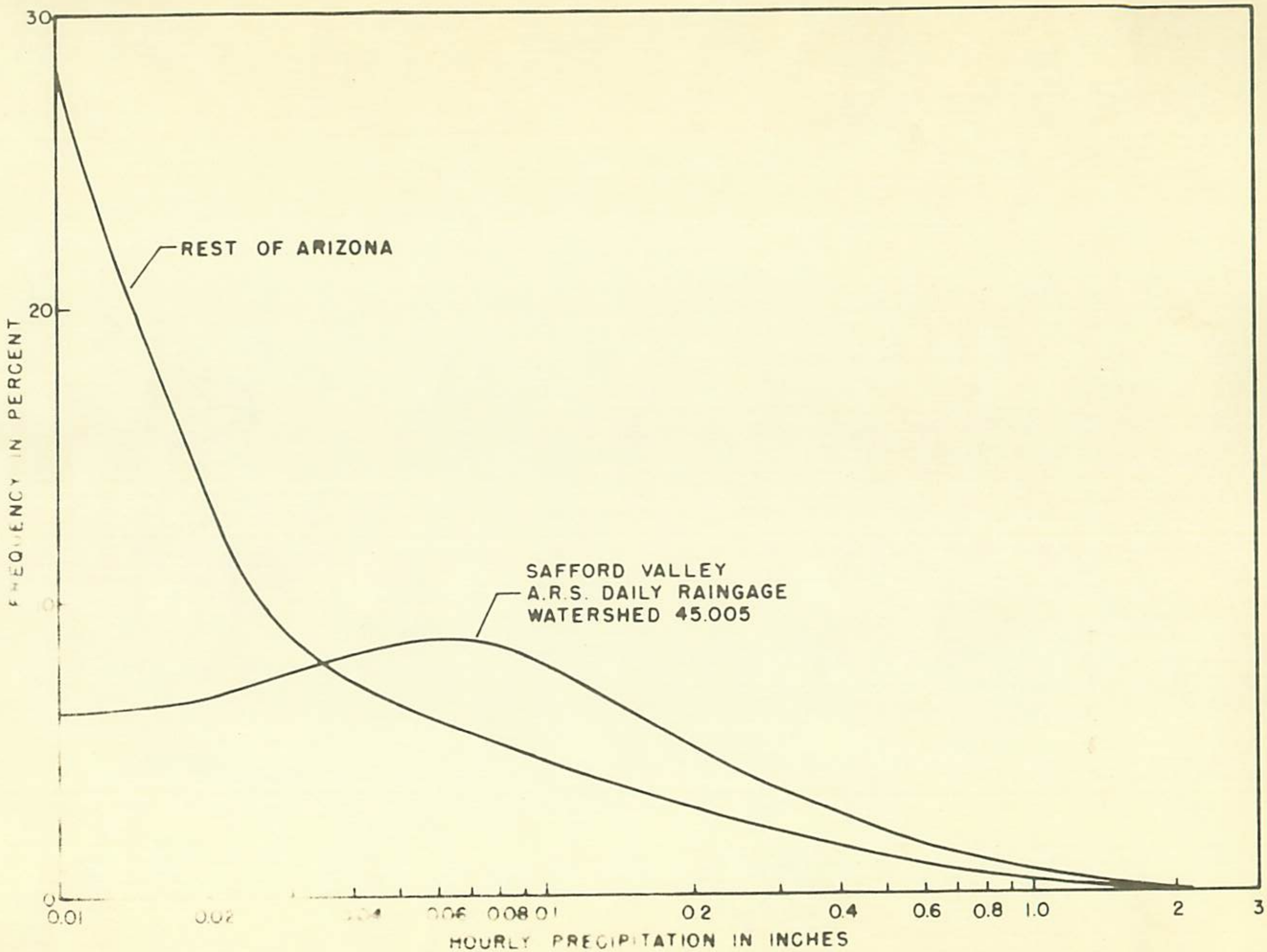

Figure 2. July-August Hourly Precipitation Frequency Distributions 1946-1965
From Dr. Sellers University of Arizona.

