NAT'L INST. OF STAND & TECH
A11106 977873

NBS
PUBLICATIONS

NBS SPECIAL PUBLICATION 400-75

U.S. DEPARTMENT OF COMMERCE/National Bureau of Standards

Semiconductor Measurement Technology:

A FORTRAN Program for Analysis of Data from Microelectronic Test Structures

QC
100
.U57
400-75
1983
C.2

NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards[1] was established by an act of Congress on March 3, 1901. The Bureau's overall goal is to strengthen and advance the Nation's science and technology and facilitate their effective application for public benefit. To this end, the Bureau conducts research and provides: (1) a basis for the Nation's physical measurement system, (2) scientific and technological services for industry and government, (3) a technical basis for equity in trade, and (4) technical services to promote public safety. The Bureau's technical work is performed by the National Measurement Laboratory, the National Engineering Laboratory, and the Institute for Computer Sciences and Technology.

THE NATIONAL MEASUREMENT LABORATORY provides the national system of physical and chemical and materials measurement; coordinates the system with measurement systems of other nations and furnishes essential services leading to accurate and uniform physical and chemical measurement throughout the Nation's scientific community, industry, and commerce; conducts materials research leading to improved methods of measurement, standards, and data on the properties of materials needed by industry, commerce, educational institutions, and Government; provides advisory and research services to other Government agencies; develops, produces, and distributes Standard Reference Materials; and provides calibration services. The Laboratory consists of the following centers:

Absolute Physical Quantities[2] — Radiation Research — Chemical Physics — Analytical Chemistry — Materials Science

THE NATIONAL ENGINEERING LABORATORY provides technology and technical services to the public and private sectors to address national needs and to solve national problems; conducts research in engineering and applied science in support of these efforts; builds and maintains competence in the necessary disciplines required to carry out this research and technical service; develops engineering data and measurement capabilities; provides engineering measurement traceability services; develops test methods and proposes engineering standards and code changes; develops and proposes new engineering practices; and develops and improves mechanisms to transfer results of its research to the ultimate user. The Laboratory consists of the following centers:

Applied Mathematics — Electronics and Electrical Engineering[2] — Manufacturing Engineering — Building Technology — Fire Research — Chemical Engineering[2]

THE INSTITUTE FOR COMPUTER SCIENCES AND TECHNOLOGY conducts research and provides scientific and technical services to aid Federal agencies in the selection, acquisition, application, and use of computer technology to improve effectiveness and economy in Government operations in accordance with Public Law 89-306 (40 U.S.C. 759), relevant Executive Orders, and other directives; carries out this mission by managing the Federal Information Processing Standards Program, developing Federal ADP standards guidelines, and managing Federal participation in ADP voluntary standardization activities; provides scientific and technological advisory services and assistance to Federal agencies; and provides the technical foundation for computer-related policies of the Federal Government. The Institute consists of the following centers:

Programming Science and Technology — Computer Systems Engineering.

[1]Headquarters and Laboratories at Gaithersburg, MD, unless otherwise noted;
mailing address Washington, DC 20234.
[2]Some divisions within the center are located at Boulder, CO 80303.

NBS Special publication

Semiconductor Measurement Technology:

A FORTRAN Program for Analysis of Data from Microelectronic Test Structures

NATIONAL BUREAU
OF STANDARDS
LIBRARY
SEP 1 5 1983

Richard L. Mattis

National Engineering Laboratory
National Bureau of Standards
Washington, DC 20234

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, Secretary
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director

Issued July 1983

Library of Congress Catalog Card Number: 83-600548

National Bureau of Standards Special Publication 400-75
Natl. Bur. Stand. (U.S.) Spec. Publ. 400-75, 66 pages (July 1983)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1983

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402
Price $5.50
(Add 25 percent for other than U.S. mailing)

Table of Contents

List of Figures

Disclaimer

Certain commercial equipment, instruments, or materials are identified in this report in order to adequately specify the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Bureau of Standards, nor does it imply that the material or equipment identified is necessarily the best available for the purpose.

Semiconductor Measurement Technology:
A FORTRAN Program for Analysis of Data from Microelectronic Test Structures

Richard L. Mattis
Semiconductor Devices and Circuits Division
National Bureau of Standards
Washington, DC 20234

Abstract

A computer program, STAT2, is described which performs the following functions: reads data as a two-dimensional array; calculates mean, sample standard deviation, and median; identifies outliers; calculates replacement values for outliers; makes gray-tone, numerical and contour data maps on a line printer; makes a numerical map on the user's terminal; makes a histogram on a line printer; constructs a data base for examining correlations among various data sets; and searches the data base for correlations using several selective keys. The emphasis in this document is on program usage, and detailed descriptions of the commands are given. Data input requirements are addressed. Guidance regarding several types of program modifications is provided.

Key words: computer program; correlation coefficient; data management; outlier; process validation wafer; statistical analysis; test structures; two-dimensional arrays; wafer map.

Introduction

This publication is intended to serve as a Program Manual for a computer program titled STAT2 [1]. As such, it contains information on program usage, installation, and internal structure. Program usage is treated in the sections titled Program Operation Overview, Command Syntax, and Command Descriptions. Program installation is treated in sections titled Program Installation and Logical Unit Assignments. Program internal structure is treated in sections titled Data Array Definition, Data Base Structure, Addition of New Input Data Formats, and Addition of New Commands. STAT2 was originally written to run on an Interdata 7/32 minicomputer, and this publication supersedes NBS Internal Report 82-2492 which described that program [2].

STAT2 is used to analyze test data in which each data value is associated with a test site in a two-dimensional coordinate space. The test data are stored in a two-dimensional array where the subscripts associated with each data value represent the row-column location of the test site. There are some restrictions on the data array as explained in the section titled Data Array Definition.

The program can be used to analyze data from a variety of sources. The intended application, however, is the analysis of measurements from microelectronic test structures for characterizing an integrated circuit fabrication process [3-4]. A paper which discusses this application [3] has been included as Appendix V. Test structures are microelectronic devices which are fabricated by the same process used to fabricate integrated circuits. They can be used to measure selected material or process parameters by means of electrical tests. Test structures are typically fabricated on a circular silicon wafer in a pattern of test sites which is periodic in x and y over the wafer. On the wafer, there may be one or more row-column locations or test sites at which the pattern is interrupted and a different set of test structures or circuits has been inserted. Such a site is called an untested site.

In an integrated circuit process, data taken from test structures can be used to identify which parameters accurately predict or determine the degree of process control; to establish the value and range of these parameters for a given process lot; and to determine how these parameters vary across an integrated circuit die, across a wafer, from wafer to wafer, and from lot to lot. Test results must be obtained and interpreted in a timely fashion in order to be used for correcting or improving the process.

This publication describes a computer program which reads data as a two-dimensional array; calculates mean, sample standard deviation, and median; identifies outliers; calculates replacement values for outliers; makes gray-tone, numerical and contour data maps on a line printer; makes a numerical map on the user's terminal; constructs a data base for examining correlations among various data sets; and searches the data base for correlations using several selective keys. These techniques can provide the user with a relatively fast analysis capability for characterizing an integrated circuit process through the determination of the magnitudes of baseline parameters and their variation over the wafer for "properly" fabricated devices. It is assumed that the process being characterized is in sufficient control to produce a high percentage of "properly" fabricated test structures and that defective structures which are encountered are mainly the result of gross defects introduced by handling, by lithography voids, or by similar process irregularities.

An important aspect of the analysis of test data is the identification of test results from defective structures or defective measurements which do not accurately represent the parameter being measured. Such an incorrect data value is called an outlier. It is necessary to exclude outliers from the population of data values in order to make a more accurate statistical estimate of the parameter. A test site whose data value has been determined to be an outlier is called an excluded site. Other test sites are called included sites.

STAT2 is written in FORTRAN to run on a Digital Equipment Corporation VAX-11/780 computer under revision 3.0 of the VMS operating system. The program is large, requiring approximately 500 KB of memory, but because of the virtual memory feature of VMS, the program has not been divided into overlays. Guidelines for dividing the program into overlays are given in the section titled Program Installation. The part of the program that produces the gray-tone map is written for a Printronix P300 line printer/plotter (600 lines per minute). A version of STAT2 also exists which runs on an Interdata 7/32 (now Perkin-Elmer) minicomputer under revision 4.3 of the OS32MT operating system [2].

Data Array Definition

A typical data array to be analyzed by STAT2 is shown in figure 1. The data values are contained in an array known as DATA which is dimensioned 32 by 32. The first subscript represents the row number of the test site and the second subscript represents the column number. In figure 1, actual test sites exist at the locations indicated by a number. The numbers are site numbers and represent the serial order in which sites are tested. No data were taken at the other points, represented by colons; these are considered nonexistent sites. In this example, the sites numbered 11 and 34 are untested sites, and these site numbers in figure 1 are surrounded by parentheses. Even though these sites are not tested, they must still be assigned a site number. The requirements for such a DATA array to be processed by STAT2 are (1) row 1 must not be empty, (2) column 1 must not be empty, (3) no row or column may have fewer than three test sites, and (4) no row or column may have a nonexistent site between two test sites. Requirements (3) and (4) are imposed by the algorithm which calculates replacement values for sites which have been identified as outliers and excluded.

The location of test sites is available to STAT2 through the STEND (for STart-END) array. STEND is an integer array dimensioned 32 by 2 where STEND(I,1) is the column number of the first (leftmost) test site in row I and STEND(I,2) is the column number of the last (rightmost) test site in row I. Elements of STEND representing rows where there are no test sites must be 0. The STEND array which describes the test site locations in figure 1 is shown in the right portion of the figure.

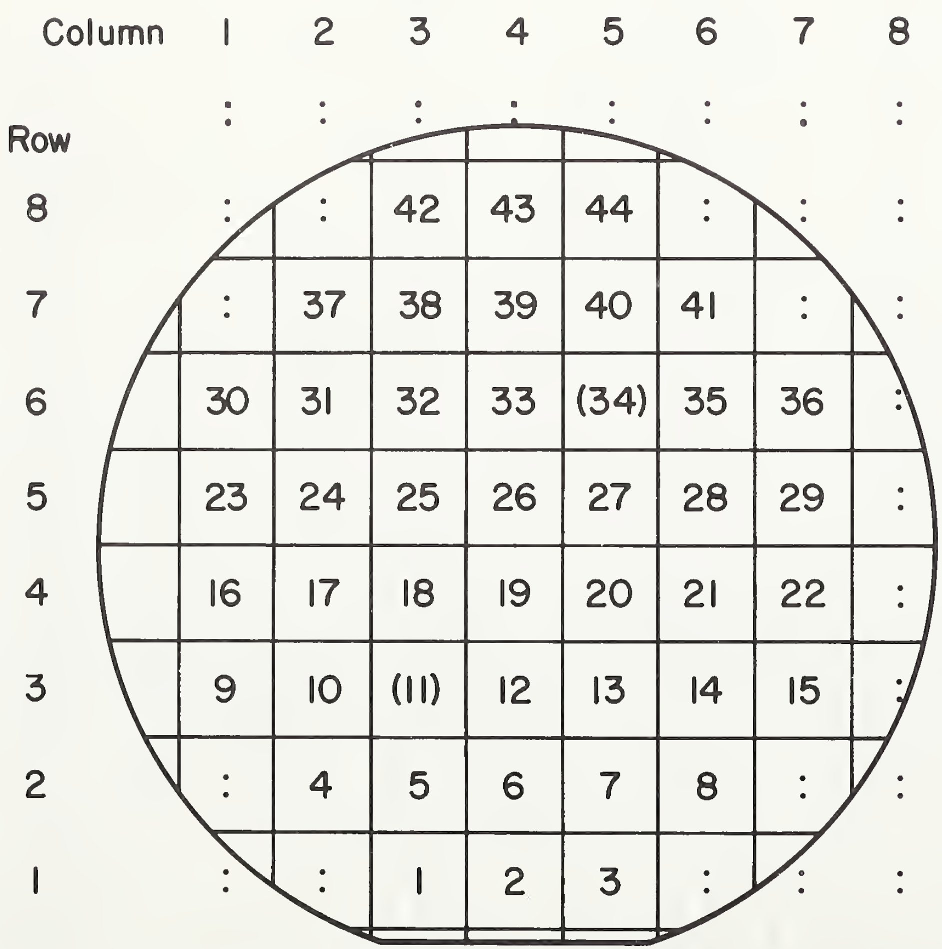
An important attribute of the points in the DATA array is the topological type. Each test site is classified according to whether it is an interior site, site on left boundary, site on an upper right corner, etc. Topological types are used by the outlier replacement algorithms and the mapping subroutines. The topological types are discussed in greater detail in the introductory comments to subroutine ITYPE beginning on page II-50 of Appendix II. Several commands give the type of data points along with other information.

Definition of Statistical Parameters

Several statistical terms used in this publication need to be defined. MEAN represents the simple arithmetic average of all data values associated with all included sites. SIGMA represents the sample standard deviation of this same set of data values. K is a multiple of SIGMA calculated by the XOL command wherein data values which are more than K*SIGMA from MEAN are declared to be outliers. The value of K satisfies the equation

$$\int_{-\infty}^{K} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\,dx = \frac{1 + (1 - p)^{1/n}}{2}\,, \tag{1}$$

where n is the number of included sites and p is the probability that at least one "good" test site may be excluded along with the outliers [5,6] under the assumption that the data values follow a normal distribution.



STEND Array

I	STEND (I,1)	STEND (I,2)
1	3	5
2	2	6
3	1	7
4	1	7
5	1	7
6	1	7
7	2	6
8	3	5
9	0	0
⋮	⋮	⋮
32	0	0

Figure 1. Example of test site locations. (The figure represents a silicon wafer with the major flat at the bottom. The square grid on the wafer represents a periodic pattern of test sites. Row and column numbers are indicated to the left of and above the wafer. Numbers within the wafer are site numbers. Parentheses around site numbers 11 and 34 indicate that these are untested sites. The STEND array for this DATA array is shown in the right part of the figure.)

The sample correlation coefficient, r, is a measure of the similarity of the spatial variation of two sets of data. When the paired observations (x_1,y_1), (x_2,y_2), ..., (x_n,y_n) are taken on two quantities, if a large value of x implies a large value of y, then the quantities are said to be positively correlated. If a large value of x implies a small value of y, then the quantities are said to be negatively correlated. If a large value of x implies nothing about y, then x and y are said to be uncorrelated. The measure of correlation is the correlation coefficient, ρ, which is estimated by the statistic r:

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right] \left[\sum_{i=1}^{n} (y_i - \bar{y})^2\right]}}, \tag{2}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of x and y, respectively, over the n points [7]. Note that r must take on values in the range [-1,1].

Program Operation Overview

This section describes the capabilities of STAT2 without regard to the details of command syntax. Command names are given parenthetically so that the user can relate this description to the discussion of individual commands in a later section. The example relates to data from a microelectronic test structure.

When STAT2 is run, the user assigns an input data file (ASG), then reads an array of data (REA), also called a data set, for examination. If there are untested sites, the user can exclude them at the outset (XIP). The user can calculate the statistics relating to all test sites (PRS) and draw a character histogram (DIS) showing data value distribution. Data values corresponding to a short- or open-circuited device can be removed from the population by excluding sites having data values less than some lower bound (XLT) or greater than some upper bound (XGT). If the test sites on the periphery differ from the interior sites, it may be desirable to exclude them (XPP). If at some time the user wants to put a particular excluded site back in the population, he may do so (IIP). He may also want to put all sites back into the population (RES) and try a different exclusion procedure. After known outliers (such as shorts or opens) and untested sites have been excluded, the user may search the remaining data values for outliers (XOL). In some instances a user may want to specify a particular multiple of the standard deviation (ENN), see what test sites lie farther than that amount from the mean (LNS), and exclude those sites (XNS). At any time the user may list the sites which are excluded (LXP) or he may list the characteristics of any individual test site (LIP) or of all the sites (LAP).

