CONF-810706--16

LA-UR -81-1964

TITLE: SEQUENTIAL TESTS FOR NEAR-REAL-TIME ACCOUNTING

MASTER

AUTHOR(S): Donald D. Cobb

SUBMITTED TO: 22nd Annual Meeting of the Institute of Nuclear Materials Management, San Francisco, California. July 13-15, 1981

By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes.

The Los Alamos Scientific Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy.

University of California

LOS ALAMOS SCIENTIFIC LABORATORY
Post Office Box 1663 Los Alamos, New Mexico 87545
An Affirmative Action/Equal Opportunity Employer

Form No. 836 R3
St. No. 2629

UNITED STATES
DEPARTMENT OF ENERGY

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

SEQUENTIAL TESTS FOR NEAR-REAL-TIME ACCOUNTING*

Donald D. Cobb
Safeguards Systems Group
Los Alamos National Laboratory
Los Alamos, New Mexico

ABSTRACT

Statistical hypothesis testing is used in the analysis of nuclear materials accounting data for evidence of diversion. Sequential hypothesis testing is particularly well suited for analyzing data that arise sequentially in time from near-real-time accounting systems. The properties of selected sequential tests adapted for this application are described.

I. INTRODUCTION

Near-real-time accounting systems are being developed for various nuclear facilities.[1] The characteristic feature of these systems is the collection of materials-balance data sequentially in time at a frequency commensurate with goals for timely detection without disrupting normal operation of the process.

The purpose of near-real-time accounting systems is to detect anomalies, possibly resulting from diversion, in a timely fashion and then to localize them for investigation. A two-step procedure is used to accomplish these objectives.[2] The detection step involves testing the data using a statistical hypothesis test. The assessment step, which is triggered by a positive indication from the detection step, involves analyzing the available data using additional tests and suitable graphic displays for localization and evaluation of anomalies. Thus the decision problem is divided into two phases: detection and assessment.

Because the data are collected sequentially, statistical techniques developed for sequential decision-making can be invoked. Consider a series of N materials balances obtained sequentially in time from measurements of net transfers of nuclear material across an accounting area boundary and net changes of the inventory within an accounting area boundary. A possible test procedure would be to test each materials balance separately. However, a problem arises with this, or any other repeated fixed-length test procedure, if we want to guarantee a fixed false-alarm probability over the N materials balances. If each fixed-length test were applied with false-alarm probability α, the overall false-alarm probability after N tests generally would be greater than α because for fixed-length tests the number of decisions is equal to the number of tests.[3]

Alternatively, a sequential test, that is a test of variable length, such as the sequential probability ratio test (SPRT), could be applied to the accounting data. The fundamental property of sequential tests is that they are applied at each time step, but only one decision is made based on all the available data. The primary advantage of sequential tests is that they require less data on the average than a fixed-length test to make the same decision with specified error probabilities. They have the disadvantage of not guaranteeing that a decision will be made at a specified time.

If decisions are required at fixed times, then two problems may arise with using sequential tests: first, there must be some means of terminating the test and forcing a decision, if a decision has not been reached by the end of the specified time interval; second, if a decision of "no diversion" is reached before the end of the time interval, the test must be restarted with the penalty of possibly increasing additional false alarms.

After reviewing the properties of fixed-length hypothesis tests, sequential tests are described that are being developed for near-real-time accounting systems.

II. FIXED-LENGTH TEST

Let x be a measured materials balance that is normally distributed with mean μ and variance σ^2. We want to test x under the two composite hypotheses, $H_0: \mu \le 0$ and $H_1: \mu > 0$. This is accomplished by comparing x with a test threshold Z_α such that,

$$\begin{aligned} &\text{(i)} \quad \text{if } x < Z_\alpha, \text{ accept } H_0 \ ; \\ &\text{(ii)} \quad \text{if } x \ge Z_\alpha, \text{ reject } H_0 \ . \end{aligned} \tag{1}$$

*Work supported under the US Department of Energy, Office of Safeguards and Security research and development program.

The test threshold is given by

$$Z_\alpha = k_\alpha \sigma \quad , \tag{2}$$

where k_α is obtained from the relationship*

$$\alpha = 1 - \Phi(k_\alpha) \quad . \tag{3}$$

The probability β of a miss is given by

$$\beta = \Phi[(Z_\alpha - \mu)/\sigma] \quad , \tag{4}$$

and the detection probability, or power of the test, is given by $1 - \beta$.

