ESARDA
Bulletin

Impression of the Venice Symposium 1984

B.W. Hooton
U.K. Atomic Energy Authority

The 6th ESARDA symposium was held in Venice from 14th to 18th May. I suspected that the attendance and contributions would be influenced by the attractions of the city, its history, art, palaces and restaurants, but in fact the attendance of 215 was slightly down on the 227 at Versailles in 1983. We had a good opportunity to appreciate the unique infra-structure of the city and some of us even experienced one of the minor floods which seem to make St. Mark's Square even more appealing, although it was a reminder of the perpetual threat to the city's very existence.

The meeting opened with invited papers that were understandably political in character. The second session was in a similar vein, so much so that some of the technical experts wondered when the real work would begin. But safeguards is political and it is essential for all involved, including the scientific practitioners, to appreciate the political nature of safeguards and to understand its objectives.

The meeting soon entered the technical sessions with poster presentations in the magnificent Palladian refectory being just as popular as the oral sessions. A short break on the third day gave many of us the opportunity to visit some of the local historic islands. This organised tour was a little disappointing in that we were just too late in the day to see much of the traditional glass blowing on Murano. However, it was more than compensated by our visit to Burano where the art of lace making is still practiced, and we were able to enjoy a magnificent informal seafood dinner at one of the local restaurants.

The theme of the symposium, International Collaboration, Need and Benefits, is a continuous theme with the key words cost effectiveness, co-operation and efficiency being carried forward from our 5th Symposium in Versailles. Many of the papers presented were the results of international collaboration.

The first paper by Dr. Blix stressed the importance of safeguards as an advanced confidence building procedure which is still not fully understood by the general public. He also stated some of the political difficulties which affect the implementation of safeguards.

Mr. Audland's paper dealt with Euratom's role in promoting international collaboration in the field of safeguards. He pointed out the importance of nuclear power to the economic development of Europe and I am sure that his remarks would apply equally to the rest of the world. Financial constraints within the Community and also at the IAEA have necessitated the development of a cost-conscious safeguards approach; it was pleasing to hear Mr. Audland's strong support for R & D activities.

The final paper in the opening session by Mr. Noè stated the importance that the Italian Government has always attached to the implementation of safeguards and the Italian efforts to attain this goal through international collaboration.

Having set the scene, several of the other invited papers addressed some of these more general aspects of safeguards.

The words adversarial and cooperation were contrasted by several speakers who stressed that they were not necessarily conflicting. Safeguards is not an international police activity since direct penalties would be virtually impossible to impose, although Mr. Gmelin referred to the right to apply sanctions in cases of infringements of Euratom safeguards. There is, of course, every incentive to consider safeguards as a cooperation and this approach is attractive because it should lead to a reduced cost. However, as one speaker from the floor pointed out, the concept of cooperation with some States may be difficult to accept by others. I think a fair analogy was drawn to our attention in the auditing of financial accounts, where the adversial and cooperation aspects do co-exist.

A fuel cycle approach to safeguards is a new approach which may show promise but these concepts take time to develop, and they will certainly need very detailed examination. There seems to be an acceptable logic to this approach which appeared in Mr. Petit's paper and was described as inspection goals which should depend on a safeguarding approach designed for the fuel cycle as a whole according to its pecularities in order to achieve, with a given amount of resources, the greatest possible deterrence effect on the State. We must remember, however, as I think Dr. von Baeckmann

pointed out, that any new approaches must be consistent with legislation.

Reference was made to the fact that we have been active for some 25 years and Mr. Harry's presentations on certified reference materials represents the type of activity that we should expect in a mature industry. The production of standards to support our work is overdue. It tends to be given a lower priority for funding, perhaps because it is such a long process and the direct benefits are not so readily perceived. Nevertheless we depend very much on measurements, they represent a major part of our effort and we surely need the standards to support it. Metrology is far from trivial and the production of standards is a very demanding and time consuming exercise, but we must be prepared to pay the price.

As usual there was no shortage of contributions on NDA techniques; neutron systems and gamma ray methods accounting for most of the papers. There were several poster session presentations on neutron coincidence counting. This is, of course, an instrument used by inspectors, but we still hear of new ideas for reducing uncertainties arising from neutron multiplication effects. I am sure we would all like to see progress in this area.

The gamma ray methods are also maturing, largely due to improvements in germanium counters, electronic stability and computing. These developments have resulted in systems which are transportable, robust and which have adequate computing power for on-the-spot analysis at a very acceptable cost. There was an interesting question on how far gamma-ray methods may advance in the next few years and I agree with the comments that were made. We cannot get much better results on the resolution since we are at a plateau almost dictated by solid state physics. Better branching ratios and half-life data may help to improve the interpretation of the results, but we should be in a position, now, in 1984, where the results of our efforts should appear in practical instruments, and indeed we have many examples of progress. Pu isotopic determination by gamma ray measurements was shown as an integral part of a calorimeter in a poster presentation by the

The General Director of IAEA, Dr. H. Blix, giving his presentation

The presentation of the ESARDA Chairman, Dr. B.W. Hooton

Mound Laboratory, and there were nine other papers on gamma ray measurements as applied to isotopic determination or spent fuel investigations. Gamma ray measurements also form the basis for new approaches to enrichment plant safeguards. There were three very detailed and interesting papers, two of which were poster sessions, on the determination of the UF_6 enrichment in a gas centrifuge plant. The successful application of international safeguards to reprocessing plants must be an important goal for the IAEA and the plant operators and the results presented should make a significant contribution. The K-edge densitometer is proving to be very useful for solutions, the results are good, certainly in the range of 50 to 300 g/l Pu. This is another example of old ideas suddenly coming into real practical use as detector technology advances.

Information systems play an important role in nuclear material management at the national level and in the evaluation of safeguards in Luxembourg and Vienna. Vast quantities of data are now being produced and efficient analysis is essential. The theme of the 1985 ESARDA meeting will address this very problem. The paper on Information Systems at Venice illustrated the importance and difficulty of the work. The block diagram of the IAEA system ISIS showed the com-

plexity of the system and very much illustrated the magnitude of the task. The report of their three years of experience showed the results of considerable investment and it was interesting to see the gradual reduction in the scale of capital investment and staff deployed as well as a reduction in the running cost per job unit.

Mr. Maxwell's paper on Transit Control was perhaps of no particular concern to individual States, apart from security, but it certainly concerns the IAEA and Euratom. It is crucial to their responsibilities for the timely recognition of any potential diversion and to this extent we should lend an ear to their difficulties, even if it does mean, in terms of

paper work, asking the receiver to make up for the shortcomings of the shipper. The operators problem is to reconcile shipment documents as they are generated. This must be done to help the Safeguards Authorities to deal with, in the case of Euratom, more than 70,000 transactions per annum. It is clear that without a databank and efficient information analysis system the task would be impossible.

Statistical methods have always been a difficult subject for the non-specialist, but I expect that safeguards will have to come to terms with near-real-time accountancy in the not too distant future. There were two papers in the final session dealing with a
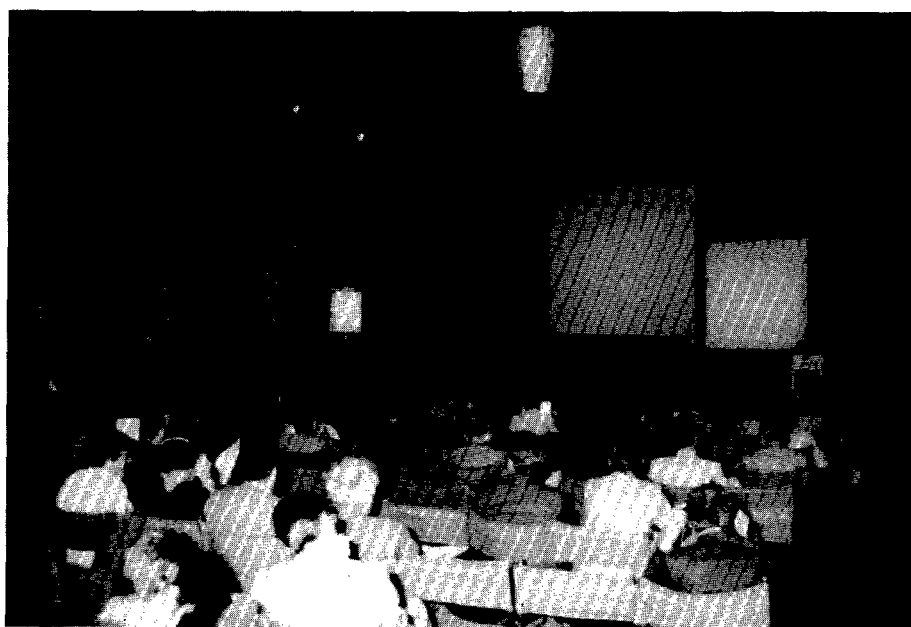
The conference room

Open air discussion during intermissions

comparison of statistical tests for NRTA. They both stressed the importance of loss patterns in developing optimal tests. This fact has not always been properly emphasised in the early model developments. The study of the properties of near-real-time accountancy models represents a good example of international collaboration for workers in a number of countries have contributed in this area in recent years.

Dr. von Baeckmann's account of the IAEA Support Programmes contained many examples of international collaboration. The formal support programmes are with 9 countries and Euratom but many other nations contribute in a variety of ways. His list of IAEA interests contained few surprises, the exception being their interest in heavy water safeguards under INFCIRC/66 type agreements.