In some applications, the user may be interested in the functional form of the data variation over the wafer surface. The user may fit the data from the included test sites to a plane (FPL) or to a quadratic function (FQD).

The user may also subtract the plane (SPL) or the quadratic function (SQD) from the DATA array and examine the residuals.

A useful output of STAT2 is a data map of which three types are available. The maps give a graphic representation of the variation of the measured parameter over the wafer surface. In order for the map to appear continuous, replacement values must be calculated by interpolation or extrapolation for those test sites which have been excluded (AXP and AIP). A numerical display can be printed on the command terminal (PLT) prior to making the map so the user can experiment with different scaling parameters. A numerical map (MP1), gray-tone map (MP2), or contour map (MP3) can be made in a variety of sizes and scales, and a label can be placed on it if desired. A histogram can also be drawn (HIS).

If the present data set is one of many on which correlation studies are to be performed, the user may write a sample of the data set in a data base (WDB). He may then assign a new input file to be read (ASG) and repeat the process until all data sets have been processed.

After a data base has been established, the user may list all or a portion of its contents (LDB). He may select a particular entry or sample as a reference sample (GET) and calculate a correlation coefficient of the reference sample and a selected group of other samples (SDB). If any entries in the data base are in error or are no longer needed, they may be marked deleted (DEL).

At any time, the user may insert a comment (asterisk in first character position) which appears in the printed output. If the user repeatedly executes the same command sequence, he may create a file of these commands and execute them (MAC). The user may pause (PAU) and terminate STAT2 (END). A help facility (HEL) prints information about STAT2, or about a particular command, on the user's terminal.

The output of STAT2 is logged to a file named STAT2.LOG which may be printed using the VMS PRINT command after the run terminates. Output also appears on the user's terminal. The main difference between the two types of output is (1) command syntax error messages, HEL output, and PLT displays go only to the user's terminal, and (2) maps made by MP1 and MP2 go only to the log file. Maps made by MP3 go to a metacode file which must undergo further processing as explained in the description of the MP3 command.

Data Base Structure

A group of commands in STAT2 provide for constructing and using data bases. A data base is a pair of direct access files which contains a sample of data from several or many data sets. The user can search for correlations among data sets by performing calculations of a sample correlation coefficient [7] for one or more samples against other samples. The greater the number of data values included in the sample, the smaller the uncertainty in the calculated sample correlation coefficient. As the sample size increases, however, disc storage requirements and computation time go up. It has been found empirically that for a data set containing 95 points, a sample of 13 points produces sample correlation coefficients of acceptable uncertainty.

Suehle discusses sample size for test chip data [8]. The nature and use of a data base are presented in this description.

(1) Data Base File Format

A data base may be constructed for any collection of data sets which may be expected to be related and for which the samples of wafer data stored in the data base are all selected in the same way. Examples of collections of data sets which may be expected to be related are (1) all runs of a particular test pattern, (2) all data from a given run of wafers, and (3) all data from a particular set of test structures. A data base consists of a label file containing 80-byte records and a data file containing 24-byte records. The two files are linked together by bidirectional pointers. A sample label file and data file are shown in figures 2 and 3. A data base entry consists of one record in the label file and (NDAT+1) records in the data file where NDAT is the number of data values in the sample. A data base may contain up to 99,999 label records and up to 999,999 data records.

The label record contains data in eleven fields. Each data base entry is identified by an entry number in the ENT field. The entry number is the same as the record number of the direct access label file. The entry number is the identifier by which the entry is known, and it is used by the GET, DEL, and LDB commands. The DLT field is normally an ASCII space, but an asterisk is placed there when an entry is marked deleted (by a DEL command). The PTR field contains the record number of the header record in the associated direct access data file. The NDAT field contains the number of data values in the data file per entry. The next five fields -- PAT, LOT, WAF, DEV and PCODE -- give the pattern number, lot number, wafer number, device number and parameter code relating to the data. The remaining two fields give the date and time at which the entry was written.

The data themselves are in the data file preceded by a header record. The header contains 'H' in the first character position and the entry number. The data records, their number given by NDAT in the label record, follow the header. Each data record contains a 'D' in the first character position, the entry number, the data value, and a one-byte EXC field. If that particular data value came from an excluded test site, an asterisk is written in the EXC field; otherwise that field contains an ASCII space.

The first label record and first two data records in a data base are title records which contain a user-assigned title and the sampling plan code (defined below). The first label record also contains the record number of the last record in each of the data base files so that the next entry added to the data base can be written to the proper record locations.

(2) Data Base Creation

The user must decide how many data points are to be included in the sample and from what row-column locations on the wafer they are to be taken. This is called the sampling plan. Having decided on a sampling plan, the user must make the necessary software changes in STAT2 if it is a new sampling plan, and write the first record of the label file and the first two records of the data file. Four sampling plans are presently available, with test

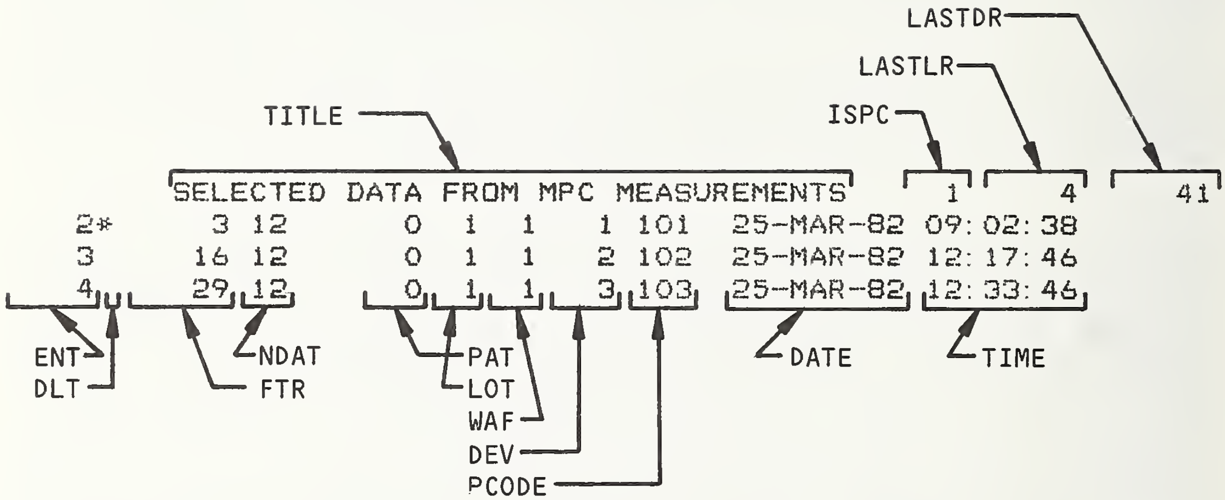


Figure 2. Sample label file. (The first record contains the data base title, sampling plan code (ISPC), entry number of the last entry in the data base (LASTLR), and record number of the last data record in the data base (LASTDR)).

```
T   SELECTED DATA FROM M
T     PC  MEASUREMENTS
H      2
D      2   0.149342E+02
D      2   0.134277E+02
D      2   0.136344E+02
D      2   0.148304E+02
D      2   0.136341E+02
D      2   0.141438E+02
D      2   0.147735E+02
D      2   0.143607E+02
D      2   0.144221E+02
D      2   0.146835E+02
D      2   0.147407E+02
D      2   0.149635E+02
H      3
D      3   0.230731E+02         Indicates an
D      3   0.733821E-03*          Excluded Site
D      3   0.228913E+02
D      3   0.230301E+02
D      3   0.228697E+02
D      3   0.232465E+02
D      3   0.233275E+02
D      3   0.233326E+02
D      3   0.233813E+02
D      3   0.236023E+02
D      3   0.237082E+02
D      3   0.473316E+04*
H      4                        Header Record
D      4   0.322554E+02
D      4   0.314867E+02
D      4   0.315397E+02
D      4   0.320440E+02
D      4   0.313571E+02
D      4   0.323289E+02
D      4   0.325107E+02         Data Records
D      4   0.324519E+02
D      4   0.329771E+02
D      4   0.326444E+02
D      4   0.329844E+02
D      4   0.321846E+02
```

Figure 3. Sample data file. (The first two records contain the data base title. The entry number is contained in the header record and preceding t data value in the data record.)

Table 1. Site Row-Column Locations for the Four Available Sampling Plans.

Sample Number	Row-Column Locations for Sampling Plan Code 0	Code 1	Code 2	Code 3
1	1,6	2,4	3,5	2,6
2	2,4	3,2	5,3	3,4
3	2,8	3,6	5,5	3,8
4	4,4	4,4	5,7	5,2
5	4,8	5,1	7,3	5,4
6	5,2	5,3	7,5	5,8
7	5,6	5,5	7,7	5,10
8	5,10	5,7	9,3	7,2
9	6,4	6,4	9,5	7,4
10	6,8	7,2	9,7	7,8
11	8,4	7,6	11,5	7,10
12	8,8	8,4	-	9,4
13	9,6	-	-	9,8
14	-	-	-	10,6
15	-	-	-	-

site locations as indicated in table 1. A fifth sampling plan (Code 4) samples every location within the bounds of the STEND array, thereby including all test sites in the correlation coefficient calculation. Other sampling plans to include up to 256 sites can be added, and the required software changes are not difficult. Refer to comments in subroutine DBØ2 beginning on page II-158 of Appendix II.

To create a label file having record length of 80 bytes and a data file having record length of 24 bytes, compile, link, and run program CRDB, listed in Appendix III. Enter a data base description or title, up to 36 characters, as requested, and enter the sampling plan code. After running CRDB, rename FORØ17.DAT to a suitable name for the label file, and rename FORØ18.DAT to a suitable name for the data file. After this renaming, the data base may be written to.

(3) Writing Entries in a Data Base

If data base operations are to be performed, the data base files must be assigned before STAT2 is run using the VMS ASSIGN command. Assign the label file specifier to FORØ17 and the data file specifier to FORØ18.

To write an entry in the data base (1) assign an input file using ASG, (2) read a data set using REA, (3) perform data point exclusions, (4) calculate replacement values for excluded sites using AXP, and (5) type the WDB command as directed in the section titled Command Descriptions. After the data base entry has been written, it is read and displayed in readable form on the user's terminal and logged to the line printer.

(4) Examining a Data Base

The data base can be examined by using the LDB command as explained in the section titled Command Descriptions. The output contains all the information in the label record plus the data values themselves if requested.

(5) Defining a Reference Data Sample

When searching for correlations among data base entries, it is necessary to define an entry, identified by its entry number, with which other entries are to be compared. This is done by the GET command as explained in the section titled Command Descriptions.

(6) Searching a Data Base

A search of a data base for entries which correlate with the reference data sample is initiated by the SDB command. In general, the user does not want to seek correlations between the reference data sample and all entries in the data base.

If any of the data values in either of the two data samples came from an excluded site (as indicated in the EXC field of the data record), those values are not used in sample correlation coefficient calculations.

The sample correlation coefficients thus calculated are based on the data samples and are intended to indicate possible correlations (depending on the magnitude of the sample correlation coefficient). The user should examine data maps of the correlating data sets to evaluate the correlation more fully.

(7) Deleting an Entry from a Data Base

An entry may be marked "deleted" by using the DEL command. Such an entry is still physically in the data base and is recognized by LDB commands but is ignored by SDB commands and produces an error when referenced by GET and DEL commands. Although the delete operation cannot be undone, the user can rewrite the entry in the data base.

Command Syntax

A command consists of a three-character mnemonic followed in some cases by a parameter list. Parameters when present are from 1 to 8 in number. A space or comma optionally preceded or followed by one or more spaces is used as a delimiter between mnemonic and parameter and between parameters. Alphabetic characters (except E), semicolons, terminal commas, and successive commas are illegal and result in a request for repeated command entry. The entire command must not occupy more than 72 characters; excess characters are ignored. Considerable freedom is available in the format of the parameters themselves. They must be no more than 20 characters in length and may be expressed in I, F, or E format as may be convenient. A minus sign is legal as the first character in a parameter or as the first character after an 'E'. An optional plus sign is also legal in these two locations. An exponent must be one or two digits plus sign if present.

Exceptions to these rules are represented in the ASG, MAC, and HEL commands. In ASG and MAC, the first (and only) parameter is replaced by a file specifier. In HEL, the parameter is replaced by a three-character mnemonic.

There are two commands (LDB and SDB) for which certain parameters can be represented legally by a minus sign only. These parameters are called privileged parameters and are explained in the discussion of LDB and SDB.

An asterisk as the first character in a command line causes that line to be interpreted as a comment. This comment is logged to the user's terminal and to the log file.

Command Descriptions

The STAT2 command set consists presently of 39 commands. These commands are now discussed individually. The symbols P1, P2, ... represent parameters which would be entered as numerical values as explained above. A summary of the commands is given below and is repeated in Appendix I.

AIP - Alter an individual point.
ASG - Assign input data file.
AXP - Alter excluded points.
DEL - Delete data base entry.

DIS - Display distribution.
END - Terminate STAT2 execution.
ENN - Set N to a specified value.
ERM - Error message switch.
FPL - Fit DATA array to a plane.
FQD - Fit DATA array to a quadratic function.
GET - Define data base entry as reference sample.
HEL - Help request.
HIS - Draw a histogram.
IIP - Include an individual point.
LAP - List all points.
LDB - List data base entries.
LIP - List an individual point.
LNS - List points beyond N*SIGMA from mean.
LXP - List excluded points.
MAC - Execute command macro.
MP1 - Draw numerical map of DATA array.
MP2 - Draw gray-tone map of DATA array.
MP3 - Draw contour map of DATA array.
PAU - Pause STAT2 execution.
PLT - Draw character display of DATA array.
PRS - Print statistics.
REA - Read input data file.
REM - Set or reset remote mode.
RES - Restore all points to included status.
SDB - Search data base for correlations.
SPL - Subtract fitted plane from DATA array.
SQD - Subtract fitted quadratic function from DATA array.
WDB - Write data base entry.
XGT - Exclude points greater than a value.
XIP - Exclude an individual point.
XLT - Exclude points less than a value.
XNS - Exclude points beyond N*SIGMA from mean.
XOL - Exclude outliers.
XPP - Exclude peripheral points.

In the command descriptions, frequent references are made to error conditions. An error condition occurs when a command as entered cannot be executed. The error condition may arise due to a command syntax error, an illegal value of a parameter, a failure to execute another command which must be issued prior to the present command, or some other cause. When remote mode is enabled (see the REM command description), such as when STAT2 commands are being read from a command file by the MAC command, an error condition causes STAT2 to pause. This is done because subsequent commands would, in many cases, be rendered meaningless or misleading until the error condition is corrected. A message indicating an error condition is preceded by three asterisks. On the other hand, warning messages are intended to inform the user of possibly unintentional consequences of a command without interrupting the execution of the command. Warning messages are preceded by three exclamation points and may be disabled by the ERM command. An explanation of all error messages is given in the section titled Error Messages.

(1) Miscellaneous: END, PAU, ERM, REM, MAC, and HEL

END - End. Terminate STAT2 execution.

PAU - Pause. Pause STAT2. Execution is resumed by the VMS CONTINUE command.

ERM,P1 - Error Message Enable/Disable. Allow warning messages to be enabled (P1 <> 0) or disabled (P1 = 0). Such messages are initially enabled when STAT2 is run. The messages most frequently deal with attempts to exclude data points which have already been excluded. All warning messages are given in the section titled Error Messages.

REM,P1 - Remote Enable/Disable. Enable Remote Mode when P1 is nonzero. Disable Remote mode when P1 is zero. Remote Mode is initially disabled when the program is started. When Remote Mode is enabled, any error condition causes STAT2 to pause. Remote Mode is enabled whenever STAT2 is directed from a MAC command file and is left enabled when exiting from such a command file.