Figure 1 shows the power curve for a fixed-length hypothesis test with $\alpha = 0.05$. Figure 2 shows power curves obtained by applying the fixed-length test to each materials balance of a series of N materials balances while adjusting the individual test thresholds to maintain the overall α fixed at 0.05. As N increases the detection probability decreases for fixed α. Two comments are in order concerning this result. The desired test thresholds for a series of fixed-length tests generally must be found by computer simulation (i.e., by trial and error). More importantly, the fact that we obtain poorer decisions while accumulating more data over the same time interval violates our intuition. We propose to develop sequential tests for near real-time systems to overcome the problems associated with using fixed-length tests.

III. THE SEQUENTIAL PROBABILITY RATIO TEST (SPRT)

A. The Wald Test

The SPRT was first derived by Wald in the early 1940's as a sequential formulation of the likelihood-ratio test. The Wald test is a test of the simple hypotheses, $H_0: \mu = 0$ and $H_1: \mu = \mu_1$, where μ_1 is a positive constant. For independent, identically distributed data the appropriate test statistic is the CUSUM,

$$C_n = \sum_{i=1}^{n} x_i \quad . \tag{5}$$

The decision rule is given by:

$$\begin{aligned} &\text{(i)} \quad \text{if } C_n \le ZL, \text{ accept } H_0; \\ &\text{(ii)} \quad \text{if } C_n \ge ZU, \text{ reject } H_0; \\ &\text{(iii)} \quad \text{if } ZL < C_n < ZU, \text{ make no decision.} \end{aligned} \tag{6}$$

Note that there are two boundaries, ZL and ZU, and three decision regions: accept H_0; reject H_0; and make no decision, that is, defer a decision until more data are taken. The test is applied at each time step, but one and only one decision (to accept or reject H_0) is reached. The specific forms of the test boundaries for independent data are given by

$$ZU = \frac{n\mu_1}{2} + \frac{\sigma^2}{\mu_1} \ln\left(\frac{1-\beta}{\alpha}\right) \quad \text{and}$$

$$ZL = \frac{n\mu_1}{2} + \frac{\sigma^2}{\mu_1} \ln\left(\frac{\beta}{1-\alpha}\right) \quad , \tag{7}$$

where α and β are the desired false-alarm and miss probabilities. These test boundaries are illustrated as functions of n in Fig. 3, where the no-decision region is denoted by $H_{1/2}$.

To apply the test, the user chooses values of α, β, and μ_1, and the decision rule given by Eq. (6) is repeated at each time step until a decision is reached. The specified values of α and β actually are bounds on the realized values α' and β', given by

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-\frac{1}{2}t^2} dt \quad .$$

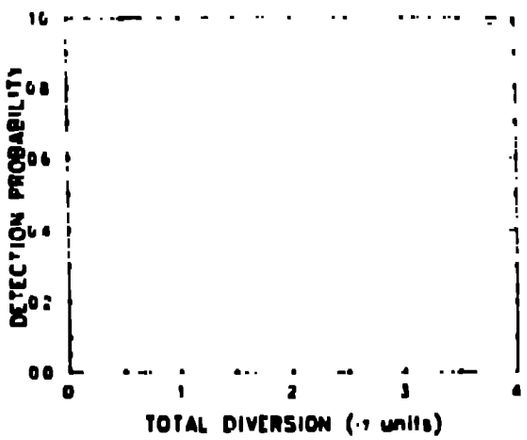

Fig. 1. Power curve for the single materials balance hypothesis test.

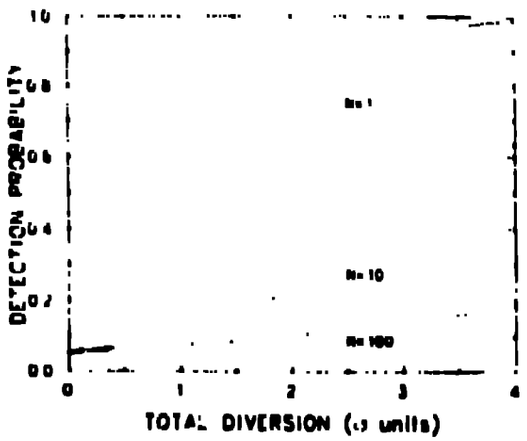

Fig. 2. Power curves for a series of N fixed-length hypothesis tests.