Mr. Cuypers described the purpose of the new PERLA facility at Ispra. It will be a performance, calibration and training laboratory and I am sure we all look forward to the gradual build up of its capabilities through the 1980s. We should be grateful that these additional resources are to become available but we must recognise in a very real sense that its success will depend very much on our own active collaboration.

The ESARDA symposium has always presented a valuable occasion at which the latest thinking, both political and technical could be reviewed. I think the Venice meeting was no exception, it gave us some new political views and demonstrated that technology is still progressing to the extent that new concepts and approaches may prove to be feasible. Advances in technology, particularly the impact of advances in data processing will be the theme of our next symposium, in Liège.

PHONID Devices : Uranium Samples Non-Destructive Measurements

P. Dell'Oro
Commission of the European Communities,
JRC-Ispra, Italy

F.J.G. Rogers
NMACT, Harwell, U.K.

PHONID-3 device was recently employed in two sucessful campaigns at Springfields Nuclear Power Development Laboratories (SNPDL) for the assessment of different types of LEU materials, namely pellets and oxide powders.

It is worth noticing that it was the first time that a PHONID device was employed with LEU; the satisfactory results were for this reason even more welcomed.

In this paper we mean to briefly review the stages of PHONID device series and its fields of application. The second part of this article will deal with the future : either future applications or future PHONID machines.

PHONID's (Photo Neutron Interrogation Device) physical principle is the interrogation of the fissile material of the sample by means of a primary photo-neutron source (^{124}Sb-gamma - Be - 27 keV neutrons).

The sample is introduced into a measuring cavity together with the neutron detectors which count the fast neutrons of the fissions induced in the sample by the primary flux. A calibration curve, specific for every type of material, relates the fission neutron counts to the ^{235}U mass present in the sample.

PHONID-1 was a laboratory made machine /1/ and was operating for about
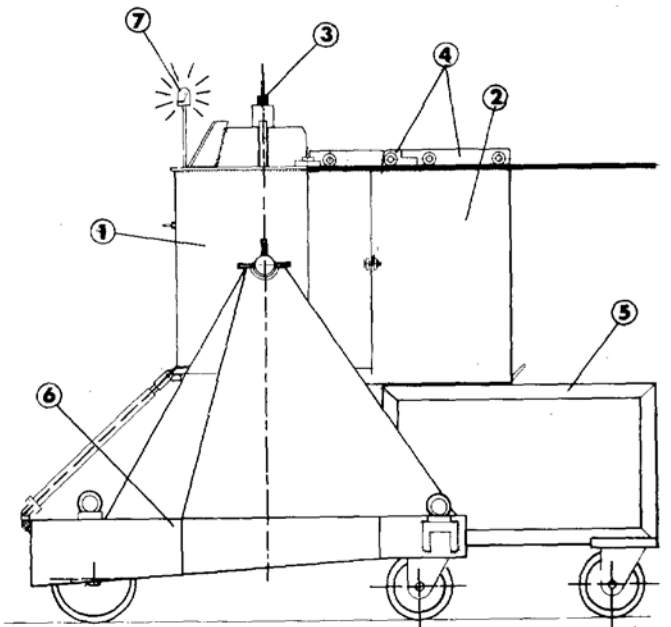
1) Source block 2) Measuring Cavity 3) Source holder
4) Fe cavity walls 5) Transport cavity cart 6) Transport source cart
7) Neutron warning lamp

Fig. 1 Phonid 1
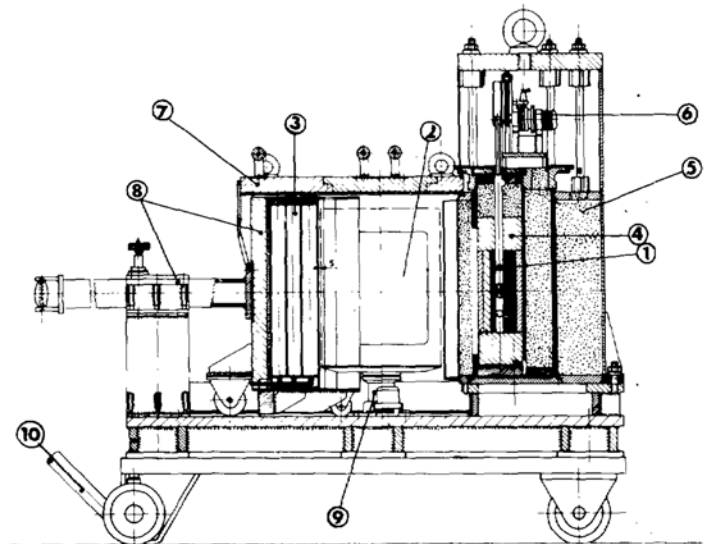
1) Sb sources 2) Sample and measuring cavity 3) He counters
4) Source block 5) Pb shield 6) Source mechanism 7) Fe cavity walls
8) Movable counter bank   9) Turn table   10) Transport device

Fig. 2 Phonid 2

Fig. 3 Phonid 3

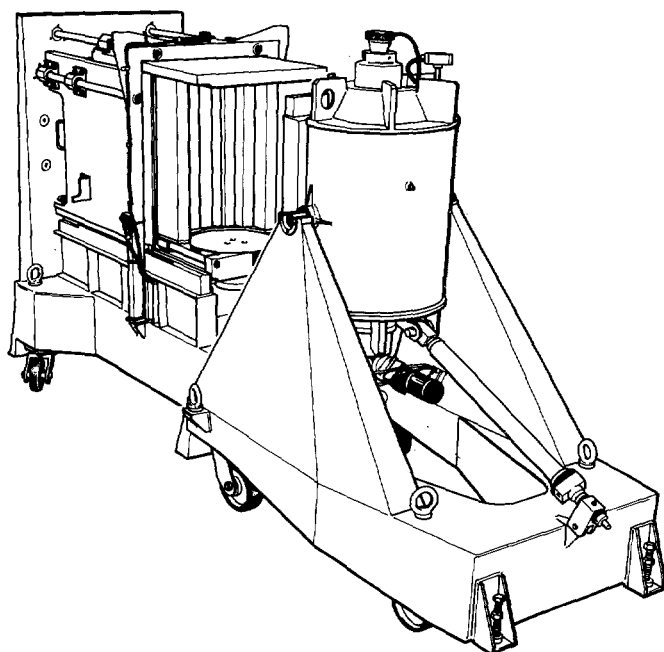| Calibr. Curve No.	Material	Enrich. %	^{235}U Mass Range g	
			From	To
1	U metal	93	20	4700
2	UF$_4$ powder	93	1000	3500
3	U-Al cermets	92	10	700
4	U-Al cermets	92	14	700
5	U$_3$O$_8$ powder	80	70	3500
6	U$_3$O$_8$ powder	93	1800	5500
7	UO$_2$ powder	43	2600	4500
8	UO$_2$ powder	59	2000	7200
9	U$_3$O$_8$ powder	20	900	2300
10	U$_3$O$_8$ powder	43	2000	5200
11	U$_3$O$_8$ powder	83-93	3800	8000
12	UO$_2$ pellets	44-59	800	4400
13	Th-U oxide	93	40	300
14	Th-U oxide + C	93	20	300

Table 1 Calibration curves for HEU.

Enrichment %	No. of Measured Items	Average Difference (Decl. - Calc.)
3.1	6	±0.4 %
11.2	6	±0.8 %
19.6	4	±1.8 %

Table 2 UO$_2$ powder samples in glass jars measurements against appropriate calibration curve.

Sample Enrichment Range		Calibration from Samples of Enrichment %	No. of Measured Items	Average Difference (Decl. - Calc.)	Container
From %	To				
2.6	3.1	2.5	9	± 2.2 %	Jig
3.5	4.4	4.4	4	± 2.5 %	''
5.1	5.6	5.6	12	± 1.4 %	''
10.2	11.2	10.2	17	± 1.7 %	''
12.0	12.3	13.3	13	± 3.3 %	''
14.3	15.6	15.6	14	± 3.1 %	''
2.5	15.6		70	± 3.7 %	''
5.1	5.6	5.6	12	±11.0 %	Plastic Bag

Table 3 UO$_2$ pellets. Measured against pellet calibration.

Enrichment %	No. of Measured Items	Average Difference (Decl. - Calc.)
3.0	2	±1.1 %
4.2	5	±1.2 %

Table 4 UO$_2$ powder in iron drum measured against appropriate calibration curve.

three years at NUKEM-Hanau (FRG) fabrication plant. Figure 1 shows this device. PHONID-2 is still at NUKEM, but some troubles with the electronics, mainly the pre-amplifiers, delayed the programme for a while. Now the updating and the general revision of this machine is in progress. New preamplifiers have been purchased and we hope to relaunch this machine soon. During the many measuring campaigns carried out at NUKEM with PHONID-1 and PHONID-2, only HEU was employed.

Almost all the principal materials of NUKEM production were taken into consideration, their calibration curves built and many measurements and inventories carried out. As an order of magnitude of this work done by PHONID-1 and PHONID-2 we can say the following : the calibration curves are about twenty, many of them could fit a whole family of materials with similar characteristics; the measurements were thousands and almost the whole NUKEM production was taken into consideration (except only long fuel elements already assembled and rare exotic stuff that represents only a few per cent of the inventory).

The results were generally good, despite some problems due to lack of experience, and to electronic failures. There were also problems due to the great variety of sample containers. If standard containers are not adopted, the accuracy of results is poorer.

