

MECHANIZATION OF THE PROCESSING OF CLIMATOLOGICAL DATA

By

Robert F. Dale^{1/}

The processing of all kinds of data by punched cards is not new, but the rapidly growing electronic computer field has been so much in the news the last few years that when faced with a large tabulation or computing jobs most organizations give some consideration to the use of mechanical methods for their accomplishment. I shall not try to speculate on why I was invited to give a paper on the mechanical processing of data before this group since there are many more qualified machine people available. In fact, at the California Snow Surveys meeting at Yosemite National Park in January, I learned that Mr. Irving Ingerson, in charge of the Engineering and Data Services Branch of the California Department of Water Resources, was giving serious consideration to an elaborate mechanical hydrologic data processing program using the electronic computers, and that several in his Branch had taken courses in the various models and makes of computers attempting to select the equipment best suited to their needs. Looking over the Proceedings of the April 1955 Western Snow Conference, I found excellent possibilities for using punched card methods in computing the precipitation probability presented in Francis Blanchard's "Operational Economy Through Applied Hydrology,"^{2/} and later, also at the January Snow Survey meeting, learned that Mr. Blanchard, East Bay Municipal Utility District, has at least a nodding acquaintance with mechanical processing from a correlation problem a few years ago. The California Forest and Range Experiment Station is placing data from the Snow Laboratory on punched cards to facilitate computation. Thus a discussion of mechanical data processing here is something like carrying owls to Athens, but perhaps there are some with less experience than I who might be interested in a few of Weather Bureau's experiences in the data processing field.

Although there was early interest in the United States in the possibilities of using punched card equipment for weather statistics, almost from its development for the 1890 national census, no funds were available for the testing of machine techniques in climatology until the advent of the "make work" projects of the middle thirties.^{3/} In 1934, a PWA project was initiated to prepare a long-needed atlas of ocean climates, and of the 5 1/2 million observations taken from 1880 to 1933 more than 2 million were punched into cards for machine sorting and summarization. In 1936 work was started on the punched card compilation and analysis of surface and upper air observations made at about 400 airway weather stations, and by 1942 20 million airway observations taken from 1928 to 1941 had been placed on punched cards. These cards were summarized to obtain the Airway Meteorological Atlas, Normal Flying Weather for the United States, Temperature Frequencies in the Upper Air, and Low Visibility Airport Wind Rose Summaries. The boost given climatology to answer the many questions which arose during World War II is well known. The Weather Bureau's WPA punched card project was placed at the disposal of the armed forces and, working under Army and Navy funds, together with installations of the Army Air Force, by the end of the war the joint New Orleans machine unit contained about 80 million punched cards which had been used in solving many strategic and tactical problems.

Impressed by the mechanized catalyst for the rapid development of new kinds of statistics and climatological methods, the joint New Orleans Machine Unit was continued after the war, and a program of current punching of weather data from all weather stations was initiated, that in the western United States beginning in July 1948.

Both surface and upper air weather data were placed immediately on 80 column punched cards at commissioned Weather Bureau stations, Regional Weather Records Processing Centers were established to place the observational data from substations on cards, and also to check, process, and publish the climatological observations from all stations. Navy and Air Force weather station data were placed on punched cards currently at New Orleans. Although there have been the inevitable changes

1/ State Climatologist, U. S. Weather Bureau, San Francisco, California

2/ Blanchard, Francis B., Operational Economy through Applied Hydrology. Proceedings 23rd Annual Meeting, Western Snow Conference, 1955, 35-48.

3/ Machine Methods of Weather Statistics. USAF, USWB, and USN, Fourth Edition, December 1949, New Orleans, Louisiana.

in codes, procedures, and organization with increasing experience in mechanized data processing methods, including the move and expansion of the Joint Machine Unit from New Orleans to Asheville, North Carolina, as the National Weather Records Center, the program to place weather observations on punched cards currently is now well established.

The data processing equipment used in the Weather Bureau ranges from the simplest mechanical card punch at field observational stations to the high speed electronic digital computers used in the Joint Weather Bureau, Air Force, and Navy Numerical Forecasting Unit at Suitland, Maryland, where the computer is already producing successful daily circulation forecasts which are transmitted by facsimile over the United States, including our Forecast Centers at San Francisco and Los Angeles. It may be interesting to note, however, that the machine processing operations at the San Francisco Weather Records Processing Center for the immediate checking, processing, and publication of the climatological data for the 11 western states, Alaska, and Hawaii, are by electro-mechanical type equipment, and that this equipment is adequate for most of the routine operations. The same is true for the other two WRPC's, that at Kansas City, Missouri, servicing the Central States and that at Chattanooga, Tennessee, the Eastern States.