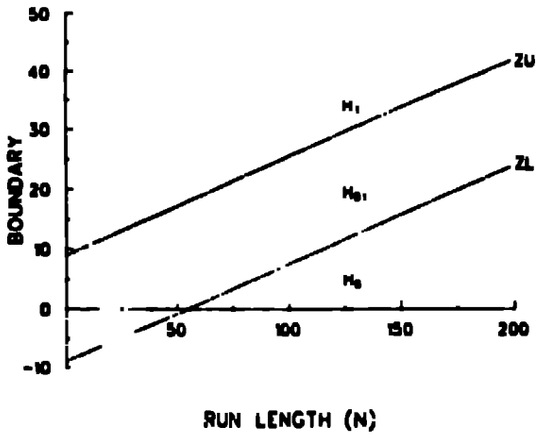

Fig. 3. Wald test boundaries.

$$\alpha' + \beta' \le \alpha + \beta ,$$

$$\alpha' \le \frac{\alpha}{1-\beta}, \quad \text{and} \quad \beta' \le \frac{\beta}{1-\alpha}. \qquad (\ldots)$$

[illegible]. Extension of the Wald Test to Composite Hypotheses

The Wald test was originally developed to test the simple hypotheses, $H_0: \mu = \mu_0$ and $H_1: \mu = \mu_1$. The hypotheses of interest for materials accounting are the composite hypotheses, $H_0: \mu \le \mu_0$ and $H_1: \mu > \mu_1$. We have extended the Wald test to treat composite hypotheses with suitable choices of α, β, and μ_1. The [illegible] μ_0 establishes a tolerable [illegible] $\mu \le \mu_0$ where a loss is [illegible]. The choice of the parameters α, β, and μ_1 can be determined by comparison with a fixed-length test of the CUSUM of N material balances.

$$\text{Fix} \quad \text{if } C_N \ge k_\alpha \sigma_N(0), \text{ accept } H_1 ,$$

$$\text{Fix} \quad \text{if } C_N \le k_\beta \sigma_N(0), \text{ reject } H_1 ,$$

where $\sigma_N(0)$ is the standard deviation of the CUSUM for a fixed-length N-[illegible] data. This fixed-length test is a test of the desired composite hypotheses, and it is equivalent to a single materials balance over all N time periods. Thus, the fixed length test of the CUSUM is equivalent to the usual test of a single materials balance over all N time periods.

The power curves of the Wald test and the corresponding fixed-length CUSUM test can be matched by choosing α, β, and μ_1 as follows

$$\beta = \alpha \quad \text{and} \quad \mu_1 = k_\alpha \frac{\sigma_N(0)}{N} , \qquad (\ldots)$$

where α and N are specified by the user. Power curves for the two tests are compared in Fig. 4. For all practical purposes, they have the same power.

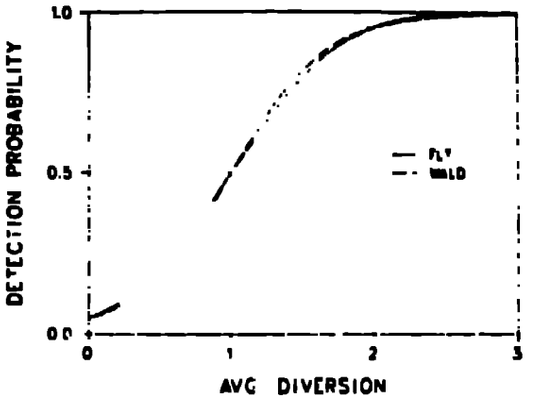

Fig. 4. Comparison of power curves for the Wald test and the corresponding fixed-length cusum test.

The advantage of the Wald test over the fixed-length test lies in the fact that fewer data points are required on the average to make the same quality (α,β) decision; therefore, more timely decisions are possible. This is illustrated by Fig. 5, which shows the ratio of the average run length of the Wald test to the fixed-length CUSUM test. This ratio is always less than unity; for example, if $\mu = \mu_1$, it is 50%, indicating that the Wald test requires only about half as much data on the average to reach the same decision as the fixed-length cusum test.

[illegible]. Extension of the Wald Test to Correlated Data

The Wald test can be extended to treat correlated data in evaluating the likelihood ratio at each time step and comparing the value of the likelihood ratio with the Wald boundaries. In the general case, simple analytic expressions for the test statistic, such as the CUSUM, and for the test boundaries, [illegible], are not available. However, the CUSUM is merely a [illegible] consequent sequential test based on the CUSUM can be [illegible].

We have derived generalized boundaries for a Wald test of correlated data

$$\mu_U = \frac{m_1}{2} + \frac{\sigma_N(0)}{m_1} \ln \left(\frac{1-\alpha}{\alpha}\right) \quad \text{and}$$

$$\mu_L = \frac{m_1}{2} + \frac{\sigma_N(0)}{m_1} \ln \left(\frac{\alpha}{1-\alpha}\right) , \qquad (\ldots)$$

where $\sigma_N(0)$ is the CUSUM standard deviation at each time step and the parameters α, β, and μ_1 are given by Eq. (10). To apply this generalized Wald test, the user must supply values of α and μ_1, and must evaluate the CUSUM and its error variance at each time step.

