

LA-10373-MS

UC-15
Issued: May 1985

LA--10373-MS

DE85 015367

An Assessment of the Measurement Control Program for Solution Assay Instruments at the Los Alamos National Laboratory Plutonium Facility

A. S. Goldman

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Los Alamos Los Alamos National Laboratory
Los Alamos, New Mexico 87545

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

JB

EXECUTIVE SUMMARY

The purpose of this report is to document and review an established measurement control program (MCP) over a 27-month period for four solution assay instruments (SAIs) at the Los Alamos Plutonium Facility (PF) and to suggest criteria and practices that will aid DOE facilities in implementing effective MCPs for non-destructive assay (NDA) instruments.

This report is the second of a series designed to provide guidance on MCPs in the six areas mentioned in DOE Order 5630.2. The first report¹ dealt with the MCP on balances at the PF, and this report focuses on SAIs. Data collected from the existing Los Alamos MCP for SAIs from January 1982 through March 1984 provided an opportunity to analyze results and make recommendations. The Los Alamos MCP can provide a model for MCPs at DOE facilities having NDA instruments whose characteristics are similar to SAIs.

The current PF MCP provides daily accuracy tests and weekly precision and randomness tests. This report reviews and suggests improvements for carrying out this program and briefly describes the SAI.

Our major findings concerning the PF MCP for SAI instruments include the following:

- (1) The MCP for the SAIs provides comprehensive data over a 27-month period and is judged to be performing satisfactorily.
- (2) Technicians enter data into a logbook and the accountability computer. The observed data entry error rate is ~17%; however, only 1% of the errors change decisions made in MCP tests.
- (3) Small, persistent biases were observed; however, from an accounting point of view, they are not significant.
- (4) The assumption of normally distributed measurements is demonstrated to be valid.
- (5) Agreement is generally found among estimates of measurement variance for the SAIs that are obtained from three sources:

- (a) repeated measurements during a "short" interval,
 - (b) day-to-day measurements, and
 - (c) propagated error formulas based on counting statistics.
- (6) Randomness tests, performed on a weekly basis, indicate that instruments are performing properly; however, a significant trend (nonrandomness) over the past 6 months is developing in one instrument.

Based upon these results, the following improvements are recommended for the PF MCP for SAIs:

- (1) Direct data transmission from the instrument to the computer should continue to be a consideration.
- (2) Biases should be tested monthly for significant differences from zero. When differences occur, an investigation is warranted to determine the cause.
- (3) A daily precision test based on current measurement data should be investigated to see if it would provide more timely decisions.
- (4) The present comprehensive data review should be extended to include the semi-annual calibration data.
- (5) Procedures for crosschecking and verifying MCP data entries should be examined further.
- (6) Data should be tested periodically for normality.

In addition to this report, DOE facilities should refer to any of the quality control texts given in the references and ANSI Standards such as N15.36 when establishing their MCPs.

CONTENTS

ABSTRACT	1
I. INTRODUCTION	2
II. THE SOLUTION ASSAY INSTRUMENT	4
III. CURRENT MCP	6
A. Management Structure	6
B. Monitoring Calibration	6
C. Data Management	7
D. Statistical Tests	8
IV. FINDINGS AND SUGGESTIONS FOR MCP IMPROVEMENTS	25
A. Transmission of Data	25
B. Normality	28
C. Accuracy Test and Bias Accuracy	29
D. Precision Tests	34
E. Tests for Randomness	35
F. Other Tests	36
V. CONCLUSIONS	37
ACKNOWLEDGMENTS	39
REFERENCES	39

AN ASSESSMENT OF THE MEASUREMENT CONTROL PROGRAM
FOR SOLUTION ASSAY INSTRUMENTS AT THE
LOS ALAMOS NATIONAL LABORATORY PLUTONIUM FACILITY

by

A. S. Goldman

ABSTRACT

This report documents and reviews the measurement control program (MCP) over a 27-month period for four solution assay instruments (SAIs) at the Los Alamos National Laboratory Plutonium Facility. SAI measurement data collected during the period January 1982 through March 1984 were analyzed. The sources of these data included computer listings of measurements emanating from operator entries on computer terminals, logbook entries of measurements transcribed by operators, and computer listings of measurements recorded internally in the instruments. Data were also obtained from control charts that are available as part of the MCP. As a result of our analyses we observed agreement between propagated and historical variances and concluded instruments were functioning properly from a precision aspect. We noticed small, persistent biases indicating slight instrument inaccuracies. We suggest that statistical tests for bias be incorporated in the MCP on a monthly basis and if the instrument bias is significantly greater than zero, the instrument should undergo maintenance. We propose the weekly precision test be replaced by a daily test to provide more timely detection of possible problems. We observed that one instrument showed a trend of increasing bias during the past six months and recommend a randomness test be incorporated to detect trends in a more timely fashion. We detected operator transcription errors during data transmissions and advise direct instrument transmission to the MCP to eliminate these errors. A transmission error rate based on those errors that affected decisions in the MCP was estimated at 1 per cent.

I. INTRODUCTION

Each facility under the auspices of the Department of Energy (DOE) is required to implement and maintain a measurement control program (MCP) for all instrumentation used to measure special nuclear material (SNM).

Order 5630.2 presents an MCP implementation in six areas and permits DOE facilities to set up their own programs for compliance. We feel that the presentation of the programs established at Los Alamos National Laboratory can serve as a guide for an MCP at other facilities.

This discussion is limited to only one part of the six areas. We suggest that a reference manual be made available that would present methodology for implementing Order 5630.2.

DOE Order 5630.2 specifies that, "where practical, an MCP should address the following six areas:

- (1) Scale and Balance Program. All scales and balances shall be maintained in good working condition and calibrated pursuant to an established control program.
- (2) Analytical Quality Control Program. Data from routine testing shall be analyzed statistically to determine and maintain the accuracy and precision of the methods.
- (3) Sample Variability Control Program. The uncertainty (variance) associated with each sampling method shall be determined, minimized, and maintained on a current basis.
- (4) Control Program for Volume, Temperature, and Pressure Measurements. The precision and accuracy of volume, temperature, and pressure determinations shall be obtained by appropriate techniques, and periodic checks of such calibrations shall be made.
- (5) Calibration Program for Nondestructive Assay Measurements. Nondestructive assay (NDA) instrumentation shall be calibrated with appropriate standards and monitored periodically to ensure proper function within established control limits. Assay predictions periodically should

be compared with more accurate measurements of the content of typical materials.

- (6) Sample Exchange Program. Facilities should, to the extent practicable, participate in extra laboratory control programs such as the Safeguards Analytical Laboratory Evaluation Program (SALE)* to provide external demonstration of the adequacy of an internal quality control program."

At Los Alamos, the responsibility of implementing and maintaining a laboratory-wide measurement control program covering these six areas is assigned to the Operational Security/Safeguards Division. This report, written by the Safeguards Systems Group of the Energy Division, is intended to support their work and is funded by the US DOE, Office of Safeguards and Security, as part of the latter's overall efforts to improve the quality of MCPs at all DOE facilities.

ANSI Standard N15.36, "Nondestructive Assay Measurement Control and Assurance" (1983) states:

The goal of a Measurement Control Program is to evaluate individual measurement uncertainties on a continuing basis, to control the magnitude of these uncertainties, and to test and assure the stability of the measurement process.²

The Los Alamos National Laboratory Measurement Control Program is based on the following criteria quoted from The Laboratory Manual (with some deletions for brevity):

- (1) Accuracy and Precision - Each measuring instrument will be analyzed using appropriate statistical techniques to determine its accuracy and precision. The accuracy and precision of each measuring instrument used will conform to a set of parameters specifically predetermined for its use.

*The New Brunswick Laboratory (NBL) SALE program has been terminated. Revised NBL programs are being developed to emphasize more specific DOE measurement, standards, calibration, and evaluation issues.

- (2) Calibrations - All instruments and devices used for special nuclear materials accountability measurements will be calibrated and records maintained in accordance with current and/or applicable American National Standards Institute (ANSI) standards.
- (3) Standards - All standards used to calibrate or evaluate the performance of instruments or devices used to measure special nuclear materials will be traceable to a national measurement and standards program such as NBS or SALE.
- (4) Records - Adequate records will be maintained of all calibration data, maintenance information, and performance history to allow for the determination of biases and trends in instrument performance. Control charts of accuracy, precision, biases, and trends will be maintained and evaluated in accordance with established procedures.³

In meeting these criteria, preparation and maintenance of reference materials (RMs) standards, periodic measurements of these RMs, and data analysis including control charts for each NDA instrument are an important part of an MCP to confirm that instruments are functioning properly. This report describes these procedures as they are performed at the Los Alamos PF for SAIs, reviews and analyzes measurement control data collected on four SAIs over a 27-month period, and suggests areas where this MCP could be improved. Although this report is confined to SAIs, features of the MCP are readily usable for many NDA instruments.

II. THE SOLUTION ASSAY INSTRUMENT

The SAIs located at the Los Alamos National Laboratory PF are used to assay process-generated plutonium solutions.⁴ The instrument uses high-resolution gamma spectroscopy to count gamma rays coming from ²³⁹Pu. The gamma-ray signal is corrected for transmission losses and dead time and is then converted to plutonium concentration.

There are four SAIs at the PF. Instruments IO2 and IO3 were installed in 1978, and IO4 and IO5 were installed in 1981. Minor differences exist among these SAIs primarily because of differences in construction.

SAI instruments are calibrated using plutonium solution primary RMs having known concentrations that include the range of values found under operational conditions. Both the RMs and process samples are dispensed in 30-mL bottles that are then placed in a sample chamber. The calibration procedure determines a constant multiplier for converting observed transmission corrected gamma-ray counts⁴ to a plutonium concentration in the sample. These procedures are carried out under conditions approximating the production environment.

Because of the handling problem and slow evaporation through the plastic bottles, it is not feasible to maintain liquid calibration standards for use in the periodic checks to confirm instrument calibration that are required by the MCP. Instead, plutonium metal foils (secondary reference materials) are mounted in sample bottles and used for MCP instrument checks. Immediately following an instrument calibration, the foil is carefully counted and the instrument response is recorded. This establishes the foil as a secondary reference source, which can be used for periodic confirmations of the instrument calibration. Because the foil is highly stable, it can be used for much longer periods than the solution standards.

Possible sources of error in SAI measurements using the RM include

- background variation,
- counting statistics,
- technician-to-technician differences,
- uncertainty in the RM value, and
- instrument performance changes.

Sources of error that are either difficult to measure or not found in the MCP and may be found in the process measurements include

- transmission measurements,
- variable isotopic compositions,
- impurity gamma interference,
- sampling homogeneity, and
- weighing or volumetric errors.

An MCP should include a sample variability control program in accordance with item (3) in DOE Order 5630.2. It is this part of an MCP that deals with estimation of variances from the sources given above.

SAI measurements may be used in the computation of estimates of inventory terms in the materials balance equation. Estimates of variances of these measurements can be obtained from the MCP; however, as noted above, these variances are smaller than those that might be obtained from process measurements. Therefore, we suggest that a program be run concurrently with the MCP for determining the variance of process measurements using the SAIs.