Following the daily punched cards for a climatological substation through the routine machine checking to preparation of the final format for printing the Climatological Data will provide an insight into the first uses of the punched cards at the WRPC. After the observational data have been placed on cards, the cards are first run through the collator which automatically checks the chronological order (days) and compares the maximum and minimum temperatures with the current and previous day's temperature at time of observation, discarding the inconsistent cards which are compared with the original forms for correction. The cards are then run through the calculating punch which punches into "work" columns on each card the departures from the previous day's maximum and minimum temperatures as well as the departures from the previous month's mean maximum and mean minimum. These cards are then "arrayed," not as a true statistical array from which the "array" has evolved, by machine sorting in a geographic code (and also by AM and PM observation stations) and machine tabulating from each daily card the maximum temperature, the departure from the previous month's mean maximum, the change since yesterday, and the same for the minimum temperatures. Precipitation data are also printed with the tabulator wired to print indicators for discrepant combinations of precipitation and type of weather as well as "Duplicate" should two cards have been accidentally punched. These "arrays," presenting a tabular synoptic picture, are professionally reviewed for any marked changes or departures in temperatures atypical of the adjacent stations in the area. Although the temperature difference from neighboring stations may be real, each such case is investigated manually for possible punching or observational error.

The cards are then sorted into chronological order within stations and run through the calculating punch to compute the heating degree days from the daily maximum and minimum temperatures, and this derived value is punched into the same card. The cards are then listed in the tabulator to provide a preliminary monthly record of all data by stations, and a summary card is punched for the monthly totals. The tabulator is also wired to print numerical indicators when discrepant combinations of data are tabulated, and these preliminary daily listings are visually checked against the original observational forms, especially for the precipitation entries. The daily cards are again run through the tabulator for listing extremes and dates with a monthly summary card punched. A run of the monthly totals cards is made through the calculating punch to prepare means. These monthly means cards are collated with the monthly extreme cards and punched into a single monthly summary card. The monthly summary cards are then collated with the respective normal cards, and the departures from normal precipitation and temperature are punched in the summary cards. The completed summary cards are then ready for the final tabulation run directly on the format which is printed by photo offset process as Table 2 in the Climatological Data. The daily temperature and precipitation tables are printed directly on the respective table formats from the daily cards.

We have described the manual side of the operation only in the review of the "arrays" and "preliminary listing," but there is considerably more of this in the WRPC processing operations than indicated above. Suffice it to say that the machines cannot run by themselves, and with increased mechanization we can expect increased maintenance problems and needs for more highly qualified operators. It may have been noted also that no mention was made of verifying machines. These are not used in Weather Bureau operation since the punched card editing processes described above screen out all large errors on review, whether committed by the observer or the card punch operator. The verifying machines at best can detect only those errors in transcription from the written record to the punched card, and it has been demonstrated that only 75 to 80 percent of the errors made in punching are caught by the verifying machines.^{4/} Time is not sufficient to run

^{4/} Bosen, Julius F. Mechanical Computation. World Meteorological Organization Commission for Climatology, Second Session, Washington, January 1957. Chapter IV, Item 8.

through the checking and editing procedures for the hourly observation stations to the final preparation of frequencies of joint ceiling and visibility conditions, wind speed and direction, and various psychrometric data summaries published in the Local Climatological Data Supplement, which is also done at the San Francisco WRPC.

One illustration of the adaptability of the machines and the inventiveness of experienced operators is worthy of note, however. It is the practice of the authors of the narrative summaries in the Climatological Data to prepare rough work maps showing the precipitation and temperature anomalies over the State. Computation of the percent of normal precipitation for all of the California stations took considerable time, and Mr. Johnson, in charge of the machine unit of the WRPC, was asked if he could prepare these percent departures from normal precipitation as a part of one of the routine machine runs. The following month he presented us with the percent of normal monthly precipitation, as well as the regular temperature and precipitation departures and the name of the station tabulated in approximately the correct geographical position on tabulating paper cut out to the shape of the State!

It seems probable that if no further use of the punched cards were made after these briefly described functions of quality control, summarization, and publication, the machine methods would have paid for themselves, but after this first use the punched card is still as good as new and is then available for whatever other problems of a climatological or hydrologic nature may arise. The file of substation punched cards, current since July 1948, is far from adequate for most climatological work, however, such as frequency studies of precipitation intensities and durations. What is being done with long record climatological stations?

The Weather Bureau has not been able to obtain funds for placing historical climatological data on punched cards, but has been encouraging climatological research with the punching of climatological data before 1948 by other agencies when such a project can be shown to be the most efficient way of accomplishing their specific problems or research activities. Most of the cooperative weather data punching projects to date have been with the State Colleges and Universities in support of agricultural research, 23 States presently being engaged in these punching contracts. Many already have prepared climatological bulletins for their States as valuable by-products.

Through the cooperative agreement the Weather Bureau furnishes the blank card stock and makes the climatological records available to the University, as well as helping with the initial editing and getting the project under way. The University furnishes copies of the completed punched cards to the National Weather Records Center, in return for which the University may request duplicates of any cards already available at the NWRC in a one-to-one ratio to those furnished. A few States have also entered a cooperative trust fund agreement providing the Weather Bureau with funds for back log punching at the NWRC or the WRPC of the climatological records desired in its work, also obtaining an equal return of cards from the Weather Bureau. If the University makes any use of the cards this is a very advantageous arrangement, and of course the Weather Bureau profits by obtaining the back log card file which increases the length and, correspondingly, the value of the climatological record on cards.