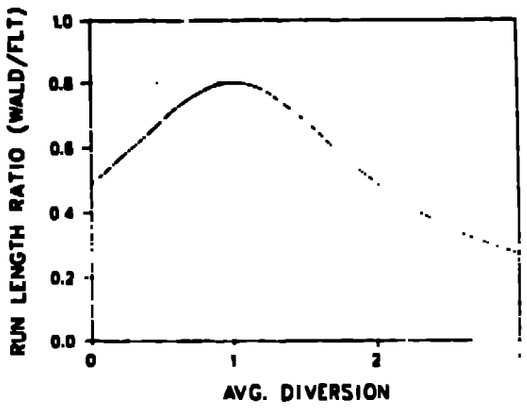

Fig. 5. Ratio of average run lengths of the Wald test and the fixed-length cusum test.

Tables I and II compare the performance of the fixed-length CUSUM test (N = 100) and the Wald test. The data for the Wald test are computer-generated using the boundaries of Eq. (11). Table I shows results for independent data. Table II shows results for data with constant, pairwise correlation (correlation coefficient $\rho = 0.25$). The maximum permitted run length in the computer simulation was 5000 data points.

The effect of serial correlation is evident by comparing the tables. Run lengths of sequences of correlated data that cross boundaries (in less than 500 points) tend to be short; 50% of the decisions are made in run lengths of four or less. There is also a significant probability ($P_{01} \approx 0.1$) that no decision is reached in 500 points. Thus, the probability that a decision will not be made in a reasonable length of time is significant when the data are correlated positively. Conversely, negative correlation enhances the performance of the tests.

These results are illustrated by the run-length distributions shown in Figs. 6 and 7 for $\rho = 0$ and $\rho = 0.25$, respectively, and fixed detection probabilities 0.5 and 0.95. The run-length distributions for independent data are rather broad, but are essentially zero for n > twice the average run length. The run-length distributions for correlated data are sharply peaked at small values of n, but have nonzero tails that persist to very large values of n.

TABLE I

FIXED-LENGTH TEST AND THE WALD TEST: $\rho = 0$

	FLT			WALD				
μ	P_1	P_0	N	P_1	P_0	P_{01}	Avg N	Med N
0	0.05	0.95	100	0.05	0.95	0.00	52	47
0.0804	0.20	0.80	100	0.16	0.84	0.00	75	58
0.1445	0.50	0.50	100	0.50	0.49	0.00	92	69
0.2487	0.80	0.20	100	0.83	0.17	0.00	76	59
0.3290	0.95	0.05	100	0.96	0.04	0.00	53	43

TABLE II

FIXED-LENGTH TEST AND THE WALD TEST: $\rho = 0.25$

	FLT			WALD				
μ	P_1	P_0	N	P_1	P_0	P_{01}	Avg N	Med N
0	0.05	0.95	100	0.03	0.87	0.10	11	3
0.4077	0.20	0.80	100	0.10	0.63	0.27	18	4
0.8367	0.50	0.50	100	0.31	0.32	0.37	19	4
1.2418	0.80	0.20	100	0.63	0.11	0.26	18	4
1.5695	0.95	0.05	100	0.87	0.03	0.10	12	3

D. Terminating the Wald Test

If we want to guarantee that a decision is made by the end of a specified time period (say N balances), then we must terminate the sequential test. We want to know the power of the terminated test and the risk of a false alarm if we terminate the test. A termination procedure that can be justified mathematically[4] is to terminate the Wald test using the fixed-length CUSUM test of N balances, i.e., the single materials balance test.

The modified decision rule is: (i) apply the Wald test for $n \leq N$ using the boundaries given by Eq. (11); (ii) if no decision is reached after $n = N$, apply the fixed-length decision rule given by Eq. (9). This procedure guarantees that a decision to accept or reject H_0 will be reached in any specified number N of material balances.

Tables III and IV compare the performance of the fixed-length CUSUM test and the terminated Wald test. Bounds on α and β are still maintained, and fewer data are still required for detection, on the average, but a decision is now guaranteed after any specified number N of materials balances, that is, after any specified time period.

E. Page's Test

Another sequential test similar to the Wald test is the so-called Page's test.[5] The decision rule is given by:

(i) if $C_n \geq n\frac{\mu_1}{2} - \ln \alpha$, reject H_0; (12)

(ii) if $C_n \leq n\frac{\mu_1}{2}$, accept H_0;

(iii) if otherwise, make no decision.

The upper boundary of this test is identical to a Wald test with β set to zero, but the lower boundary of Page's test does not correspond to any choice of α and β. Therefore, the properties of Page's test generally must be determined by computer simulation. A two-sided Page's test, consisting of two one-sided tests, does have well-defined behavior and has been proposed for use in safeguards.[7] The two-sided Page's test is the same as a two-sided Wald test with $\beta = 0$. (Setting $\beta = 0$ places the acceptance region at infinity.)