III. CURRENT MCP

A. Management Structure

Measurement control procedures that include preparation of RMs, measurements of RMs, instrument calibration, analysis of data with control charts and statistical tests, and data archiving are coordinated by a management structure that incorporates several groups. The analytical chemistry group (CHM-1) is responsible for preparing the 30-mL bottles of plutonium solution that are used as calibration standards, and the plutonium metal technology group (MST-13) is responsible for preparing the plutonium foils. The PF nuclear materials management group (MST-10) is responsible for calibrating instruments, periodically checking the calibration, and reporting the calibration information to a central data base. The nuclear materials accountability group (OS-2) has overall responsibility for the design and implementation of the laboratory-wide MCP.

B. Monitoring Calibration

Calibration checks of each SAI instrument are performed daily by MST-10 personnel (instrument technicians) through measurements on plutonium foils. The measurements are recorded in a logbook,

entered into a computer terminal, and immediately analyzed by a computer program that determines whether the instrument is functioning properly.

The data from the MST-10 calibration checks are collected by OS-2 in a data base. Data are frequently monitored for anomalies that could indicate deviation from MCP procedures or a malfunctioning instrument. In addition, OS-2 reviews control charts and performs statistical tests to confirm that the instruments are in control.

C. Data Management

All of the MCP data are originated by MST-10 instrument technicians through the daily measurement of a plutonium foil to confirm instrument calibration. A measured value for the foil and its estimated standard deviation are displayed as digital output on the instrument. The standard deviation is estimated primarily from counting statistics. The technician submits these values to a computer through terminal entry. In addition, the technician lists the measured value, measurement errors, and other pertinent information (obtained from the digital output display) about the measurement in a logbook. The technician has the responsibility of transcribing the numbers, but OS-2 has the responsibility of maintaining the computer data base of measurement control information.

A terminal display of instrument MCP data enables OS-2 to obtain an almost instant check on technician input. This type of review is useful for auditing relatively inexperienced personnel or reviewing recent measurement data on a suspected instrument.

Entries in logbooks provide a comparison with MCP data records. Two of the SAIs are equipped with internal tapes that record measurement control data, which may also be compared with the other records.

An instrument status report is issued daily that lists all instruments and their current status. A written report from MST-10 is sent to OS-2 if an instrument failure leads to repair.

End-of-month reports provide a data printout of all tests conducted during the month and serve as "historical" hard copy records. Additionally, all MCP data are archived on tape. These data are issued to generate control charts for the accuracy and precision tests. The accuracy control charts for the SAIs over the 27-month period January 1982 through March 1984 are given in Figs. 1-4. The accuracy, precision, and randomness tests will be explained in the next section.

D. Statistical Tests

MCP data entered into the OS-2 computer by the MST-10 instrument technicians are subjected to several statistical tests to confirm that the instrument is functioning properly, for example, that the calibration of the SAI instruments has not changed.

Data subjected to statistical testing are examined in several ways, including internal checks found within two of the SAIs. If the SAI fails any of the MCP tests, the SAI cannot be used for accountability related measurements until the SAI passes the test. The technician must correct any obvious errant data entries or rerun the test to correct the situation.

The procedure for an SAI accuracy check is to compare a measured amount of plutonium with the accepted value for the same plutonium foil. The accuracy test procedure for the Los Alamos PF was explained by Severe and Thomas⁵ in a 1976 report as part of the OSS-sponsored dynamic materials accountability (DYMAL) program facility MCP. A similar test is recommended for all DOE facilities.

The accuracy test is performed every working day on every SAI instrument. The technician inserts a plutonium foil into the SAI. The instrument performs the necessary measurements and computations and produces a digital display of the measured amount of plutonium and an estimate of the standard deviation based on propagation of counting statistics.

Let σ denote the true but unknown standard deviation of the measured plutonium values. The propagation of error estimate

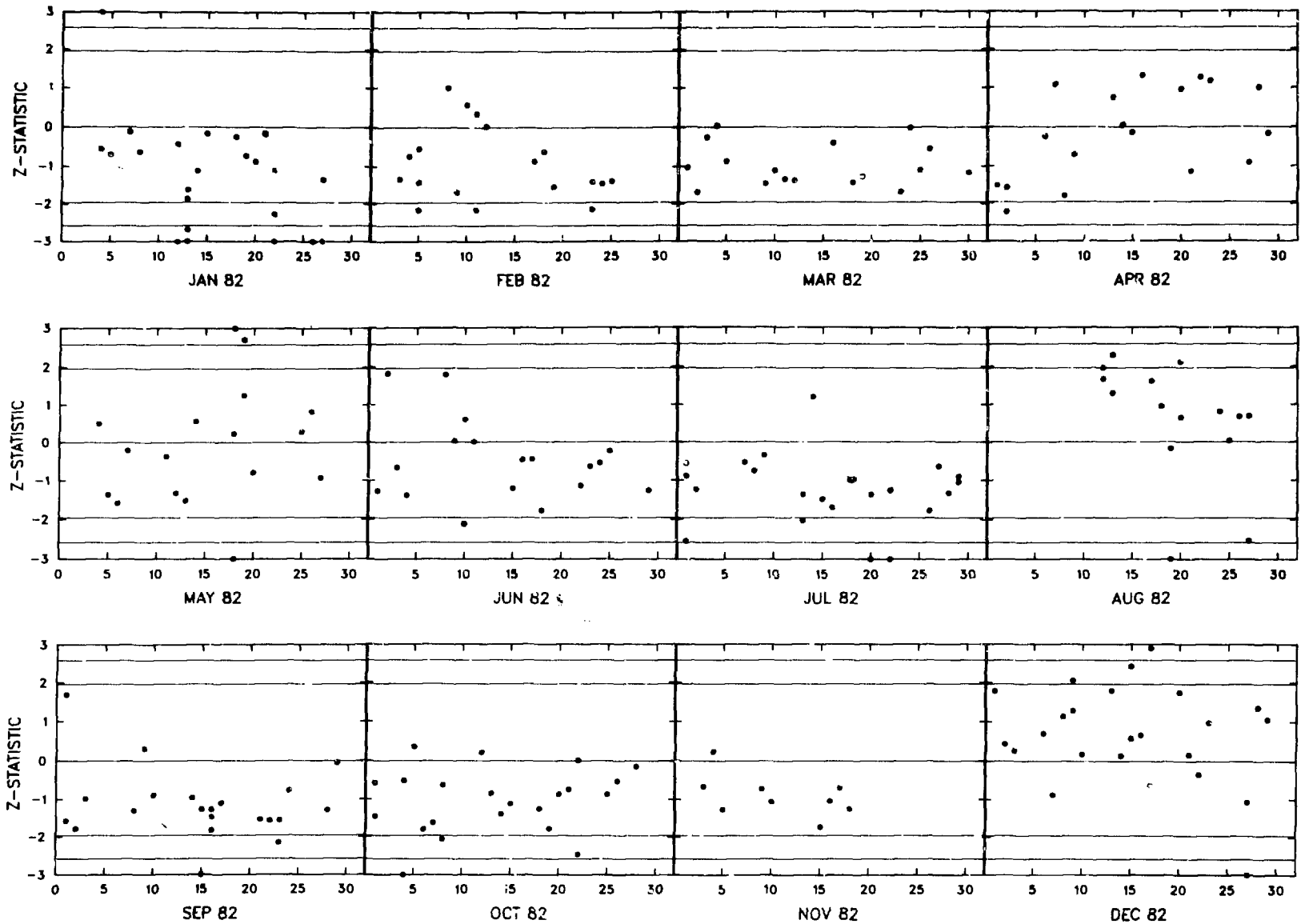


Fig. 1. Accuracy test control charts over a 27-month period for solution assay instrument I02. Accuracy test = (measured minus standard) divided by measurement standard deviation.

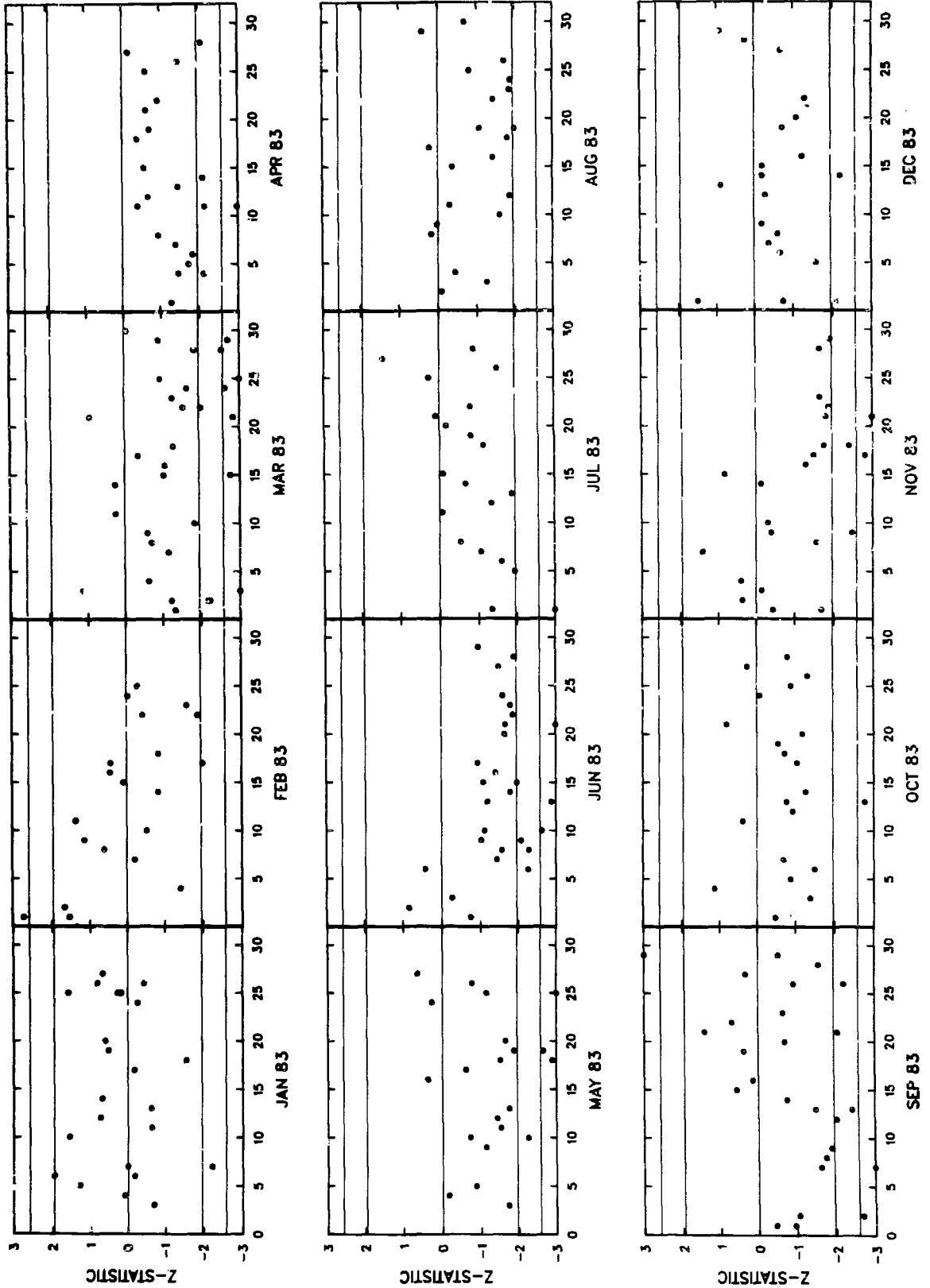


Fig. 1 (cont).

