Conf-940748--49

LA-UR-94-2292

Title: MULTIVARIATE DIAGNOSTICS AND ANOMALY DETECTION FOR NUCLEAR SAFEGUARDS

Author(s): Tom Burr, James Jones, and Larry Wangen

Submitted to: 35th Annual Meeting of the Institute of Nuclear Materials Management, Naples, Florida, July 17-20, 1994

Los Alamos
NATIONAL LABORATORY

Form No. 836 R5
ST 2629 10/91

Multivariate Diagnostics and Anomaly Detection for Nuclear Safeguards*

Tom Burr
Safeguards Systems Group (NIS-7), MS E541
Los Alamos National Laboratory
Los Alamos, NM 87545 USA

James Jones
UCLA Mathematics Department
520 Kelton Apts, 306
Los Angeles, CA 90024 USA

Larry Wangen
International Atomic Energy Agency
Wagramerstrasse 5, P. O. Box 100
A-1400, Vienna, Austria

ABSTRACT

We first review recent literature that applies multivariate Shewhart and multivariate cumulative sum (Cusum) tests to detect anomalous data. These tests are used to evaluate residuals obtained from a simulated three-tank problem in which five variables (volume, density, and concentrations of uranium, plutonium, and nitric acid) in each tank are modeled and measured. We then present results from several simulations involving transfers between the tanks and between the tanks and the environment. Residuals from a no-fault problem in which the measurements and model predictions are both correct are used to develop Cusum test parameters which are then used to test for faults for several simulated anomalous situations, such as an unknown leak or diversion of material from one of the tanks. The leak can be detected by comparing measurements, which estimate the true state of the tank system, with the model predictions, which estimate the state of the tank system as it "should" be. The no-fault simulation compares false alarm behavior for the various tests, whereas the anomalous problems allow us to compare the power of the various tests to detect faults under possible diversion scenarios. For comparison with the multivariate tests, univariate tests are also applied to the residuals.

INTRODUCTION AND MOTIVATION

For process control and other reasons, new and future nuclear reprocessing plants are expected to be increasingly more automated than older plants. As a consequence of this automation, the quantity of data potentially available for safeguards may be much greater in future reprocessing plants than in current plants. These data will consist of control data and physical and chemical measurements of process inputs and outputs during plant operations. It can also include traditional data from inventories and transfers of nuclear materials. Not only will more process variables be monitored, but data collection will be more frequent than in the past. If workable methods of authenticating and analyzing these data can be developed, they should be useful for safeguards.

Recent developments from two different, but related, fields are applicable to the use of these data for safeguards: 1) process fault detection and diagnosis results from the chemical engineering field and 2) quality control assessment methods from the statistical and quality control field. For this report, our main interest in the fault-detection literature is the concept of data redundancy provided by measurements and system models. Our main interest in the quality control field is a recent multivariate version of a fault detection statistic.

Process Fault Detection and Diagnosis

For many years the chemical industry has been considering process control issues that are directly applicable to reprocessing plant safeguards. A significant amount of recent work has been done in the area of process fault detection and diagnosis.[1] A fault or abnormal condition occurs when some state of the chemical process, e.g., temperature, pressure, or mass of plutonium, is outside of acceptable limits. Fault detection is based on the availability of *redundant* process information. The methods consist of two general categories based on the kind of redundant information possessed about a process. Both categories assume the availability of measurement data. "Measurement-based" methods rely on redundant measurements and include methods that use historical data as target values or set-points against which new measurements can be compared. "Analytical-based" methods involve redundant

*This work supported by the US Program for Technical Assistance to IAEA Safeguards.

information obtained from a mathematical model* of the process. The concept is illustrated in Fig. 1 where model estimates are compared with measurements to determine the presence or absence of a fault with some desired degree of confidence.

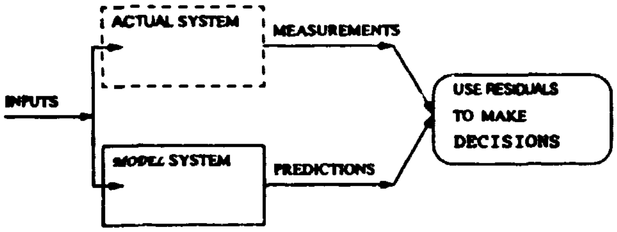

Fig. 1. Process fault detection.