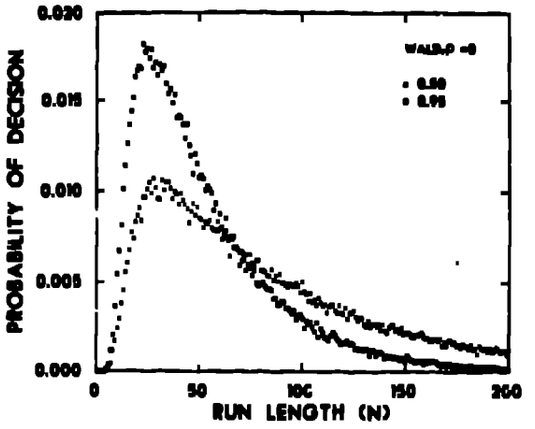

Fig. 6. Run-length distributions for the Wald test applied to uncorrelated data ($\rho = 0$).

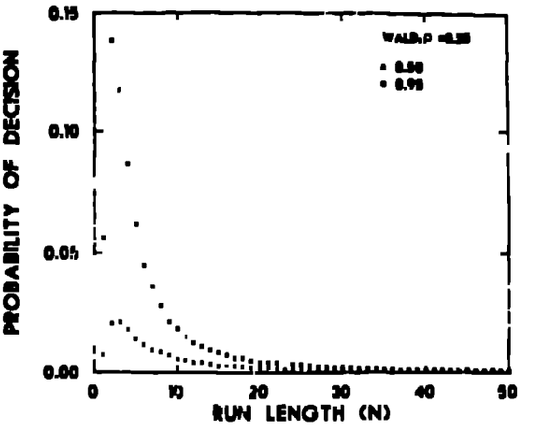

Fig. 7. Run-length distributions for the Wald test applied to serially correlated data ($\rho = 0.25$).

Page's test is interesting because the problem of restarting the test after the lower boundary has been crossed is considered.[6] This problem arises because, in the practical situation, we want to detect a change from the condition that H_0 is true to the condition that H_1 is true, and we want to detect this change as rapidly and as accurately as possible. There currently is no mathematically rigorous treatment of this problem; however, one procedure, formulated by Page,[6] consists of restarting the test each time the lower boundary is crossed (i.e., accepting H_0 on subsets of the data) and continuing to search for a rejection of H_0 as additional data are collected. Rejection of H_0 is based on testing all contiguous subsequences of the data to find the subsequence that produces the largest change in the CUSUM value. Unfortunately, there is currently no way, other than by computer simulation, of implementing Page's procedure with guaranteed *a priori* error probabilities.

The Test of Power One

An SPRT has been developed by Robbins and Siegmund[8] specifically for testing the composite hypotheses, $H_0: \mu \le 0$ and $H_1: \mu > 0$. This so-called test of power one guarantees that $\beta = 0$ for any value $\mu > 0$ and that α is bounded for $\mu \le 0$. The power curve of this test jumps discontinuously from the specified value of α at $\mu = 0$ to unity at $\mu = \delta$ (an arbitrarily small positive quantity) and is unity for all positive values of μ.

The test of power one was originally developed for independent, normally distributed data by applying the law of the iterated logarithm to certain classes of test functions.[8] The decision rule is given by:

(i) if $C_n < Z^*$, make no decision;

(ii) if $C_n \ge Z^*$, reject H_0. (13)

There is only one boundary and it separates the rejection and no decision, or normal operating, regions. There is no acceptance region. Thus, the test never terminates and never needs to be restarted, except following a rejection.

The boundary of the power one test also can be derived from an extension of the likelihood ratio. This has the advantage that there is a more direct connection with Wald's SPRT so that the error probabilities and specific functional form of the test boundaries can be investigated. This approach starts with a ratio of suitable

TABLE III

FIXED-LENGTH TEST AND THE TERMINATED WALD TEST: $\rho = 0$

	FIT			TWALD			
μ	P_1	P_0	N	P_1	P_0	Avg N	Med N
0	0.05	0.95	100	0.06	0.94	49	43
0.1806	0.20	0.80	100	0.21	0.79	41	37
0.1445	0.50	0.50	100	0.49	0.51	67	60
0.2487	0.80	0.20	100	0.78	0.22	65	50
0.3290	0.95	0.05	100	0.94	0.06	50	43

TABLE IV

FIXED-LENGTH TEST AND THE TERMINATED WALD TEST: $\rho = 0.25$

	FIT			TWALD			
μ	P_1	P_0	N	P_1	P_0	Avg N	Med N
0	0.05	0.95	100	0.05	0.95	16	3
0.4073	0.20	0.80	100	0.20	0.80	33	6
0.8347	0.50	0.50	100	0.49	0.51	63	12
1.2618	0.80	0.20	100	0.79	0.21	35	7
1.6695	0.95	0.05	100	0.95	0.05	18	3

weighted density functions so that the form of the likelihood ratio is given by:

$$L(x) = \frac{\int_{H_1} w_1(\theta)\, p_1(x,\theta)\, d\theta}{\int_{H_0} w_0(\theta)\, p_0(x,\theta)\, d\theta} \quad , \qquad (14)$$

where p_0 and p_1 are density functions under H_0 and H_1, respectively, and w_0 and w_1 are weight functions over the parameter spaces of H_0 and H_1.