MAC,file-specifier - Execute Command Macro. Go to the specified file and begin reading and executing STAT2 commands from that file. Control is returned to the user terminal after executing the last command in the file. During execution of the command macro, remote mode is enabled so that any error (except a warning error) causes STAT2 to pause. It is illegal to call a macro command file from another macro command file.

HEL,mnemonic - Type Help Message. When HEL or HELP is followed by a STAT2 delimiter and a command mnemonic, a description of that command is printed on the user terminal. HELP COM produces a list of all STAT2 commands. HELP SYN gives the rules of STAT2 command syntax. HELP followed by anything else or by itself produces a default message which explains the types of help available. Messages produced by HELP are not printed on the log file.

(2) Data Input: ASG and REA

ASG,file-specifier - Assign Input File. This command closes the input file and assigns a new input file. ASG is followed by a delimiter, then the file specifier of the input file. An error condition is produced if the file specifier is not syntactically correct or if it refers to a nonexistent file.

REA,P1,P2 - Read Data. Read data from a disc file. Three data formats can be read, details of which are given below. Addition of the capability to read other formats is not difficult and is treated in the introductory comments to subroutine REA beginning on page II-38 of Appendix II.

When an REA command is issued, two arrays are read into the computer memory, the STEND array and the DATA array. The organization of these two arrays is discussed in the section titled Data Array Definition.

For format 1, read by an REA,1,0 command, the full 32-by-32 DATA array is available. This format is designed so that the file may consist of 18-byte records. It is also designed to be space-efficient. The FORTRAN WRITE statements used to generate a data file in format 1 are given below. The

STEND array is contained in the first 16 records. Following that are the data values, one per record, in the order of the site numbers in figure 1. The file ends after the value for the last site number.

```
      Format #1
C
      DO 900 I=1,29,4
      K=I+3
      WRITE (11,800) (STEND(J,1), J=I,K)
      WRITE (11,800) (STEND(J,2), J=I,K)
800   FORMAT (1X,4(I2,1X))
900   CONTINUE
C
      DO 909 I=1,32
      ICOL1=STEND(I,1)
      ICOL2=STEND(I,2)
      IF (ICOL1.EQ.0) GOTO 999
      DO 909 J=ICOL1,ICOL2
      WRITE (11,901) DATA(I,J)
901   FORMAT (1X,E13.6)
909   CONTINUE
999   CONTINUE
C
```

For format 2, read by an REA,2,0 command, the measurement locations are restricted to a 16-by-16 subarray of the DATA array. As seen from the listing below, the STEND array (for 16 rows) is contained in the first two records of the file. Data for the 256 possible test sites follow in the next 64 records. The first record of data following the STEND array contains values from row 1, columns 1 through 4. The next three records contain values from row 1, columns 5 through 16. The next four records contain values from row 2, and so on. Data from untested sites are included in the file as zero values, making format 2 inefficient with respect to disc storage space.

```
      Format #2:
      INTEGER STEND(16,2)
      DIMENSION DATA(16,16)
C
C     WRITE THE STEND ARRAY
C
      WRITE (LU,10) (STEND(I,1), I=1,16)
      WRITE (LU,10) (STEND(I,2), I=1,16)
10    FORMAT (1X,16(I2,1X))
C
C     WRITE THE DATA ARRAY.
C
      DO 110 I=1,16
      DO 120 J=1,16,4
      WRITE (LU,130) DATA(I,J), DATA(I,J+1), DATA(I,J+2), DATA(I,J+3)
130   FORMAT (1X,4(E13.6,2X))
120   CONTINUE
110   CONTINUE
```

Format 3 is designed to accommodate a full 32-by-32 DATA array, to store the data in a space-efficient manner, and to allow multiple data sets to be stored in a single direct access file. The data sets must be related to a single STEND array which appears only once in the data file. The file format is shown in figure 4. The STEND array is contained in the first 16 records in the same format as in format 1. Record 17 contains an integer in (1X,I4) format giving the number of data sets in the file. The data sets follow, each set preceded by a flag record containing an asterisk and an integer in (1X,A1,1X,I4) format, where the integer is the serial number of the data set within the file. The data values are formatted and ordered in the same manner as in format 1. The flag records are used to verify positioning within the file. The second parameter of the REA command specifies the data set within the file which is to be read. This must be an integer between 1 and the number of data sets in the file. The second parameter must be present to read formats 1 and 2 also, but in these cases that parameter is ignored. Note that the file organization must be relative and file access must be direct as defined within the VMS file structure.

(3) Statistics: ENN, PRS, DIS, LAP, LNS, and LIP

ENN,P1 - Enter N. Enters the value of N, the user-specified multiple of the standard deviation, for use by PRS, DIS, LNS, and XNS commands. P1 is the value of N. A negative P1 produces an error condition. If ENN is not executed, N has the default value of 1.0.

PRS - Print Statistics. Prints the following statistics of the included test sites (that is, the test sites currently included in the population): maximum, minimum, MEAN, median, SIGMA, percent standard deviation, N times SIGMA, and number of included and excluded sites. An error condition occurs if no data file has yet been read. If there are fewer than two included sites, the calculations done by PRS are not meaningful and an error condition occurs. If the median value occurs repeatedly in the DATA array, the algorithm which calculates the median (subroutine MEDCAL) cannot converge to the median; in this case the median is approximated, and a warning message is printed indicating a lower and upper bound of the median. For real data, these two bounds are usually very close together.

DIS - Display Distribution. Divides the range of values of the included sites into 50 equal intervals or bins and places each site in the appropriate bin. It then prints a string of 50 characters (bounded by exclamation points), each character representing the number of sites in the corresponding bin. A minus sign signifies zero sites and an 'X' signifies ten or more sites in a bin. Beneath this character histogram, double quotes are placed at the bin containing the mean value and at the bins containing the N*SIGMA limits. These limits often would fall outside of the display. When that happens, they appear at the end position of the display; however, the bin numbers at which they would appear are printed beneath the display as IBMIN and IBMAX. Bin numbers outside the range 1 to 50 indicate that one or both of the N*SIGMA points are off the scale of the display. DIS gives a compact display of the manner in which the data values of the included sites are distributed between the minimum and maximum values. DIS cannot be executed unless PRS has been executed with no intervening site exclusions or reinsertions or change in N value.

```
 4  3  2  1
 7  8  9 10
 1  1  2  3        STEND array contained in
10 10  9  8        first 16 records
 4  0  0  0
 7  0  0  0
 0  0  0  0
 0  0  0  0
      .
      .
      .
 0  0  0  0
 0  0  0  0
  37               Number of data sets contained in the file
*     1            Flag record for first data set.
 0.428431E+01
 0.427305E+01      Data in first data set.
 0.430013E+01
      .
      .
      .
 0.427419E+01
*     2            Flag record for second data set.
 0.273193E+02
 0.274012E+02      Data in second data set.
 0.283955E+02
      .
      .
      .
*     3
      .
      .            Additional data sets.
      .
*    37            Flag record for last data set.
 0.592734E-01
 0.632993E-01      Data in last data set.
 0.601832E-01
      .
      .
      .
 0.573480E-01
```

Figure 4. Example of an input data file in format 3.

LAP - List All Points. Gives the row-column location, data value, topological type, and status of each test site. Topological types are discussed in the section titled Data Array Definition. The status indicates whether a particular site is presently included or excluded.

LNS - List Beyond N*SIGMA. Lists all included sites which have data values outside the range MEAN ± N*SIGMA where N is defined by the ENN command. The list includes row-column location, data value, and topological type. This command is legal only when SIGMA is current; that is, there have been no exclusions or insertions or change in N value since the last time SIGMA was calculated by PRS. If SIGMA is not current, an error condition occurs.

LIP,P1,P2 - List Individual Point. Gives the data value, topological type, and status of the test site at row P1 and column P2. If this location is not an actual measurement site, an error condition occurs.

(4) Test Site Exclusion: XOL, XPP, XNS, XGT, XLT, XIP, LXP, IIP, and RES

XOL,P1 - Exclude Outliers. Uses an iterative method to identify and exclude outliers [3]. All sites having data values beyond a multiple K of SIGMA are excluded, where the multiple K is a function of the number of included sites and of P1, the probability that one or more good sites might be excluded along with the outliers. P1 must be in the range of 0.05 to 0.90 inclusive or an error condition occurs. A value of 0.20 for P1 has been found to be suitable with data from microelectronic test structures [3]. When XOL is invoked, STAT2 makes one or more passes through the data set, prints the statistics current for that pass, the K value, and the number of sites excluded on that pass. The process stops when on a given pass no sites are excluded. Note that XOL alters the value of N set by the ENN command and sets it to the final value of K.

This command and the commands which follow relating to exclusions of test sites require that REA be executed previously. If any of these commands would attempt to exclude a site which is already excluded, a warning message is logged.

XPP - Exclude Peripheral Points. Excludes all sites which are located on the boundary of the test site space.

XNS - Exclude Beyond N*SIGMA. Excludes all sites which have data values outside the range MEAN ± N*SIGMA where N is defined by the ENN command. This command is legal only when SIGMA is current; that is, there have been no exclusions or insertions since the last time SIGMA was calculated by PRS. If SIGMA is not current, an error condition occurs.

XGT,P1 - Exclude if Greater Than. Excludes all sites having data values which are greater than the value given by P1.

XLT,P1 - Exclude if Less Than. Excludes all sites having data values which are less than the value given by P1.

XIP,P1,P2 - <u>Exclude an Individual Point</u>. Excludes the site at row P1 and column P2 in the DATA array. If there is no site at that location, an error condition occurs.

LXP - <u>List Excluded Points</u>. Lists all sites which have been excluded, giving their row-column location, value, and topological type.

IIP,P1,P2 - <u>Reinsert an Individual Point</u>. Reinserts or includes a previously excluded site located at row P1 and column P2. If there is no site at that location, an error condition occurs. If the site to be reinserted has not been excluded, a warning message is printed.

RES - <u>Reset</u>. Restores all test sites to included status. An REA command must be executed prior to RES.

(5) Data Value Replacement: AXP and AIP

AXP - <u>Alter Excluded Points</u>. Performs interpolation or extrapolation of data values from nearby sites to calculate a replacement value for each excluded site for plotting purposes. The replacement is calculated by 1 of 15 algorithms depending on the topological type and the proximity of other excluded sites. For a site having a given topological type, a "first choice" algorithm is selected to calculate the replacement value. If this algorithm would use data values from other excluded sites in the calculation, it cannot be used. A "second choice" algorithm is then selected. For sites having some topological types, a third or fourth choice algorithm is also attempted, if necessary. When the replacement is made, a message is printed giving the site coordinates, topological type, old and new values, and the algorithm used. When a replacement cannot be made, a message so stating is also printed. Note that AXP alters the DATA array. The array can be restored only by reissuing the REA command. Additional information on the mechanism of data value replacement is given in the program listing in the introductory comments to subroutines ALTO and RALGA beginning on pages II-57 and II-63 of Appendix II.

AIP,P1,P2,P3 - <u>Alter Individual Point</u>. Assigns the site at row P1, column P2 the data value P3. The site must already be excluded; otherwise, an error condition occurs. This command is intended to be used judiciously when none of the replacement algorithms is able to calculate a replacement value.

(6) Functional Fits: FPL, SPL, FQD, and SQD

FPL - <u>Fit to a Plane</u>. Finds the equation of the plane which gives the least squares fit to the included test sites. The coefficients of the equation of the plane and the residual standard deviation are printed. REA must be executed prior to FPL. This command uses several routines from CMLIB, a library of mathematical software supported by NBS [9].

SPL - <u>Subtract Plane</u>. Takes the plane, the equation of which was calculated by FPL, and subtracts it from the DATA array leaving the residuals in the DATA array. FPL must have been executed prior to SPL with no intervening exclusions or reinsertions of sites; otherwise, an error condition occurs. If a MAP (MP1, MP2, MP3, PLT) or histogram (HIS) of the residuals is requested with the autoscale option, PRS must be executed after SPL and prior to the map or histogram command in order that the maximum and minimum values might pertain to the residuals and not to the original DATA array. The SPL operation cannot be undone, so the original DATA array can be recovered only by reissuing the REA command.

FQD - Fit to a Quadratic Surface. Finds the equation of the quadratic surface which gives the least squares fit to the included test sites. The coefficients of the equation of the quadratic surface and the residual standard deviation are printed. REA must be executed prior to FQD. This command uses several routines from CMLIB [9].

SQD - Subtract Quadratic Surface. Takes the quadratic surface, the equation of which was calculated by FQD, and subtracts it from the DATA array leaving the residuals in the DATA array. FQD must have been executed prior to SQD with no intervening exclusions or reinsertions of test sites; otherwise, an error condition occurs. If a map or histogram of the residuals is requested with the autoscale option, PRS must be executed after SQD and prior to the map or histogram command in order that the maximum and minimum values might pertain to the residuals and not to the original data array. The SQD operation cannot be undone, so the original DATA array can be recovered only by reissuing the REA statement.

(7) Data Display: PLT, MP1, MP2, MP3, and HIS

PLT,P1,P2,P3 - Plot on Terminal. Produces a numerical display on the user's terminal in which a number represents the data value at each test site. The range of values between P2 (the lower bound) and P3 (the upper bound) is divided into eight equal intervals. Data values falling in these intervals are represented by the characters 1 for the lowest interval through 8 for the highest interval. Test sites having values greater than P3 are represented by '+', and sites having values less than P2 are represented by '-'. Excluded sites are represented by ':'. Following this character display, a table is printed showing the upper and lower bounds and the number of sites which fall into each of the eight bins.

The above discussion of P2 and P3 as the lower and upper bounds of the display range applies only if P1 is 0. If P1 is positive, the range is autoscaled to the maximum and minimum values of the included sites. If P1 is negative, the range is autoscaled to the MEAN ± K*SIGMA limits where K is the multiple of SIGMA calculated by the XOL command on its last pass. In either case P2 and P3 must still be present, but their values are ignored. No '+' or '-' characters should appear in the display if either autoscale is used. Either PRS (P1 > = 0) or XOL (P1 < 0) must be executed prior to PLT with no intervening site exclusions or reinsertions or change in N value.

The maps produced by MP1, MP2, and MP3 are not produced on the user's terminal. PLT provides a preview of what such a map will look like and enables the user to adjust the scaling to produce the desired pattern of numbers, shades, or contours.

MP1,P1,P2,P3,P4,P5,P6,P7 - Draw Numerical Map. Plots the DATA array as a numerical map using the numbers 1 through 8. P1 is the width and P2 is the height of the map in inches. The allowed range of P1 and P2 is determined by the characteristics of the line printer, which can be specified in subroutine SIZE which begins on page II-94 in Appendix II. The maximum map size using the present line printer is approximately 13 inches wide by 9.5 inches high. P4 is used to provide a label for the map. If P4 is nonzero, the user is asked to enter a label up to 72 characters long. The label is printed above the map. If P4 is zero, no label is requested or printed, but the space above the map is still reserved. A label may also be entered from a MAC command file by placing it immediately following the MP1 command. P5 is used to optionally specify the format of the numbers in the map key. If P5 is nonzero, the user is asked to enter up to five characters representing a FORTRAN format. This may be an E or F format. If P5 is zero, the numbers in the key print in F10.5 or E12.5 format depending on their magnitudes. A format may also be entered from a MAC command file by placing it immediately following the MP1 command or, if present, immediately following the map label. If P3 is positive, the plot is automatically scaled between the minimum and maximum included data values. The minimum and maximum must be current, however, with no site exclusions or reinsertions or change in N value since the last PRS command. If P3 is negative, the range is autoscaled to the MEAN ± K*SIGMA limits where K is the multiple of SIGMA calculated by the XOL command on its last pass. If this option is used, an XOL command must have been issued with no intervening site exclusions or reinsertions or change in N value. If P3 is zero, then the scaling is based on P6, the lowest value represented by the number 1, and P7, the highest value represented by the number 8. P7 must be greater than P6, and P6 and P7 must be present even if P3 is nonzero, although they are then ignored. The map is drawn on a new page. On the following page a key is drawn giving the range of values represented by each number, the number of test sites having values which lie within the range denoted by each number, and the mean, sample standard deviation, and median of the included sites. On the plot, the locations of included test sites are represented by an 'X' or by a '+' or '-' in the event that autoscaling is not used and the data values are greater than or less than the scaling values given by P6 and P7. Excluded sites are not represented by one of these symbols but by the appropriate number. A sample map is shown in figure 5 with the accompanying key in figure 6.