If a process is operating with no faults, we expect that all of the redundant information will be consistent except for the unavoidable presence of modeling and measurement errors. Fault detection methods are usually designed to estimate residuals, i.e., deviations between what is expected and what is observed, based on this redundant information. So measured values are compared with model predictions of the same variable. For example, nuclear reprocessing plants periodically calculate the amount of nuclear material that is unaccounted for (MUF). In this case, a combination of measurements and materials-balance relationships constitute the redundant data. If the MUF is too large, as judged by our knowledge of the uncertainty in its calculation, a possible fault has occurred and some action is required.

Cusum for Testing Residuals

Having compared the redundant data provided by the model predictions and the measurements and obtained the residuals, the next stage in fault detection is determining when the residuals are "large" enough to indicate a process fault. At this stage, monitoring residuals is a version of outlier detection that is used in statistical process control. If a residual is judged to be an outlier, then the goal is to isolate the location and determine the cause of the fault. For safeguards, appropriate action might be to investigate whether a diversion has occurred and, if so, to determine where and how it occurred, and to identify appropriate follow-up action. Obviously fault diagnosis is more difficult than fault detection because it requires more detailed knowledge of the processes and more numerous and specific measurements.

Many different outlier tests for fault detection have been proposed in the literature.[1] They include maximum likelihood ratio, sequential likelihood ratio, cumulative sum (Cusum), Bayesian and the univariate z or student t tests. Multivariate residuals with multinormal distributions can be individually evaluated by univariate tests or by multivariate tests, such as tests that use the Mahalanobis distance. In this paper we investigate recently proposed multivariate Cusum[2] and Shewhart tests for evaluation of residuals from a three-tank system involving transfers of nuclear material between the tanks and the environment. It may be advantageous to use multivariate Cusum tests for testing the long sequences of data (several hundred successive multivariate observations) expected to be available in future automated reprocessing plants. We know of no published attempt to monitor a vector-valued residual as we are proposing here. In safeguards, Page's test is the most commonly used univariate Cusum test.

We now review a few candidate outlier tests that could be used to monitor a multivariate residual time series.

Individual Outlier Tests

The univariate test statistic for residual p of the residual vector $\mathbf{r}_t$ is

$$Z_{tp}=\frac{r_{tp}}{\sigma} \quad , \tag{1}$$

where σ is the known standard deviation of the pth residual. Test $E(r_{tp}) = 0$ versus $E(r_{tp}) \neq 0$ (E denotes expected value) by using quantiles of the standard normal distribution. To correct for multiple testing (there are p variables), use the conservative Bonferroni method,[3] which replaces the desired false alarm rate α by α/p.

Individual outlier tests as just described will fail to detect slow sustained anomalies. Therefore, the univariate Cusum test might be preferred. One univariate Cusum test for evaluating a particular scalar residual is

$$S_t=\max\left\{0, S_{t-1}+r_{tp}-h\sigma\right\} \ . \tag{2}$$

Outlier Tests

The vector-valued residual $\mathbf{r}_t$, r_{tp}, is expected to be zero, $\mathbf{H}_0$: $E(r_{tp}) = 0$ versus the alternative hypothesis $\mathbf{H}_1$: $E(r_{tp}) \neq 0$. The univariate test statistic for variable p of r_{tp} is

---

* Knowledge-based models, such as expert systems, are also suggested as a type of process model for use in fault detection. In this report we are only considering mathematical models of the process.

$$z_{tp}(\alpha) = \frac{r_{tp}}{\sigma/\sqrt{n}} \quad . \tag{3}$$

The standard deviation σ is known. The variable n is the number of samples used in the calculation ($n = 1$ for this work) and E denotes the expected value. The critical values to which these test statistics are to be compared come from the normal distribution with a mean of 0 and a standard deviation of 1, N(0,1). The user specifies what significance level (α value) will be used to signal a fault, depending on the number of false alarms to be tolerated. For uncorrelated multivariate normal distributions, if we wish to maintain the same overall significance level for detecting a fault, tests for individual residuals use the Bonferroni method.[3] This method replaces α by α/p to account for the multiple tests, where p is the number of individual z values being tested. If the standard deviation is not known but must be estimated, critical values from the student t distribution are used.

The univariate Cusum version of this test for evaluating a particular scalar, r_{tp}, versus a target value, a, is

$$S_t = \max\left[0, S_{t-1} + \left(r_{tp} - a\right) - k\sigma\right] \quad . \tag{4}$$

In this case, σ is the standard deviation of r, $k>0$, and $S_0=0$. To test for a decrease in the mean, the same equation can be used if we define $r_t = -r_t$.