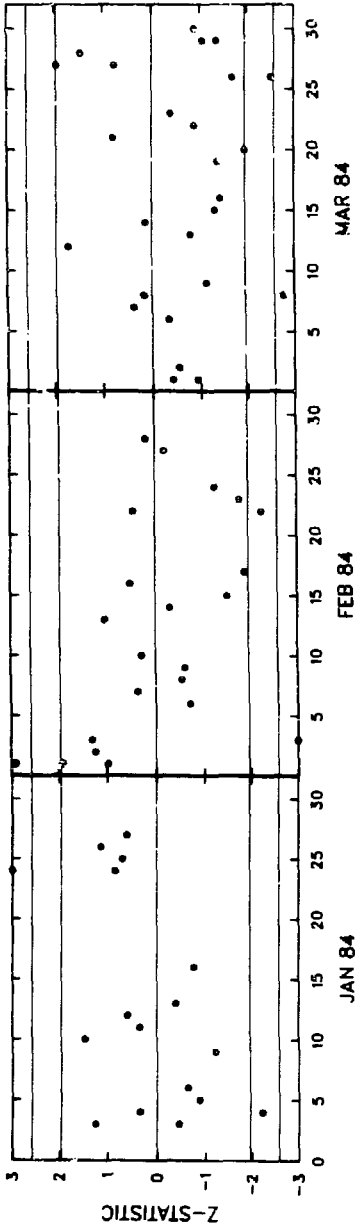


Fig. 1 (cont).

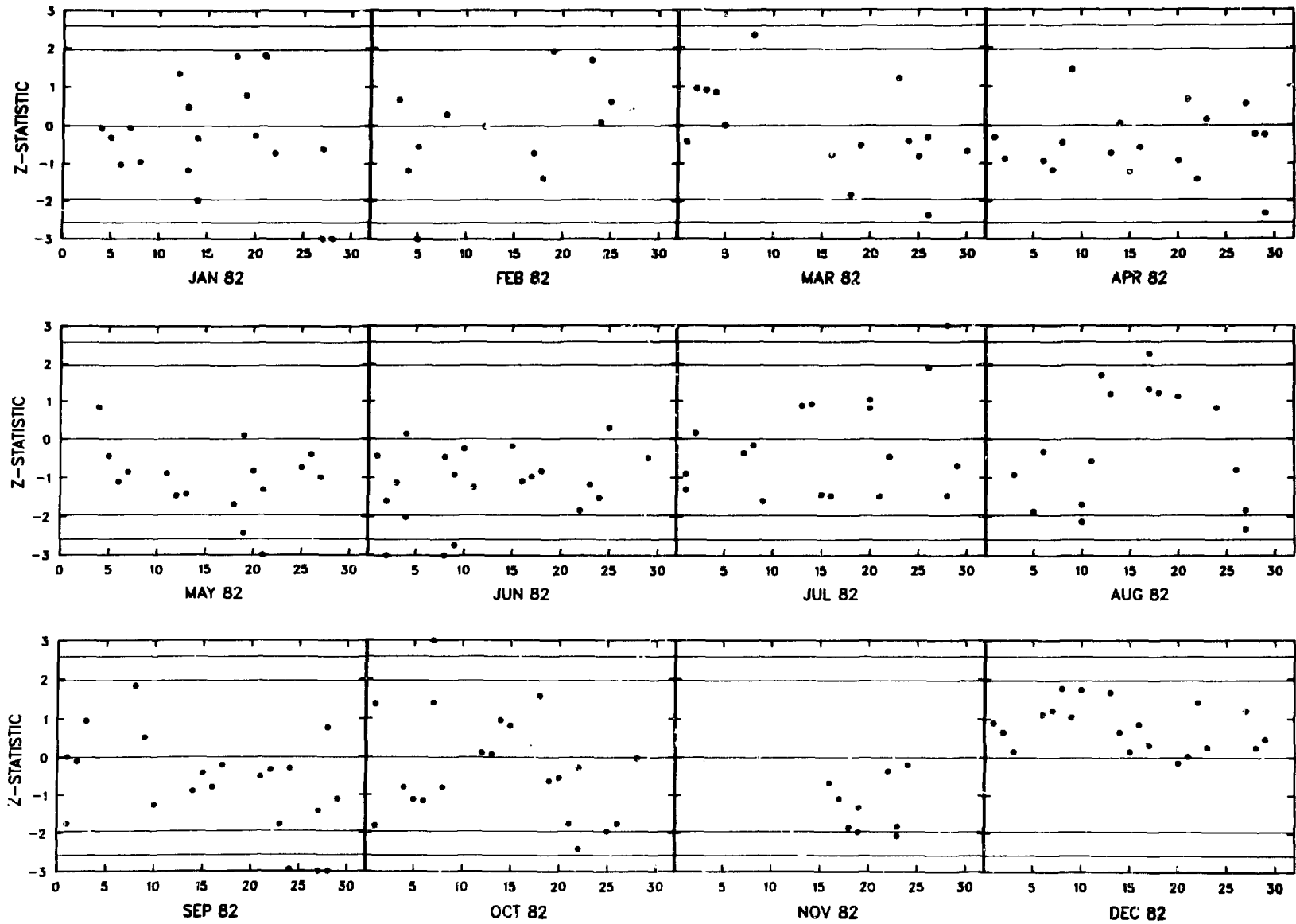


Fig. 2. Accuracy test control charts over a 27-month period for solution assay instrument IO3. Accuracy test = (measured minus standard) divided by measurement standard deviation.

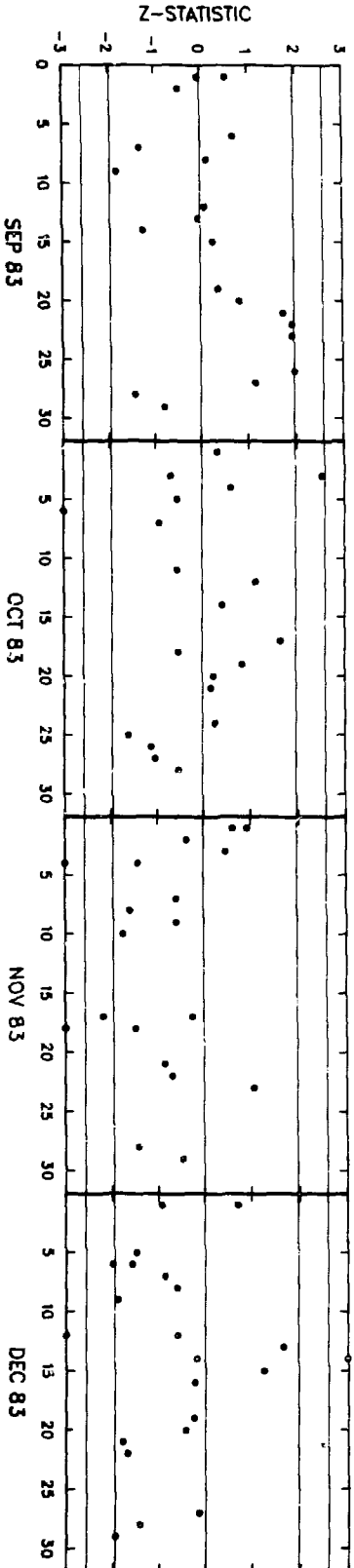
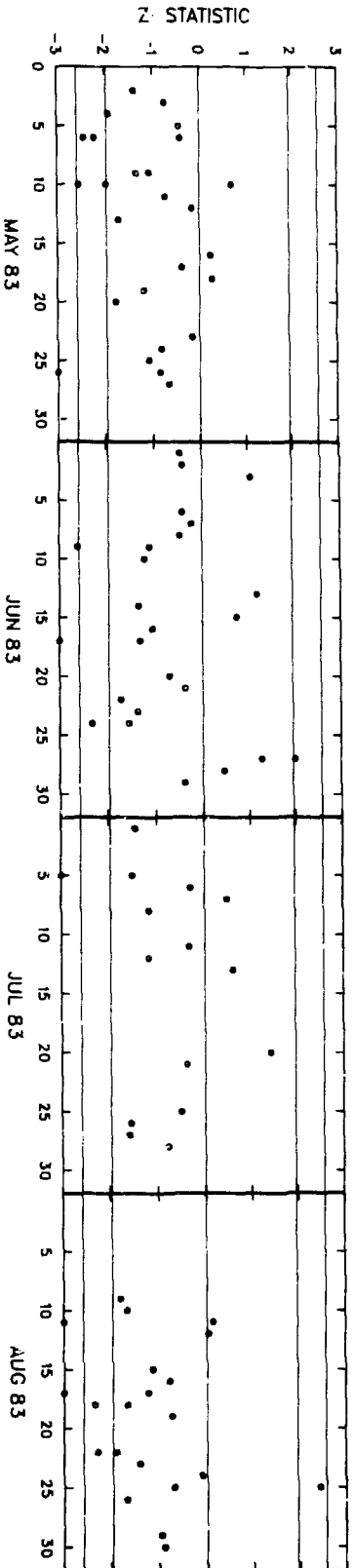
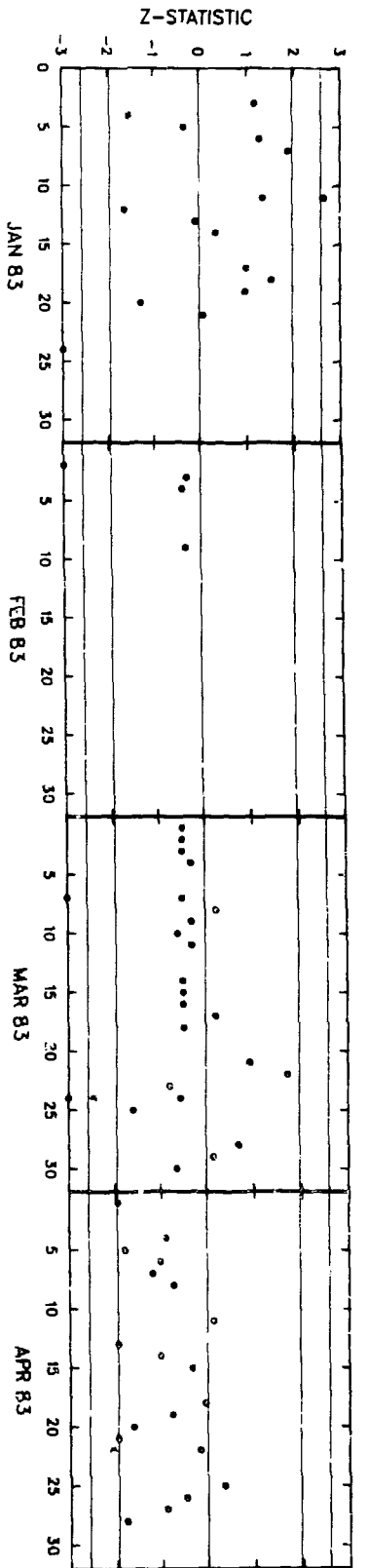


Fig. 2 (cont).