A numerical map cannot be used to represent data containing more than about 27 rows. This limitation is imposed by the number of lines on a line printer page and the need to have at least one row of numerical symbols calculated by interpolation between rows containing symbols representing actual test sites.

The map produced by MP1 is intended to be applicable to any line printer and is therefore a portable map. The maps produced by MP2 and MP3 have special hardware and software requirements.

MP2,P1,P2,P3,P4,P5,P6,P7 - Draw Gray-Tone Map. Plots the DATA array as a gray-tone map with an eight-level gray scale. P1 is the width and P2 is the height of the map in inches. The width must not be greater than 13 inches nor the height greater than 8 inches. The minimum width and height in inches are given by 0.2 times the number of columns and 0.2 times the number of rows, respectively, of data in the map. This is to ensure that there are a sufficient number of picture elements (pixels) to make a meaningful plot.

```
                    X8888888888888888888X
                    777788888888888888888
                    777777777777777777777
                    666666667777777777777
          X777766665X556666666X666666666X666777777X
          77766666655556666666666666666666666677777
          66666665555555556666666666666666666666677
          66665555555555555555666666666666666666666
X776666655X555555555X555555555X555555666X666666666X666667777X
6666665555555555555555555555555555555555555555556666666666667
5555555555555555544445555555555555555555555555555555666666666
5555555555544444444444444444444444444444444455555555555555555
X444444444X444444444X444444444X444444444X444445555X555555555X
3333333444444444444444444444444444444444444444455555555555555
3333333333333333333444444444444444444444444444444555444444444
2222222222223333333333333333444444444444444444444444444444444
X111111111X222223333X3333333333333444444X444444444X444444444X
2222222111222222223333333333333333333333333333333333333333333
2222222221122222222222223333333333333333333333333333333333333
2222222221112222222222222222223333333333333322222222222233333
X222222221X111111111X112222222X222222233X222222222X222222222X
2222221111111111111111111222222222222222332222222221222222222
1111111111111111111111111111222222222223322222222212222222222
1111111111111111111111111111111122222223322222222222222222222
X111111111X111111111X111111111X111222223X322222222X211111111X
          11111111111111111111112222222222222222211
          11111111111111222222222222222222222221111
          11111111122222222222222222222222222111111
          X111111222X222222222X222222222X221111111X
                    222222222222222222222
                    222222222222222222222
                    112222222222222222222
                    X1111111111111222222X
```

Figure 5. Sample numerical map.

	PARAMETER VALUE			# SITES
8'S	24.07178	TO	24.27020	3
7'S	23.87335	TO	24.07178	3
6'S	23.67492	TO	23.87335	4
5'S	23.47650	TO	23.67492	6
4'S	23.27808	TO	23.47650	8
3'S	23.07965	TO	23.27808	4
2'S	22.88123	TO	23.07965	7
1'S	22.68280	TO	22.88123	13
SITES INCLUDED				48

X X X X INCLUDED SITES WITHIN SCALE

+ + + + INCLUDED SITES ABOVE SCALE

- - - - INCLUDED SITES BELOW SCALE

SAMPLE MEAN	23.28777
SAMPLE STD DEV	0.44481
SAMPLE MEDIAN	23.24650

Figure 6. Key which accompanies map of figure 5.

Parameters P3, P4, P5, P6, and P7 have the same meaning in the MP2 command as they have in the MP1 command. As with MP1, the map and an accompanying key are drawn on separate pages. On the map, the locations of included test sites are represented by an 'X' or by a '+' or '-' in the event that autoscaling is not used and the data values are greater than or less than the scaling values given by P6 and P7. Excluded sites are not represented by one of these symbols but by the appropriate gray tone. A sample map is shown in figure 7 with the accompanying key in figure 8.

MP3,P1,P2,P3,P4,P5,P6 - Draw Contour Map. Plots the DATA array as a circular contour map. This command uses a plotting package written by the National Center for Atmospheric Research (NCAR) [10], plus several routines from CMLIB [9]. P1 is the diameter of the map in inches which under the present revision must be 8. As with parameter P3 of the MP1 and MP2 commands, P2 of MP3 determines the type of scaling to be used, and P5 and P6 are the minimum and maximum values to be plotted when P2 is 0. P4 is again a flag for entering a plot label up to 72 characters in length. Excess characters are ignored. P3 specifies the number of contour lines in the map. P3 must be a positive integer in the range 5 to 40, inclusive. PRS (P2 >= 0) or XOL (P2 < 0) must be executed prior to MP3. A sample contour map is shown in figure 9.

In the maps made by MP1 and MP2, all plotted numbers or shades lie within a region bounded by test sites so that interpolation can be used to determine the number or shade appropriate to a particular location on the map. In MP3, however, a circular boundary is drawn outside the region bounded by the sites. In order to extend the contour lines to the circular boundary, extrapolation algorithms are used. This can produce patterns of lines around the edge of the map which do not accurately represent the spatial variation of the data in that region. Also, the contours as drawn are processed by a smoothing algorithm. Without smoothing, the lines would be piecewise linear over distances equivalent to the distance between sites, creating a jagged map. The smoothing enhances the appearance of the map by curving the lines, but the curvature which is introduced does not always accurately represent the behavior of a measured parameter between test sites. In spite of these effects, the objective of a visual display which conveys the general nature of the data variation over the wafer is still achieved.

The contour interval is given underneath the map. The values represented by the contour lines are given in labels on every fourth contour line to three significant digits. When mapping data sets consisting of values having magnitudes less than 1 or greater than 1000, the contour labels on the map itself may be off by some power of ten. Then a scale factor will be given beneath the map. In the process of forcing the contour interval to a "nice" value, that is, representable by a minimum number of significant digits, subroutine MP3 may produce a map having somewhat more or fewer contour lines than was specified in P3.

When MP3 is executed, the map does not go to STAT2.LOG; rather a metacode file named IOPØ2Ø.DAT is produced. A metacode translator named MCVAX must be invoked to produce the map from the metacode file after STAT2 has terminated. The translator is also part of the NCAR plotting package.

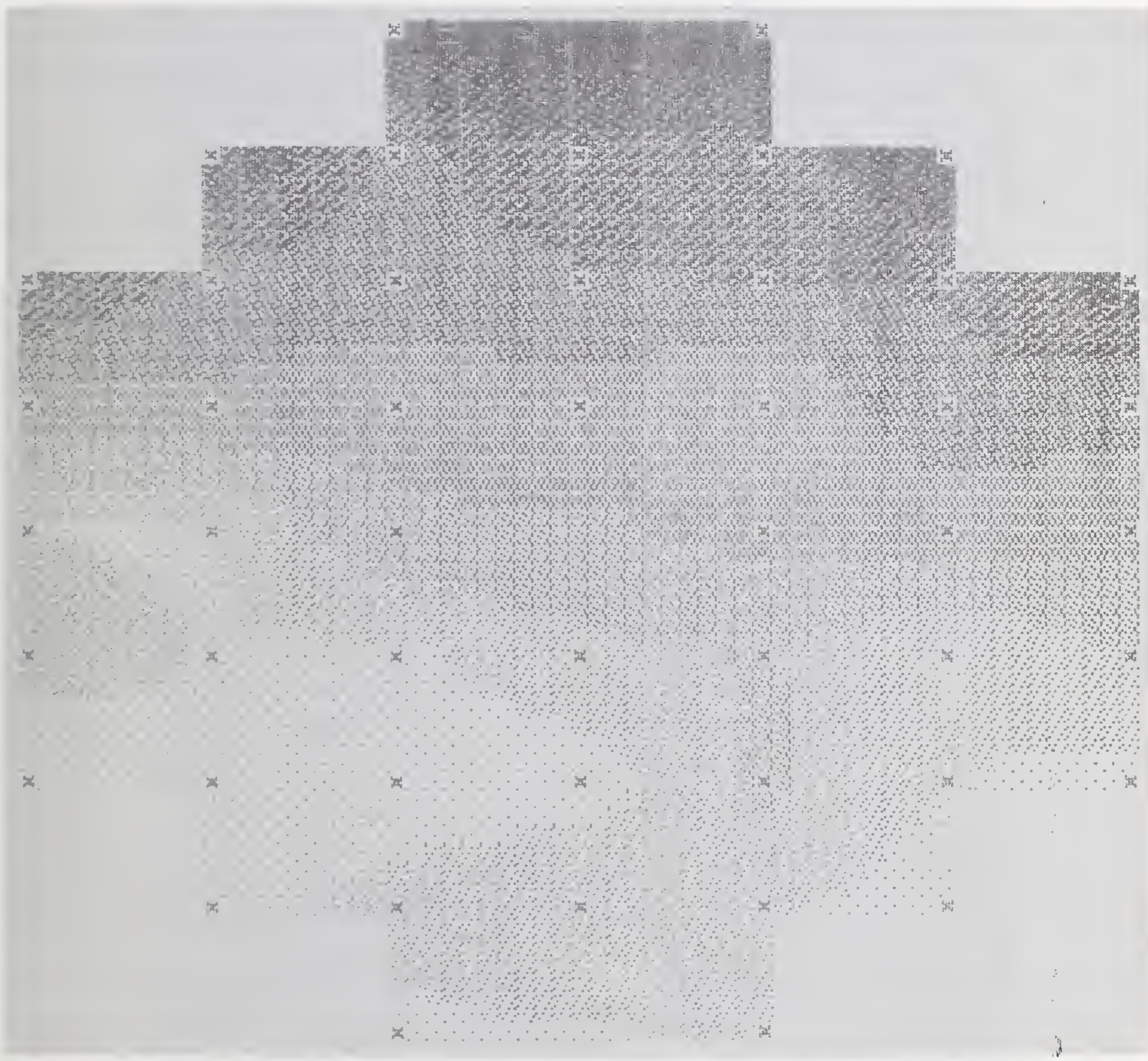


Figure 7. Sample gray-tone map.

PARAMETER VALUE			# SITES
24.07178	TO	24.27020	3
23.87335	TO	24.07178	3
23.67492	TO	23.87335	4
23.47650	TO	23.67492	6
23.27808	TO	23.47650	8
23.07965	TO	23.27808	4
22.88123	TO	23.07965	7
22.68280	TO	22.88123	13

SITES INCLUDED 48

x x x x INCLUDED SITES WITHIN SCALE

+ + + + INCLUDED SITES ABOVE SCALE

\- - - - INCLUDED SITES BELOW SCALE

SAMPLE MEAN 23.28777

SAMPLE STD DEV 0.44481

SAMPLE MEDIAN 23.24650

Figure 8. Key which accompanies map of figure 7.

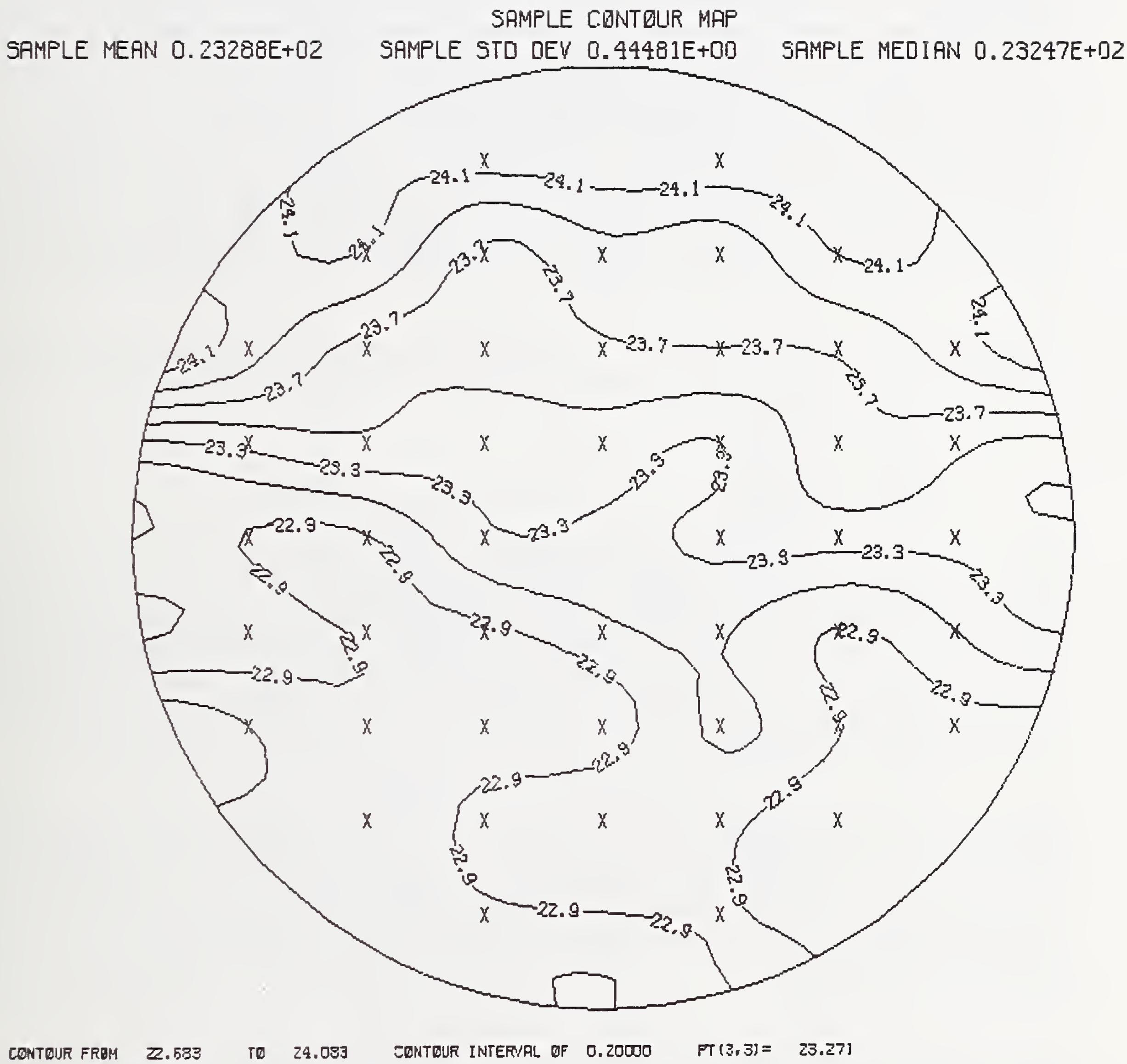


Figure 9. Sample contour map.

The contour map is offered as an option knowing that such maps are used in the electronics industry as a data analysis tool [11]. Because NBS cannot make a commitment to support the NCAR plotting package, the level of support available for MP3 may not be as great as for the rest of the program in the long term.

HIS,P1,P2,P3,P4 - Draw a Histogram. Draws a histogram to display the distribution of data values for the included sites. The histogram is 50 character positions high and so can accommodate up to 50 sites in a bin after which a "cell fraction" or normalized display is produced. P1 is the scaling mode and behaves in the same manner as P1 of the PLT command. P2 is the number of bins in the histogram, which can be any integer in the range 2 to 50, inclusive. P3 and P4 are, respectively, the lower and upper bounds of the histogram for the case P1 = 0. When P1 is nonzero, P3 and P4 must be present but are ignored. The histogram labeling includes number of sites in each bin, fraction of sites in each bin, cumulative distributions going from minimum to maximum and from maximum to minimum values, bin boundaries, and number of sites beyond the bin boundaries.

Either PRS (P1 >= 0) or XOL (P1 < 0) must be executed prior to HIS with no intervening site exclusions or reinsertions or change in N value.