The multivariate Shewhart test, Hotelling's T^2, is

$$T_t = +\left[(\mathbf{r}_t - \mathbf{t})'\Sigma^{-1}(\mathbf{r}_t - \mathbf{t})\right]^{1/2} \quad , \tag{5}$$

where Σ is the covariance matrix of the $\mathbf{r}_t$ under no fault conditions and $\mathbf{t}$ is the *target vector*, which is assumed to be $\mathbf{0}$. T, a positive scalar, is compared to the Shewhart control limit defined by the desired false alarm rate and, if it is larger, an alarm is sounded. In the present application, the target vector $\mathbf{t}$ is zero. The T_t are compared to user-specified critical values from the chi-squared distribution with p degrees of freedom.

Clearly, T_t tests only the residual at time t. A simple scalar-valued cumulative sum of the T_t is defined as follows:

$$(COT)_t = \max\left[0, (COT)_{t-1} + T_t - k\right] \tag{6}$$

$$COT_0 = 0 \quad .$$

$(COT)_t$ is compared to the critical value h specified to give a desired average run length or false alarm rate. The value h can be found by simulation. Here k (>0) is a parameter that is adjusted to obtain a statistic with desired properties.

In analogy to the univariate Cusum,* Crosier[2] describes a vector-valued Cusum ($\mathbf{S}_t$) as follows: Define a scalar quantity C_t, the length of $\mathbf{S}_{t-1} + \mathbf{r}_t$, according to Eq. 7:

$$C_t = \left[\left(S_{t-1} + r_t\right)' S^{-1}\left(S_{t-1} + r_t\right)\right]^{1/2} \quad . \tag{7}$$

Specify

$$\mathbf{S}_t = \mathbf{0} \qquad \text{if } C_t \le k$$

$$\text{or} \quad \mathbf{S}_t = (\mathbf{S}_{t-1} + \mathbf{r}_t) - \mathbf{k} \qquad \text{if } C_t > k \tag{8}$$

where the vector $\mathbf{k}$ is

$$\mathbf{k} = (\mathbf{S}_{t-1} + \mathbf{r}_t)k/C_t \quad . \tag{9}$$

By construction, the vector $\mathbf{k}$ is in the same direction as $\mathbf{S}_{t-1} + \mathbf{r}_t$ but of shorter length. Thus the effect of Eq. 8 is to shrink the cumulative sum vector $\mathbf{S}_t$ towards the origin along its direction. Using Eq. 9, Eq. 8 can be written as

$$\mathbf{S}_t = (\mathbf{S}_{t-1} + \mathbf{r}_t)(1 - k/C_t) \quad , \tag{10}$$

where, as above, the *target vector* that would normally be subtracted from $\mathbf{S}_{t-1} + \mathbf{x}_t$ is taken to be $\mathbf{0}$. In this procedure $\mathbf{S}_0 = \mathbf{0}$ and $k > 0$. Upon calculation of $\mathbf{S}_t$, a scalar is calculated

$$Y_t = \left(\mathbf{S}_t'\Sigma^{-1}\mathbf{S}_t\right)^{1/2} \quad , \tag{11}$$

which is tested against the critical value, h, specified to give the desired false alarm rate.

Three Tank System

We applied the above multivariate Cusum tests to simulated data from a three-tank system (Fig. 2) containing nitric acid, plutonium, and uranium. More details can be found in Ref. 4. The dynamics are described by a system of coupled differential equations based on total mass balances and on individual mass balances for each chemical species for each tank [Eq. (12)].

[Time rate of
change of mass] = [Mass in] − [Mass out] . (12)

---

* The univariate Cusum to detect an increase in the mean is $S_t = \max(0, S_{t-1} + (x_t - a) - k\sigma)$. Where a is the target value, for the mean, σ is the standard deviation of the x's, $k > 0$, and $S_0 = 0$. For a decrease in the mean the same equations can be used if we define $x_t = -x_t$.

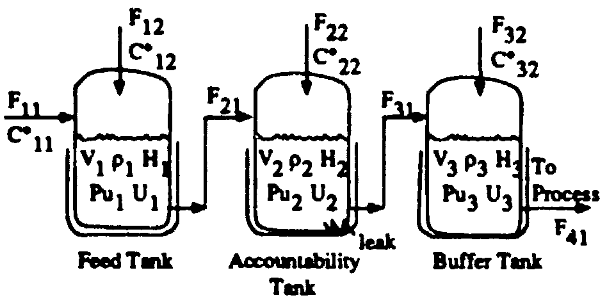

Fig. 2. A three-tank system with three non-reacting constituents.