TABLE 2. Amount of estimation in Location 63 Precipitation Data.

Year	No. of Gages Tot. No.	No. of Gages No. Full Yr.	Duration % Est. Aver.	Duration % Est. Std. Dev.	Depth % Est. Aver.	Depth % Est. Std. Dev.	No. of Gages with 10% or over est.
Period for which partial amount of estimation is noted in records.							
1955	44	24	11.57	23.88	8.18	16.38	12
1956	44	42	9.61	11.97	8.50	12.61	18
1957	44	41	6.34	12.09	4.77	9.31	9
1958	43	43	5.26	6.49	5.74	7.76	12
1959	46	43	7.26	10.11	5.80	6.42	12
1960	60	46	11.85	18.03	7.98	12.87	25
1961	67	55	8.49	12.13	4.81	7.41	23
1962	73	63	11.12	13.39	7.71	9.31	33
1963	83	73	3.69	6.06	2.95	5.44	17
1964	84	79	3.15	5.45	3.10	5.45	13
1965	85	82	4.55	9.77	3.16	6.49	11
1966	92	85	9.60	7.90	2.64	3.96	45
Period of complete estimate tally in record.							
1967	94	94	8.82	11.43	5.09	6.93	32
1968	98	92	3.37	5.25	2.40	3.95	14
1969	98	98	8.08	7.04	5.15	5.01	36
1970	99	94	12.57	9.77	5.59	6.37	61
1971	98	94	7.82	6.35	3.14	3.97	36
1972	98	94	7.28	7.43	4.13	5.36	29

estimates have been noted. The second part shows the tally for years 1967 through 1972, in which all estimates were noted as the data were processed. As Table 2 indicates, there is a considerable amount of estimation in the record. Data for 1968 had, on the average, the least amount of estimation with 3.37% of the duration and 2.40% of the depth estimated. These amounts may represent about the least amount of estimation that can be expected with the type of manual/mechanical rainfall recording system now being operated. With this type of system there is simply an irreducible amount of opportunity for human error or mechanical malfunction of the old equipment.

DATA COLLECTION AND PROCESSING DESIGN

Data Needs

Fundamentally, the research objectives or investigation of the professional staff dictate the kinds of information collected. Many researchers are directly involved in or one step removed (through their supportive staff) from the measurements being made and data collected. In other types of research, the professional is at the end of an extensive funnel of measurement, collection and data processing. The data needs of the hydrologic researcher are of this latter type.

The hydrologist using precipitation data for hydrologic modeling must depend on information collected over large areas and long periods of time. The researcher, not able to span these dimensions himself, must rely on some organization to collect the information for him. As these dimensions increase, especially the individual point measurements in space, this organization becomes increasingly complex. The researcher also becomes dependent upon the reliability and accuracy of the system. Thus, the processed data must contain internal descriptions of measurement and processing methods and accuracies. The collection and processing systems need to be constructed to maintain established data standards or strive to meet some self-imposed standards.

In addition, these processing systems need to maintain a facility to accommodate changing requests for information. The tendency is toward more measurements, more accuracy and additional types of measurements. The system needs to accommodate corrections to the data as well as restructing of its components while maintaining the data bank's continuity and flexibility.

Data Processing System Design

From past experience, the Center has seen how the processing systems dictate the form and accessibility of the data they produce. This system has produced a large volume of card and paper output, with all event data stored on a series of magnetic tapes for further computer use. For precipitation data, the format and organization of the magnetic tapes has remained constant. Such consistency is essential for any data bank but may restrict future adjustment to changes in instrumentations or techniques. For this reason, system design must be considered. The processing system design can either facilitate or drastically curtail modifications and maintainance.