A sample histogram is shown in figure 10.

(8) Data Base Operations: WDB, LDB, GET, SDB, and DEL

WDB,P1,P2,P3,P4,P5 - Write to Data Base. Writes an entry in the data base where P1 is the pattern number, P2 is the lot number, P3 is the wafer number, P4 is the device number, and P5 is the parameter code to be associated with the entry. These must be non-negative integers having a maximum value of 999 for P1, P2 and P3, or 9999 for P4 and P5. The WDB command must be preceded by REA with no intervening GET or other WDB commands. After the entry has been written, the label and data information are printed on the user's terminal and the log file.

LDB,P1,P2,P3 - List Data Base. Produces a listing of data base entries beginning with entry number P1 and ending with entry number P2. If P3 is zero, only the label information is printed; if P3 is nonzero, the data values in the sample are also printed. P2 must not be less than P1 or an error condition occurs. If P1 is greater than the largest entry number, an error condition occurs. If P2 is greater than the largest entry number, a warning message is printed.

P2 is a privileged parameter; that is, it can optionally be represented by only a minus sign. When P2 is a minus sign, LDB lists all records beginning at entry number P1 through the last entry.

When P3 is such that data values are also displayed, the data values immediately follow the label information, six data values per line in E-format. An asterisk immediately following a data value indicates that the particular site was excluded.

```
 INTERVAL      CUM.    1-CUM.   CELL    NO.                    NUMBER OF OBSERVATIONS
 MID POINT    FRACT.   FRACT.   FRACT.  OBS.
                                              0        10        20        30        40        50
------------------------------------------    +---------+---------+---------+---------+---------+
 22.7357       0.104    1.000    0.104    5    +++++
 22.8415       0.292    0.896    0.188    9    +++++++++
 22.9474       0.375    0.708    0.083    4    ++++
 23.0532       0.458    0.625    0.083    4    ++++
 23.1590       0.479    0.542    0.021    1    +
 23.2648       0.521    0.521    0.042    2    ++
 23.3707       0.646    0.479    0.125    6    ++++++
 23.4765       0.667    0.354    0.021    1    +
 23.5823       0.771    0.333    0.104    5    +++++
 23.6882       0.833    0.229    0.063    3    +++
 23.7940       0.875    0.167    0.042    2    ++
 23.8998       0.875    0.125    0.000    0
 24.0056       0.917    0.125    0.042    2    ++
 24.1115       0.958    0.083    0.042    2    ++
 24.2173       1.000    0.042    0.042    2    ++

HISTOGRAM LOWER BOUND  =  0.226828E+02        POINTS BELOW BIN  1  =   0
HISTOGRAM UPPER BOUND  =  0.242702E+02        POINTS ABOVE BIN 15  =   0
BIN WIDTH  =   0.105827E+00
```

Figure 10. Sample histogram.

GET,P1 - Get Entry Number. The data base entry number P1 is defined as the reference data sample. P1 must be a legal entry number not marked deleted or an error condition occurs. If GET is successful, the characteristics of the reference data sample are displayed.

SDB,P1,P2,P3,P4,P5,P6 - Search Data Base. Correlates the reference data sample with other entries in the data base. Parameters P1 through P5 are privileged parameters, such that they can optionally be represented in the command by a minus sign. If these are all minus signs, then correlation coefficients for all entries in the data base are calculated. When P1 is not a minus sign but is a numerical value, then correlation coefficient is calculated only for those data base entries for which the pattern number is P1. Similarly, P2 set to a numerical value restricts correlation coefficient calculation to entries for which the lot number is P2. Similarly, P3, P4 and P5 can be used to specify a particular wafer number, device number, and parameter code, respectively, for which correlation coefficient is to be calculated. Considerable flexibility is thus available for selectively searching a data base for possible correlations.

The sample correlation coefficient is calculated for all entries which are selected by parameters P1 through P5. Those entries for which the correlation coefficient is greater than or equal to P6 are printed. P6 provides a way of screening out entries for which correlation is poor. If P6 is zero, all results of calculations are printed.

If any of the data values in the reference data sample or the samples being searched represent excluded sites, these data values are not included in the correlation coefficient calculation. The printed output of SDB includes the number of site pairs included in each calculation. There is also a summary printed of the number of data base entries searched, number of entries selected for correlation coefficient calculation (matches), and number of results printed. A GET command must be executed prior to the SDB command.

DEL,P1 - Delete. Marks the data base entry number P1 as deleted by placing an asterisk in the DLT field of the label record. Such an entry is physically in the data base and is recognized by LDB commands but is not recognized by SDB commands, causes an error condition when referenced by a GET command, and prints a warning message when referenced by a DEL command. If DEL executes successfully, the entry which has been marked deleted is displayed.

Error Messages

STAT2 contains numerous error messages intended to inform the user of error conditions and to prevent an error from causing the program to terminate abnormally. There are also warning messages which inform the user of various conditions but which do not interrupt program execution. Error messages begin with three asterisks, whereas warning messages begin with three exclamation points. In this section, the various messages are given in alphabetical order along with a brief description of each message. Lower case letters in the messages represent numerical values which would appear in the actual messages.

*** ASG CALLED IN FCALL
The command processor has been altered so that ASG commands are not properly processed.

*** COMMAND CONTAINS MORE THAN 8 PARAMETERS
More than nine delimiters have been detected on the command line.

*** COMMAND HAS TOO FEW PARAMETERS
The number of parameters following the mnemonic is less than the required number.

*** COMMAND HAS TOO MANY PARAMETERS
The number of parameters following the mnemonic is more than the required number.

*** CORRELATION COULD NOT BE DONE FOR ENTRY NUMBER i
Correlation coefficient could not be calculated because too many data values in the pair of data samples represented excluded sites.

*** DATA FILE CONTAINS ONLY i DATA SETS
The data set requested by a REA,3,P2 command does not exist. The set number exceeds the number of sets in the file.

*** DATA POINT NOT YET EXCLUDED
The AIP command cannot be applied to a site which has not first been excluded.

!!! DEFAULT HELP FILE COULD NOT BE OPENED
This error probably results from the help message files having excessive protection or being in the wrong subdirectory.

*** DELIMITER CANNOT HAVE MORE THAN ONE COMMA
A delimiter has been found to contain more than one comma.

*** DENOM IS 0.0000 IN R
This error is produced by an SDB command when all data values in either of the respective samples are equal.

*** END OF COMMAND INPUT FILE
A command file has reached end-of-file without MAC having been executed. This would probably result from an attempt to initiate STAT2 directly from a command file.

!!! END OF MACRO FILE
The last command in a macro command file has executed and control is returned to the user terminal.

*** ENTRY NUMBER i MARKED DELETED
An entry referenced by GET cannot be accessed because it is marked deleted.

!!! ENTRY NUMBER i MARKED DELETED
An attempt was made to delete an entry already marked deleted.

*** ENTRY NUMBER i NOT FOUND
An entry number which does not exist has been referenced by GET or DEL.

*** ERROR IN READING {LABEL, HEADER, DATA} RECORD i
A data base command was not able to successfully read a label, header, or data record.

*** ERROR IN STEND ARRAY
The STEND array read from an input file in format 1 is illegal. Refer to comments in subroutine VSTEND beginning on page II-54 of Appendix II for further discussion.

*** ERROR IN WRITING {LABEL, HEADER, DATA} RECORD i
The WDB or DEL command was not able to successfully write a label, header, or data record.

*** EXPONENT TOO LARGE IN PARAMETER i
The ith parameter has an exponent whose magnitude is greater than 32.

*** FORMAT ERROR
The input data file is not in the proper format or has an access method which differs from the access method expected by the REA command. This message is printed whenever an error occurs in a READ or REWIND statement associated with the REA command.

*** FORMAT TYPE MUST BE IN THE RANGE 1 to 3
The first parameter of the REA command has been given an illegal value.

*** GET MUST PRECEDE SDB
A reference data sample must be specified by GET before SDB can be executed.

*** HEADER FIRST BYTE IS "b" RATHER THAN "H".
The pointers which link the label file with the data file have been corrupted. Refer to the section titled Data Base Structure. The "b" represents the actual character which appears in the first byte position.

*** HEL CALLED IN FCALL
The command processor has been altered so that HEL commands are not properly processed.

*** IFLAG = i IN B2INK
An error has occurred in subroutine B2INK so that a contour map cannot be drawn. This error should not occur and probably indicates a problem with portability or a user modification.

*** IFMT = i IN REA
An illegal format number has been detected in subroutine REA. This probably means that line labeled 14 was not properly modified when a new format was added.

*** ILLEGAL CHARACTER IN POSITION i OF PARAMETER j
The command contains an illegal character such as an alphabetic character or punctuation symbol.

*** ILLEGAL NEGATIVE PARAMETER
One of the first five parameters of the SDB or WDB commands has been given a negative value.

*** ILLEGAL NESTED MAC COMMAND
A command file, invoked by a MAC command, cannot contain a MAC command.

*** ILLEGAL P VALUE IN NORPPF
This message originates in a subroutine called by XOL but is not expected to occur so long as DELTA is in the range 0.05 to 0.90 inclusive.

*** ILLEGAL TERMINAL COMMA
The command ends with a comma.

*** ILLEGAL VALUE FOR DELTA
The parameter of the XOL command as specified lies outside the range 0.05 to 0.90.

*** INPUT FILE CONTAINS {INSUFFICIENT, TOO MUCH} DATA
An input file in format 1 does not have the proper number of data values to satisfy the requirements of the STEND array.

*** INPUT FILE NOT YET OPENED
REA cannot be executed until an input data file has been assigned by ASG.

*** INVALID RANGE, RANGE 1 > RANGE 2
The range of entry numbers of the LDB command has the first entry number greater than the last entry number.

*** ISITES = i--SIGMA CANNOT BE CALCULATED
There are fewer than two included sites, making calculation of SIGMA impossible.

*** LABEL ENTRY = i, {HEADER, DATA} ENTRY = j.
The pointers which link the label file with the data file have been corrupted. Refer to the section titled Data Base Structure.

*** LABEL FILE HAS ILLEGAL SAMPLING PLAN CODE
The sampling plan code in the first record in the label file has an illegal value. Refer to the section titled Data Base Structure.

!!! LAST ENTRY IS NUMBER i
The second parameter of the LDB command refers to a nonexistent entry number.

*** MAC CALLED IN FCALL
The command processor has been altered so that MAC commands are not properly processed.

*** MATRIX APPEARS SINGULAR
An FPL or FQD command is unable to execute because of a singularity in a matrix which must be inverted. This condition should not occur when operating on real data.

!!! MEDIAN IS BETWEEN i AND j
The algorithm of subroutine MEDCAL did not find a unique median data value and an approximation is being reported.

*** MUST EXECUTE PRS TO GET CURRENT STATISTICS
A message produced by several commands whenever the maximum, minimum, and standard deviation values have not yet been calculated for the present population and N value.

*** NCELL MUST BE BETWEEN 2 AND 50
The number of cells or bins in the histogram must be in the range 2 to 50, inclusive.

*** NEGATIVE VALUE OF N NOT AllOWED
An illegal value of N has been specified in the ENN command.

*** NO CHARACTERS SHOULD FOLLOW MNEMONIC
The command should have no parameters associated with it, but it has been represented as other than the three-letter mnemonic.

*** NOCNLN MUST BE BETWEEN 5 AND 40
The number of lines requested for a contour map must be in the range 5 to 40, inclusive.

!!! NO DATA BASE {LABEL, DATA} FILE OPENED
These messages appear when STAT2 is started if the VMS ASSIGN command has not been used to properly specify the data base files.

*** NO DATA FILE HAS YET BEEN READ
An input data file must be read by the REA command before the requested operation can be executed.

*** NO DATA POINT AT SPECIFIED ROW-COLUMN
A row-column location at which there is no test site has been specified by a XIP, LIP, IIP, or AIP command.

*** NO ENTRIES IN REQUESTED RANGE
The range of entry numbers specified in the LDB command contains no entries.

*** NO FIT MADE FOR CURRENTLY INCLUDED SITES
The SPL or SQD command must be preceded by an FPL or FQD command, respectively, with no intervening site exclusions or reinsertions.

*** NO VALUES TO PLOT HISTOGRAM
A range has been specified which contains no data values.

*** NSUBR = i IN FCALL
A new command has been added without properly modifying subroutine FCALL.

*** NUMBER OF SITES EXCEEDS MAXIMUM OF 256
Sampling plan 4, which includes all sites in the sample, has been applied to a DATA array containing more than 256 sites.

*** PARAMETER i HAS MORE THAN 20 CHARACTERS
The ith parameter of the command has been represented by more than 20 characters.

*** PARAMETER i HAS NO DIGITS FOLLOWING E
The exponent of an E-format number has been omitted.

*** PARAMETER i HAS NO DIGITS PRECEDING E
The base of an E-format number is missing. This message also appears if a parameter consists of only a decimal point or plus sign or if a nonprivileged parameter consists only of a minus sign.

*** PARAMETER i MUST BE AN INTEGER
A numerical value i has been entered for a parameter where an integer is required.

*** PARAMETER MUST BE AN INTEGER > + 1
A parameter of GET, DEL, or LDB as entered is not an integer greater than 1 and so is not a legal entry number.

*** PARAMETER MUST BE A POSITIVE INTEGER
One of the two parameters of the LDB command or the second parameter of the REA command has been given an illegal value.

*** PARAMETER VALUE TOO LARGE
In a WDB command one of the parameters exceeds the maximum allowable value. The limit is 999 for P1, P2 and P3 and 9999 for P4 and P5.

*** PARAMETER 6 MUST BE IN THE RANGE 0 TO 1
The sixth parameter of the SDB command has been given an illegal value.

*** PAUSED ON ERROR IN REMOTE MODE
An error has occurred while in remote mode, causing STAT2 to pause.

!!! PLOT DIAMETER SET TO 8 INCHES
Under the present revision, contour map diameter must be 8 inches. If another diameter is specified, the diameter is forced to 8 inches and this message is printed.

*** PLOT WIDTH OR HEIGHT TOO BIG OR TOO SMALL
An improper size has been entered for width or height in the MP1 or MP2 commands. See the appropriate command description for the allowed range of values.

*** RECORD FIRST BYTE IS "b" RATHER THAN "D".
The pointers which link the label file with the data file have been corrupted. Refer to the section titled Data Base Structure. The "b" represents the actual character which appears in the first byte position.

*** RLABEL ERROR: LREC = i, ENT = j
The label record number and entry number do not agree.

!!! SITE AT ROW i COL j ALREADY INCLUDED
The IIP command is being applied to an included site.

!!! SITE AT ROW i COL j PREVIOUSLY EXCLUDED
A command has been issued which would exclude a site which has already been excluded. The status of the site is not affected.

*** SPACE OR COMMA MUST FOLLOW MNEMONIC
Either a command has been entered for which required parameters are not present, or a character other than comma or space has been used between the mnemonic and the first parameter.

*** STEND ARRAY FAULT IN MP3
The STEND array is such that column 1 is empty.

*** TOO MANY CHARACTERS IN EXPONENT OF PARAMETER i
A parameter exponent contains more than two digits.

*** TYPE = i AT SITE (j, k)
A topological type outside the range -1 to 9 has been detected by subroutine AXP. This condition probably is caused by portability problems.

*** TYPE = i IN MP1 OR MP2
A topological type outside the range -1 to 9 has been detected in subroutine MP1 or MP2. This error should not occur and may be caused by portability problems.

*** UNABLE TO OPEN FILE
The file specified in a MAC or ASG command cannot be opened.

*** UNDEFINED MNEMONIC
The command mnemonic does not correspond to any of the legal mnemonics.

*** UPPER SCALE VALUE < = LOWER
If autoscale is not enabled in a histogram or mapping command, then the upper scale value must be larger than the lower scale value.