Table 1 gives a rough summary of the exercise carried out at NUKEM by PHONIDs.

By reutilizing PHONID-1's source block, PHONID-3 was built by JRC-Ispra. Though the physical principle is the same as that of its predecessors, PHONID-3 is a very

modern device with features that enable the operator to carry out up to two hundred measurements per day. The measuring cavity is accessible from any side and easy to be loaded. The overall neutronic efficiency is fairly high and a counting time of 200 s is enough to give a good statistics.

Figure 3 shows PHONID-3 machine. It is mounted on two separate carts, and though heavy (3 tons total) and cumbersome, one person can operate it for movings and assemblings.

After the laboratory tests, PHONID-3's first destination was SNPDL where many LEU samples are in stock.

As we have said at the beginning of this article, two campaigns (summer and autumn 1983) have given very good results. Three types of materials have been measured at Springfields : UO_2 powder in glass jars, UO_2 pellets in a jig and UO_2 powder in plastic bags contained in an iron drum. Tables 2, 3 and 4 summarize the results of these measurements /2/.

As far as the future of PHONID series is concerned two different trends are in progress : development of little and "portable" machines for canned uranium measurements and construction of non movable installations.

For the former problem PHONID-4 device has already been designed and the construction of three samples of this device is in progress. Within 1984 it will be accomplished and the first laboratory tests done. The leading idea of this design is that a great many uranium samples measured by the former PHONIDs (say about 80% of the total) are canned samples of very compact dimensions (say 10-12 cm diameter and 20 cm height). In these conditions it is not worthwhile operating such a cumbersome and heavy device as the three PHONIDs build up at present.

PHONID-4 will weigh about 180 kg, but can be broken down into about twenty main pieces of less than 20 kg each. The source block is meant to fit different shaped measuring cavities (circular or square) but always of very compact dimensions. These cavities can receive cans of about 15 cm diam. or, removing top and bottom, long fuel elements can be driven through them for scanning measurements.

The second trend of future PHONID development is the construction of heavy highly specialized, fixed installations. A mock-up of this type is being planned by the JRC-Ispra and its construction will start within the current year.

The main task of this instrument will be to test new components such as neutron sources, detector systems and new materials. A scanning system will be incorporated. Long fuel elements will be scanned through the cavity by means of a computerized stepping motor.

Its dimensions are very large, hence the idea of mounting it on a cart has been abandoned. This device and also the future ones of this kind, are best designed if cheap materials, i.e. concrete instead of lead, are employed. The problem of squeezing down costs is an urgent matter for PHONID but it is not likely that great steps will be accomplished in the future if electronics and detectors keep the present incidence in the whole construction. The right way of reducing the economical weight of such devices is using them more in order to squeeze down the specific cost of the measurement.

References

/1/ BIRKHOFF et al., Euratom Internal Report 1661 (Ispra) 1975

/2/ P. Dell'ORO et al., Field experience of non-destructive assay of low enriched uranium by means of photoneutron interrogation. Proceedings of the ESARDA Sixth Annual Symposium, Venice, 1984

Safeguards Effectiveness Assessment : An Alternative Approach

M.J. Canty and G. Stein
KFA Jülich

A. Rezniczek
UBA Aachen

Introduction

Current attempts to systematize safeguards effectiveness assessment stem from the desire of the IAEA to plan and carry out its expanding inspection activities in an efficient, objective manner. Specifically, a scheme is being sought which will enable (i) a priori assessment of new approaches, both for generic as well as for specific fuel cycle facilities, and (ii) a posteriori evaluation of safeguards activities already implemented.

One methodology which has been proposed to meet these requirements /1/ is SEAM (Safeguards Effectiveness Assessment Methodology), an approach based essentially on a very detailed diversion path analysis. In SEAM the system to be evaluated is first of all defined in terms of safeguards-relevant facility characteristics, quantified safeguards goals inferred from Art. 28, INFCIRC/153 and a set of model, or actual, inspection activities. The assessment then entails a classification of all conceivable diversion paths according to the type of material involved, its location, concealment methods used and the technical complexity (from the divertor's standpoint) of the path as a whole. If a given diversion path is attempted, one or more anomalies may arise in connection with routine inspection activities. SEAM measures the system's effectiveness in terms of the associated anomaly detection probabilities, first at the diversion path level (degree of path coverage, probability bar charts) and then at the global level ("aggregate measure").

FRG experts have taken the standpoint on a number of occasions that SEAM does not provide a valid means of characterizing the effectiveness of international safeguards /2/. Before making more general criticism and proposing an alternative it is perhaps worthwhile to restate two purely technical objections to the methodology as it now stands.

First of all, anomaly detection is confused with diversion detection. Anomalous situations arising in the course of inspections will in general merely serve as triggers for further investigation. In fact, safeguards activities designed to maximize anomaly detection sensitivity alone are necessarily prone to

false alarms. Detailed consideration of the system's capability for alarm resolution would constitute an essential, yet hitherto neglected, component of SEAM.

Second, the SEAM approach to establishing overall figures of merit lacks internal consistency. On the one hand the standpoint is taken that inspections are carried out in an adversary environment in which hypothetical diversion paths with zero or low detection probabilities represent deplorable weaknesses. On the other hand, proposed aggregate measures of effectiveness invariably average over such paths as if a dedicated divertor would show no preference for them. The truism that a chain is as strong as its weakest link finds no counterpart in the assessment procedure.

Design vs implementation assessment

Figure 1 is an attempt to sketch the evolution of a facility-oriented safeguards approach through the various phases of design, negotiation and implementation.

At the design phase, safeguards experts, on the basis of internal detection guidelines, try to develop as complete and credible a safeguards system as possible. Emphasis is placed on timely and sensitive diversion detection capability, whereby diversion path analysis plays an important role.

At the negotiation stage, the interests of the facility operator and of the state in which

the facility is located are also represented. Various articles of INFCIRC/153 guarantee the rights of NPT states to unhindered development of peaceful uses of nuclear energy, and the emergent safeguards system, necessarily a compromise between the legitimate interests of both parties to the safeguards agreement, reflects the influence of these provisions.

There follows the Implementation phase in which practical, often conflicting requirements further restrict the safeguards activities actually carried out. Here in particular, the Agency has identified the need for a regular reassessment of its ongoing, routine inspection activities. The object of this reassessment is to permit more efficient resource allocation, to be able to identify areas for research and development and, perhaps most importantly, to make public results and conclusions in an objective, non-discriminatory way. The SEAM approach to dealing with this problem is to carry over, unmodified, precisely the same criteria used at the initial design stage and apply them for a posteriori effectiveness assessment. The existence of an intervening process of negotiation and compromise, the boundary conditions imposed by finite inspection resources and efficient commercial plant operation, the intangible but very important deterrent effect of the consequences of treaty violation, all these
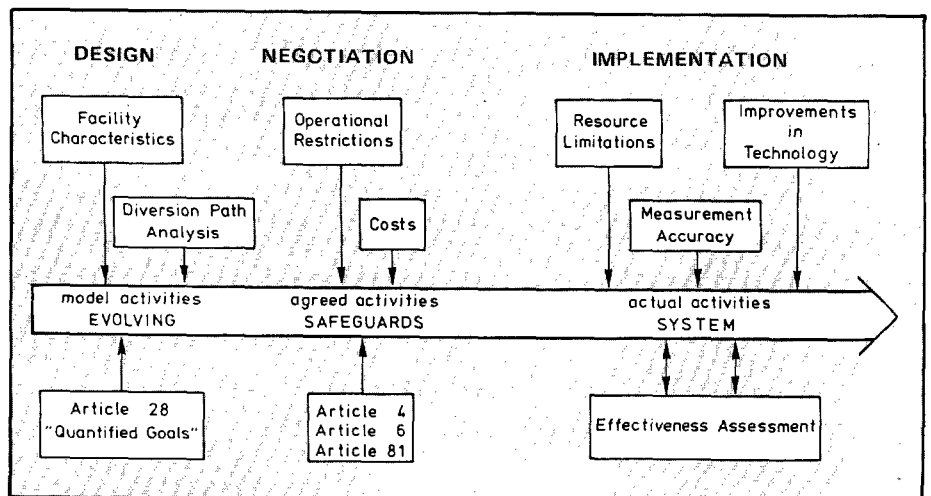

Fig. 1. The three phases of safeguards systems evolution

aspects are ignored. Instead, the actual inspection activities are again compared with the original set of diversion paths, and the associated anomaly detection probabilities alone measure the system's worth. The diversion paths derived in SEAM are hypothetical and open-ended, so it is always possible to find a large number of paths with no anomaly detection probability whatsoever, particularly in real implementations. One is then forced to the conclusion, aggregate measures notwithstanding, that all practicable safeguards activities are "ineffective". The methodology is thus self-defeating and is not, in our opinion, a profitable framework for implementation assessment. An alternative approach will be proposed in the following section.

Assessment based on the revision principle

An "alarm-response" philisophy appears to lie at the roots of SEAM. Alarm-response control mechanisms presuppose an internal authority which has the power to intervene directly in the process when an alarm occurs, a good example being provided by nuclear material physical protection systems. Such systems are evaluated according to the extent to which alarm thresholds can be reduced, often knowingly accepting a high false alarm rate to achieve high sensitivity. It is tacitly assumed that the response measures will always be potent enough to resolve the alarm unambiguously and influence the process correctively. Seen from this standpoint, the way in which SEAM goes about assessing safeguards effectiveness becomes more understandable, albeit no more justifiable. Containment/surveillance measures now assume a stand-alone status as response triggers which is far removed the accountancy-supportive role normally attributed them, while the closed, verified material balance is treated merely as another type of anomaly generator.

