Adaptation of the IBM Card System to Analysis of Sugar-Beet Varietal Data

GEORGE K. RYSER AND BLISS H. CRANDALL[1]

Introduction

Analyses of sugar beet varietal data require many computations which must be recopied and summarized before the data can be presented in final form. This work is not only time consuming, but the recopying often brings in errors which are difficult to eliminate. The comprehensive and dependable analyses made possible by punched card equipment offer many advantages. Data recorded on a card with permanently punched holes are very flexible. The information can be classified rapidly and transcribed accurately, completely or in part.

A standard IBM card is $3^1/_4$ inches wide by $7^3/_8$ inches long with 80 columns and ten vertical positions in each column for recording of numerical

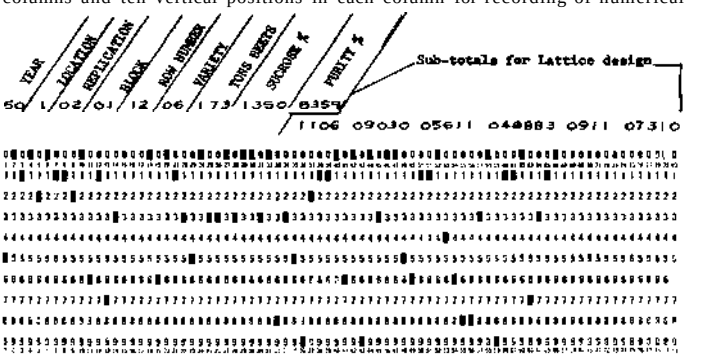

Figure 1.—IBM Card Showing Layout, Interpretation and Typical Punched Card.

information (Figure 1) . For alphabetical information two additional vertical positions above the zero are used and each letter is represented by two holes in each column. The significant use of the punched card has been summarized as follows (1) :[2]

> "Just as punched holes in a roll of paper can produce music, so punched holes in IBM cards can produce reports. Written records can be read only by the eye, but punched card records can be read by machines as well. Such records are sorted and summarized by machines at higher speed and with greater accuracy than would be possible by any manual process. The cards, therefore, once they are punched and verified, are the basis of all printed reports."

[1] Agent, Division of Sugar Plant Investigations, Bureau of Plant Industry, Soils and Agricultral Engineering, Agricultural Research Administration, U. S. Department of Agriculture, and Professor and Director, Statistical Laboratory, Utah Agricultural Experiment Station, Utah State Agricultural College, Logan, Utah, respectively.

[2] Numbers in parentheses refer to literature cited.

To study the feasibility of using the IBM card system for analyzing sugar beet varietal data, the facilities at the Utah State Agricultural College were utilized for data obtained in 1950 and 1951. To save trips between the sugar-beet laboratory at Salt Lake City, and Logan, Utah, some of the cards were punched with machines at the University of Utah. The Service Bureau of the International Business Machines Company at Salt Lake City also assisted with some of the work which was paid for on a "job" basis.

Description of Equipment and General Operation

The IBM equipment of the Utah State Agricultural College includes the following machines: 1. card punch, 2. verifier, 3. interpreter, 4. sorter, 5. tabulator or accounting machine, 6. reproducing summary punch, 7. collator, and 8. multiplier. The first seven were used in connection with the sugar beet data. There would be an advantage in having the new IBM statistical machine which sorts and tabulates at the same time, although this machine has its primary use with enumeration rather than measurement data.

1. The card punch machine enables the operator to record data by means of punched holes in the IBM card. This is similar to a typing operation.

2. With the verifier a different operator repeats the punching operation. If a mistake has been made the verifier does not agree with the punched holes, the machine stops and a correction must be made. Cards should always be verified.

3. The interpreter prints information from the punched holes in either of two positions to the upper margin of the card. The present machine is limited to printing 60 digits on a single line and hence the printing cannot be directly above the 80 respective columns (see Figure 1). There is a new punching machine which punches and prints in one operation with smaller print directly above the punched columns.

4. By means of the sorter the cards can be arranged in any order desired providing suitable code numbers have been punched in the card and dials on the machine can show visual card counts at the same time.

5. The tabulator or accounting machine gives either a listing with totals and sub-totals or a tabulation of totals, depending upon how the control panel is wired and the switches set.

6. The reproducing summary punch machine reproduces all or any portion of the punched data from one group of cards to another. This machine also punches totals or sub-totals to new summary cards as recorded by the IBM tabulator. This punch machine may also transfer information from master cards to detail cards and compare the punched results for accuracy. A special device may be added which makes possible the punching of holes from suitable marks made on the card with a special lead pencil.

7. The collator is used in a number of ways. One use is to intersperse summary cards on top of detail card groups so that totals can be "gang-punched" by the reproducing machine into detail cards. This

operation was found particularly useful for punching sub-totals on detail cards for tabulation of "y" and "b" values (2) for obtaining the "W" values (quantities required for the analysis of variance of lattice designs) .

8. The multiplier performs many operations which are difficult if not impossible with the other machines described. The usual operations, addition, subtraction and multiplication, in any combination of two or more amounts punched in a single card are readily accomplished by this machine with the answer punched in the same card. Division may also be included with the newer machine, whereas with the older ones division may be done by multiplication with reciprocals.