*** WDB ILLEGAL AFTER GET OR WDB OR BEFORE REA
An attempt has been made to write a data base entry under improper conditions. WDB is intended to be executed no more than once after a given REA command with intervening data point exclusions.

*** WRONG FILE TYPE -- RE-ISSUE ASG
An attempt was made to read an input data file in format 3 but the file is not a direct access file. The ASG command must be entered again.

*** XOL K VALUE IS NOT CURRENT
The autoscale option of MP1 or MP2 with P3 negative or of MP3 with P2 negative or of PLT or HIS with P1 negative has been invoked, but subroutine XOL has not first been executed with no intervening site exclusions or reinsertions or change in N value.

Addition of New Input Data Formats

Users may want to add the capability to read additional input data formats using the REA command. Instructions for doing so are given in the program listing in the introductory comments to subroutine REA, beginning on page II-38 of Appendix II.

Addition of New Commands

Users may find it desirable to add one or more new commands to STAT2. Instructions for adding new commands are contained in the listing of STAT2 in Appendix II beginning on pages II-6, II-11, and II-16.

The CORTABLE Program

The SDB command as described above permits a user to calculate the correlation coefficient of a reference sample and a selected group of other entries. It is often desirable to correlate a group of entries against each other in all possible combinations. CORTABLE is a FORTRAN program which does this. A listing of CORTABLE is given in Appendix IV. When CORTABLE is run, the user enters up to 20 entry numbers. Nonexistent, duplicate, or deleted entries are not allowed. CORTABLE prints the label information associated with each entry as the entry number is typed. An entry number of 0 signals CORTABLE that all entry numbers have been entered. CORTABLE then creates two triangular displays, one containing all the correlation coefficients, and one containing the number of data pairs used in the correlation coefficient calculation. These displays are written to FORØ14.DAT which may be printed after the run terminates. Before CORTABLE is run, the user must assign the data base label and data file specifiers to FORØ17.DAT and FORØ18.DAT, respectively. Sample displays produced by CORTABLE are shown in figure 11.

Program Installation

The eight pieces of software needed to make STAT2 run as described are (1) STAT2.FOR, the FORTRAN source representation of STAT2; (2) the help message files; (3) CRDB.FOR, the stand-alone program for creating a data base; (4) CORTABLE.FOR, the stand-alone program for making a correlation table; (5) NCAR.OLB and (6) UTILITY.OLB, the NCAR plotting software libraries; (7)

TABLE OF CORRELATION COEFFICIENTS

	3	4	5	6	7	8	9	10	11	12
3	1.00									
4	-0.70	1.00								
5	-0.87	0.91	1.00							
6	0.02	-0.02	-0.03	1.00						
7	0.02	-0.25	-0.03	-0.04	1.00					
8	-0.05	0.25	0.03	0.02	-1.00	1.00				
9	0.00	0.02	-0.01	0.05	-0.04	0.07	1.00			
10	0.80	-0.54	-0.82	0.03	-0.47	0.46	0.04	1.00		
11	1.00	-0.70	-0.87	0.02	0.02	-0.05	-0.02	0.80	1.00	
12	-0.70	1.00	0.91	-0.02	-0.25	0.25	0.02	-0.54	-0.70	1.00

NUMBER OF POINTS USED TO CALCULATE CORRELATION COEFFICIENT

	3	4	5	6	7	8	9	10	11	12
3	47									
4	46	49								
5	46	48	49							
6	47	49	49	50						
7	46	48	48	49	49					
8	45	47	47	48	47	48				
9	47	49	49	50	49	48	50			
10	47	49	49	50	49	48	50	50		
11	47	49	49	50	49	48	50	50	50	
12	47	49	49	50	49	48	50	50	50	50

Figure 11. Sample output of CORTABLE showing sample correlation coefficients (above) and number of point pairs in the sample (below).

MCVAX.EXE, the NCAR program for printing plot files; and (8) CMLIB.OLB, the CMLIB library needed for the functional fits and, along with the NCAR software, for making contour maps. If any of the last four pieces is not available, contour plots cannot be made. Subroutine MP3 should then be replaced with a stub which simply prints a "command not available" message and returns. Similarly, if CMLIB is not available, the FPL and FQD subroutines should be replaced with stubs. CORTABLE is an optional program which, if available, can be compiled and linked by itself to create an executable task. CRDB is necessary for working with data bases. It needs only to be compiled and linked by itself to create an executable task.

The help message files must all be contained in the same directory or subdirectory, must have a HLP extension, and must have three-character filenames which match the command mnemonics. COM.HLP and SYN.HLP must also be present. As written, STAT2 looks for help messages in DRB1:[USERLIB.STAT2HELP]. This is specified on page II-23 in the DATA statement of subroutine HELP. Users at other installations would have to alter this.

The STAT2 program itself must be altered with respect to the help message file subdirectory, then compiled, then linked together with NCAR.OLB, UTILITY.OLB, and CMLIB.OLB. STAT2 should then be ready to run.

When configured with the NCAR libraries and CMLIB, STAT2 requires approximately 570 KB of memory to run. A virtual memory operating system can accommodate a program of this size with no difficulty. In other environments, overlays may have to be used. In an earlier version of STAT2 [2], the program was broken into a main segment and seven overlays. Depending on user needs, large portions of STAT2 could be removed and replaced with stubs. For example, a user may not want all three map options or functional fits or data bases.

Logical Unit Assignments

The logical unit assignments used by STAT2 are given below. The logical unit number precedes the description. As the parenthetical notes indicate, not all logical units may be needed for a given set of operations.

- 5. Command input, normally a CRT terminal. When a MAC command is entered, logical unit 5 is assigned to the specified command file.
- 6. A CRT terminal for command echoes and other output.
- 10. Help message file.
- 11. Input data file (not used if only data base operations are to be done).
- 14. STAT2.LOG, the program log file, to be sent to the line printer by the PRINT command following STAT2 program termination.
- 17. Data base label file (used only for data base operations).
- 18. Data base data file (used only for data base operations).
- 19. Scratch file used for constructing the gray-tone map. Record length must be 132 bytes (not used if maps are not being generated).

Acknowledgments

The author is indebted to Ruth Zucker for writing the contour map code, to Larry J. Till for writing an earlier version of CORTABLE, to Ronald C. Frisch for the original gray-tone map, to Geoffrey McFadden who brought to his attention the idea of modifying the CONREC subroutine in the NCAR plotting package for circular plotting, to Ronald F. Boisvert for subroutines B2INK and B2VAL, to Colleen H. Ellenwood for work on the histogram, to Charles P. Reeve for assistance with the functional fits, and to Loren W. Linholm, John F. Barkley, Kenneth F. Galloway, and Walter S. Liggett, Jr. for careful readings of and helpful suggestions regarding the manuscript. The manuscript was prepared with proficiency by E. Jane Walters and Jo S. Halapatz.

References

1. Linholm, L. W., Mattis, R. L., Till, L. J., Frisch, R. C., and Reeve, C. P., Analysis of PVW Data, Semiconductor Progress Briefs, NBSIR 81-2230, W. M. Bullis, ed., pp. 4-6, March 1981.

2. Mattis, R. L., Till, L. J., and Frisch, R. C., A Computer Program for Analysis of Data from Microelectronic Test Structures, NBSIR 82-2492 (June 1982).

3. Linholm, L. W., Mattis, R. L., Frisch, R. C., and Reeve, C. P., Characterizing and Analyzing Critical Integrated Circuit Process Parameters, *Semiconductor Silicon/1981*, H. R. Huff, R. J. Kriegler, and Y. Takeishi, eds., pp. 906-920 (ECS, Pennington, New Jersey, May 1981).

4. Carver, G. P., Linholm, L. W., and Russell, T. J., The Use of Microelectronic Test Structures to Characterize IC Materials, Processes, and Processing Equipment, *Solid State Technology* 23 (9), 85-92 (September 1980).

5. Odeh, R. E., and Evans, J. O., The Percentage Points of the Normal Distribution, Algorithm AS70, Applied Statistics, *J. Royal Statistical Soc.* 23 (1), 96-97 (1974).

6. Snedecor, G. W., and Cochran, W. G., *Statistical Methods,* 6th edition, chapter 2 (Iowa State University, 1967).

7. *Ibid.*, chapter 7.

8. Suehle, J. S., An Experimental Study of Wafer Level Acceptance Testing Based on Sampling Test Chip Data, Master's Thesis, University of Maryland, 1982.

9. Core Mathematics Library, available from Center for Applied Mathematics, Scientific Computing Division, National Bureau of Standards, Washington, DC 20234.

10. The SCD Graphics Utilities, Technical Note NCAR-TN/166+IA, G. R. McArthur, ed., Scientific Computing Division, National Center for Atmospheric Research, Boulder, CO. For information on availability, contact Ray Bovet, NCAR, (303) 497-0610.

11. Perloff, D. S., Gan, J. N., and Wahl, F. E., Dose Accuracy and Doping Uniformity of Ion Implanation Equipment, *Solid State Technology* 24 (2), 112-120 (February 1981).

Appendix I

Command Summary

Following is an alphabetical list of the command mnemonics, each accompanied by a phrase describing the function and the page number at which the command description begins. In the phrases, point refers to a test site.

AIP - Alter an individual point, p. 19.
ASG - Assign input data file, p. 14.
AXP - Alter excluded points, p. 19.
DEL - Delete data base entry, p. 30.
DIS - Display distribution, p. 16.
END - Terminate STAT2 execution, p. 14.
ENN - Set N to a specified value, p. 16.
ERM - Error message switch, p. 14.
FPL - Fit DATA array to a plane, p. 19.
FQD - Fit DATA array to a quadratic function, p. 20.
GET - Define data base entry as reference sample, p. 30.
HEL - Help request, p. 14.
HIS - Draw a histogram, p. 28.
IIP - Include an individual point, p. 19.
LAP - List all points, p. 18.
LDB - List data base entries, p. 28.
LIP - List an individual point, p. 18.
LNS - List points beyond N*SIGMA from mean, p. 18.
LXP - List excluded points, p. 19.
MAC - Execute command macro, p. 14.
MP1 - Draw numerical map of DATA array, p. 20.
MP2 - Draw gray-tone map of DATA array, p. 21.
MP3 - Draw contour map of DATA array, p. 24.
PAU - Pause STAT2 execution, p. 14.
PLT - Draw character display of DATA array, p. 20.
PRS - Print statistics, p. 16.
REA - Read input data file, p. 14.
REM - Set or reset remote mode, p. 14.
RES - Restore all points to included status, p. 19.
SDB - Search data base for correlations, p. 30.
SPL - Subtract fitted plane from DATA array, p. 19.
SQD - Subtract fitted quadratic function from DATA array, p. 20.
WDB - Write data base entry, p. 28.
XGT - Exclude points greater than a value, p. 18.
XIP - Exclude an individual point, p. 19.
XLT - Exclude points less than a value, p. 18.
XNS - Exclude points beyond N*SIGMA from mean, p. 18.
XOL - Exclude outliers, p. 18.
XPP - Exclude peripheral points, p. 18.

Appendices II, III, and IV

are located in microfiche form

in the pocket of the inside back cover.

Originally published in *Semiconductor Silicon/1981*, H. R. Huff, R. J. Kriegler, and Y. Takeishi, Eds., pp. 906-920 (Electrochemical Society, Pennington, N.J., May 1981).

CHARACTERIZING AND ANALYZING CRITICAL INTEGRATED CIRCUIT PROCESS PARAMETERS*

by

L. W. Linholm, R. L. Mattis, and R. C. Frisch
Electron Devices Division
and
C. P. Reeve
Statistical Engineering Division

National Bureau of Standards
Washington, DC 20234

ABSTRACT

Microelectronic test structures are frequently used to measure the degree of process control in developmental integrated circuit processes. Test results from these structures must be obtained and interpreted in a timely fashion in order to be used for correcting or improving the process. This paper describes techniques for determining and displaying critical process parameters in forms convenient for characterizing the intrawafer variation of these parameters.

INTRODUCTION

With the increasing complexity of integrated circuits, it is becoming more difficult for both the manufacturer and user to fully characterize circuit performance. Functional testing alone is an impractical approach for evaluating complex circuits. As a result, greater utilization is being made of microelectronic test structures which provide clear and unambiguous test results for characterizing the integrated circuit fabrication process (1).

In a developmental integrated circuit process, test structures are used to identify which parameters accurately predict or determine the degree of process control; to establish the value and range of these parameters for a given process lot; and to determine how these parameters vary across an integrated circuit die, across a wafer, from wafer to wafer, and from lot to lot (2-5). Test results must be ob-

*This work was conducted as a part of the NBS program on Semiconductor Measurement Technology. Portions of this work were supported by the Air Force Wright Aeronautical Laboratory and by the Defense Nuclear Agency. Not subject to copyright.

tained and interpreted in a timely fashion in order to be used for correcting or improving the process.

This paper describes analytical techniques for identifing test results from nondefective structures, estimating parameter correlations, and presenting results graphically. These techniques can provide the user with a relatively fast analysis capability for characterizing an integrated circuit process through the determination of the magnitudes of baseline parameters and their variation over the wafer for "properly" fabricated devices. It is assumed that the process being characterized is in sufficient control to produce a high percentage of "properly" fabricated test structures and that defective structures which are encountered are mainly the result of gross defects introduced by handling, by lithography voids, or by similar process irregularities.

A laboratory-based minicomputer-controlled test system is used both to measure selected structures found on a process validation wafer (PVW) and to analyze the resulting data. After identifying and excluding test results from test structures considered to be defective, the mean, median, and standard deviation are calculated for the remaining data set. Further analysis is done to identify possible correlations between critical process parameter data sets. These sets are then displayed as wafer maps to provide graphical illustrations of parameter variation over the wafer.

In the next section the data analysis techniques will be described. An example will then be presented where the techniques are used to analyze a serious process problem.

DATA ANALYSIS TECHNIQUES

The characterization and analysis of a given parameter is performed with a computer program named STAT2. STAT2 is an interactive program, written primarily in FORTRAN, which can analyze a set of data for one parameter. The analysis includes (1) calculation of mean, median, and standard deviation of all data within the set; (2) identification and removal from the data set of test results from structures suspected of containing defects; (3) entry of a 13-point sample of the data set into a data base for use in determining correlations with other data sets; and (4) production of a wafer map in which the parameter variations are displayed as a gray-tone illustration.

To characterize baseline electrical parameters, it is necessary to identify test results from defective structures or defective measurements which did not accurately represent the parameter being measured. The inclusion of data from these structures would result in an incorrect determination and analysis of baseline electrical parameters. Data are initially excluded from the main data set if they could easily be identified as coming from a defective test structure, e.g., a structure with an open or short between test points. Identification of defective structures in the remaining data base is very difficult without either additional electrical or visual information which requires additional time to obtain or precise fault simulation models which provide an accurate description of the interactions between fault occurrence and measured test results.

After excluding data from the main set for reasons previously described, the remaining measurement data (y_1, y_2, ... y_N) are assumed to be normally distributed with a relatively high occurrence of outliers (up to 20 percent). The outliers are occasionally of a large magnitude and are not necessarily distributed symmetrically about the mean. A datum y_i is rejected as an outlier if:

$$|y_i - \tilde{\mu}| > K \sigma, \tag{1}$$

where $\tilde{\mu}$ and $\tilde{\sigma}$ are the mean and standard deviation calculated from measurements at the included sites (those sites which have not already been identified as outliers), and K is a value to be determined. In order to determine K, the experimenter must specify p, the probability with which he is willing to reject at least one "good" value from N included sites. The value of K satisfies the equation

$$\int_{-\infty}^{K} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = \frac{1 + (1 - p)^{1/N}}{2}, \tag{2}$$

involving the standard normal distribution (6). The value of K is numerically computed using an algorithm for the percent point function of the standard normal distribution (7). Knowing K, outliers are identified using eq (1) and excluded. If any points are excluded, new values of $\tilde{\mu}$ and $\tilde{\sigma}$ are calculated based on the new population, N', of included sites. A new value of K is calculated for N' (p is held constant). The procedure is repeated until no new outliers are identified. The number of iterations required depends on the selected value of p. In this work, p = 0.2 was determined to be a reasonable value based on experience using realistic data. Further techniques for robust outlier detection can be found elsewhere (8,9).