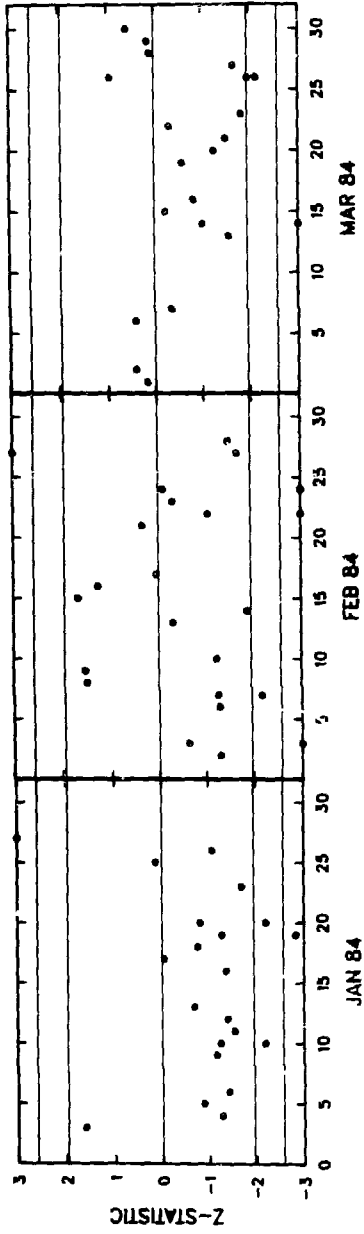


Fig. 2 (cont).

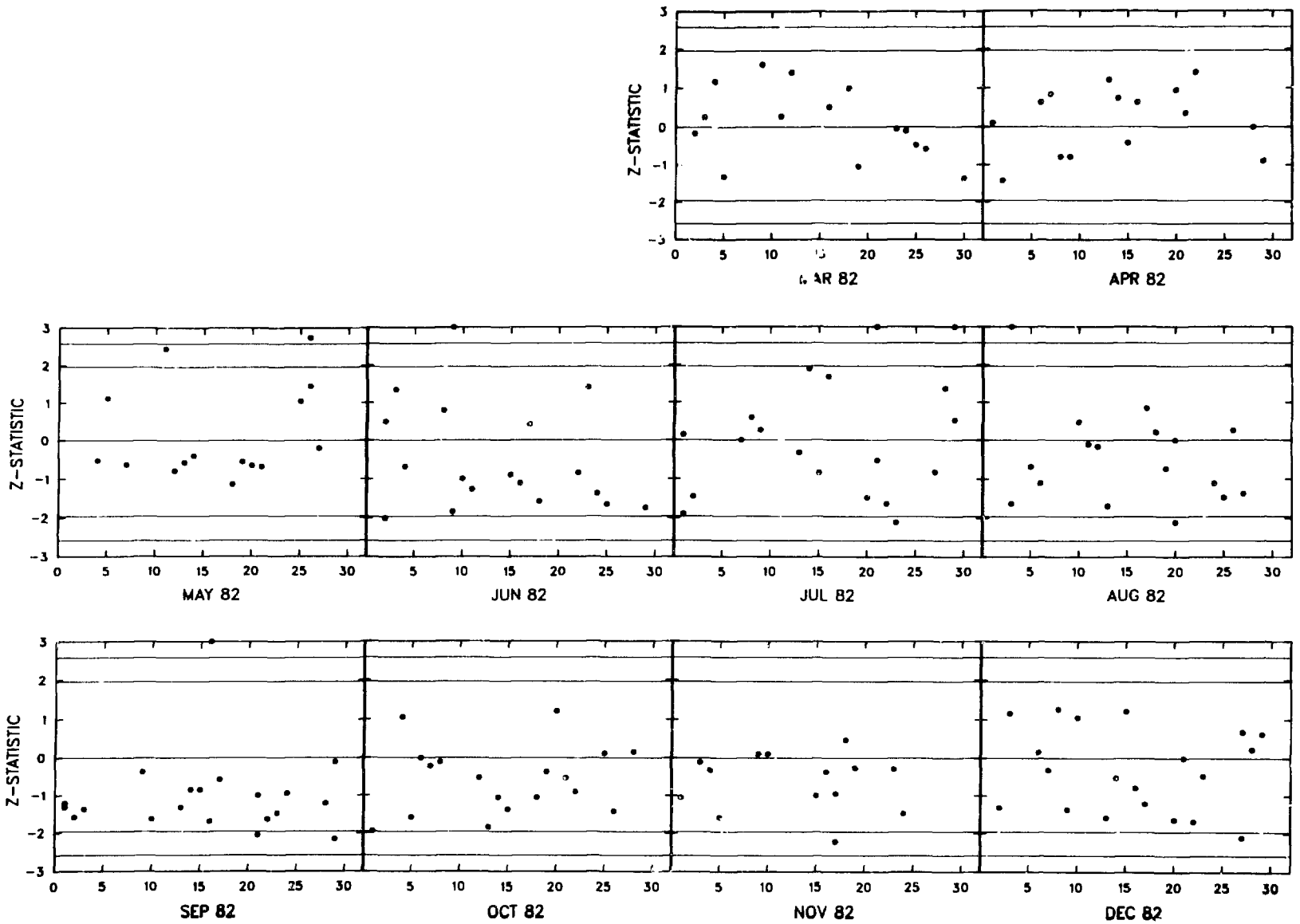


Fig. 3. Accuracy test control charts over a 27-month period for solution assay instrument IO4. Accuracy test = (measured minus standard) divided by measurement standard deviation.

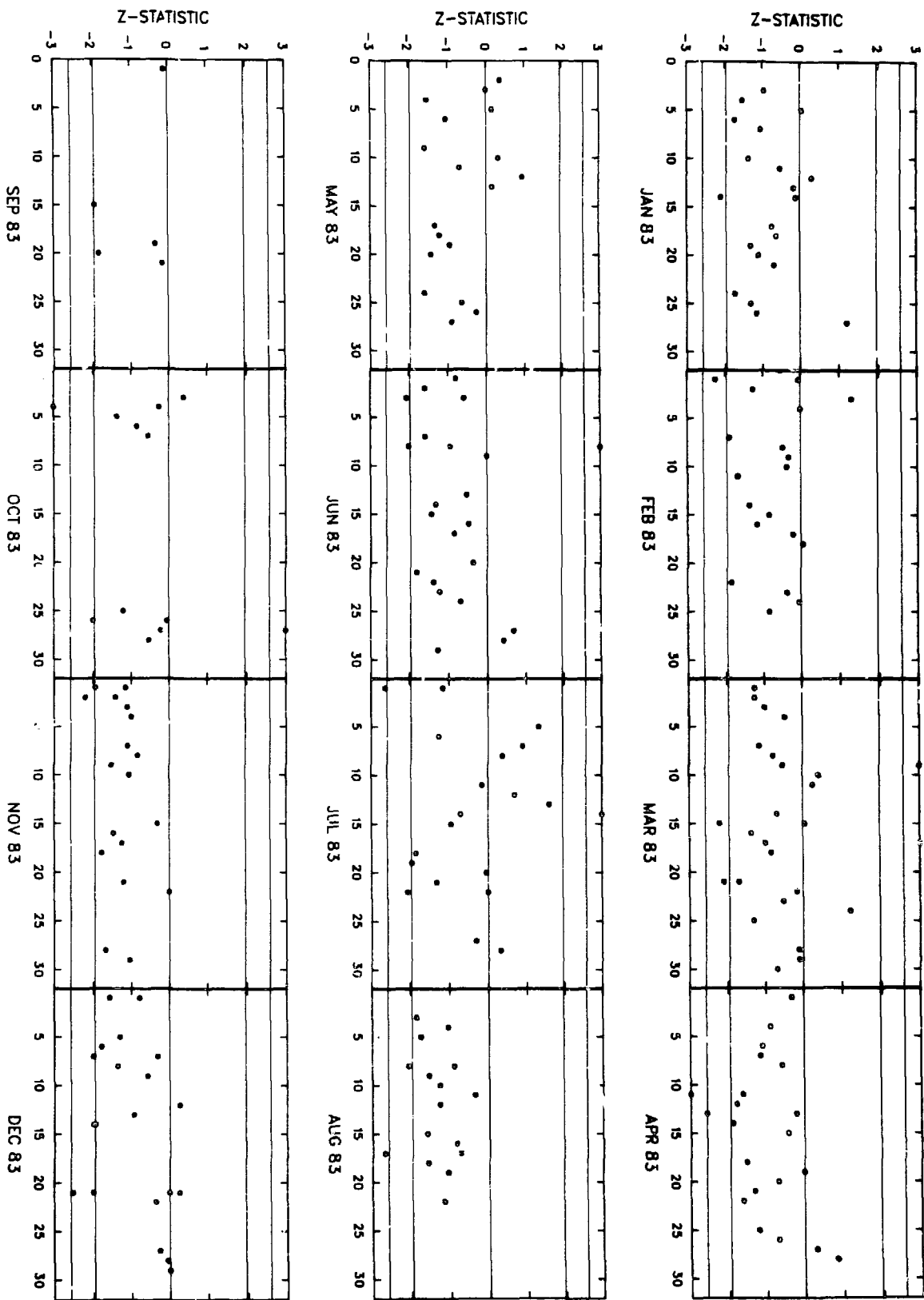


Fig. 3 (cont) .

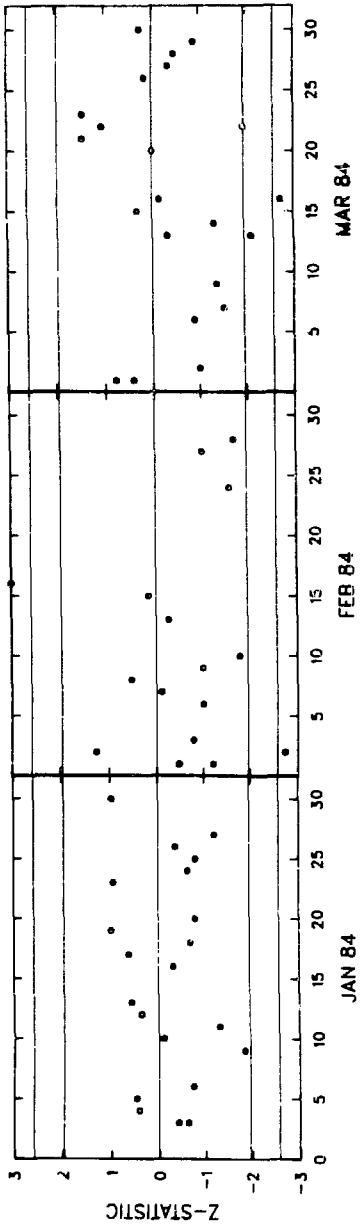


Fig. 3 (cont).