For normally distributed, independent data, the likelihood ratio of Eq. (14) can be evaluated explicitly for various choices of weight functions. In particular, if we choose w_1 to be gaussian on the half plane $\theta > 0$ and w_0 to be a delta function at $\theta = 0$, and if we set $\beta = 0$, we obtain a test boundary that has the same form as derived by Robbins and Siegmund[8] for the power-one test. We have derived a general form of this boundary that is useful for correlated data as well. It is given by:

$$ZU = \sigma_c(n)\, [(1 + \frac{1}{n}) \ln (\frac{1 + n}{4\alpha^2})]^{1/2} \quad , \qquad (15)$$

where α is the desired value of the false-alarm probability. This boundary is shown in Fig. 8 for $\alpha = 0.05$. To apply this test, the user must specify α and must evaluate the CUSUM and its error variance at each time step.

Tables V and VI show comparisons of the fixed-length CUSUM test for N = 100 with the power-one test. The false-alarm probability is bounded, and the power of the test is equal to unity. The average run length of the test decreases as the average diversion (μ) increases. However, as was seen for the Wald test, there is a significant probability that the power-one test will not terminate for very long sequences (>5000 in the examples) if the data are serially correlated. This is also illustrated by Figs. 9 and 10 showing run-length distributions of the power-one test for uncorrelated and correlated data, respectively, and fixed detection probabilities 0.5 and 0.05.

The single materials-balance test over the entire period (i.e., the CUSUM test of N balances) that was used for terminating the Wald test can also be used to terminate the test of power one. Examples of the performance of the terminated power-one test are given in the next section.

IV. APPLICATIONS

A. Measurement-Error Model for Materials Balances

We have simulated sequences of materials-balance data using previously developed mathematical models of measurement errors.[5] The variance of each materials balance is given by:

$$\sigma^2_{MB} = \sigma^2_{\Delta I} + \sigma^2_{\Delta T} \quad , \qquad (16)$$

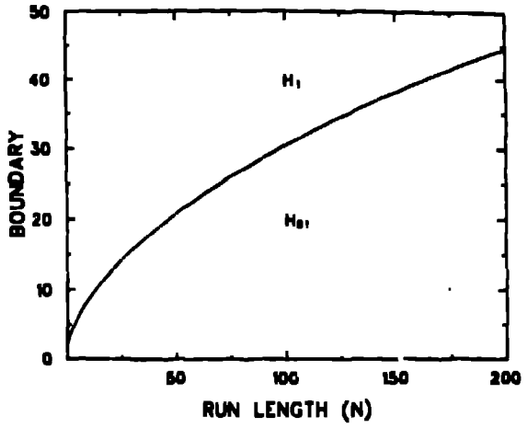

Fig. 8. Boundary of the power-one test

where $\sigma^2_{\Delta I}$ and $\sigma^2_{\Delta T}$ are the variances of the measured inventory change and net transfer, respectively, and we have assumed that inventory measurements are not correlated with transfer measurements.

For materials balance data with constant, pairwise correlations among the transfer measurements, the form of the CUSUM variance $\sigma^2_c(n)$ can be determined by specifying the following parameters:

(i) the ratio r of the net-transfer variance to the inventory-change variance,

$$r = \frac{\sigma^2_{\Delta T}}{\sigma^2_{\Delta I}} \quad ; \qquad (17)$$

(ii) the correlation coefficient ρ between pairs of transfer measurements; and

(iii) a scale factor that we choose without loss of generality by setting $\sigma^2_{\Delta I} = 1$.

The CUSUM variance in terms of these parameters is given by

$$\sigma^2_c(n) = \frac{1}{1 + r} [1 + n(1 - \rho)r + n^2 \rho r] \quad , \qquad (18)$$

where $\sigma^2_{\Delta I} = 1/(1 + r)$. This form of the CUSUM variance is equivalent to that previously given in Ref. 5, where the CUSUM variance was scaled by $\sigma^2_{\Delta I}$. Positive values of ρ arise because systematic errors cause the transfer measurements to be correlated.

B. Terminated Power-One Tests

Tables VII-XII show results of applying the terminated power-one test to sequences of materials-balance data for various choices of the measurement-error-model parameters r and ρ.