The alarm-response interpretation is incorrect. International safeguards as set out in INFCIRC/153 are based on the principle of revision. It is characteristic of revisions that they are carried out by an external authority and that they do not directly interfere with the process being examined. At predetermined intervals, comparisons are made and conclusions are drawn. At most, indirect influences are present through deterrent or preventative aspects that may be associated with the revision. In international safeguards, comparison is made between official reports having to do with material movements and inventories, operators' records and the nuclear material itself. The technical conclusion of the revision is "a statement in respect of each material balance area of the amount of material unaccounted for over a specific period, giving the limits of accuracy of the amounts stated." (Art. 30,

INFCIRC/153). To the extent that records and actual material movements may have to be compared when the movements occur, information may be gathered at irregular times, and even quasi-continuously. However, the drawing of conclusions takes place after the fact and is based on the extent to which safeguarded nuclear material is present and accounted for.

An alternative effectiveness assessment methodology will now be described which can be applied at the implementation stage and which is fully consistent with the revision interpretation of safeguards.

Certainly a safeguards system cannot be viewed as a bundle of technical measures all on a par with one another. Rather there exist (as is implied by the word "system") a multitude of logical relationships between its components. The assessment methodology must take account of this structure and make it transparent. Our starting point for a formal representation is the material accountancy system. The basic criteria for an effective safeguards implementation are taken to be

1. The accountancy verification procedures were logically complete :
   - Verification measures must exist for inventory and flow at all key measurement points defined for each material balance area in the relevant Facility Attachment. In other words, each component of the MUF equation is independently verifiable.
   - Agreed inspection activities were carried out over each balance period sufficient to verify the operator's inventory change and material balance reports.
   - The reported material unaccounted for does not significantly exceed the accountancy verification goal set for the

respective MBA. (For item facilities, the MUF would of course be expected to be zero.)

2. Additional activities (if needed) were performed to meet timeliness goals. (These are of course highly material and facility-specific.)

As a framework for normal representation of the assessment procedure, suitable for computer manipulation and standardization according to facility type, the predicate calculus was chosen. This symbolic language permits the manipulation of purely descriptive concepts and frees one from the necessity to translate all evaluations into numerical form. Moreover, the basic spirit of revision (comparison-conclusion) can easily be captured in the logical implication structure : if A and B then C.

The kind of "bottom-up" reasoning involved in an effectiveness assessment is sketched in Fig. 2. Here a small portion of the problem space involved is shown along with the if-then implication structure leading to the (desired) conclusion that the accountancy verification goal (AVG) was attained. Similar deductive schemes can be built up for the detection time goals. A simple implementation in the logic language DEDUC /3/ for a hypothetical LWR safeguards implementation is indicated in Tables 1 through 4. The object structure (domain of discourse) in Table 1 consists of the classes NUCMAT (nuclear material), REPORTS, RECORDS, GOALS, etc. The predicates in Table 2 define the attributes which these objects (or combinations of objects) are allowed to have. A typical implication reflecting the logical structure of the safeguards approach being assessed might read :

if IC(cf,int1) and
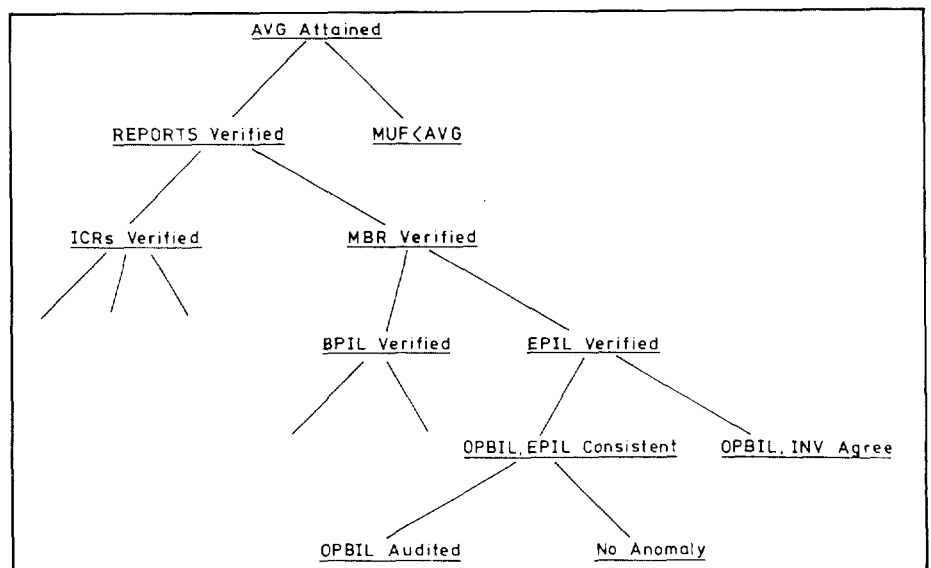NO ANOMALY(sea11,int1,int2)
then IC(cf,int2)

Fig. 2. Problem space for effectiveness assessment

NUCMAT	Nuclear Material
MUF	Material Unaccounted For
SRD	Shipper-Receiver-Difference
AVG	Accountancy Verification Goal
KMPMAT	KMP Material
INV	Inventory
FF	Fresh Fuel
CF	Core Fuel
SF	Spent Fuel
FLOWMAT	Flow Material
RECFF	Fresh Fuel Receipt
RECFF01	Fresh Fuel Receipt in 1. Period
RECFF12	Fresh Fuel Receipt in 2. Period
RECFF23	Fresh Fuel Receipt in 3. Period
RECFF34	Fresh Fuel Receipt in 4. Period
SHPSF	Spent Fuel Shipment
SHPSF01	Spent Fuel Shipment in 1. Period
SHPSF12	Spent Fuel Shipment in 2. Period
SHPSF23	Spent Fuel Shipment in 3. Period
SHPSF34	Spent Fuel Shipment in 4. Period
REPORTS	Reports
ICR	Inventory Change Report
RECEIPTS	Receipts
RECPT01	Receipts in 1. Period
RECPT12	Receipts in 2. Period
RECPT23	Receipts in 3. Period
RECPT34	Receipts in 4. Period
SHPMENTS	Shipments
SHPMNT01	Shipments in 1. Period
SHPMNT12	Shipments in 2. Period
SHPMNT23	Shipments in 3. Period
SHPMNT34	Shipments in 4. Period
MBR	Material Balance Report
BPIL	Beginning Physical Inventory Listing
BPIL FF	Beginning Physical Inventory Listing Fresh Fuel
BPIL CF	Beginning Physical Inventory Listing Core Fuel
BPIL SF	Beginning Physical Inventory Listing Spent Fuel
EBIL	Ending Book Inventory
SRD	Shipper-Receiver-Difference
AEBIL	Adjusted Ending Book Inventory Listing
EPIL	Ending Physical Inventory Listing
EPIL FF	Ending Physical Inventory Listing Fresh Fuel
EPIL CF	Ending Physical Inventory Listing Core Fuel
EPIL SF	Ending Physical Inventory Listing Spent Fuel
MUF	Material Unaccounted For

RECORDS	Records
CMR	Crane Movement Record
STDS	State Transfer Documents
STDS01	State Transfer Documents for 1. Period
STDS12	State Transfer Documents for 2. Period
STDS23	State Transfer Documents for 3. Period
STDS34	State Transfer Documents for 4. Period
OPBIL	Operators Book Inventory Listing
POWOPREC	Power Records
T	Time
BT	Beginning Time
BPIVT	Beginnin Physical Inventory Verification Time
INT1	Interim Time 1
INT2	Interim Time 2
INT3	Interim Time 3
EPIVT	Ending Physical Inventory Verification Time
IT	Interim Time
BPIVT	Beginning Physical Inventory Verification Time
INT1	Interim Time 1
INT2	Interim Time 2
INT3	Interim Time 3
EPIVT	Ending Physical Inventory Verification Time
ET	Ending Time
BPIVT	Beginning Physical Inventory Verification Time
INT1	Interim Time 1
INT2	Interim Time 2
INT3	Interim Time 3
EPIVT	Ending Physical Inventory Verification Time
TINT	Time Intervals
BALPER	Balance Period
DTIMFF	Detection Time Fresh Fuel
DTINSF	Detection Time Spent Fuel
INSPINT	Inspection Interval
CSDEVICE	C/S-Devices
CSON FF	C/S on Fresh Fuel
CSON CF	C/S on Core Fuel
SEAL1	SEAL on Reactor Core
CAMERA1	Camera on Reactor Core
CSON SF	C/S on Spent Fuel
CAMERA2	Camera on Spent Fuel Pont
GOALS	Safeguards Goals
AVG	Accountancy Verification Goal
DTIM FF	Detection Time Goal Fresh Fuel
DTIM SF	Detection Time Goal Spent Fuel

Table 1

The following conventions were chosen for assigning predicates to the objects listed in Table 1. Objects are in brackets.