The data sets are then analyzed to identify possible spatial correlations between sets. When the paired observations (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) are taken on two quantitites, if a large value of x implies a large value of y, then the quantities are said to be positively correlated. If a large value of x implies a small value of y, then the quantities are said to be negatively correlated. If a large value of x implies nothing about y, then x and y are said to be uncorrelated. The measure of correlation is the correlation coefficient, ρ, which is estimated by the statistic r:

$$r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^{n} (x_i - \bar{x})^2\right]\left[\sum_{i=1}^{n} (y_i - \bar{y})^2\right]}}, \tag{3}$$

where $\bar{x}$ and $\bar{y}$ are the sample means of x and y, respectively, over the n points (8). Note that r must take on values in the range [-1,1].

The statistic, r, is calculated from a 13-point data sample from paired sets to serve as a screen or indicator of possible correlation between parameters measured on the same wafer or on different wafers. The data contained in the 13-point sample are from the selected test sites shown in figure 1. A set of 13 was determined to be a reasonable compromise between keeping sufficient information to characterize the spatial parameter variation and minimizing data storage requirements.

Often it is of interest to know whether the computed value of r is significantly different from zero (or some other number). If the (x,y) pairs are from a bivariate normal population, then a confidence interval can be computed using the Fisher z-transformation [10]. Consider the variance-stabilizing transformation function

$$f(r) = \frac{1}{2}\ln\left(\frac{1 + r}{1 - r}\right) \tag{4}$$

and its inverse

$$g(z) = \frac{e^{2z} - 1}{e^{2z} + 1} . \tag{5}$$

The value $z = f(r)$ is approximately normally distributed with variance $1/(n - 3)$, thus a $100(1 - \alpha)$ percent confidence interval for z can be constructed of the form

$$(z_l, z_u) = \left(z - \frac{k}{\sqrt{(n - 3)}} , z + \frac{k}{\sqrt{(n - 3)}}\right) \tag{6}$$

where subscripts l and u represent the lower and upper bounds of the confidence interval, and k is the $(1 - \alpha/2)$ critical point of the standard normal distribution. For example, for the case of a confidence interval of 99 percent, 99 percent of the points in a normal distribution lie within 2.58 standard deviations of the mean; therefore, k in this example would be 2.58. Using the inverse transformation g, the 100(1 - α) percent confidence interval for ρ can be constructed as

$$(r_l, r_u) = [g(z_l), g(z_u)] \ . \tag{7}$$

For example, for a calculation based on 13-point pairs which yields r equal to 0.70, the 99-percent confidence interval for ρ would be [0.051, 0.933]. It may be concluded that the correlation is statistically significant because the interval does not include zero.

Data from sets containing possible correlation are displayed as wafer maps. The wafer map provides a graphic illustration of the spatial parameter variations over the wafer. A map, shown in figure 2, uses an eight-level gray scale to represent parameter values. The height and width of the display as well as the maximum and minimum values to be plotted can be selected. The map is made on a line printer, each data point being represented by a 5 by 7 dot symbol. In between data point locations, other symbols are placed with the shade of gray determined by interpolation, thus producing a continuous wafer map. Each actual data point location is represented on the map by an "x," if the parameter value is greater than the maximum plotted value by a "+", and if the parameter value is less than the minimum plotted value by a "-".

By using these techniques, it is possible to quickly examine large quantities of test data. Analysis of selected data sets can lead to the identification of previously unknown process problems or a hypothesis as to the cause of known problems. The analysis can also serve as a guide for the selection of other measurement techniques requiring more time or specialized analysis equipment. In some cases the identity and physical nature of process problems can be determined.

AN EXAMPLE

This technique was used to analyze data obtained on test pattern NBS-16 (11). This pattern, shown in figure 3, was designed to evaluate a developmental CMOS/SOS silicon gate process. It was implemented into a commercial manufacturing facility as a process validation wafer (PVW) (12,13), a wafer consisting only of identical test patterns. Ninety-five NBS-16 test patterns were fabricated on each 3-in. (76.2-mm) diameter silicon-on-sapphire PVW. One PVW accompanied each production run and was subsequently tested in order to determine the value and range of critical process parameters.

The measurement system used to test the PVWs consists of a laboratory-based minicomputer and associated electrical test instruments. The minicomputer is configured with 544 kilobytes of memory, two 10-megabyte disc drives, two floppy disc drives, a nine-track dual density magnetic tape drive, a system console, several CRT and hard-copy terminals, a line printer, a digital plotter, and a multiuser operating system. The test instruments consist of (1) an automatic wafer prober, (2) a bipolar current supply with 1-μA resolution, (3) two bipolar voltage supplies with 1-mV resolution, (4) an autoranging five-digit digital voltmeter with 1-μV resolution, (5) an autoranging three-digit picoammeter with 1-pA resolution, (6) eight 20-channel scanners, (7) a six-channel autoranging analog-to-digital converter, (8) two digital-to-analog converters, (9) 16 single-pole, single-throw relays, and (10) a digital thermometer with 0.1-K resolution (for reading wafer chuck temperature). All these instruments are digitally programmable. The configuration of the test system is shown in figure 4.

After testing is completed, test results are analyzed using the techniques previously described. Table 1 is a list of the sample correlation coefficients for the 13-point samples from selected parameters on one PVW. From this information, an unexpected correlation is observed between metal-to-n^+ contact resistance and n^+ sheet resistance. These parameters were determined from data taken on a four-terminal contact resistor (14) and a cross-bridge resistor (15), respectively, that were located in adjacent areas of test pattern NBS-16. The magnitude of the sample correlation coefficient, r = 0.76, suggests that the high metal-to-n^+ contact resistance is a function of n^+ sheet resistance or phosphorus concentration.

Sample correlation coefficients between these parameters and other selected parameters were also examined. Because both metal-to-n^+ and metal-to-p^+ contact resistors are adjacent devices, and because the contact window is defined in the same photolithographic process for both structures, variations or problems with contact window photolithography, etching, and subsequent thermal processing are likely to result in similar parameter variations for these structures. Since no apparent correlation was determined, r = 0.01, it was concluded that these processing steps were properly performed. Also, since the contact resistor test structure is a four-terminal kelvin-type structure with current taps separated from voltage taps, the effects of probe-to-probe-pad contact resistance or the series resistance of the epi layer or metal layers connecting the probe pads to the voltage taps do not affect the measurement.

Wafer maps for metal-to-n^+ contact resistance and n^+ sheet resistance were produced and are shown in figure 5. A wafer map of metal-to-p^+ contact resistance is shown in figure 2. Based on the correlation between metal-to-n^+ contact resistance and n^+ sheet resistance and lack of correlation between metal-to-n^+ contact resistance and other parameters, the variation in phosphorus concentration was considered to be the likely cause of metal-to-n^+ contact resistance variation.

The phosphorus concentration of the measured structures was controlled by a two-stage phosphorus implant. The first implant was intended to dope the majority of the epi island region. The second was intended to increase the dopant concentration at the island surface in order to decrease contact resistance in source and drain regions. Both implants were made through a gate oxide which covered the epi island.

It was concluded that due to variations in the gate oxide thickness which were unaccounted for in the process design, the peak of the phosphorus implant varied between the silicon and silicon dioxide depending upon the oxide thickness. This caused significant variations in the amount of phosphorus reaching the silicon surface during the implant and caused the observed variation in metal-to-n^+ contact resistance.

To further support this conclusion, subsequent capacitance measurements on a p-type MOS capacitor were made with a manual test system. The results of these measurements indicated that the gate oxide thickness was greatest in the areas of lowest phosphorus concentration.

Based on the calculated correlation coefficient and associated wafer maps, of test results from a single wafer, specific parameters were identified and further analysis was performed which led to the identification of a serious process problem. The identification and analysis of this processing problem was possible only because both parameter magnitude and test site location were recorded and analyzed in a manner that allowed the rapid spatial correlation of these parameters. Such correlations require enough data to obtain statistically significant results; they cannot be reliably obtained from measurements at only two or three test structures per wafer, as is often done at "drop-in" sites.

SUMMARY

In order to be able to characterize the performance of an integrated circuit process, it is necessary to determine the baseline electrical parameters of the process. The example presented shows that significant variations in these parameters can occur across a wafer.

Statistical correlation techniques and graphical parameter mapping are important tools for analyzing critical parameter variations and identifying process problems in a timely manner from measurements on a single PVW.

ACKNOWLEDGMENTS

The authors are indebted to Larry J. Till for assistance in implementing some of the correlation coefficient software, and to Thomas J. Russell, Martin G. Buehler, and W. Murray Bullis for many helpful suggestions in preparation of the manuscript.

REFERENCES

1. Carver, G. P., Linholm, L. W., and Russell, T. J., The Use of Microelectronic Test Structures to Characterize IC Materials, Processes, and Processing Equipment, Solid State Technology 23 (9), 85-92 (September 1980).

2. Wahl, F. E., and Perloff, D. S., Techniques for the Evaluation and Display of VLSI Process Uniformity, Proc. Microelectronics Measurement Technology Seminar, San Jose, California, February 6-7, 1979, pp. 17-26.

3. Perloff, D. S., Hasan, T. F., and Blome, E. R., Real-Time Monitoring of Semiconductor Process Uniformity, Solid State Technology 23 (9), 81-86 (February 1980).

4. Russell, T. J., Maxwell, D. A., Reimann, C. T., and Buehler, M. G., A Microelectronic Test Pattern for Measuring the Uniformity of an Integrated Circuit Fabrication Technology, Solid State Technology 22 (8), 71-74 (February 1979).

5. Ham, W. E., Test Data Reduction, IEEE Trans. Manufacturing Technology MFT-5, 24-29 (1976).

6. Snedecor, G. W., and Cochran, W. G., Statistical Methods, 6th edition, chapter 2 (Iowa State University, 1967).

7. Odeh, R. E., and Evans, J. O., The Percentage of Points of the Normal Distribution, Algorithm AS 70, Applied Statistics, J. Royal Statistical Soc. 23 (1), 96-97 (1974).

8. Gross, A. M., Confidence Interval Robustness with Long-Tailed Symmetric Distributions, J. Am. Stat. Assoc. 71, 409-416 (June 1976).

9. Devlin, S. J., Gnanedesikan, R., and Kettenring, J. R., Robust Estimates and Outlier Detection with Correlation Coefficients, Biometrika 62, 531-545 (1975).

10. Snedecor, G. W., and Cochran, W. G., Statistical Methods, 6th edition, chapter 7 (Iowa State University, 1967).

11. Linholm, L. W., Semiconductor Measurement Technology: The Design, Testing, and Analysis of a Comprehensive Test Pattern for Measuring CMOS/SOS Process Performance and Control, NBS Spec. Publ. 400-66 (to be published).

12. Reynolds, F. N., Lawson, R. W., and Mellor, P. J., Specifications for Integrated Circuits in Telecommunications Equipment, Proc. IEEE 62, 223-230 (1974).

13. Buehler, M. G., and Sawyer, D. E., Microelectronic Test Patterns and Production of Custom ICs, Circuits Manufacturing 17 (2), 46-56 (February 1979).

14. Thurber, W. R., and Buehler, M. G., Semiconductor Measurement Technology: Microelectronic Test Pattern NBS-4, NBS Spec. Publ. 400-32 (April 1978).

15. Buehler, M. G., and Thurber, W. R., An Experimental Study of Various Cross Sheet Resistor Test Structures, J. Electrochem. Soc. 124, 645-650 (1978).

Table 1. Sample Correlation Coefficients for Selected Process Parameters

Wafer: NBS-16, A10

Original Sample Size: 13

	A	B	C	D	E	F	G
A	1						
B	-0.80	1					
C	-0.13	0.11	1				
D	-0.01	-0.12	0.01	1			
E	-0.65	0.57	-0.03	-0.13	1		
F	0.28	-0.30	-0.10	0.76	-0.41	1	
G	-0.13	0.12	-0.68	-0.61	0.16	-0.50	1
H	-0.29	0.01	0.60	-0.07	-0.13	-0.23	-0.27

Parameter

A	p-channel threshold voltage
B	n-channel threshold voltage
C	metal-to-p^+ contact resistance
D	metal-to-n^+ contact resistance
E	p^+ sheet resistance
F	n^+ sheet resistance
G	metallization linewidth
H	polysilicon sheet resistance

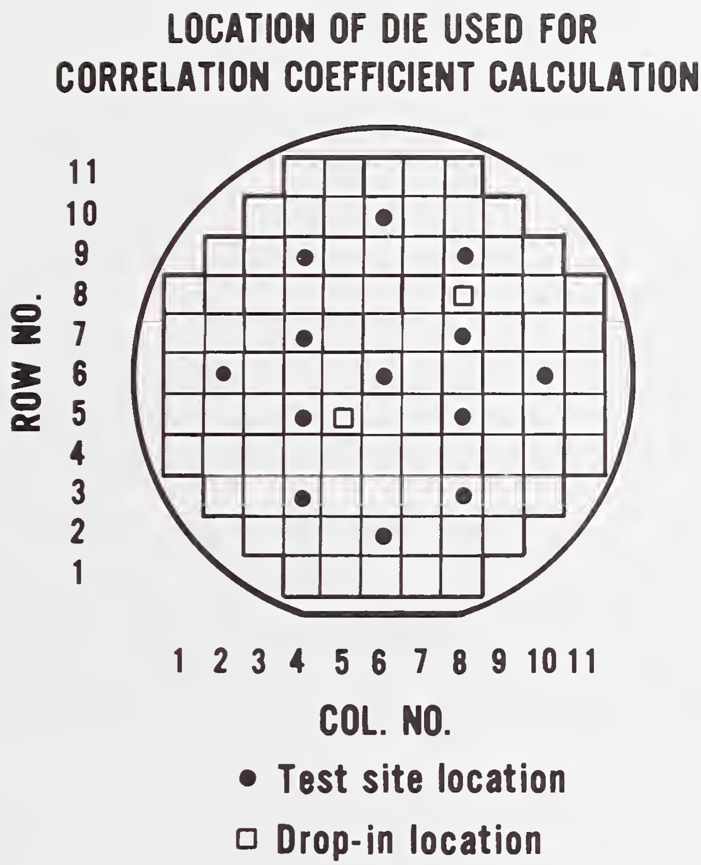


Figure 1. Location of test sites used for determining correlation coefficient for the 13-point sample and location of the "drop-in" sites which contained test patterns other than NBS-16.

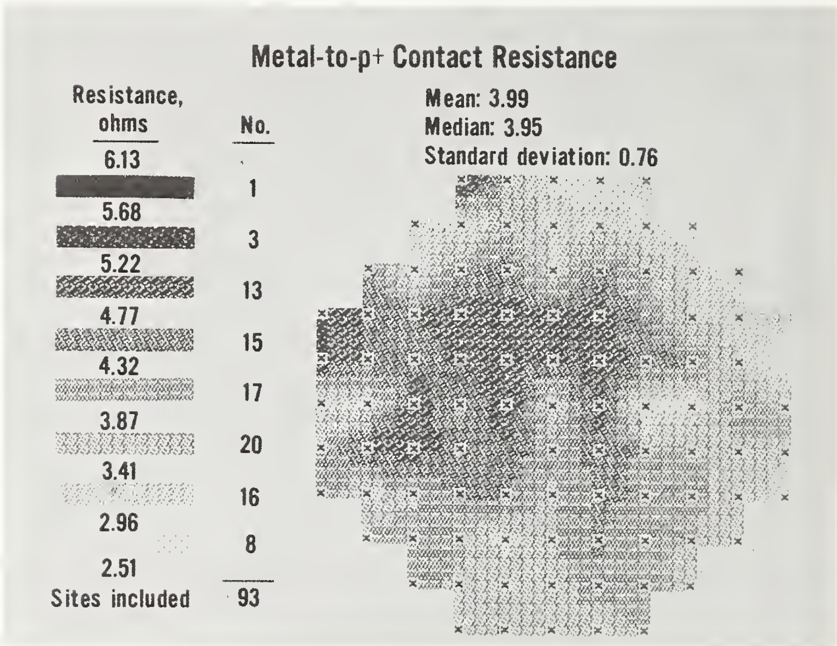


Figure 2. Metal-to-p^+ contact resistance computer-generated (eight-level) gray scale wafer map showing test site location and intrawafer parameter variation.