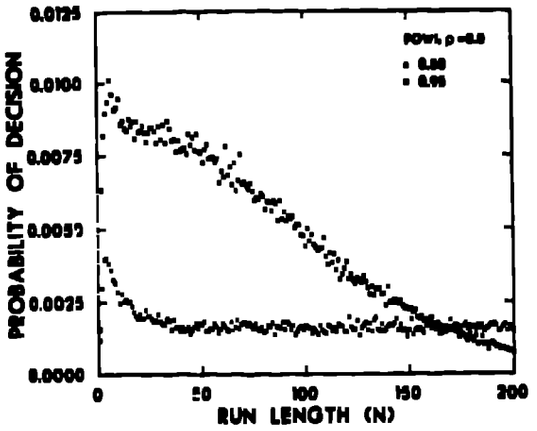

Fig. 9. Run-length distributions for the power-one test with $\rho = 0$.

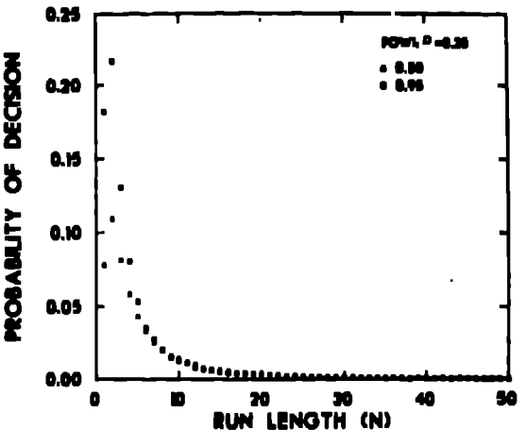

Fig. 10. Run-length distributions for the power-one test with $\rho = 0.25$.

This procedure is shown to be robust in the sense that it performs well over a wide range of possible measurement errors.

This approach, based on the CUSUM statistic, generally is robust with respect to different diversion strategies as well. Test statistics other than the CUSUM may perform better in certain circumstances.[5] Generally speaking, for any measurement-error structure (i.e., any values of r and ρ), the optimal strategy to defeat near-real-time systems, optimal from the diverter's point of view, is some combination of block diversion on a relatively short time scale, corresponding to the so-called inventory-error-dominated regime, and protracted diversion over a relatively long time scale, corresponding to the so-called transfer-error-dominated regime. The optimal test procedure to counter this mixed diversion strategy must still be developed.

The terminated power-one test can also be used with other statistics besides the CUSUM. These statistics must be normally distributed at each time step and be describable as a Wiener process.[8] Because the power of the test is always guaranteed, at least in principle, to be unity, the main reason for using test statistics other than the CUSUM is to enhance the ability of the test to respond rapidly to a change in conditions from H_0 to H_1. Our preliminary work in this area indicates that a statistic, based on Page's procedure, which looks for the maximum change in the CUSUM, may be better than the CUSUM itself. Other statistics derived for specific diversion strategies can also be used with the power-one test.

C. Detection and Assessment

The materials balance data are tested using a two-step detection and assessment procedure.[2] In step one, the power-one test, which has no lower boundary for $n \leq N$, may be used to take a single pass through the data looking for evidence of diversion. This is the detection phase. As long as the upper boundary is not crossed, we continue to collect more data and apply the power-one test up to $n = N$, at which time the test is terminated using the fixed-length CUSUM test for $n = N$ to force a decision. Note that a decision having well-defined error probabilities can be obtained at any desired time by this procedure.

This decision is guaranteed to have at least as good error probabilities as the single materials balance test, and, in general, the decisions will be more timely in that fewer data are required. Thus, one of the main objections to multiple materials balances, that is, uncontrolled error probabilities, is eliminated.

TABLE V

FIXED-LENGTH AND POWER ONE TESTS: $\rho = 0$

	FLT			POWI			
μ	P_1	P_0	N	P_1	P_{01}	Avg N	Med N
0	0.05	0.95	105	0.03	0.97	240	94
0.0804	0.20	0.80	100	0.98	0.01	1560	1380
0.1445	0.50	0.50	100	1.00	0.00	330	280
0.2487	0.80	0.20	100	1.00	0.00	130	110
0.3290	0.95	0.05	100	1.00	0.00	76	64

TABLE VI

FIXED-LENGTH AND POWER ONE TESTS: $\rho = 0.25$

	FLT			POWI			
μ	P_1	P_0	N	P_1	P_{01}	Avg N	Med N
0	0.05	0.95	100	0.01	0.99	15	7
0.4077	0.20	0.80	100	0.05	0.95	34	8
0.8347	0.50	0.50	100	0.20	0.80	56	8
1.2018	0.80	0.20	100	0.50	0.50	25	6
1.6695	0.95	0.05	100	0.78	0.22	16	5