Analysis of Sugar Beet Data

Most of the sugar beet varietal data were first copied on large sheets of ruled paper numbered to 80 columns, corresponding to the 80 columns of the IBM card. In this way each line represented an IBM card or plot. Balanced lattice designs with 25 varieties were analyzed. The punching and tabulating work for one of these designs was completed in approximately three hours with the equipment at the Utah State Agricultural College. This included five categories or fields of information, number of beets after thinning, number beets harvested, tons per acre, percent sucrose and percent purity. The number of columns for the setup required two sets of cards. The lattice design required a listing by variety with variety and final totals, tabulation by replication, tabulation by block within replication, and accumulated totals for each variety, and tabulation by variety with final totals for y and b values. The W values were calculated from the formula (y — rb) (2) . This operation plus calculating all sums of squares required approximately one hour additional work with a conventional calculating machine for each lattice design.

After these tabulations were made for 1951 data some interesting information on sodium content (by use of the Amalgamated Sugar Company's flame spectrophotometer) was obtained for the same varietal material. The IBM Service Bureau at Salt Lake City was able to follow the procedure and tabulate the sodium information. The cost was $18 per hundred cards or a total of $27 for the test involving 150 cards. This included listing each plot of each variety and tabulating the totals together with tabulations for y and b values, as above.

Discussion

The size or the number of experiments controls whether the IBM card system offers an economical method. The sorters can handle 450 to 650 cards per minute and the tabulator 80 to 150 cards per minute. The machines can be adapted to a large number of statistical problems. They can readily be used in the analysis of variance of randomized blocks and Latin-square designs, or for lattice designs, particularly the balanced lattice.

With the use of the IBM machines the biggest problem is to collect the information on the IBM cards. The method which would involve the least copy work would be the best. The most efficient procedure would

be to plan the details from the beginning with punched cards designed for the specific data to be collected. These cards can be pre-punched for each plot and the randomization with restrictions imposed by the design accomplished by the machine, including the detailed printing of the field plan.

Cards for each plot could be prepunched for each type of data to be taken and used in the field for recording with a graphite pencil directly to the card. This method would eliminate the time-consuming transfer of data and the elimination of errors which invariably creep in when data are copied. Multiple cards per plot create no problem because the machines process the cards rapidly and all data can be brought together from separate cards easily, providing suitable control or identification digits are prepunched in the cards.

During the past year in many projects at the Utah Agricultural Experiment Station this system has been used with excellent success. The basic plan is to record raw data on cards at the time they are observed and then all subsequent computations and summaries can be completed by machine. This requires careful planning in advance and often provision must be made for recording data which cannot be anticipated.

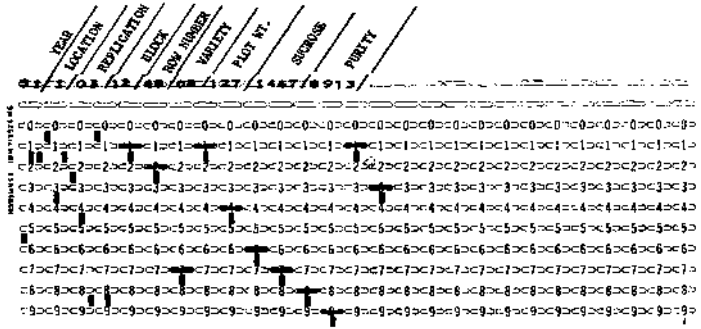

Figure 2.—IBM card showing special graphite pencil marks for plot weight, sucrose and purity. The electrographic pencil marks make it possible for the machine to punch holes corresponding to the penciled data. In the above card the coded information for year, location, replication, block and row number and variety was punched in advance.

If recording directly to cards is not feasible the next best procedure would be to organize the field reports so that the data can be key punched directly from these original records. Once the data have been recorded in the form of punched holes, complete control of accuracy is easily maintained.

Most of the sugar beet varietal data were recopied to the large special sheets of ruled paper with 80 columns corresponding to the 80 columns

on the IBM card. This recopy work, however, requires a good deal of time with corresponding chances for mistakes, but on the other hand it makes the card punching operation simple.

The experimental designs tabulated by Cochran and Cox, pages 304-305 and 360-361 (3), were found to be extremely useful and facilitated the coding of plot data.

The Service Bureau in Salt Lake City was asked for a comparative cost estimate *on* the lattice design problem with 750 cards (five lattice designs of 150 cards each). The price quoted was $8 per hundred cards. As the number of cards increases, the time required per unit is considerably reduced.

Summary

The advantages of the IBM system in analysis of sugar beet varietal data are greater accuracy with a large saving of time as compared with tedious copy work, sorting and hand recording. The method would be more practical if data from similar experimental designs from several locations could be run at the same time. The time involved in setting up the machines is much greater than that of running additional cards.

Literature Cited

(1) Anonymous—IBM Operators Guide
1951. Form 22-8485-0—Copyright by International Business Machines Corporation, New York, N. Y.

(2) GOULDEN, C. H.
1944. A Uniform Method of Analysis for Square Lattice Experiments. Scientific Agriculture, 25: 118-136.

(3) COCHRAN and Cox
1950. Experimental Designs, New York, John Wiley and Sons.