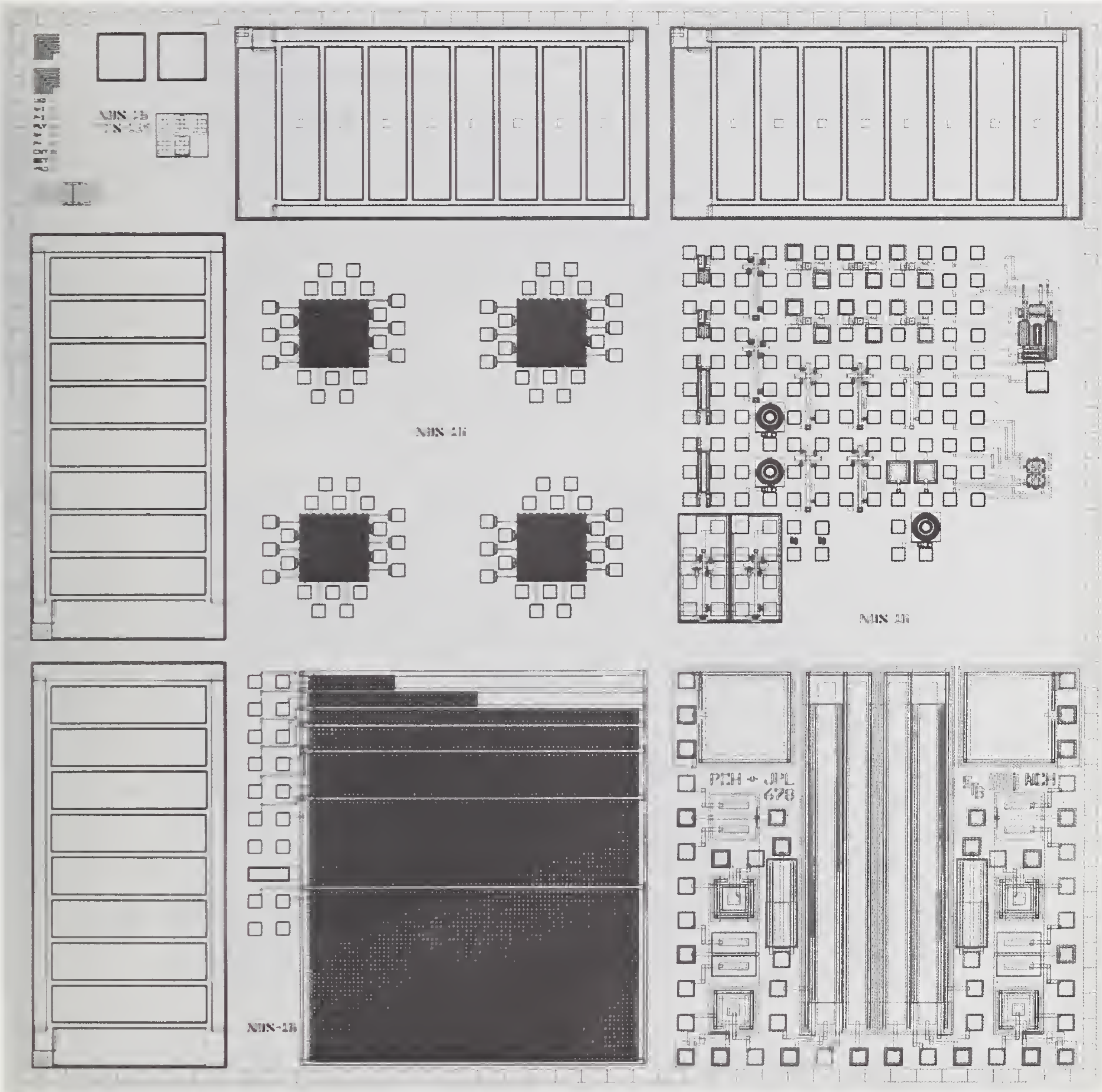


Figure 3. Computer-generated plot of test pattern NBS-16.

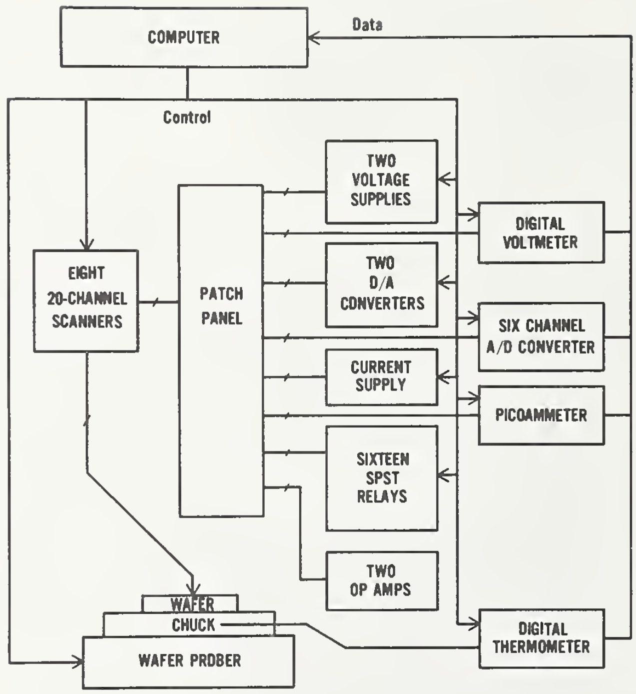


Figure 4. Block diagram of computer-controlled electrical test system.

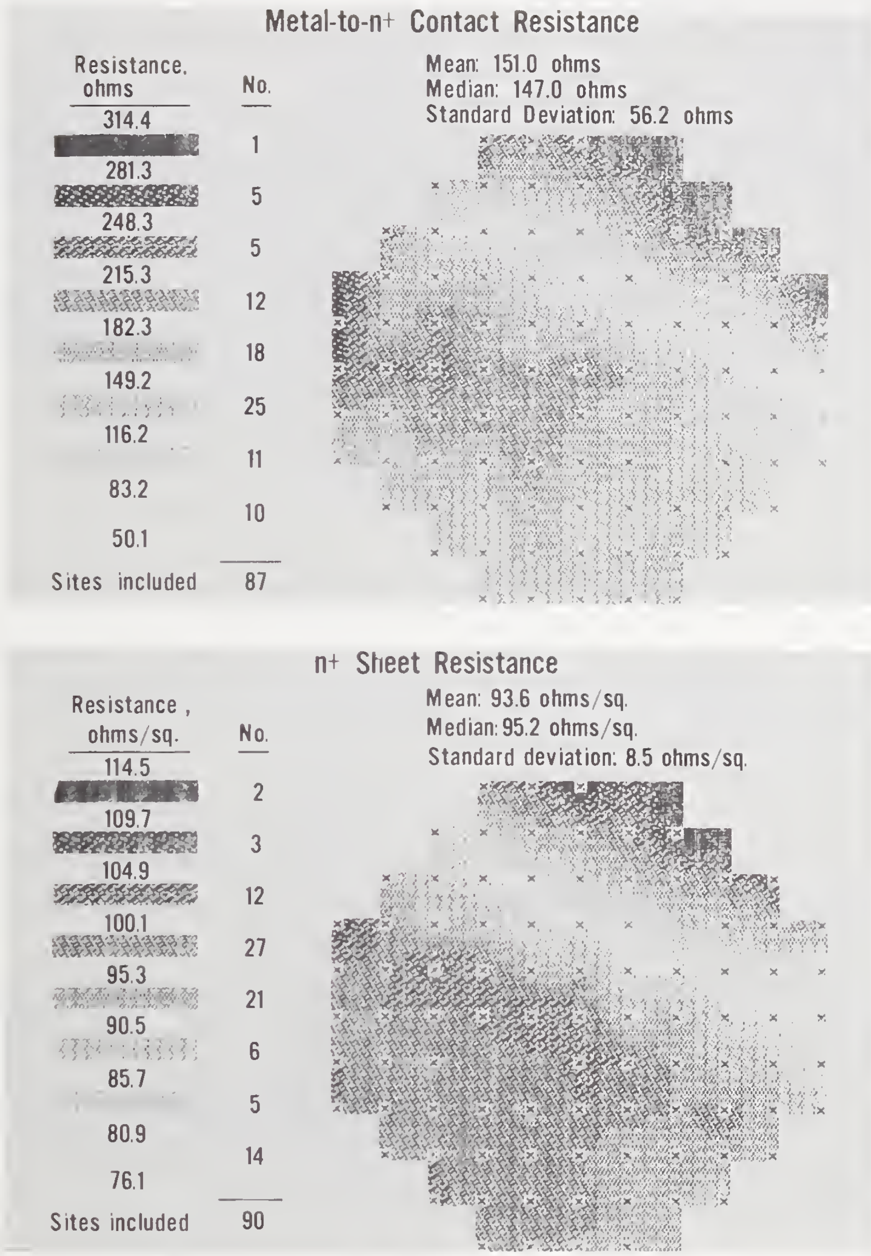


Figure 5. Wafer maps of metal-to-n^+ contact resistance (top) at 87 test sites and n^+ sheet resistance (bottom) at 90 test sites for an NBS-16 process validation wafer containing 95 test sites. In both maps, the scale or gray tone boundaries were selected such that the upper bound of the darkest gray tone was the largest resistance value, and the lower bound of the lightest gray tone was the smallest resistance value. The "x" symbols on the maps represent the locations of nondefective test sites. Test results from these sites were used to calculate mean, median, and standard deviation and also to produce the wafer map.

FEDERAL INFORMATION PROCESSING STANDARD SOFTWARE SUMMARY

01. Summary date — Yr. 82, Mo. 11, Day 01

02. Summary prepared by (Name and Phone) — Richard L. Mattis (301) 921-3621

03. Summary action — New [x] Replacement [ ] Deletion [ ]

Previous Internal Software ID

04. Software date — Yr. 82, Mo. 11, Day 01

05. Software title — *Semiconductor Measurement Technology:* A FORTRAN Program for Analysis of Data from Microelectronic Test Structures

06. Short title

07. Internal Software ID — STAT2

08. Software type

- [ ] Automated Data System
- [x] Computer Program
- [ ] Subroutine/Module

09. Processing mode

- [x] Interactive
- [ ] Batch
- [ ] Combination

10. Application area

General:

- [ ] Computer Systems Support/Utility
- [x] Scientific/Engineering
- [ ] Bibliographic/Textual
- [ ] Management/Business
- [ ] Process Control
- [ ] Other

Specific: Analysis of Integrated Circuit Process Validation Wafer Test Data

11. Submitting organization and address

Semiconductor Devices and Circuits Division
National Bureau of Standards
Washington, DC 20234

12. Technical contact(s) and phone

Richard L. Mattis
(301) 921-3621

13. Narrative

The program performs the following functions: read data as a two-dimensional array in several formats; calculate mean, standard deviation, median; identify outliers; calculate replacement values for outliers; make gray tone, numerical and contour data maps on a line printer; make a numerical map on the user's terminal; make a histogram on a line printer; construct a data base for making correlations among various data sets; search said data base using several selective keys. The package includes a set of help files callable from the program.

14. Keywords

Correlation coefficient; data management; outlier; process validation wafer; statistical analysis; two-dimensional arrays; wafer map.

15. Computer manuf'r and model	16. Computer operating system	17. Programing language(s)	18. Number of source program statements
DEC VAX 11/780	VMS 3.0	FORTRAN	About 12000

19. Computer memory requirements	20. Tape drives	21. Disk/Drum units	22. Terminals
500 Kb	None	Disk needed for map scratch file, data input, help file	1 video or hard copy

23. Other operational requirements

24. Software availability — Available [x] Limited [ ] In-house only [ ]

25. Documentation availability — Available [x] Inadequate [ ] In-house only [ ]

A user's manual is available as NBS Spec. Publ. 400-75

26. FOR SUBMITTING ORGANIZATION USE

Standard Form 185
1974 July
U.S. Dept. of Commerce—NBS
(FIPS. Pub. 30)

If you would like to receive information on updates of STAT2, fill out the bottom portion of this page and mail to:

Richard L. Mattis
Semiconductor Devices and Circuits Division
National Bureau of Standards
Washington, DC 20234

To: Richard L. Mattis
Semiconductor Devices and Circuits Division
National Bureau of Standards
Washington, DC 20234

Please put me on your mailing list to receive information on updates of the STAT2 computer program.

Name: ______________________________

Company: ______________________________

Address: ______________________________

City/State/Zip Code: ______________________________

Announcement of Semiconductor Measurement Technology
List of Publications 72 - 1962-1982

Semiconductor Materials and Processes Division
National Bureau of Standards
Bldg. 225, Room A305
Washington, DC 20234

Please send a copy of your latest "Semiconductor Measurement Technology, List of Publications 72."

Name ______________________________

Company ______________________________

Address ______________________________

City ____________________ State ____________ Zip Code _____

NBS-114A (REV. 2-8C)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET (See instructions)	1. PUBLICATION OR REPORT NO. NBS SP 400-75	2. Performing Organ. Report No.	3. Publication Date July 1983

4. TITLE AND SUBTITLE

Semiconductor Measurement Technology: A FORTRAN Program for Analysis of Data from Microelectronic Test Structures

5. AUTHOR(S)
Richard L. Mattis

6. PERFORMING ORGANIZATION (If joint or other than NBS, see instructions)

NATIONAL BUREAU OF STANDARDS
DEPARTMENT OF COMMERCE
WASHINGTON, D.C. 20234

7. Contract/Grant No.

8. Type of Report & Period Covered
Final

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (Street, City, State, ZIP)

Same as in item 6 above.

10. SUPPLEMENTARY NOTES

Library of Congress Catalog Card Number: 83-600548

[X] Document describes a computer program; SF-185, FIPS Software Summary, is attached.

11. ABSTRACT (A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here)

A computer program, STAT2, is described which performs the following functions: reads data as a two-dimensional array; calculates mean, sample standard deviation, and median; identifies outliers; calculates replacement values for outliers; makes gray-tone, numerical and contour data maps on a line printer; makes a numerical map on the user's terminal; makes a histogram on a line printer; constructs a data base for examining correlations among various data sets; and searches the data base for correlations using several selective keys. The emphasis in this document is on program usage, and detailed descriptions of the commands are given. Data input requirements are addressed. Guidance regarding several types of program modifications is provided.

12. KEY WORDS (Six to twelve entries; alphabetical order; capitalize only proper names; and separate key words by semicolons)
computer program; correlation coefficient; data management; outlier; process validation wafer; statistical analysis; test structures; two-dimensional arrays; wafer map.

13. AVAILABILITY

[X] Unlimited
[ ] For Official Distribution. Do Not Release to NTIS
[X] Order From Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402.
[ ] Order From National Technical Information Service (NTIS), Springfield, VA. 22161

14. NO. OF PRINTED PAGES
66

15. Price
$5.50

☆ U.S. GOVERNMENT PRINTING OFFICE : 1983 O - 410-910

USCOMM-DC 6043-P80

APPENDIX II, III & IV

U.S. Department of Commerce
National Bureau of Standards

Washington, D.C. 20234
Official Business
Penalty for Private Use $300

POSTAGE AND FEES PAID
U.S. DEPARTMENT OF COMMERCE
COM-215

SPECIAL FOURTH-CLASS RATE
BOOK