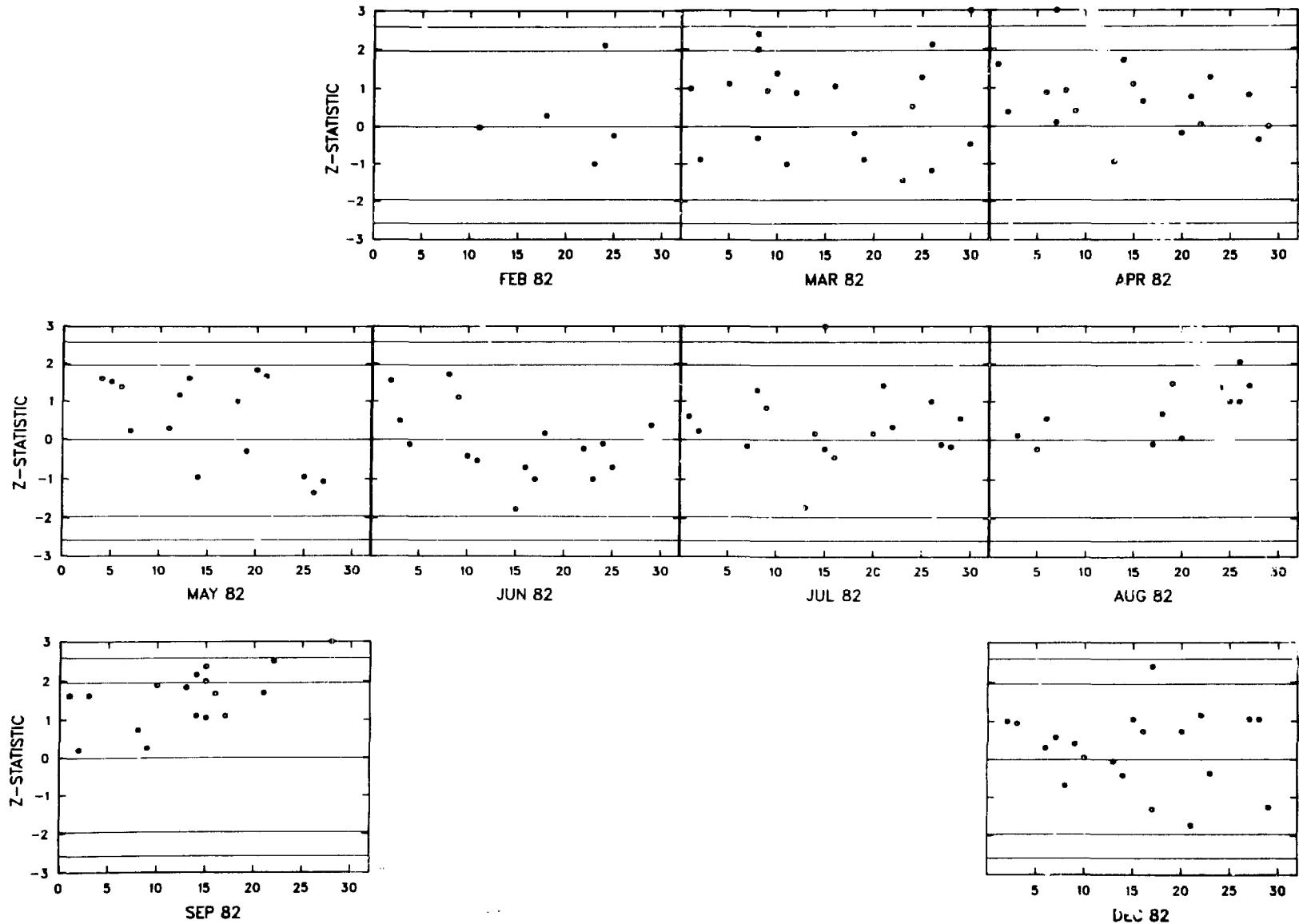


Fig. 4. Accuracy test control charts over a 27-month period for solution assay instrument IO5. Accuracy test = (measured minus standard) divided by measurement standard deviation.

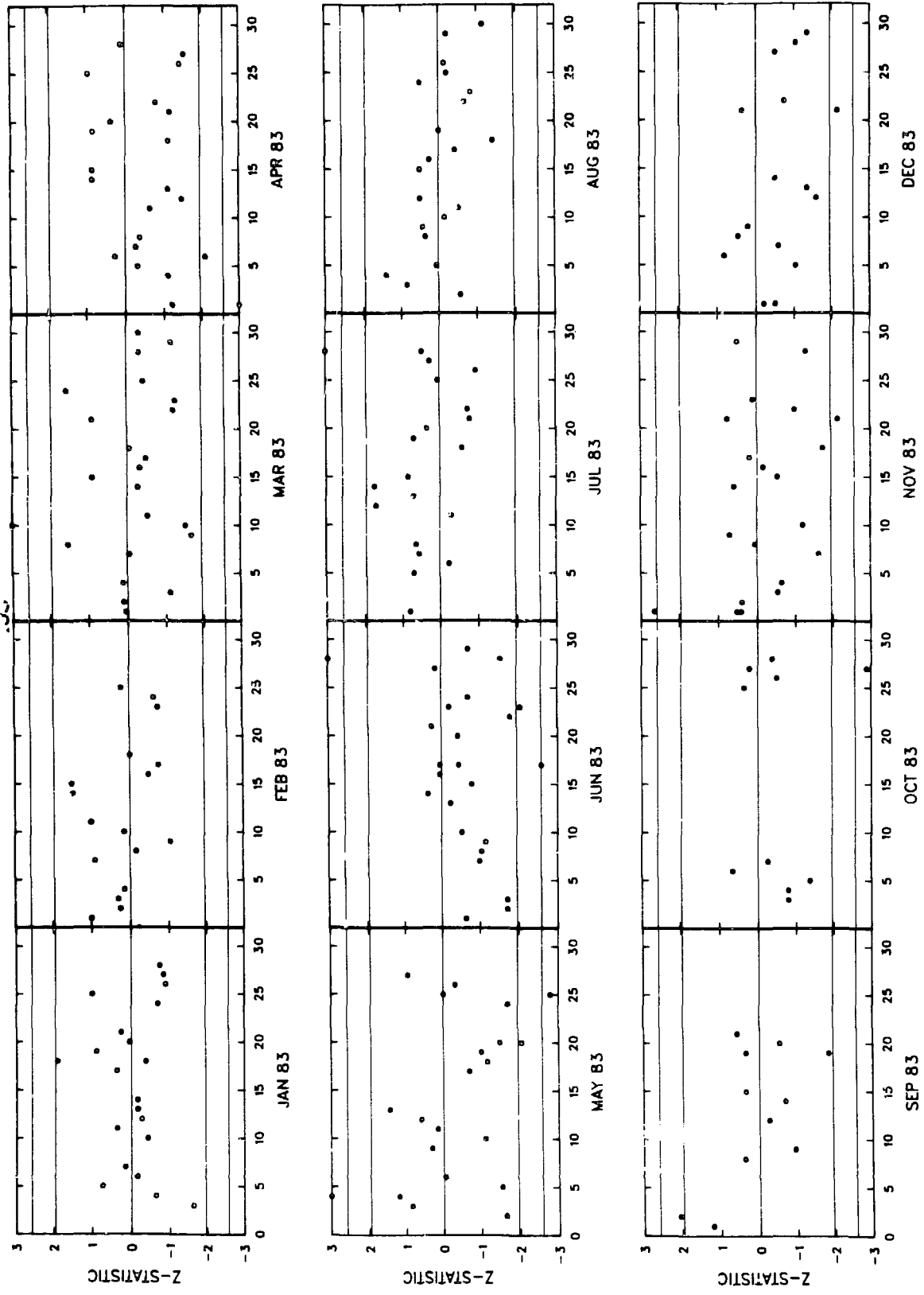


Fig. 4 (cont).

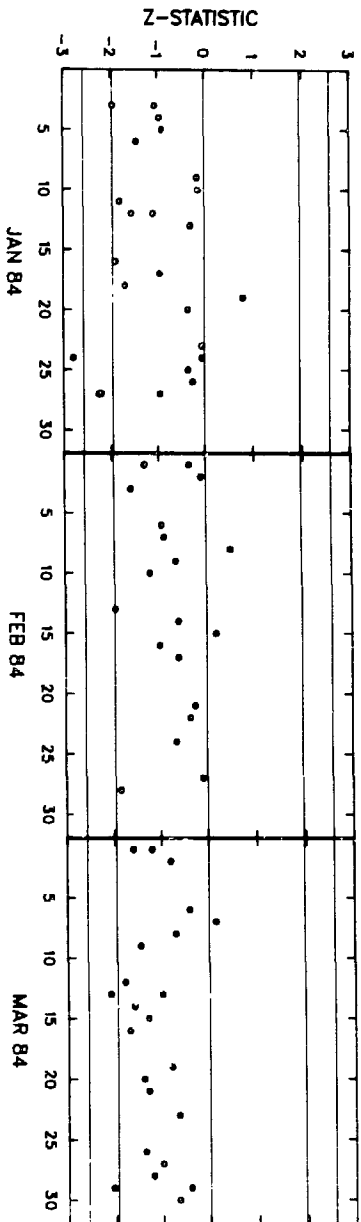


Fig. 4 (cont).

of σ of the measured result is denoted by $\hat{\sigma}$ and is the denominator in the accuracy test

$$z = \frac{x - x_0}{\hat{\sigma}},$$

where x is the measured value of the plutonium in the reference foil and x_0 is the accepted standard value. This standard is assumed to be valid for approximately 6 months. The random variable z does not have a normal distribution because $\hat{\sigma}$ is a variable. However, we have found that histograms of z for the SAIs closely approximate normal distributions that have means different from zero and standard deviations approximately equal to one. The nonzero means are related to measurement bias. (Bias is discussed in part C of this section.) Because of the normality, we conclude that $\hat{\sigma}$ is an acceptable substitute for σ in the z test statistics.

If the test statistic falls either outside action limits of ± 2.58 once or outside warning limits of ± 1.96 twice in succession, the instrument is considered out of control. Corrective action must be taken before the instrument can be used for accountability measurements.

The control charts for accuracy tests on the four SAIs are given in Figs. 1-4. These charts represent a collection of data from January 1982 through March 1984. The central line passes through $z = 0$, and the outer lines are the warning and action limits. The process is "in control" when points are scattered randomly about 0 and when no more than 1% of the points fall outside ± 2.58 and no more than 5% fall outside ± 1.96 .

An estimator of σ can be obtained from historical data. This estimator is referred to as the "sample standard deviation" and is defined by

$$s = \sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1) ,$$

where \bar{x} is the mean of the n measured observations. We assume $n \leq 30$, which enables us to use the z statistic as before. The denominator of z would be s instead of $\hat{\sigma}$. [If $n > 30$, then z is distributed as Student's "t" with $(n-1)$ degrees of freedom.]