For given input flows, initial tank volumes and initial concentrations of nitric acid, plutonium, and uranium, the differential equations are solved to give the outputs, i.e., the volumes and concentrations in the tanks at various times. When density is a linear function of concentration, the equations for Tank 1 are

$$
\begin{aligned}
dV_1/dt &= F_{11} + F_{12} - F_{21} \qquad (13)\\
dH_1/dt &= [(H^o{}_{11}\, F_{11} + H^o{}_{12}\, F_{12}) - (F_{11} + F_{12})\, H_1] / V_1\\
dPu_1/dt &= [(Pu^o{}_{11}\, F_{11} + Pu^o{}_{12}\, F_{12}) - (F_{11} + F_{12})\, Pu_1] / V_1\\
dU_1/dt &= [(U^o{}_{11}\, F_{11} + U^o{}_{12}\, F_{12}) - (F_{11} + F_{12})\, U_1] / V_1
\end{aligned}
$$

with analogous equations for Tanks 2 and 3. The density of each tank solution is determined from empirical relationships between densities and concentrations of nitric acid, plutonium, and uranium.[5,6] The system of equations is solved by the Euler method for modeling flexibility. The implementation of the model requires that initial volumes and concentrations are known from measurements (including error) at the start of each simulation. Densities are then estimated by using the empirical relationships mentioned above applied to the estimated concentrations. Flow rates are also required for model simulations, and these are assumed to be available from measurements at 6-min. intervals. Thus the flow rates used in the model F_{11}, F_{12}, etc., are changed after every six minutes of simulation time. The volume, density, and concentrations of plutonium, uranium, and nitric acid in each tank are the model predictions that are compared to measured values to give the residuals that are tested for possible faults. The measured values are obtained by adding zero-mean Gaussian random errors having variances determined by the percent standard deviations given in Table I to the known true values. The measured values are assumed to be available every 30 min. Thus after each 30 min. of elapsed time, the residuals are calculated and tested by the above Cusum tests. Then the model values are updated to the new measured values and the simulation continues until all Cusum tests have alarmed. Because measured values at the end of each 30-min. period are used both to calculate the residuals for that period and to serve as starting values to generate the next period's predictions, there will be negative serial correlation in the residuals (except for the residuals for densities in each tank, because of the way density is estimated). The presence of serial correlation makes our sequential tests perform differently than if the residuals were not correlated. We will discuss this effect in the results section.

Table I. Percent Standard Deviations Assumed for Measured Flow Rates, Volumes, Concentrations, and Densities

Flow rates	2.0
Volume	0.2
Nitric acid	1.0
Plutonium	0.2
Uranium	0.4
Density	0.2

No-fault Simulation

Application of the above Cusum tests requires knowledge of the covariance matrix Σ for no-fault conditions. To prevent creating problems because of a changing covariance matrix as we step along in time, the no-fault simulations were performed under steady-state conditions. One thousand 20-h simulations were performed as outlined above, and several estimates of the covariance matrix Σ, corresponding to different elapsed times, were compared and found to be the same, within errors attributable to using a finite number of simulations. Thus one value of Σ was used for all of the testing. After acquiring Σ, a sufficient number of no-fault simulations were performed and the residuals tested to find a value for h giving an average run length of 200, that is, an average false alarm rate of 0.5%. For all tests except the COT, k values of 0.5 were used. For COT, k was set to $p^{1/2}$ following Crosier.[2]

Anomalous Simulations

Two basically different diversion scenarios were simulated; the first is equivalent to a slow leak from one or more of the tanks and the second is similar to the first except the removed liquid solution is replaced with water or water plus nitric acid and uranium. These removals were assumed to occur at a constant rate and to continue until an alarm was sounded. Several different loss rates were tested.

RESULTS AND DISCUSSION

No-fault Simulations

Figure 3 shows the distribution of run lengths giving an average run length (ARL) of 200 for the univariate, principal components univariate, COT, and Crosier Cusum tests. These distributions are based on 10 000 alarms for each test, but only 1000 run lengths are plotted in the histograms. Mean and median run lengths are given in Table II. These "sample means" can be considered to be within 3 of the true means in all cases and in most cases are within 1 of the true means. That is, in the worst case (longest true average run length), the standard deviation of the sample average run length is about 1.5, so an approximate 95% confidence interval (CI) for the true average run length would extend from the sample mean (mean of the 10 000 run lengths for each case) plus or minus 3. Most of the standard deviations of the sample means are much less than 0.5, so a CI extending from the sample mean plus or minus 1 would have greater than 95% coverage probability.