The flow of punched cards from the WRPC's and cooperating Universities is only a portion of the data accumulating in punched card form at the National Weather Records Center in Asheville, North Carolina. Observations from Air Force and Navy bases over the world, from marine observing programs, from Arctic and Antarctic stations, and from international data exchange programs, all contribute to the present rate of growth of our punched card libraries - 30,000,000 cards per year.

The demands for analysis of this vast store of data have shown even more remarkable growth in both quantity and sophistication. In the 1930's and early 40's, simple sorting, summing, and counting operations were the order of the day. Today the moments of a distribution are computed as a matter of course; and there is more and more demand for such mathematical-statistical techniques as power spectrum analysis, autocorrelation, and computation of orthogonal polynomials and maximum likelihood estimates of extreme values. The demands for mathematical transformation of the data and for statistical treatment come from many quarters - agencies of civil defense, atomic energy, radio communications, aviation, public health, agriculture, air-conditioning and heating, water supply, and all branches of the Defense Department; in fact, every conceivable area in which weather and climate are factors. Of the Weather Bureau's analytical work at Asheville, more than 90 percent is performed for such outside agencies and interests.

The machine facilities at NWRC have multiplied both in number and capability, although they have by no means succeeded in keeping pace with the demand. In 1947, there were 3 type 405

tabulators and a few sorters at New Orleans; in 1957, Asheville is busy with 2-shift operations on 11 Type 407 tabulators; 5 Type 101 Statistical Machines; a host of sorters, collators, reproducers; Type 604 and 607 Electronic Calculators; a Card-Programmed Electronic Calculator; and 3 magnetic-drum electronic computers. Using the same data library, the Air Weather Service uses for its data processing program a giant Type 705 computer.

We have dealt here primarily with cards since this is the system with which we are most familiar, although we realize there is an increasing variety of data processing systems employing other media as data input, such as perforated paper tapes, magnetic tapes, and microfilm. The 80-column card has been adopted as the "common denominator" medium, however, by the World Meteorological Organization, and almost all existing systems use the 80-column card or provide the means to convert this to the necessary media.^{4/}

Surveying the progress, now that we have a file of more than 300 million cards at the NWRC, Jacobs pointed out that the standard forms of punched cards are not suitable for use in numerical weather prediction or in climatological computations where time or space is considered as the variable rather than the element ... that the appetite of the newer forms of computers is more versatile than provided for by the existing card files.^{5/} The variable time spectra required in climatic analyses have also posed problems in machine processing; for example, the hourly and daily forms of climatological data are not always sufficient for some needs and may be in more detail than necessary for others. Jacobs claims that the hourly frequencies of weather elements for our airport stations show that differences in frequencies between one hour and the next are seldom statistically significant, or that summarizing three consecutive hourly observations frequently gives us no more information than could have been obtained from one. In this case the processing load could have been reduced by 2/3. On the other hand the time spectrum required for other studies can range from seconds to periods of days. Studies of wind gusts, temperature and pressure variation, and precipitation intensity are not always adequately defined by 24 equally spaced time intervals during the day, but would best be described by a continuous transcription in a form suitable for automatic processing by equipment such as one of the analogue computers. The next best form would be a digital transcription of significant points. Thus, if not already obvious, the most important problem in the processing plan starts right at the beginning — the development of the basic data forms and adopting the best time spectra for the work which is anticipated for the cards or data input media. If cards are to be used, the design of the cards and observational forms should be in close conformance. The data punched should be in identical order with the data as entered in their originally recorded form to minimize errors and punching time in the first manual phase of the operation. It is significant to note that the Air Force, Navy, and Weather Bureau were able to agree upon common observational forms and punched cards; however, at the WRPC we still utilize a few non-standard forms used by other Federal, State, and local agencies. There are also problems in the storage of the cards. For the conventional climatological work the present method of filing chronologically within stations is the most convenient, but for other studies of a synoptic nature, especially as the punched card files grow in volume, it becomes increasingly more difficult to collate dispersed elements within the NWRC files. This same is true in any data searching imposed upon the NWRC by research agencies, and with widely separated dates and locations the job becomes entirely manual.

This last is just to illustrate that we do not yet have the complete answer ... there will be the inevitable shortcomings and problems arising in any mechanized data processing system, but with increased experience and development of higher type machine processing equipment it is believed the major difficulties will be solved or circumvented. Certainly, the machine methods have proved themselves by the saving of manual labor, the preparation of summaries, in opening new fields of meteorological research and numerical forecasting, and in the development of modern climatology.

^{4/} Bosen, Julius. Mechanical Computation. WMO Commission for Climatology, Washington, January 1957. Chapter IV, Item 8.

^{5/} Jacobs, Woodrow C. Some problems associated with modern requirements for data processing. WMO Commission for Climatology, Second Session, Washington, January 1957.