The choice as to whether to use s or $\hat{\sigma}$ in the accuracy test is based on assumptions concerning the underlying distribution of the measurements. The estimator $\hat{\sigma}$ would be used if there were fluctuations in day-to-day operating procedures for obtaining measurements or if the foil standard were to decay over the measurement period.

The counting statistics model used to compute $\hat{\sigma}$ may not take into consideration all of the important contributors to variability. In this case we would prefer s , which would be computed over a large number of data points ($n \leq 30$).

For the current MCP, we found no significantly different results when either s or $\hat{\sigma}$ was used. We conclude that (1) the process is "stationary" (i.e., the underlying measurement distribution remains unchanged in the testing period), (2) the foil standard is very stable, and (3) the counting statistics model used in the error propagation for $\hat{\sigma}$ is correct.

The purpose of a precision test is to determine reproducibility. A test similar to the one used in the PF MCP is recommended for all NDA instruments. The procedure, explained by Severe and Thomas⁵ for the SAI, is as follows. Each week a set of 15 successively repeated measurements (assays) is taken. The sample variance is computed using

$$s_w^2 = \sum_{i=1}^{15} (x_i - \bar{x})^2 / 14 ,$$

where x_i is a measured value and \bar{x} is the mean of the 15 measurements. Then s_w^2 is compared with $\hat{\sigma}_p^2$, the average of the 15 propagated variances, based on counting statistics. Because the propagated variances remain roughly constant, the ratio can be approximated by a "Chi-square divided by 14" distribution, written $\chi^2/14$. There are 15 measurements and 14 degrees of freedom. In general, the distribution is called "Chi-square divided by degrees of freedom." Values of this parameter outside the action limits of 0.29 and 2.24 or two repeated values outside the warning limits of 0.40 and 1.87 warrant investigation.

We chose $n = 15$ for practical reasons. Time and cost effectiveness studies indicated that taking any sample size greater than 15 would become burdensome to the operator.

A criterion that is often used to select n relates to the desired difference between an estimate of a parameter and its true value. If the desired difference is too small, then the sample size is prohibitively large. On the other hand, a small sample size would give an interval estimate far too wide to be practical.

We state the problem as follows: Find a sample size, n , such that the difference between s^2 and σ^2 is within p per cent of σ^2 with probability 0.95.

In mathematical terminology the problem is to find n such that

$$P[|s^2 - \sigma^2| > (p\%)\sigma^2] \leq 0.95 \quad .$$

For example, it would require $n \approx 750$ to estimate the variance correctly within 10% of σ^2 . At the other extreme, it would require $n \approx 5$ to estimate the variance correctly within 180% of σ^2 .

Figure 5 gives sample sizes required to obtain desired interval width differences in terms of $p\%$ of σ^2 with 95% confidence. Figure 5 shows that $n = 15$ would give an estimate well within 90% of σ^2 .

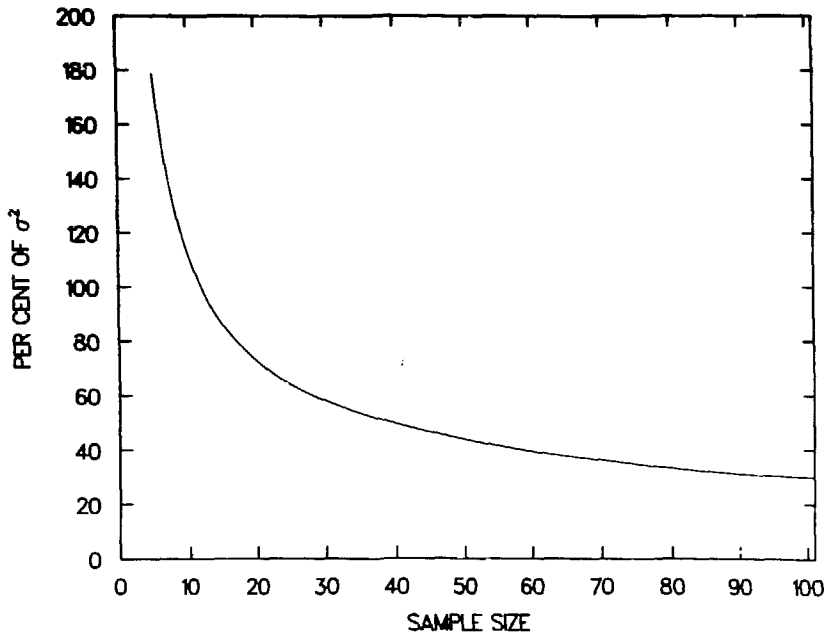


Fig. 5. Sample size required so that the variance is estimated within a prescribed percentage of $\hat{\sigma}^2$ with 95% probability.

The curve depicted in Fig. 5 is flat for large n , indicating that little can be gained from choosing a sample size of 100 instead of 50.

Ignoring time and cost considerations, Fig. 5 can be used as a guide in selecting sample size.

The purpose of a "randomness" test is to determine either unusual oscillations or trends in the data. We suggest the mean square successive difference test as a useful test for randomness. This test has been implemented in the Los Alamos Plutonium Facility.⁵

Our definition of randomness "equates" randomness with independence. The mean square successive difference test (also called the " η -test") is a test of the null hypothesis that a sequence of observations is independent. (See Ref. 6 for the mathematical development.) It is important that observations be independent because nearly all of the test statistics used in the MCP assume independent data. Therefore, the η -test is suggested as a part of an MCP for any DOE facility. The η -test is given by

$$\eta = d^2/s_w^2 ,$$

where

$$d^2 = \sum_{i=1}^{14} (x_{i+1} - x_i)^2 / (14) ,$$

14 is the sample size minus 1, and x_i and s_w^2 are defined as in the precision test. The random variable η is distributed such that warning limits are 1.21 and 2.79 and action limits are 0.92 and 3.08. These limits apply only for $n = 15$. A more general approach given in Ref. 7 is to use the z statistic defined by

$$z = \left(\frac{d^2/2}{s_w^2} - 1 \right) / \sqrt{(n-2)/(n^2-1)} ,$$

which is appropriate for $n \leq 10$. The critical values for z are warning limits of ± 1.96 and action limits of ± 2.58 .

IV. FINDINGS AND SUGGESTIONS FOR MCP IMPROVEMENTS

Our recommendations for the MCP on PF SAIs are based upon criteria given in Refs. 6-9 and ANSI Standard N15.36. Our suggestions are directed primarily at improving the MCP data input and diagnostic procedures.

A. Transmission of Data

Current procedures for entering MCP data into the OS-2 computer require that the MST-10 technician read the instrument's

measured value and enter it at a terminal. A facility that uses technicians to transmit data must define a procedure for entering measurement data into the computer. The procedure adopted at the PF calls for technicians to enter values rounded to the fourth significant digit and to report all measurements. The technician makes a transmission error if appropriate instructions are not followed. For example, errors of transmission include rounding or truncating at the wrong digit, or failure to report a measurement.

The technician makes a transcription error if the wrong number is entered in the computer. These errors are exemplified by transposing digits, misplacing decimal points, or entering the same data points twice. For the sake of convenience, we will identify transmission errors as transcription errors.

We estimate, based on an examination of data from different sources, that even though approximately 17% of the computer entries contain transcription errors, only 1% would be classified as serious. Although a direct data transmission link between the SAI

and the materials accounting computer could eliminate all of the transcription errors, tradeoffs such as cost must be considered.

Estimates of transcription error rates were obtained from three different sources of information about the measurements on the foils. These sources include (1) the instrument tapes that listed measurement parameters such as total plutonium and the variance of that estimate, (2) the history tapes that contain the same measurement parameters and the value of the test statistic calculated from these parameters, and (3) the logbooks that contain relevant counting statistics.

Although the logbooks are kept primarily to provide a maintenance record, we used them as a confirmatory check on the data entered into the MCP. We were then able to obtain an estimate of the error rate.

One hundred logbook entries were compared with corresponding MCP data entries for I02 and I03 during January through February

1984. Transcription errors consisted of differences between the compared numbers, caused primarily by omission of a digit when typing a number into the MCP data base. In summary, we examined 100 entries and found 17 errors, for a 17% error rate. However, only one error led to acceptance of an instrument as being in control when in fact the test statistic exceeded the alarm limit.

The second estimate of error comes from alarm data found on the history tapes. If the instrument fails any test, the operator must either repeat the experiment or make a correction to an obvious transcription error. We can examine successive entries for all points falling outside the warning limits for the MCP. We assume the technician made a transcription error if the times reported on the history tape differed by less than 3 minutes on successive entries. Otherwise we assumed the second entry was a rerun because it would take at least 3 minutes to repeat the trial.

A total of 537 data points were listed on the MCP for I02 over the 26-month period from January 1982 through February 1984. There were 92 values outside the warning limits (± 1.96). Of these 92 entries, we found 25 repeated within a 3-minute interval. This error rate of 27% was higher than that reported by Brouns, Smith, and Eggers.¹⁰ They found 23 occurrences of transcription errors out of 125 alarms for 18%. We emphasize that these error rates are based on alarm data only. None of these data entry errors affected the MCP because they were all corrected promptly by the technician.

The third method of estimating error was a comparison of instrument tapes with history tapes. Instruments I04 and I05 have internal tapes that record the measured values. Entries on the instrument and historical tapes were compared for March and April 1984.

A total of 103 entries were compared and 19 transcription errors were found, caused either by failure to enter data into the MCP or entry of the wrong number. Of these 19 errors, two were corrected by the technician and only one resulted in incorrectly

accepting an instrument as being in control. In summary, the error rate based on this comparison was 18% with less than 1% of consequence for the MCP.

B. Normality

All of the recommended tests are based on the assumption that data distributions are approximately normal; therefore, we suggest periodic tests of normality be conducted on each instrument. Data observed on IO2 from March through May 1983 were selected to test for normality. The analysis given in Table I supports the hypothesis that the data are from a normal distribution.

Data from Table I are displayed in Fig. 6, which illustrates the validity of the normality assumption. In addition, Fig. 6 serves as an indication that, if desired, the underlying distribution can be adjusted easily for bias. Control limits are not adjusted for bias. We will elaborate on bias adjustments in the next section.