Note from Fig. 3 that the distributions are highly skewed to the right indicating that most runs alarm at run lengths less than the average. Nearly two-thirds of the runs alarm in less than 200 tests. However, once in a while there are very long run lengths before a false alarm sounds. As always, it may be preferable to use the median run length rather than the mean, but we restrict this discussion to the mean run length. The ARLs in our simulations were sensitive to values of h in accord with the results of Crosier's work.[2]

Anomalous Situations

Results for the simulations with faults are summarized in Figs. 4 and 5 and Tables II and III. Figures 4 and 5 compare the run length distribution for a 0.3 L/h leak from tank 1 in the two cases: replace the leak with water (Fig. 4) or do not replace the leak (Fig. 5).

The run length distributions shown in Figs. 4 and 5 are similar to those for the no-fault case in that they are asymmetric, skewed to the right. These demonstrate that even though some faults are rapidly detected on the average, there are occurrences when the same fault may not be detected for an uncomfortably long time using these tests. They suggest the median may be a better measure of the tests' ability to detect material losses. Also, note the unusually large number of long run lengths using COT in Fig. 5. We have noticed this behavior of the COT in several moderately hard-to-detect anomalous cases. We believe that this unusual behavior results because of a cancellation effect in the magnitudes of the residuals that arises from the negative serial correlation in the non-anomalous and moderately hard-to-detect anomalous cases.

Table II contains results for comparing the four cumulative sum tests for the general case when liquid solution is removed from one or more tanks. Thus total mass and mass of each individual species will change because the total volume changes. Neither the concentrations of nitric acid, plutonium, and uranium nor the densities will change under this type of material loss. For all tests, the expected average run lengths are 200 for the no-fault case with 95% confidence intervals of about 196 to 204. Thus, if a loss is detectable using process fault detection and these tests, then the mean values in Tables II and III should be significantly less than 200. For the leaks from Tank 1, it appears that Croiser's vector Cusum consistently alarms at a smaller run length than do any of the other tests. The univariate test using the principal components appears to be the next most sensitive with COT being the least sensitive test.

Table III contains results for the situations when the lost liquid is replaced, either by tank solution less the plutonium or tank solution less all constituents. This activity will affect all concentrations and densities in the tank from which material is lost as well as any downstream tanks. Thus a diversion from Tank 1 will affect Tanks 2 and 3, whereas a diversion from Tank 3 will not affect the other tanks in this system. As for the loss without replacement scenario, Crosier's Cusum is the most sensitive test for loss with replacement with the two univariate Cusum tests being next and quite similar in results, whereas the COT is the least sensitive test.

It is instructive to compare the results for the two different diversion scenarios with respect to the performance of a specific test, e.g., Crosier's Cusum. The ARL to alarm for the 0.1 L/h leak from Tank 3 without replacement was 151, whereas the value for the same leak with replacement by water was 81, or nearly twofold less. We attribute this to the fact that when the diverted tank solution is replaced with water to bring the volume back into balance, all of the other measured variables have been altered by dilution. Thus the redundancy provided by the multiple measurements and the multivariate tests provides more sensitive detection when a diverter tries to be clever. This may frequently be the case as was found in our preceding work using Shewhart-type tests without cumulative sums.[4]

In conclusion, the Crosier's Cusum was the most powerful test of those investigated for the detection of diversions from a three-tank system using two different types of diversion scenarios: loss of tank liquid without replacement and loss of tank liquid solution with replacement by another liquid. We were somewhat surprised by the lack of sensitivity of the COT test for detecting losses. Apparently the directionality character of Crosier's Cusum test provides a very significant advantage, at least for this type of problem. We had expected the univariate test to be most sensitive for the loss without replacement scenario because only one variable (volume) is affected. Thus we were somewhat surprised that Crosier's Cusum exhibited superior performance for this scenario as well as for the replacement with liquid scenario, as we thought it might. This is probably a consequence of the adjustment of the critical values for the univariate tests needed to obtain the appropriate run lengths for performing multiple univariate tests. Presumably, if we monitored fewer variables, the univariate test would be superior. As a partial check of this conjecture, in one simulation we monitored only the volume in Tank 1 with a univariate Cusum test, and in that extreme case the univariate test was the best for detecting a leak from Tank 1. Finally, as we mentioned previously, successive values of these residual time sequences are correlated for all variables except for the densities in each tank. This is because measured values are used both to calculate the current residual and as initial values for modeling to calculate the next prediction, which affects the next residual. This is an interesting aspect that we are investigating. At present, we note from Table II that the negative correlation has an adverse affect on the COT. For small leak rates, the COT actually takes longer to alarm than it does in the no-leak case. It turns out that the COT test is able to detect a small gain of volume fairly well but is poor at detecting a small loss of volume.

