INTERNAL REPORT 62

PROGRESS REPORT ON AUTOMATED DATA ACQUISITION: AN EXAMPLE

ABSTRACT

Biome investigators are well skilled in collecting and recording numbers. The investigators continue through description of the notebook data. Some investigators encounter much trouble in getting the data our of their notebooks and interchanged among people in a usable format. This paper describes an effort to automate the calculations done in a specific project and automatically store the data in the information bank.

INTRODUCTION

The information bank is traditionally called to get some one else's data because the project is completed or someone is leaving town. The data is in most cases: (a) poorly documented, (b) incomplete because insufficient time has been allowed for conversion calculations, (c) sometimes not available. The author must conclude that some effort to further automate data reduction and acquisition must be available. Pearl Knopf was hired on 3 January 1973 as a data processing technician to get as much as possible "working up" of data automated so that data analysis would be finished a reasonable time after the data collection was completed, and as a consequence, the data would automatically be available for others to use.

The critical test project was available in that the aquatic coordinator had requested that the information bank assist Marleen M. Wekell in data analysis. Pearl wrote a preprocessor for a canned regression program so that Marleen only needs to prepare 5 control cards and the data. The rest of the "working up" of the data and submission to the information bank is completed by intercepting the desired estimates in the regression program calculating the desired statistics with their standard deviations. The program also punches the estimated data and the input data with the necessary labels so that the information bank will know where to file the punched deck. The program was implemented on 16 January 1973.

This method allows:

1. fast reanalysis of the data by new methods not yet available;
2. quicker analysis feedback to sampling strategy hopefully within one season;
3. quicker and easier accessibility of the data to model synthesis.

The authors realize that the following disadvantages are associated with the methodology.

1. Many investigators are not acquainted with the computer and feel uneasy using it.
2. A computer based system is technologically more complicated and safeguards must be built into the methodology to ensure safety of the information on the computer.

The information bank has developed the necessary safeguards to protect itself from losses of information.

Calculations

The program uses a

A. Linear Regression:

$$Y = B_o + X * B_1$$

where

$$Y = t/F$$
$$X = A$$

t is incubation time in hours
F is the fraction of the C^{14} utilized
A is (micrograms glucose)/(gram dry weight of sediment)

1. $B_o = (K+S)/Vmax$

K is Michaelis-Menten uptake parameter
S is natural substrate (unknown)
$Vmax$ is maximal velocity for mineralization
$(K+S)/Vmax$ is substrate turnover time

$B_1 = 1/Vmax$

The statistics:

B. Turnover, Standard Deviation of turnover, $Vmax$, and standard deviation of $Vmax$ can be calculated from the regressions equations

turnover = B_o

standard deviation of turnover = $SRS2BO$

$Vmax = 1/B_1$

Standard deviation $Vmax$ = $$-\frac{1}{(B_1)^2}^{2} * (SRS2B1)^2$$

where

SRS2B0 is the standard deviation of the intercept B_0.
SRS2B1 is the standard deviation of the slope B_1.

Program Usage

A. Punch the following cards:

Green Jobcard Col. 1 Col. 31
JØBCARD,T100,CM75000,PO. MARLEEN
REQUEST,TAP,VRN=P 894,DI,IN FILES=30-32.
BEGIN(PEARL, TAP,TAP)
(End of record card.)

B. Punch data in this way:

1st card	Col. 1-3	5-7	8	10-14	16-17	18-19
1st experiment	EXP	Experiment # (last number in column 7)	letter of experiment (if there is a letter)	DATE	YEAR	MONTH

Columns	20-21	25-29	32-36	40-43	46-50	55-59	61-80
	DAY	DEPTH	Meters Depth	TEMP	Degrees	LAKE	Spell name of lake

Data Cards for first experiment	Columns 1-10	11-20	21-30	31-40	41-50
	Value of A	Value of CPM	Value of CPM-B	Time in minutes	Value of F

Columns 51-60

Value of *T/F* if calculated
(can leave this out and program
will calculate it)

7
End of first experiment 8
9

Punch second experiment in same way

7
8
9
Last experiment
6
7
8
9