Use of computer processing often does not begin with an overall design, but begins instead out of necessity. Before 1960, time and depth values on rain gage charts were read manually and recorded on tabular paper. This task became overwhelming, requiring alternate methods of tabulating the precipitation data. With the acquisition of an analog-to-digital converter in the early 60's, charts were digitized directly to computer cards and the Center began its first use of computers to tabulate precipitation data.

The processing system soon expanded with the addition of computer programs to check the digitized data. Emphasis shifted from cards to magnetic tape as a more concise means of storing data. Programs were written to allow the processing staff to correct errors on magnetic tape and to use tape as the main input media for tabulation programs. An intensive period of computer program development was undertaken between 1964-1969 in which overall processing procedures were automated. The efficiency of the processing increased at such a pace that, by 1969, all pervious data years had been tabulated and the operation was able to process records from the current year.

Two specific designs have been very beneficial to this operation: the 'modularization' of tasks and computer programs and a unifying organization.

'Modularization' entails the isolation of tasks into units that a specialized program or operator can handle. For example, the digitizing of chart data is a specialized operation with the equipment operator responsible only for that particular task. Most of the programs are also arranged in units. The tabulation programs do not check for errors and the checking programs, in turn, do not reduce the data.

'Modularization' increases the efficiency of the process and permits easy maintenance, modification or replacement without causing major repercusions outside the unit.

Current improvements in the field of data acquisition technology illustrate the flexibility of a modular system. Our original system digitized chart data to computer cards which accumulated over the years becoming difficult to manage. Digitizers are now available to convert chart records directly to magnetic tape, and the Center is currently implementing such equipment without having to totally revamp the system. The changeover consists of retraining operators for the new digitizer and using an intermediate computer program to convert the magnetic tape to a compatible format for the series of correcting and tabulating programs already in use. The major portion of the data processing system remains intact.

Since many of the programs are written in modular units (statements grouped at a specific location to handle a specific process), these programs can be modified to handle new formats should an interim program be phased out.

Severe difficulties are encountered when programs are not arranged in process-oriented modules. For example, the major updating program was not written in this form. When it was modified to run with a new computer operating system, modifications in one section had serious effects on other sections. Such effects are difficult and time-consuming to trace and emphasize the importance of organizing programs to facilitate future maintenance and modification.

At SWWRC, the need for improvement in our processing design is recognized. Each step of the data reduction process must be self-contained, clearly performing one task without overlapping into the next one. Each program in turn must be separated into distinct components. Experience has shown that it is desireable to have input and output independent of the main body of the program. Each calculation unit should be separate, with its entry points and conditions clearly documented. Such an arrangement allows future programmers to revise individual sections without affecting the rest of the program or the ability to reorder units.

To make 'modularization' function properly, centralized information and good management are needed. When programs and tasks are specialized, especially when personnel are responsible for only one processing step, a unifying factor is mandatory to prevent operations from becoming disjoint, confused and inefficient.

Each step of the process must be fully documented. Many decisions that are made early in the processing are later programmed. Thus, it is important that all personnel have access to information about each step of the process. Efforts now are being made to centrally locate information pertinent to all aspects of data processing. Such a process cannot flow properly unless there is an overall organization--organization of structure as well as organization of information.

Past experience has shown that data processing has gained the most efficiency when there was management directly aware of and responsible for all steps of the process. This kind of supervision is especially important in detecting weak areas in the data reduction and establishing a clear line of communications. Overall supervision must also be balanced by the individual expertise in the various areas. Many times, an open ear to suggestion has improved an operation more than any high-level policy. Although the individual employee may be responsible only for his or her given task, there must be an overall feeling of organization and communication so that a sense of importance is established for each step of the process.

The acquisition, reduction and maintenance of precipitation data is an involved operation, especially with the quantity of data involved. Experience at the SWWRC has shown the importance of evaluating the quality of the data and designing a modular, but centralized processing system. With processes and organization clearly defined, the whole structure of precipitation data processing can work in a consistent and integrated fashion, producing a very valuable data resource for the hydrologic researcher.