TABLE VII

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 100$, $\rho = 0$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.06	0.94	94
0.0800	0.20	0.80	100	0.21	0.79	87
0.1637	0.50	0.50	100	0.51	0.49	73
0.2474	0.80	0.20	100	0.80	0.20	58
0.3274	0.95	0.05	100	0.95	0.05	48

TABLE VIII

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 100$, $\rho = 0.25$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.05	0.95	95
0.4057	0.20	0.80	100	0.20	0.80	81
0.8306	0.50	0.50	100	0.49	0.51	53
1.2555	0.80	0.20	100	0.79	0.21	26
1.6612	0.95	0.05	100	0.95	0.05	11

TABLE IX

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 1$, $\rho = 0$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.06	0.94	95
0.0571	0.20	0.80	100	0.22	0.78	88
0.1169	0.50	0.50	100	0.51	0.49	73
0.1767	0.80	0.20	100	0.81	0.19	58
0.2348	0.95	0.05	100	0.95	0.05	46

TABLE X

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 1$, $\rho = 0.25$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.05	0.95	95
0.2884	0.20	0.80	100	0.20	0.80	81
0.5904	0.50	0.50	100	0.49	0.51	54
0.8924	0.80	0.20	100	0.79	0.21	26
1.1807	0.95	0.05	100	0.95	0.05	11

TABLE XI

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 0.01$, $\rho = 0$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.09	0.91	87
0.0113	0.20	0.80	100	0.31	0.69	74
0.0231	0.50	0.50	100	0.67	0.33	57
0.0350	0.80	0.20	100	0.92	0.08	45
0.0463	0.95	0.05	100	0.99	0.01	40

TABLE XII

FIXED-LENGTH AND TERMINATED POWER ONE TESTS:
$r = 0.01$, $\rho = 0.25$

	FLT			TPOW1		
μ	P_1	P_0	N	P_1	P_0	Avg N
0	0.05	0.95	100	0.07	0.93	88
0.0414	0.20	0.80	100	0.24	0.76	74
0.0847	0.50	0.50	100	0.54	0.46	52
0.1280	0.80	0.20	100	0.82	0.18	30
0.1693	0.95	0.05	100	0.95	0.05	19

If the upper boundary is crossed, we move to step two, which is the assessment phase. All contiguous subsequences of the data available at the time of detection are tested to determine their significance levels. The Wald test, extended for composite hypotheses, is a convenient choice for the assessment phase. The most significant subsequence, that is, the sequence of data with the largest value of the CUSUM divided by its standard deviation, is used to estimate the time, location, and amount of loss. Note that this corresponds to Page's procedure.The results of the assessment phase are displayed graphically using alarm charts.[1,2]

Note that the assessment procedure involves searching for the maximum change in the CUSUM. In future work we hope to combine Page's procedure with the power-one test and use this new test for the detection phase. Thus, a truly integrated detection and assessment procedure could result from this approach.

ACKNOWLEDGMENT

The author acknowledges contributions to this work from several colleagues in the Safeguards Systems Group, especially J. P. Shipley, D. B. Smith, and J. T. Markin.

REFERENCES

1. D. D. Cobb, E. A. Hakkila, H. A. Dayem, J. P. Shipley, and A. L. Baker, "Development and Demonstration of Near-Real-Time Accounting Systems for Reprocessing Plants," presented at the 22nd Annual Meeting of the Institute of Nuclear Materials Management, San Francisco, California, July 13-15, 1981 (published in these proceedings).

2. J. P. Shipley, "Decision-Directed Materials Accounting Procedures: An Overview," presented at the 22nd Annual Meeting of the Institute of Nuclear Materials Management, San Francisco, California, July 13-15, 1981 (published in these proceedings).

3. R. Avenhaus, Material Accountability: Theory Verification and Applications (John Wiley & Sons, Inc., New York, 1977)

4. A. Wald, Sequential Analysis (John Wiley & Sons, Inc., New York, 1947).

5. D. D. Cobb and J. P. Shipley, "Performance Analysis of Nuclear Materials Accounting Systems," Nucl. Mater. Manage. VIII(2), 81-92 (Summer 1979).

6. E. S. Page, "Continuous Inspection Schemes," Biometrika 41, 100 (1954).

7. A. J. Woods, D. J. Pike, and D. M. Rose, "Analysis and Interpretation of Materials Accounting Data," Nucl. Mater. Manage. IX (Proceedings Issue), 230-235 (1980).

8. H. Robbins and D. Siegmund, "Confidence Sequences and Interminable Tests," Bulletin of International Statistics Institute 43, 379 (1969).