REFERENCES

1. Paul M. Frank, "Fault Diagnosis in Dynamic Systems Using Analytical and Knowledge-based Redundancy—A Survey and Some New Results," *Automatica* **26**, 459-474 (1990).
2. Ronald B. Crosier, "Multivariate Generalizations of Cumulative Sum Quality Control Schemes," *Technometrics* **30**, 291-303 (1988).
3. The Bonferroni Method can be found in many statistics books. For example, see Johnson and Wichern, *Applied Multivariate Statistical Analysis* (Prentice Hall, 1988).
4. T. L. Burr, L. E. Wangen, and M. F. Mullen, "Authentication of Reprocessing Plant Safeguards Data Through Correlation Analysis," Los Alamos National Laboratory report, in review.
5. A. Beyerlein, private communication, 1992.
6. W. B. Watson and R. H. Rainey, "Modification of the SEPHIS Computer Code for Calculating the Purex Solvent Extraction System," Oak Ridge National Laboratory report ORNL-TM-5123 (December 1975).

Table II. Mean and Median Run Lengths for Simulations Involving Leaks With No Replacement of Lost Solution

Simulation Conditions	Univariate Cusum		PC Univariate Cusum		COT		Crosier's Vector Cusum	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Tank 1 Leak 0.1	201	147	170	115	286	152	150	106
Tank 1 Leak 0.3	103	85	52	38	241	134	24	24
Tank 1 Leak 0.5	14	14	17	15	133	83	11	12
Tank 3 Leak 0.1	202	149	202	139	286	155	151	108
Tank 3 Leak 0.3	87	71	158	105	238	136	26	25
Tank 3 Leak 0.5	14	14	50	36	135	82	12	12
Leak 0.033 from each tank	202	147	201	135	243	149	185	130
Leak 0.1 from each tank	183	132	160	106	710	176	90	59

Table III. Replacement of Loss Solution: Mean and Median Run Lengths								
Simulation Conditions	Univariate Cusum		PC Univariate Cusum		COT		Crosier's Vector Cusum	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Tank 1 Leak 0.1 (Pu only)	196	145	135	93	204	145	124	89
Tank 1 Leak 0.3 (Pu only)	26	25	24	18	151	108	17	17
Tank 1 Leak 0.5 (Pu only)	8	8	9	8	71	54	8	8
Tank 3 Leak 0.1 (All species)	190	142	185	128	190	139	81	62
Tank 3 Leak 0.3 (All species)	16	16	20	19	94	72	11	11