(GOALS) can be ATTAINED between times (BT) and (ET).
ATTAINED(GOALS,BT,ET)

(NUCMAT) can EXIST at some (T).
EXIST(NUCMAT,T)

(REPORTS) can be VERIFIED for the current (BALPER).
VERIFIED(REPORTS,BALPER)

(RECORDS) can be RELevant to (NUCMAT) at some (T).
RECMREL(RECORDS,NUCMAT,T)

(REPORTS) can be RELevant to (NUCMAT) at some (T).
REPMREL(REPORTS,NUCMAT,T)

(RECORDS) can be RELevant to (REPORTS) at some (T).
RRREL(RECORDS,REPORTS,T)

A (CSDEVICE) can be RELevant to (INV) at some (T).
CSREL(CSDEVICE,INV,T)

(RECORDS) can be AUDITed at some (T).
AUDIT(RECORDS,T)

(NUCMAT) can be INSPECTed at some (T).
INSPECT(NUCMAT,T)

A (CSDEVICE) can be CHECKed at some (T).
CHECK(CSDEVICE,T)

(RECORDS) and (REPORTS) can be CONSisTaNT at some (T).
CONSTNT(RECORDS,REPORTS,T)

(RECORDS) and (NUCMAT) can AGREE at some (T).
AGREE(RECORDS,NUCMAT,T)

C/S measures on (INV) can be SUCCesSful from (BT) to (ET).
SUCCS(INV,BT,ET)

An amount of (NUCMAT) can be Less than or Equal to another amount of (NUCMAT).
LE(NUCMAT,NUCMAT)

One (TINT) can be LessThan some other (TINT).
LT(TINT,TINT)

One (T) can be LessThan some other (T).
LT(T,T)

(NUCMAT) can be Item Counted at some time (T).
IC(NUCMAT,T)

(NUCMAT) can be Non-Destructively Assayed at some (T).
NDA(NUCMAT,T)

(NUCMAT) can be IDENTified at some (T).
IDENT(NUCMAT,T)

An ANOMaly can arise between (RECORDS) and (REPORTS) at some (T).
RRANOM(RECORDS,REPORTS,T)

An ANOMaly can arise between (RECORDS) and (NUCMAT) at some (T).
RMANOM(RECORDS,NUCMAT,T)

An ANOMaly can arise for a (CSDEVICE) over an interval (BT) to (ET).
CSANOM(CSDEVICE,BT,ET)

Intermediate Predicates (introduced to speed up deduction):
BPILVFD,EPILVFD,RECVFDij,SHPVFDij,RECVFD,SHPVFD,ICRSVFD,REPSVFD,SRDVFD

Table 2

|Base implications and premises for LWR knowledge module |

I MAIN IMPLICATIONS		I	C/S LOGIC	I

I MAIN IMPLICATIONS

```
IF REPSVFD(BALPER) AND LE(MUF,AVG) THEN ATTAINED(AVG,BPIVT,EPIVT);
IF BPILVFD(BALPER)  AND EPILVFD(BALPER) AND ICRSVFD(BALPER)
    THEN REPSVFD(BALPER);
IF VERIFIED(BPILFF,BALPER) AND VERIFIED(BPILCF,BALPER) AND
    VERIFIED(BPILSF,BALPER) THEN BPILVFD(BALPER);
IF VERIFIED(EPILFF,BALPER) AND VERIFIED(EPILCF,BALPER) AND
    VERIFIED(EPILSF,BALPER) THEN EPILVFD(BALPER);
IF VERIFIED(RECPT01,BALPER) AND VERIFIED(RECPT12,BALPER)
    THEN RECVFD02(BALPER);
IF VERIFIED(RECPT23,BALPER) AND VERIFIED(RECPT34,BALPER)
    THEN RECVFD24(BALPER);
IF RECVFD02(BALPER) AND RECVFD24(BALPER) THEN RECVFD(BALPER);
IF VERIFIED(SHPMNT01,BALPER) AND VERIFIED(SHPMNT12,BALPER)
    THEN SHPVFD02(BALPER);
IF VERIFIED(SHPMNT23,BALPER) AND VERIFIED(SHPMNT34,BALPER)
    THEN SHPVFD24(BALPER);
IF SHPVFD02(BALPER) AND SHPVFD24(BALPER) THEN SHPVFD(BALPER);
IF RECVFD(BALPER) AND SHPVFD(BALPER) THEN ICRSVFD(BALPER);
```

| VERIFIED

```
IF REPMREL(REPORTS,KMPMAT,T) AND CNSTNT(RECORDS,REPORTS,T)
    AND AGREE(RECORDS,KMPMAT,T) THEN VERIFIED(REPORTS,BALPER);
```

| CNSTNT

```
IF RRREL(RECORDS,REPORTS,T) AND AUDIT(RECORDS,T)
    AND NO RRANOM(RECORDS,REPORTS,T) THEN
    CNSTNT(RECORDS,REPORTS,T);
```

| AGREE

```
IF RECMREL(RECORDS,KMPMAT,T) AND INSPECT(KMPMAT,T)
    AND NO RMANOM(RECORDS,KMPMAT,T) THEN
    AGREE(RECORDS,KMPMAT,T);
```

| INSPECT

```
IF NO EXIST(KMPMAT,T) THEN INSPECT(KMPMAT,T);
IF IC(KMPMAT,T) AND NDA(KMPMAT,T) THEN INSPECT(KMPMAT,T);
IF IC(CF,T) AND IDENT(CF,T) THEN INSPECT(CF,T);
IF IC(SF,T) AND IDENT(SF,T) THEN INSPECT(SF,T);
IF IC(SHPSF,T) AND IDENT(SHPSF,T) THEN INSPECT(SHPSF,T);
```

| SUCCS

```
IF CSREL(CSDEVICE,INV,BT) AND CHECK(CSDEVICE,ET) AND
    NO CSANOM(CSDEVICE,BT,ET) THEN
    SUCCS(INV,BT,ET);
```

C/S LOGIC

```
IF SUCCS(INV,BT,IT) AND SUCCS(INV,IT,ET) THEN
    SUCCS(INV,BT,ET);
IF IC(INV,BT) AND SUCCS(INV,BT,ET) THEN IC(INV,ET);
IF NDA(INV,BT) AND SUCCS(INV,BT,ET) THEN NDA(INV,ET);
IF IDENT(INV,BT) AND SUCCS(INV,BT,ET) THEN
    IDENT(INV,ET);
```

| DTIMSF

```
IF VERIFIED(SHPMNT01,BALPER)
    AND AGREE(OPBIL,CF,INT1) AND AGREE(OPBIL,SF,INT1) THEN
    ATTAINED(DTIMSF,BPIVT,INT1);
IF VERIFIED(SHPMNT12,BALPER)
    AND AGREE(OPBIL,CF,INT2) AND AGREE(OPBIL,SF,INT2) THEN
    ATTAINED(DTIMSF,INT1,INT2);
IF VERIFIED(SHPMNT23,BALPER)
    AND AGREE(OPBIL,CF,INT3) AND AGREE(OPBIL,SF,INT3) THEN
    ATTAINED(DTIMSF,INT2,INT3);
IF VERIFIED(SHPMNT34,BALPER)
    AND AGREE(OPBIL,CF,EPIVT) AND AGREE(OPBIL,SF,EPIVT) THEN
    ATTAINED(DTIMSF,INT3,EPIVT);
IF ATTAINED(GOALS,BT,IT) AND ATTAINED(GOALS,IT,ET)
    THEN ATTAINED(GOALS,BT,ET);
```

| RELEVANCE PREMISES

```
PREM RRREL(STDS01,SHPMNT01,INT1),RRREL(STDS01,RECPT01,INT1);
PREM RRREL(STDS12,SHPMNT12,INT2),RRREL(STDS12,RECPT12,INT2);
PREM RRREL(STDS23,SHPMNT23,INT3),RRREL(STDS23,RECPT23,INT3);
PREM RRREL(STDS34,SHPMNT34,EPIVT),RRREL(STDS34,RECPT34,EPIVT);
PREM RRREL(OPBIL,BPIL,BPIVT),RRREL(OPBIL,EPIL,EPIVT);
PREM CSREL(CSONFF,FF,T),CSREL(CSONCF,CF,T),
    CSREL(CSONSF,SF,T);
PREM RECMREL(STDS01,RECFF01,INT1),RECMREL(STDS01,SHPSF01,INT1);
PREM RECMREL(STDS12,RECFF12,INT2),RECMREL(STDS12,SHPSF12,INT2);
PREM RECMREL(STDS23,RECFF23,INT3),RECMREL(STDS23,SHPSF23,INT3);
PREM RECMREL(STDS34,RECFF34,EPIVT),RECMREL(STDS34,SHPSF34,EPIVT);
PREM RECMREL(OPBIL,INV,T);
PREM REPMREL(BPILFF,FF,BPIVT);
PREM REPMREL(BPILCF,CF,BPIVT);
PREM REPMREL(BPILSF,SF,BPIVT);
PREM REPMREL(EPILFF,FF,EPIVT);
PREM REPMREL(EPILCF,CF,EPIVT);
PREM REPMREL(EPILSF,SF,EPIVT);
PREM REPMREL(RECPT01,RECFF01,INT1),REPMREL(SHPMNT01,SHPSF01,INT1);
PREM REPMREL(RECPT12,RECFF12,INT2),REPMREL(SHPMNT12,SHPSF12,INT2);
PREM REPMREL(RECPT23,RECFF23,INT3),REPMREL(SHPMNT23,SHPSF23,INT3);
PREM REPMREL(RECPT34,RECFF34,EPIVT),REPMREL(SHPMNT34,SHPSF34,EPIVT);
```

Table 3

In English : "If the fuel elements in the reactor core were accessible and counted at the first interim inspection and the missile shield was successfully sealed from that time until the second interim inspection, then the item count for the core is also valid at the time of the second inspection."