TABLE I
COMPARISON BETWEEN OBSERVED AND NORMAL FREQUENCIES
FOR IO2 (MARCH THROUGH MAY 1983)

<u>Measurement Minus Standard Class Interval (g/L)</u>	<u>Observed Frequency (O)</u>	<u>Expected "Normal" Frequency (E)</u>
-0.04	1	1.06
-0.04 through -0.03	7	5.35
-0.03 -0.02	14	17.38
-0.02 -0.01	28	24.92
-0.01 0.00	17	17.63
0.00 0.01	5	6.05
0.01 0.02	2	1.02
> 0.02	0	0.08
	<u>74</u>	

$$\chi^2 = \sum \frac{(O - E)^2}{E} = 2.413$$

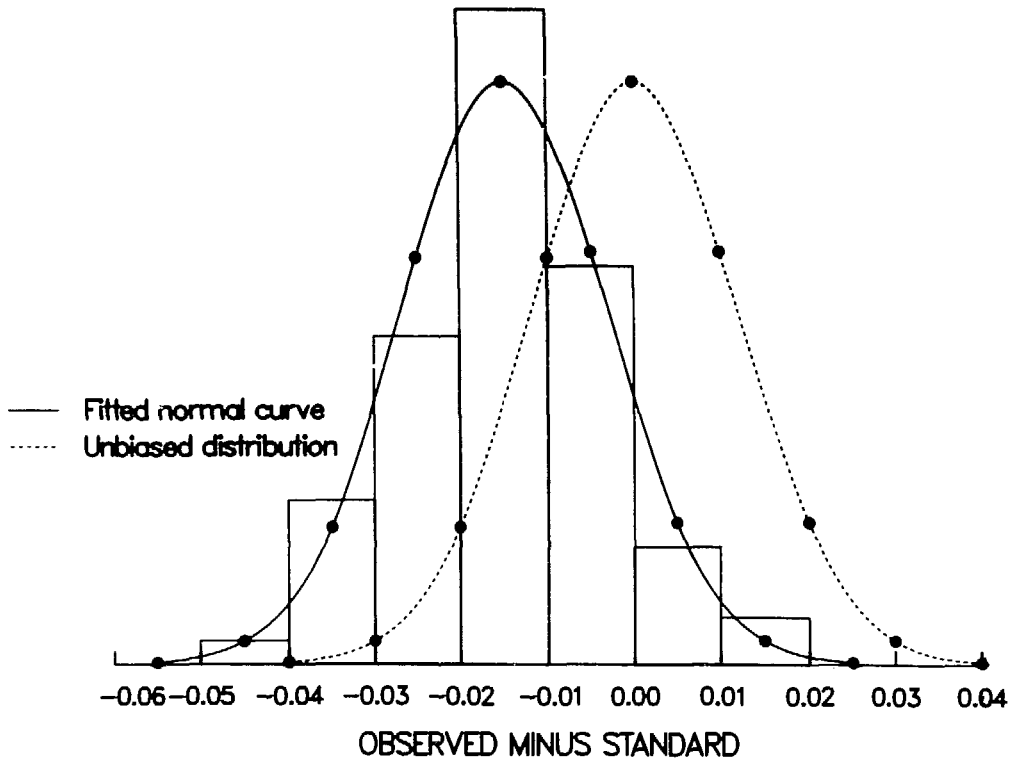


Fig. 6. Histogram and approximate normal distribution for IO₂ data from March through May 1983. Also, normal distribution adjusted for bias.

DOE facilities should test NDA data periodically for normality. Data not normally distributed might imply an instrument failure of some kind. Other consequences of non-normally distributed data are beyond the scope of this discussion but include establishing incorrect warning and action limits.

C. Accuracy Test and Bias

Accuracy of a single measurement is defined as the difference between the measured and standard values. We define bias as the mean of the difference between measured and standard values taken over a period of time. Measurement biases over the 27-month period are summarized in Table II.

TABLE II
ESTIMATES OF INSTRUMENT BIAS

<u>Instrument</u>	<u>Bias (g/L)</u>	<u>Relative Bias (%)</u>
IO2	-0.0080	0.35
IO3	-0.0078	0.34
IO4	-0.0108	0.47
IO5	-0.0021	0.09

We suggest that the period for measuring bias be 1 month. Monthly bias would then be tested for significant difference from zero by use of the Student "t" test, with degrees of freedom equal to one less than the number of observations, to compute s by using the statistic

$$t = \frac{\bar{x}_b \sqrt{n}}{s} ,$$

where s is the measurement standard deviation computed from many previous measurements (for example, over a 1-yr period), \bar{x}_b is the bias, and n is the number of observations taken during the 1-month period. If n is too large, the test could become sensitive to small departures from the underlying statistical model (such as normality of the observations) and indicate a bias. For this reason, we suggest a monthly check.

If the bias term is significantly different from zero (T greater than the tabulated value of the t statistic), then its effect, from an accounting viewpoint, should be investigated.

Plots of monthly bias estimates are given in Figs. 7-10 for all SAIs. Points appear to be randomly scattered about the average bias line in Figs. 7-9. However, as depicted in Fig. 10, IO5 shows a definite downward trend from January 1983 through March 1984. This atypicality has been investigated and corrected.

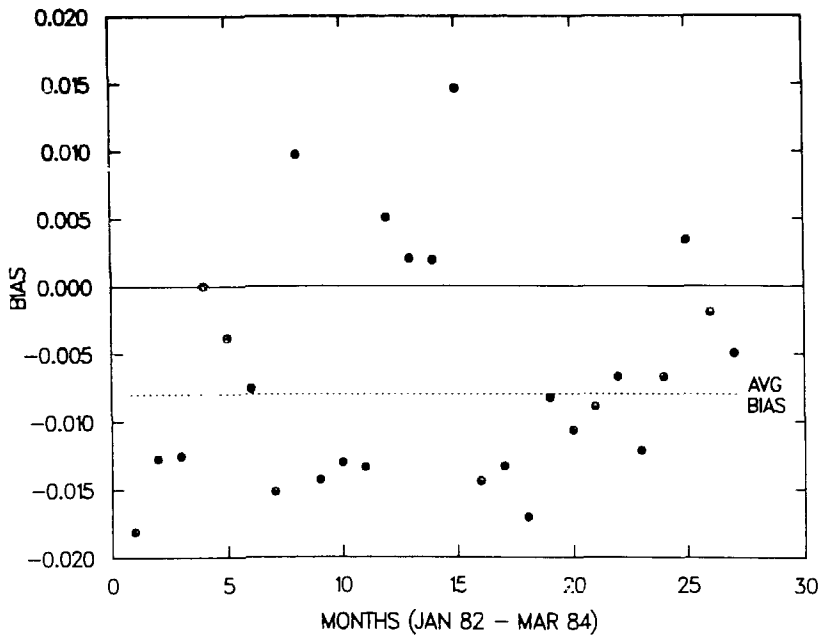


Fig. 7. Bias over a 27-month period for solution assay instrument IO2. (Bias equals average of measured minus standard.)

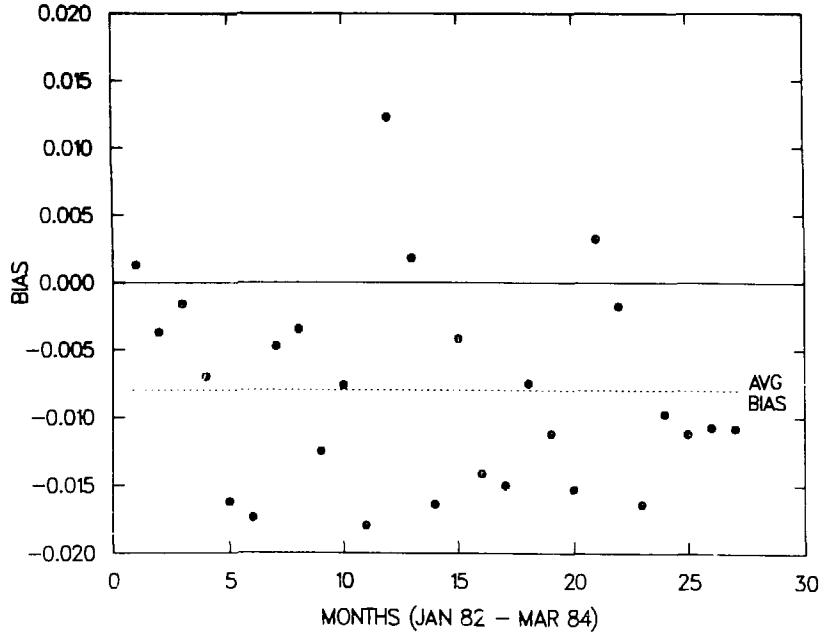


Fig. 8. Bias over a 27-month period for solution assay instrument IO3. (Bias equals average of measured minus standard.)

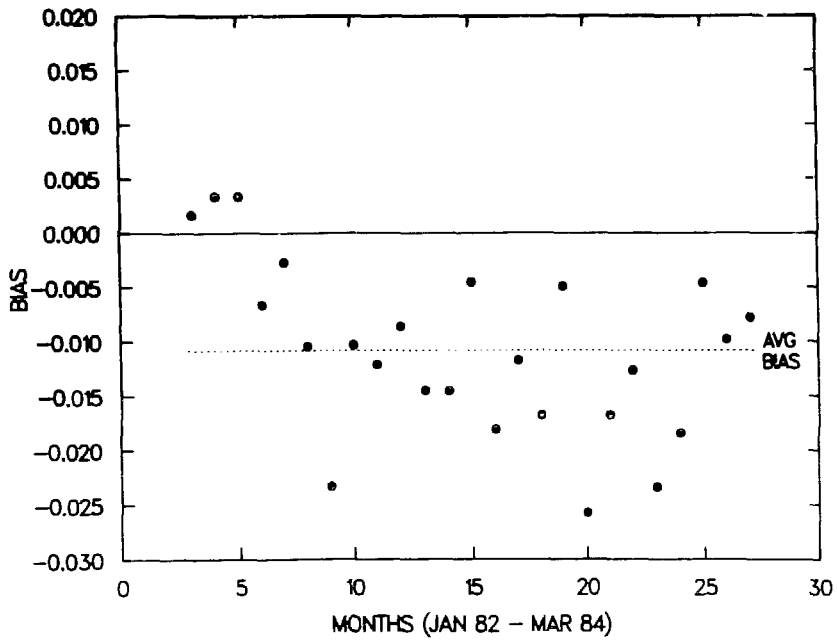


Fig. 9. Bias over a 27-month period for solution assay instrument I04. (Bias equals average of measured minus standard.)

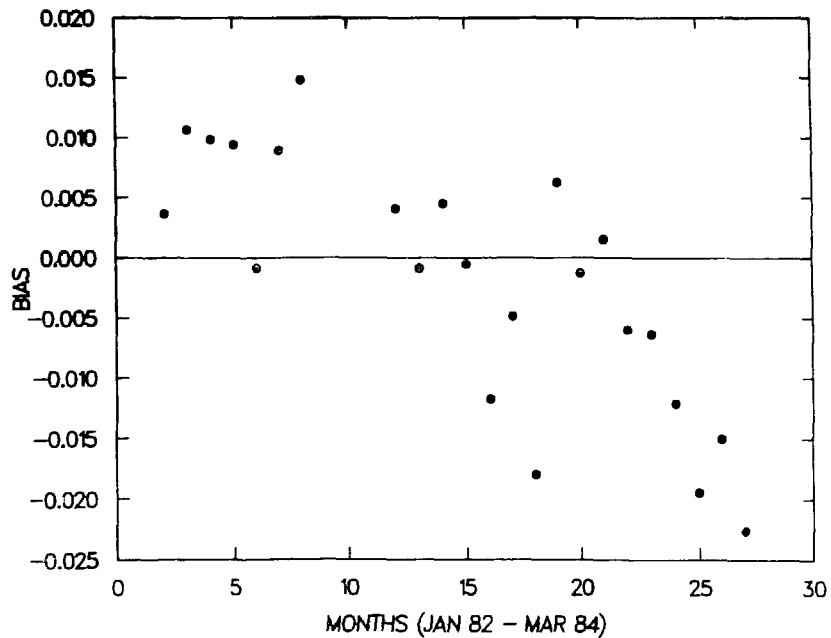


Fig. 10. Bias over a 27-month period for solution assay instrument I05. (Bias equals average of measured minus standard.) Average bias line deleted because of trend.