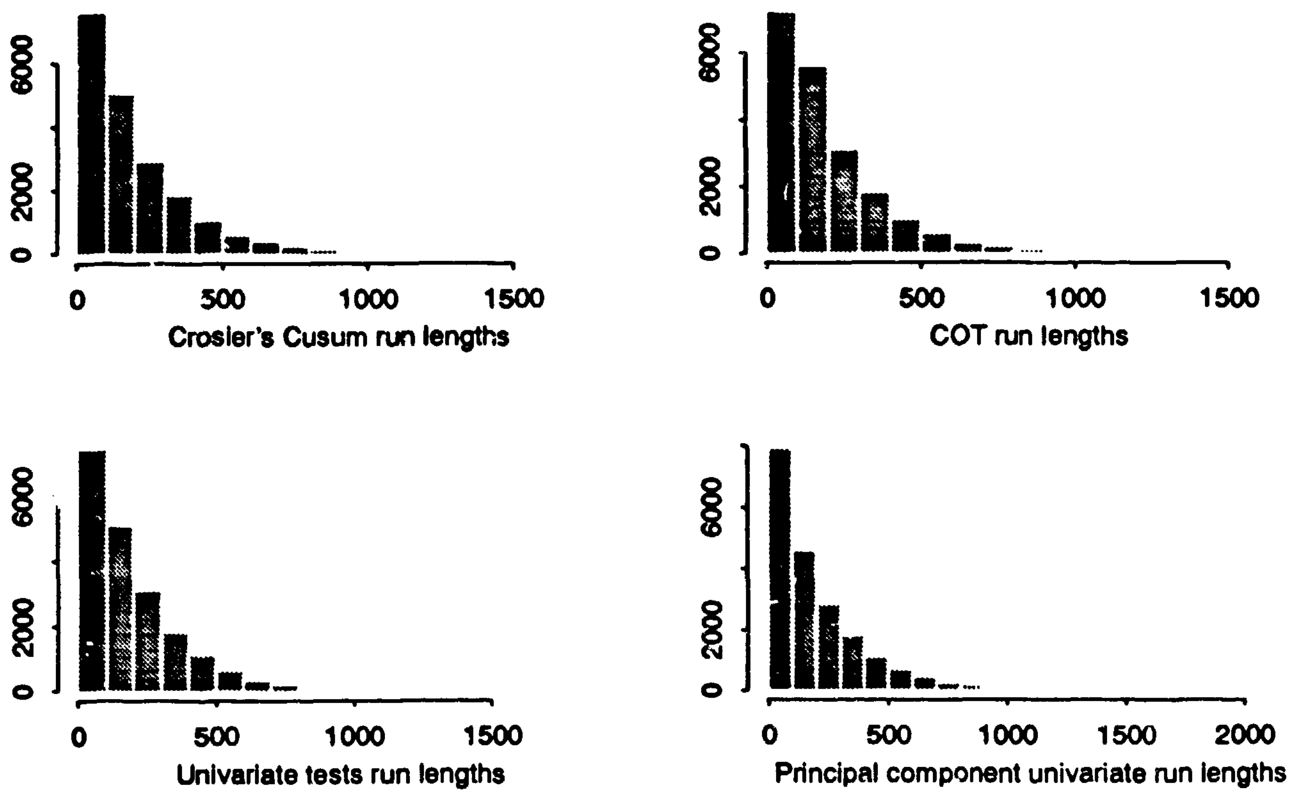

Fig. 3. Zero leak.