Table 3 gives the complete implication structure in DEDUC FOR THE LWR example. A set of premises reflecting the activities performed during a year's inspection of the facility sufficient to achieve both AVG and timeliness goals in the model are listed in Table 4.

While a formalization of the assessment procedure in terms of the predicate logic can be purely descriptive, it need not be so. To characterize detection capability, hindrance to operations, required inspection effort, etc., numerical parameters can be associated with those predicates which define the basic inspection activities (such as AUDIT, IC, NDA in the LWR model of Table 1-4). It is then possible to associate

a confidence factor with the final conclusion regarding the attainment of goals, as well as give numbers for inspection effort expended and cost to facility operations.

The requirement of transparency is met by the logical implication structure into the assessment model. All conclusions reached can easily be understood and even explained by the program itself.

Conclusion

A methodology for assessing the effectiveness of safeguards implementation has been proposed which exhibits a number of desirable properties :

a) The fundamental role of material accountancy is explicit and transparent in the assessment procedure.
b) By focusing on the logical consistency and completeness of material accountancy verification activities, the capability

of the safeguards system to establish the continuing presence of nuclear material is evaluated, rather than its sensitivity to indirect anomalies.

c) While the assessment procedure can clearly lead to the conclusion that a particular implementation was ineffective, it is not doomed to do so at the outset by virtue of open-ended diversion path analyses.
d) Safeguards aspects of interest to both parties to a verification agreement can be treated on an equal basis.

References

/1/ H. Gruemm et al., Proc INMM, San Francisco (1981) p. 150
/2/ G. Stein et al., Proc INMM. Vail (1983)
/3/ K.F. Müller-Reissmann and F. Rechenmann, in "Concepts and Tools of Computer-assisted Policy Analysis", Ed. H. Bossel, Birkhauser, Basel (1977) p. 482

Test premises for LWR knowledge module

```
PREM VERIFIED(BPIL,BALPER);              |INITIAL INVENTORY WAS VFD  |
                                         |***************************|
                                         | FIRST INTERIM INSPECTION  |
                                         |***************************|
PREM AUDIT(OPBIL,INT1),                  |OPERATOR RECORDS AUDITED   |
    AUDIT(STDS01,INT1);
PREM NO EXIST(RECFF01,INT1),             | NO RECEIPTS OR SHIPMENTS  |
    NO EXIST(SHPSF01,INT1);              | TOOK PLACE AND THIS WAS   |
PREM NO RRANOM(STDS01,RECPT01,INT1),     | CORRECTLY REFLECTED IN    |
    NO RRANOM(STDS01,SHPMNT01,INT1);     | RECORDS AND REPORTS       |
PREM NO RMANOM(STDS01,RECFF01,INT1),
    NO RMANOM(STDS01,SHPSF01,INT1);
PREM IC(CF,INT1),IDENT(CF,INT1);         | CORE FUEL WAS ACCESSIBLE  |
                                         | AND WAS INSPECTED         |
PREM IC(SF,INT1),IDENT(SF,INT1);         | DITTO FOR SPENT FUEL POND |
PREM NO RMANOM(OPBIL,CF,INT1) 95;        | OPBIL AGREES WITH MATERIAL|
PREM NO RMANOM(OPBIL,SF,INT1) 95;        | 95% STATISTICAL SAMPLING  |
                                         |***************************|
                                         | SECOND INTERIM INSPECTION |
                                         |***************************|
PREM AUDIT(OPBIL,INT2),                  | OPERATOR RECORDS AUDITED  |
    AUDIT(STDS12,INT2);
PREM NO EXIST(RECFF12,INT2),             | NO RECEIPTS OR SHIPMENTS  |
    NO EXIST(SHPSF12,INT2);              | TOOK PLACE AND THIS WAS   |
PREM NO RRANOM(STDS12,RECPT12,INT2),     | CORRECTLY REFLECTED IN    |
    NO RRANOM(STDS12,SHPMNT12,INT2);     | RECORDS AND REPORTS       |
PREM NO RMANOM(STDS12,RECFF12,INT2),
    NO RMANOM(STDS12,SHPSF12,INT2);
PREM CHECK(SEAL1,INT2),                  | C/S DEVICES INSPECTED     |
    NO CSANOM(SEAL1,INT1,INT2),          | MISSILE SHIELD SEAL IS    |
    CHECK(CAMERA2,INT2),                 | 100% OK                   |
    NO CSANOM(CAMERA2,INT1,INT2) 90;     | SUBJECTIVE VALUE FOR SF   |
                                         | POND CAMERA               |
PREM NO RMANOM(OPBIL,CF,INT2) 95,        | OPBIL CORRECTLY REFLECTS  |
    NO RMANOM(OPBIL,SF,INT2) 95;         | STATIC SITUATION. NOTE CFS|
                                         | ARE CARRIED OVERI         |
                                         |***************************|
                                         | THIRD INTERIM INSPECTION  |
                                         |***************************|
PREM AUDIT(OPBIL,INT3),                  | RECORDS AUDITED           |
    AUDIT(STDS23,INT3);
PREM NO EXIST(RECFF23,INT3);
PREM IC(SHPSF23,INT3),                   | SHIPMENT TOOK PLACE, ITEMS|
    IDENT(SHPSF23,INT3);                 | WERE COUNTED AND IDENTIFIED|
PREM NO RRANOM(STDS23,SHPMNT23,INT3),
    NO RMANOM(STDS23,SHPSF23,INT3);
PREM NO RMANOM(STDS23,RECFF23,INT3),
    NO RRANOM(STDS23,RECPT23,INT3);
PREM IC(SF,INT3),                        | SINCE SF WAS SHIPPED, THE |
    IDENT(SF,INT3);                      | SF POND WAS REVERIFIED    |
PREM NO RMANOM(OPBIL,SF,INT3) 95;
PREM CHECK(SEAL1,INT3),                  | SEAL ON MISSILE SHIELD    |
    NO CSANOM(SEAL1,INT2,INT3);          | STILL OK                  |
PREM NO RMANOM(OPBIL,CF,INT3);
                                         |***********************************|
                                         | PHYSICAL INVENTORY VERIFICATION   |
                                         |***********************************|
PREM AUDIT(OPBIL,EPIVT),                 | REPORTS WERE AUDITED      |
    AUDIT(STDS34,EPIVT);
PREM NO EXIST(RECFF34,EPIVT),            | NO RECEIPTS OR SHIPMENTS IN|
    NO EXIST(SHPSF34,EPIVT);             | FINAL QUARTER, AND THIS WAS|
PREM NO RRANOM(STDS34,EPIVT),            | CONSISTENT WITH RECORDS   |
    NO RMANOM(STDS34,RECFF34,EPIVT),     | AND REPORTS               |
    NO RMANOM(STDS34,SHPSF34,EPIVT);
PREM NO RRANOM(OPBIL,EPIL,EPIVT) 95;     | RECORDS CONSISTENT WITH   |
PREM CHECK(SEAL1,EPIVT),                 | FINAL PIL, TO 95%         |
    NO CSANOM(SEAL1,INT3,EPIVT);         | CORE NOT ACCESSIBLE, BUT  |
PREM IC(FF,EPIVT),                       | SEAL OK                   |
    NDA(FF,EPIVT),                       | VERIFICATION OF SPENT AND |
    IC(SF,EPIVT),                        | FRESH FUEL ELEMENTS       |
    IDENT(SF,EPIVT);
PREM CHECK(CAMERA2,EPIVT),               | DOPPELT-GEMOPPELT HAELT   |
    NO CSANOM(CAMERA2,INT3,EPIVT) 95;    | BESSER                    |
PREM NO RMANOM(OPBIL,FF,EPIVT) 90,       | NO DEFECTS WERE OBSERVED  |
    NO RMANOM(OPBIL,CF,EPIVT) 95,
    NO RMANOM(OPBIL,SF,EPIVT) 90;
PREM LE(MUF,AVG);                        | REPORTED MUF WAS ZERO     |
```

Table 4

A Particular Problem of Data Evaluation in Safeguards

C. Foggi

CEC, JRC-Ispra
Convenor of the RIV Working Group

Introduction

The ESARDA Working Group for Re-processing Input Verification (RIV) is performing an exercise to test the performances of the existing Isotope Correlation Techniques (ICTs). This exercise is called "Benchmark on ICT", currently abbreviated BME.

In the evaluation of the results, the participants were confronted with a particular statistical problem, consisting in the comparison of data produced by various laboratories which share a common source of errors. Since no reference was found in the literature on how to treat the particular case, a new statistical model was proposed by one of the members of the WG-RIV.

Possibly, the problem described above may present itself every time that data obtained by two or more laboratories have to be compared. Since this occurrence is not infrequent in safeguards, the reader of the ESARDA Bulletin may be interested in being informed on the case which generated the statistical problem (cfr. this article), and in learning the solution which was worked out (see next article, by Miss Neuilly).

The Benchmark Exercise

Many organizations and plant operators use the methods of isotope correlation in analyzing and evaluating the isotopic composition data of irradiated fuels.

Depending on their needs and objectives, they have accordingly developed specific "techniques" for the practical application of the method to their specific problem.