A recalibration in September 1984 has brought about a random distribution of observations.

Figures 2-4 and Figs. 7-9 display the small, persistent, statistically significant biases in all SAIs. These biases are small relative to typical assay values, may not apply to process solution measurements, and are not significant from an accounting point of view.

The SAIs are calibrated during February and September each year. Figures 2-4 indicate that these calibrations have very little effect on the accuracy test or the small biases. The SAIs are providing very stable results; however, the calibration procedure would be expected to remove even these small biases immediately after being performed. The reason for this minor discrepancy has been investigated. Calibrators had noticed only small changes in the foil standard values and did not replace them after each calibration. This difference may have caused the bias. "Correct" values have been in use since September 1984; however, at this time there is insufficient data to verify any improvements.

No bias corrections have been made on any control chart. Jaech¹¹ states:

. . . "but statistical significance is not essential to the decision concerning whether or not to make the bias correction. It is, of course, one factor in making this decision, but other considerations also play a role, e.g., the size of the bias, irrespective of its statistical significance or nonsignificance, and technical judgment as to whether or not the method is inherently biased."

Measurement bias has been considered too small to make corrections. Small additions of bias would be insignificant compared with the much larger process measurements. Even cumulative biased measurements, which go into material balance equations over a long period of time, would not have significant effects on accounting. This is due, in part, to a much larger variance in SAI process measurements than in the MCP.

An experiment was conducted to measure process variability. Two different floor operators, whose duties include making SAI measurements on process materials, measured the same quantity of process material five times on two different SAIs. The relative standard deviation was estimated to be ~5%, which is considerably larger than the ~0.7% MCP relative standard deviation.

Another reason for neglecting MCP bias effects is that ID measurements are rounded off to the nearest gram. The average bias is typically -0.008 g/L; consequently cumulative affects are negligible.

D. Precision Tests

We suggest the current precision test be continued. In addition, the following two tests should be explored:

- (1) A daily precision test that compares the variance of 15 previous measurements (excluding any weekly precision measurements) with the average propagated variance in that time period. This is the same as the current precision test except that the measurements are made over an extended period.
- (2) A 6-month goodness-of-fit test on the ratio of the sample variance to the average propagated variance for the precision data. This ratio should have a Chi-squared-over-degrees-of-freedom distribution (χ^2/df). The test statistic is

$$\sum_{i=1}^k \frac{(O_i - T_i)^2}{T_i} ,$$

where k is the number of intervals (usually specified by the analyst), O_i is the number of observed ratios in interval i , and T_i is the number of observations in the interval if the ratios are distributed as χ^2/df . The test statistic has a χ^2 distribution under the null hypothesis.

Test 1 would use the same measurements as the accuracy test and allow measurement control on precision to take place daily, which might give more timely detection information on the SAI. Test 2 may also be used as a check on the SAI. If observed and theoretical values disagree, then either something is wrong in the instrument or the propagated variance does not equal the observed one.

We have compared propagated standard deviations, $\hat{\sigma}$, with sample standard deviation estimates, s , during the 27-month period. Results are summarized in Table III. The SAIs passed the precision tests with the expected frequency during the period considered.

For the MCP, contributions of variance of SAI measurements on the RMs are mostly caused by counting statistics, which are estimated by $\hat{\sigma}$ (Ref. 4). Therefore, if $\hat{\sigma}$ does not agree with s , as determined by the precision test, we would conclude there may be a problem with the instrument.

E. Tests for Randomness

We suggest continuing the mean square successive difference test (a test for serial correlation) and the policy of conducting an investigation when two points fall consecutively outside control limits.

Visual inspection of control charts should continue to be one of the diagnostic tools for identifying peculiar data patterns. The data should appear to be randomly scattered about the central line. As an example of nonrandomness, many consecutive measurements may lie on one side of the central line. An investigation is suggested if seven consecutive points satisfying this criterion are found; however, there does not appear to be a mathematical justification for the use of seven (Ref. 7, p. 347). Because of the small bias term, this test would fail frequently.

The mean square successive difference test is performed at the PF using the same set of data as the precision test. During the 27-month testing period, the η statistic for IO2 failed

TABLE III

COMPARISON BETWEEN PROPAGATED AND 27-MONTH
ESTIMATES OF STANDARD DEVIATION

<u>Date</u>	<u>Propagated Standard Deviation (g/L)</u> $\hat{\sigma}$	<u>Measurement Standard Deviation (g/L)</u> s
<u>I02</u>		
Jan-July 82	0.0130	0.0113
Aug-Nov 82	0.0137	0.0099
Dec 82-Mar 84	0.0115	0.0108
<u>I03</u>		
Jan 82-Mar 83	0.0153	0.0146
Apr 83-Mar 84	0.0140	0.0133
<u>I04</u>		
Jan 82-Mar 84	0.0190	0.0196
<u>I05</u>		
Jan 82-Nov 83	0.0190	0.0185
Dec 83-Mar 84	0.0200	0.0147
Jan 82-Mar 84	0.0200	0.0183

4 times out of 113 or 3.54%, which compares very favorably with an expected 5%.

A third control criterion for randomness already used in the MCP is to state the process out of control if two points in succession fall outside the warning limits on the accuracy plot. This situation occurred infrequently, as expected.

F. Other Tests

There are other types of quality control tools, such as moving average charts to detect bias or cusum (cumulative sum)

charts to detect sudden shifts in data. However, the use of a multiplicity of criteria can result in an increased incidence of false alarms requiring time and expense to investigate.

To illustrate this point, suppose there are x independent tests with probability α_i that the i th test statistic will fall outside the control limits when the process is really in control. The overall probability, α , of getting a false alarm is given by

$$\alpha = 1 - \prod_{i=1}^x (1 - \alpha_i) .$$

Suppose the number of different tests are 1, 5, and 10. Set $\alpha_i = 0.05$ for all tests. For $x = 1$, $\alpha = 0.05$; $x = 5$, $\alpha = 0.22$; and $x = 10$, $\alpha = 0.40$. Thus, as the number of tests increases, the false-alarm rate increases dramatically. Usually tests are not independent and α values are exaggerated; however, the benefits of employing a limited number of criteria should be clear.

V. CONCLUSIONS

The MCP for the SAIs has been successfully applied over an extended period in an operating plutonium facility. Although this analysis of the program has identified some minor problems with instrument operation such as transcription errors and the existence of a small bias, it has also indicated that these problems have only a slight effect on materials accounting.

The ultimate goal of an MCP is to provide assurance that materials accounting data are reliable. The suggested additions to the MCP discussed in the previous sections and summarized in Table IV should be explored in view of the potential improvements they might make to the quality of accounting data under operating conditions.

TABLE IV

SUMMARY OF CURRENT AND PROPOSED MCPs AT THE
LOS ALAMOS NATIONAL LABORATORY PLUTONIUM FACILITY

Transmission of Data	Accuracy Tests	Precision Tests	Randomness	Data Should Have An Approximate Normal Distribution	Data Should Be Crosschecked
<u>CURRENT MCP</u>					
Operator transmits data from instruments to computer. Data available in near-real-time accounting code. Data are summarized monthly. Control charts analyzed by Nuclear Materials Accountability Group.	Accuracy tests made each work day. The test statistic is the difference between a measured quantity and known standard quantity divided by a propagated standard deviation.	Precision tests involve 15 successive measurements. The test statistic is obtained by taking the ratio of the sample variance and the average propagated variance.	A mean square successive difference test is conducted on 15 successive measurements.	There is no test.	Log books are occasionally crosschecked with data entries by an "accurate" source.
<u>SUGGESTED MCP</u>					
Data transmitted directly from instrument to computer.	Unchanged from above. Propagated standard deviation should be periodically compared for agreement with "long-term" standard deviation.	In addition to above: 1. The distribution of the test statistic is compared with its expected distribution by a Chi-square goodness-of-fit test. 2. Daily precision tests using 15 previous measurements from accuracy tests.	In addition to above, more than 7 successive data points on one side of the central line warrants investigation.	Data are tested periodically by a Chi-square goodness-of-fit test for normality.	Internal instrument records should be crosschecked periodically with history tapes.

ACKNOWLEDGMENTS

I wish to thank J. Tape, J. Markin, D. Smith, and R. Picard for helpful discussions and Sharon Hurdle, Edith Williams, and Charlene McHale for taking care of my transcription errors while doing the typing. I would also like to acknowledge the cooperation given by OS-2, in particular R. Marshall, N. Roberts, and F. Hsue. John Hafer and Lucille Bonner get credit for computer and graphics work. T. K. Li, J. Parker, and S. T. Hsue of Q-1 provided help in understanding the SAI. C. Bonner of MST-10 assisted in understanding and compiling the logbook entries.

REFERENCES

1. A. S. Goldman and R. S. Marshall, "An Assessment of the Measurement Control Program for Balances at the Los Alamos National Laboratory Plutonium Facility," Los Alamos National Laboratory report LA-9975-MS (December 1983).
2. "Nondestructive Assay Measurement Control and Assurance," ANSI Standard N15.36 (1983).
3. "Nuclear Materials" in The Laboratory Manual (Los Alamos National Laboratory, Los Alamos, New Mexico, January 1982), Chap. 6.
4. T. K. Li, T. Marks, and J. L. Parker, "Solution Assay Instrument Operations Manual," Los Alamos National Laboratory report LA-9820-M (September 1983).
5. W. Severe and C. Thomas, "Measurement Control Program for In-Line NDA Instruments," Nucl. Mater. Manage. VIII (Proceedings Issue), 620-633 (1979).
6. C. A. Bennett and N. L. Franklin, Statistical Analysis in Chemistry and the Chemical Industry (John Wiley and Sons, Inc., New York, 1954).
7. K. A. Brownlee, Statistical Theory and Methodology, 2nd Ed. (John Wiley and Sons, Inc., New York, 1965).
8. A. J. Duncan, Quality Control and Industrial Statistics, 3rd Ed. (Richard D. Irwin, Inc., Homewood, Illinois, 1965).

9. A. H. Bowker and G. J. Lieberman, Engineering Statistics, 2nd Ed. (Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1972).
10. R. J. Brouns, B. W. Smith, and R. F. Eggers, "Human Errors in Nuclear Material Accounting and Control Data," Nucl. Mater. Manage. XIII, 48-56 (1984).
11. J. Jaech, Statistical Methods in Nuclear Material Control, (Technical Information Center, United States Atomic Energy Commission, 1973).