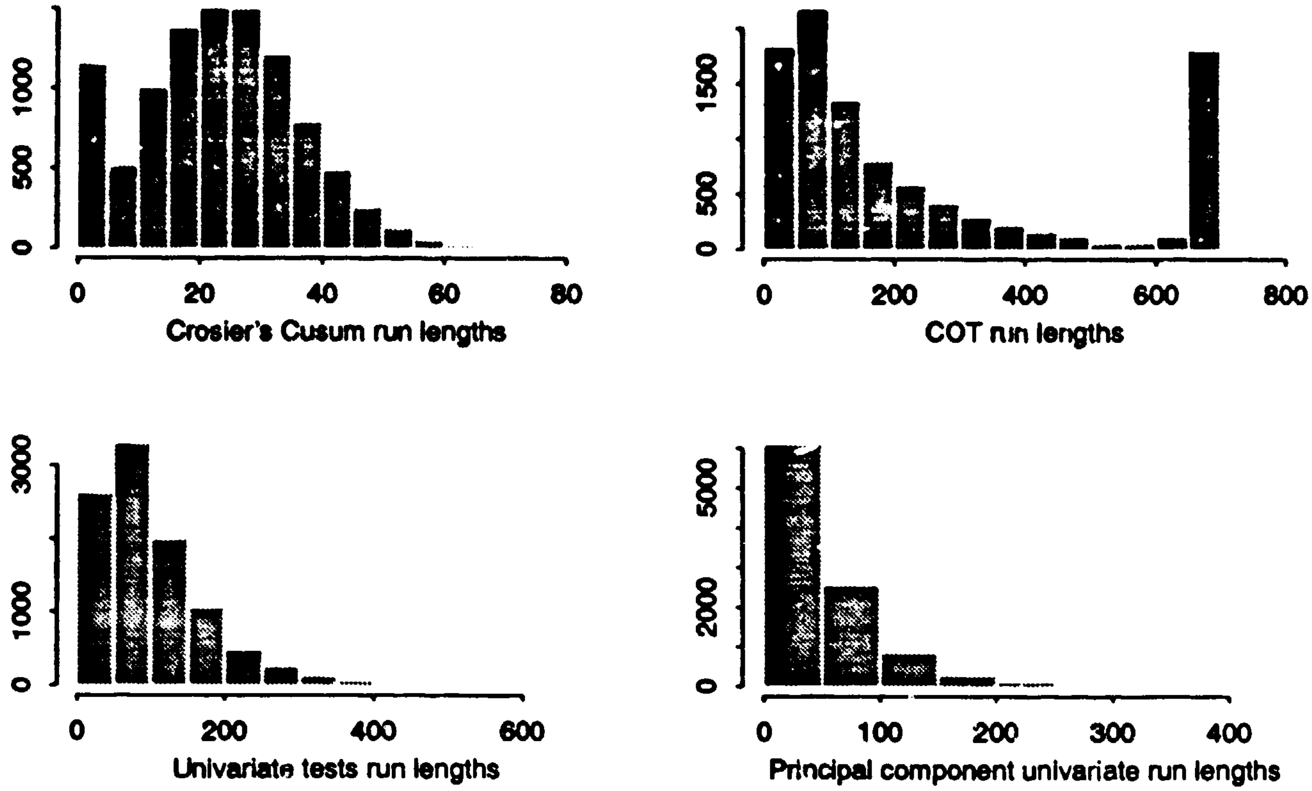

Fig. 4. Leak of 0.3 L/h, with replacement.

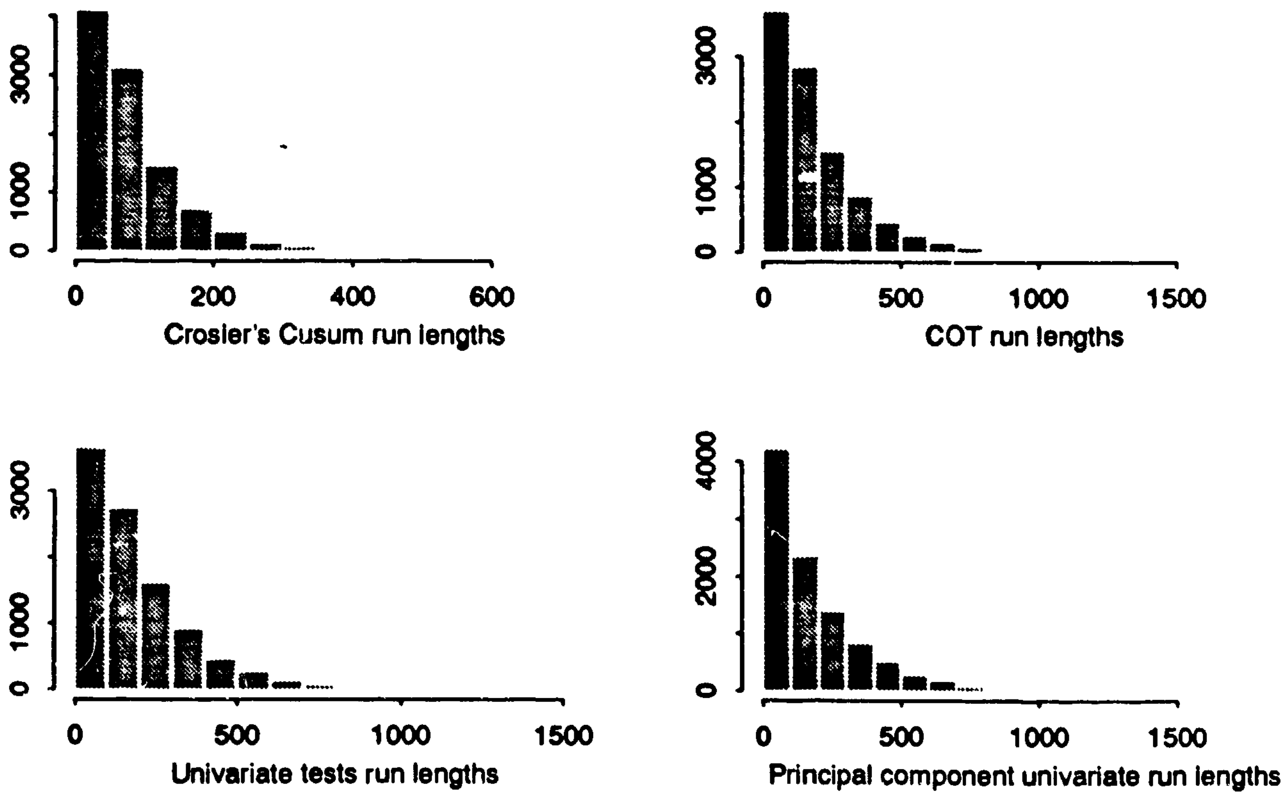

Fig. 5. Leak of 0.3 L/h, no replacement.