As a result of these studies, several techniques of isotope correlation do exist nowadays, which differ from each other in various respects, e.g. :

- the objective of their application (material balance making, verification of the material balance, verification of laboratory performance, etc.)
- the isotopic data which they use
- the logical approach in the determination of the correlations (reactor physics calculation, statistical analysis of historical data)
- their state of development.

People working with isotopic correlation methods are often asked : "how does your technique perform and how does your performance compare with that of similar techniques ?". To answer this recurrent question, an exercise was initiated by the ESARDA Working Group RIV, with the aim of testing the various techniques so far proposed on a common set of data and with a common objective.

Luckily enough, all the organizations which had developed a consistent rationale in the field of isotopic correlations joined the initiative of the WG-RIV, so that the list of participants now includes : the CEA-Cadarache (which organized the exercise), the IAEA, the Euratom Safeguards, the KfK-Karlsruhe (in association with the JRC-Karlsruhe), the CEN/SCK-Mol, the JAERI-Tokai Mura and the ANL-Argonne.

The industrial operators having legal rights to the data needed for the exercise kindly agreed to release the information to the WG. They are: KWU-Germany (as fabricator of the fuel), KWO-Germany (as owner of the reactor where the fuel was irradiated) and COGEMA-France (as reprocessor of the fuel).

In organizing the exercise four logical tasks were devised:

- to choose an appropriate set of isotopic composition data to be submitted to all participants intending to test their ICT procedure. This task was carried out by the CEA. The data provided consisted in the isotopic composition data of 53 batches of fuel elements discharged from the Obrigheim reactor, reprocessed by COGEMA, and measured upon dissolution.

- to define common objectives to be assigned to all participants. These objectives were established by the WG-RIV as follows : a) to calculate the masses of U and Pu present in each of the 53 batches and b) to detect a certain number of anomalies intentionally introduced in part of the data to simulate various types of possible errors.

- to apply the various ICT procedures proposed to the common data set. This task is to be carried out by the participants themselves.

- to compare the results obtained by the various participants and to assess the performance of their respective ICTs. This task is being carried out by the CEA.

At present, the first two tasks have been completed, whereas the remaining two are in progress. An interim report on the results so far obtained was presented by the CEA at the 6th ESARDA Symposium in May 1984 /1/.

Evaluation of the BME Results

The crucial point of the BME lies in the evaluation of the results obtained by the various participants. To meet the objectives established for the exercise, the evaluation will be based on two types of "comparison" :

- the masses of U and Pu calculated by the various participants as compared to the masses of U and Pu determined by the reprocessor

- the anomalies discovered by the participants as compared to the anomalies introduced by the organizer of the exercise.

So far, only the first comparison has been undertaken. Since the statistical problem that we are reporting originated in this comparison, we will describe this step of the work in more detail.

The procedures by which the quantities of U and Pu contained in each dissolution are determined by the plant operator and by the participants are schematized in Fig. 1.

The operator takes samples from each fuel dissolution. These are used to measure the Pu/U mass ratio in the solution (by means of a double isotopic dilution analysis), the fuel burnup (by mass spectrometry of the Nd) and the isotopic composition of the U and Pu present in the solution (by mass spectrometry). From the results obtained, and from some additional information provided by the fuel manufacturer (fresh fuel mass), the operator calculates the quantities of U and Pu dissolved. This procedure is represented in the left part of Fig. 1.

The participants, in their turn, use the U and Pu isotopic compositions measured by the operator as an input to their ICT procedure. The results of their calculations are combined with some additional information provided by the fuel manufacturer (fresh fuel mass and, in some cases, fresh fuel enrichment) and with the fuel burnup
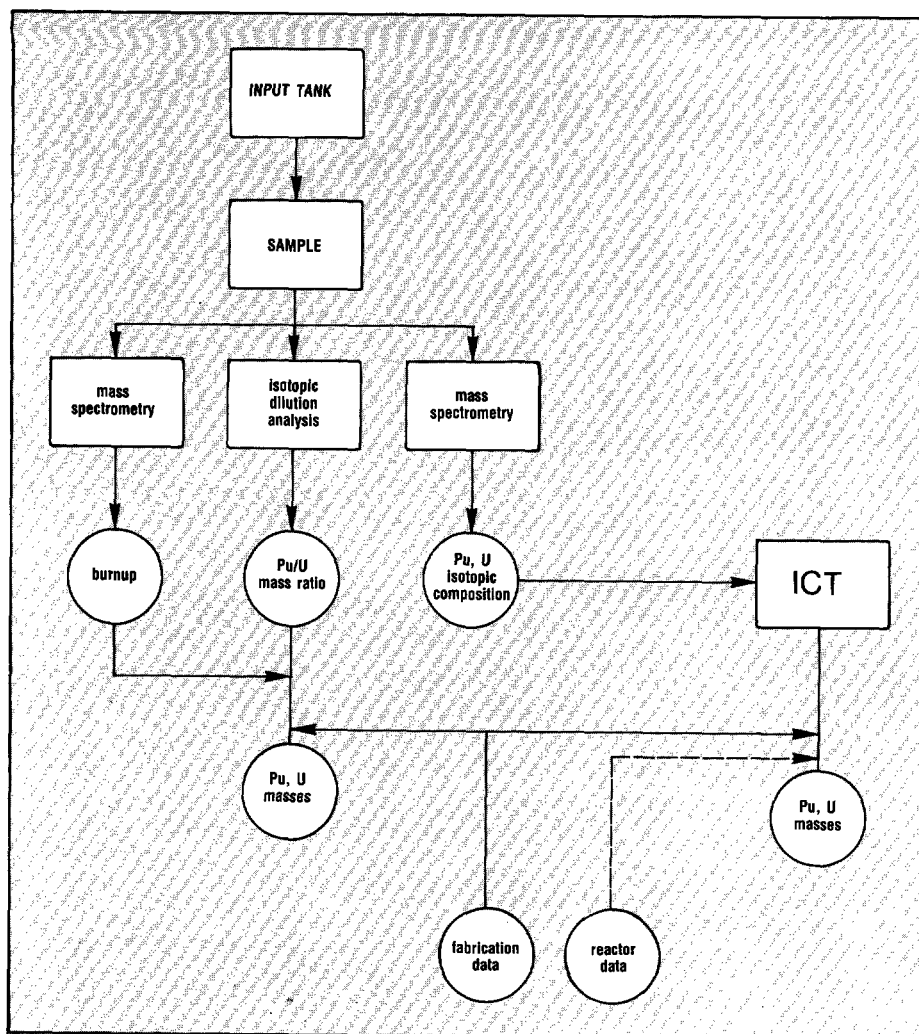
Fig. 1 - Scheme of the determination of U and Pu masses

calculated by the reactor operator, to determine the quantities of U and Pu contained in the dissolution. This procedure is represented in the right part of Fig. 1. It should be noted that some ICT procedures do not make use of the reactor burnup, since they can derive it directly from the isotopic data.

At the end of the exercise, each of the seven participants obtains values for the U and Pu masses contained in each dissolution. These must be compared with those determined by the operator. The comparison aims at characterizing each laboratory (therefore each ICT procedure) with a "precision" and an "accuracy" attained in calculating the U and Pu.

In performing this comparison, consideration must be given to the fact that the operator has obtained his results by means of a well established and proven procedure, which guarantees good accuracy.

How to make the comparison

Having decided "what" to compare, it is now necessary to define "how" to do it. The first step was to re-formulate the problem in

a form suitable for statistical treatment; this eventually led to the formulation below.

Eight data sets are given, representing the same physical quantities (U and Pu contained in the dissolutions) as determined by eight different laboratories. They must be compared. The assumption is made that one of these data sets (that of the operator) is affected by a negligible systematic error and therefore this set can be taken as a reference for the comparison. Under these conditions, the statistician is requested to calculate:

- *the "repeatability variance" of the results obtained by each of the eight laboratories. This parameter (which describes the variation among the results obtained for a given quantity by a single laboratory with constant operating conditions over a relatively short period of time) will be considered as representing the "precision" of that laboratory in using his ICT procedure. By common agreement, it will also be assumed as being the precision of "that particular ICT procedure".*

- *The "reproducibility variance" of the*

results obtained by each of the other seven laboratories, in respect of the operator's results. This parameter (which describes the variations between the results obtained by different laboratories) will be considered as representing the "accuracy" attained by that laboratory in applying its ICT procedure. Here again, the reproducibility variance will also be assumed as representing the accuracy of "that particular ICT procedure".

The problem described above is familiar to statisticians, since it is generated in a number of situations. A classic example is offered by the so-called "circular experiments", in which a given object is measured in turn by a certain number of laboratories. The results are then compared with the reference data in order to characterize the errors of the laboratories themselves or, if applicable, the error of the method generally employed by all the laboratories.

Appropriate statistical techniques are therefore available for treating such problems. One example is the well known Grubb's model, which was used in the present exercise.

All the models are, however, applicable within a well defined range of situations, and only if a certain number of conditions is met. In Grubb's model, for instance, it is required that the systematic errors of the various laboratories be either non-correlated or correlated through a linear dependence upon the value of the measured quantity.

In evaluating the results of the BME it appeared that the errors of certain laboratories were in some way correlated to each other. The reason for such interdependence probably resides in the fact that these laboratories share one or other of the following inconveniences /1/ :

- in creating correlations, they make use of historical data which do not perfectly match the fuel type used in this exercise
- in the ICT procedure, they make use of isotopic correlations which are particularly sensitive to anomalies that may possibly have occurred in the irradiation history of the fuel used for the BME.

Unfortunately, the correlated errors proved not to be dependent on the actual value of the quantity affected by the error. This fact prevented the use of the Grubb's model in its present form, and called for an adaptation to the specific conditions met in the exercise. The adaptation was worked out by Miss Neuilly, CEA, and is reported in the next article.

Considerations on the Use of the Proposed Method

The importance of the Grubb-Neuilly model is clearly shown by the numerical example

reported in the next article and by the considerations developed by Giacometti in ref. /1/.

If several laboratories are to be compared which exhibit dissimilar "qualities" (i.e. they are able to achieve dissimilar accuracies), and if there is evidence that there exists a correlation between their errors which is not dependent on the actual value of the measured quantity, incorrect conclusions may be derived by application of Grubb's model: the "best" laboratories are assigned a lower accuracy than they are able to achieve, whereas the "worst" laboratories are rewarded with a better accuracy.,

This phenomenon was noticed by Giacometti in analyzing the BME results. In fact, it was known that the ICT procedure developed by the CEA should have produced a set of results having a comparatively low systematic error, since this procedure had been calibrated and long tested on data similar to those used in the BME. Unexpectedly, the systematic error estimated by the Grubb's model was relatively high, even comparable with that of other procedures which are still in the phase of development. This inconsistency was eventually removed by the use of the Grubb-Neuilly model.

A misjudgement like that which occurred at the outset of the evaluation of the BME results, could possibly have undesired consequences in the practice of safeguards. For this reason I am convinced that the proposed statistical model may help the inspector in carrying out certain data evaluations and give him the assurance of having a better understanding of the data which he has to judge.

Reference

1. A. GIACOMETTI (CEA, Cadarache, France), The ESARDA Exercise to Test the Performance of ICT Procedures. Proc. of the 6th ESARDA Symp., Venice, 14-18 May 1984, pp. 213-222

Inter-Laboratory Comparisons : The Problem of Correlated Errors

M. Neuilly
DSMN-SCMN CEN Cadarache
CEA, France

Abstract

The methods presented in various works for analyzing the results of interlaboratory circuits alsways assume that the errors of the different laboratories are either independent or linear functions of the value of the quantity measured. If these models are applied in situations where the errors are correlated, incorrect variance estimates are obtained. This article examines this problem with regard to a circuit organized by the ESARDA Working Group for Reprocessing Input Verification (RIV) to compare certain results obtained by application of ICT procedures. An analysis method was developed to estimate the errors when they are correlated, and is described here.

Introduction

This article describes how to compare the results obtained by different laboratories for the *purpose of estimating their different* errors or their variances. The text is also applicable to any other comparison factor than the laboratory : it is understood that the term "laboratory" may be considered to designate a measurment method as well as an individual laboratory.

A situation is studied in which several objects of the same nature (e.g. samples taken from different batches of a solution) are each measured by several laboratories, (or by several different methods) with all the laboratories measuring the same quantities in every object. This is the case, for example, of quality control procedures carried out by both the manufacturer and the customer, or the case of comparisons between the results of an inspecting agency and a plant operator.

A circuit of this type may operate in two different ways :

- If all the laboratories use the same method, the purpose of the circuit is to characterize the errors due to the method itself and to estimate the repeatability variance (i.e. the variations among the results obtained for a given quantity by a single laboratory with constant operating conditions over a relatively short time period) and the reproducibility variance (i.e. the variations among the results obtained by different laboratories). International Standard ISO-5725 defines the method applicable in this case /3/.

- In other cases, the objective is to characterize not the method but the laboratories participating in the circuit. The laboratories may use different techniques, and their performance records are generally not the same. Circuits of this type are considered here.

Error Correlation Models for Different Laboratories

If q represents the number of objects involved in the comparisons, and j is a subscript identifying each of them (j = 1, 2, ..., q), and if the actual value of a measured quantity (designated μ_j) varies in a random fashion from one object to the other, it can be written in the following form :

$$\mu_j = \mu + A_j$$

where μ is the mean of the population of actual values, and A_j is a random variable with a mean value of zero and a standard deviation of σ_A.

Each quantity is measured by p laboratories designated by the subscript i (i = 1, 2, ..., p). The laboratories are the same for all the quantities measured. The result of laboratory i is indicated x_{ij}.

For laboratory i and quantity j the error $(x_{ij} - \mu_j)$ is a random variable with an expected value of B_{ij}. The result can thus be written :

$$x_{ij} = \mu + A_j + B_{ij} + \epsilon_{ij} \qquad (1)$$

where ϵ_{ij} is a random variable with an expected value of zero. All ϵ_{ij} errors are independent of one another. B_{ij} is the systematic error of laboratory i for quantity j.

In order to interpret the results of the circuit, a number of hypotheses must be made concerning the manner in which the systematic errors vary from one quantity to another.
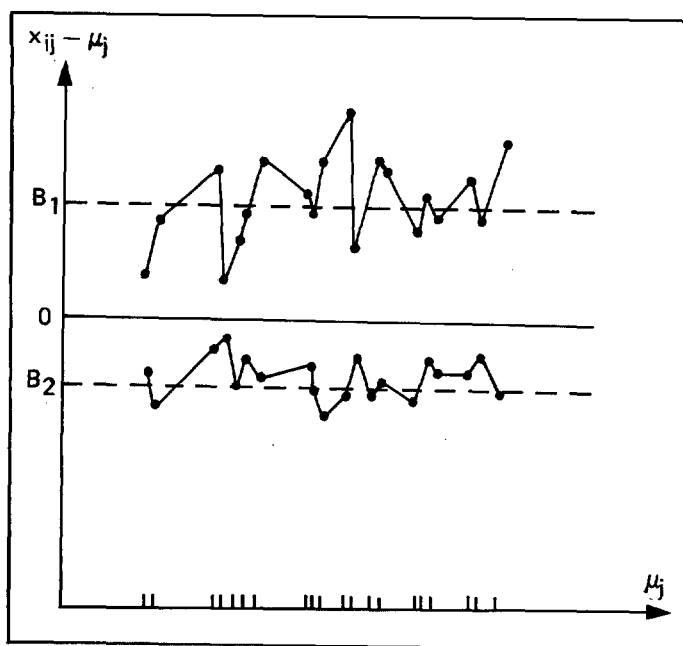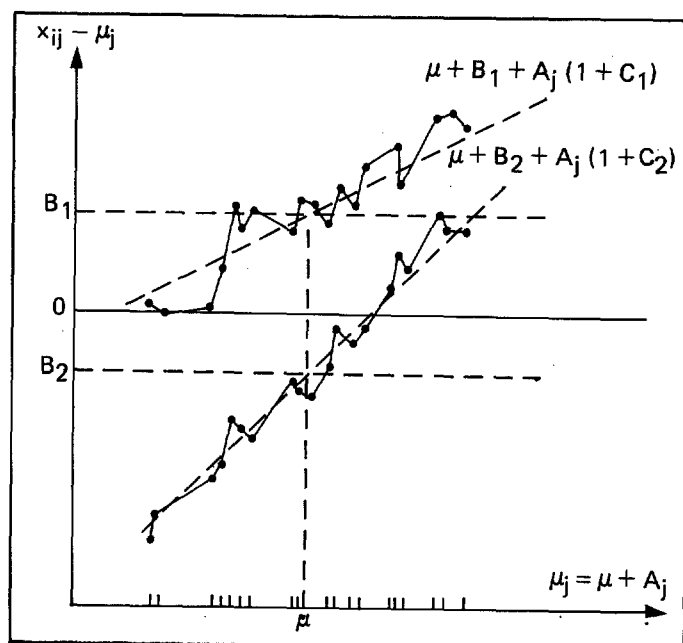
Fig. 1 - Additive Model

Fig. 2 - Linear Model

The simplest case is when the systematic errors are constant in each laboratory : this is the "additive model". Figure 1 shows the variation, for a series of quantities, of the total error for two laboratories (i = 1 and 2).
The model described in formula (1) then becomes :

$$x_{ij} = \mu_j + B_i + \epsilon_{ij} \qquad (2)$$

The errors $(x_{ij} - \mu_j)$ made by the different laboratories on a single quantity j are independent of one another.
In other cases, the systematic errors B_{ij} vary from one measured quantity to another, for example according to the actual value μ_j. If, in each laboratory i, the systematic errors are linear functions of the actual value, the model is a "linear model". Figure 2 shows the variation in the total error from one measured quantity to another for two laboratories (i = 1 and 2).
In formula (1), B_{ij} must be replaced by a function of the form :

$$B_{ij} = B_i + A_j C_i$$

and the model becomes :

$$x_{ij} = \mu + B_i + A_j(1 + C_i) + \epsilon_{ij} \qquad (3)$$

The overall errors of the different laboratories for quantity j are then :

$$x_{ij} - \mu_j = B_i + A_j C_i + \epsilon_{ij}$$

They are correlated random variables. The covariance of the errors of laboratories i and i' is equal to :

$$\text{Cov} \ (x_{ij}, x_{i'j}) = C_i C_{i'} \sigma^2_A$$

Another possibility is that the errors of different laboratories may be correlated, but independent of the actual values μ_j.
For example, in chemical analysis, the correlation may be due to a "third-element" effect. This occurs in the spectrographic analysis of uranium oxide impurities, where the results are perturbed by the presence of Cl^- ions : all samples containing Cl^- ions give erroneous results in each laboratory. Similarly, if thorium and uranium are not perfectly discriminated in a chemical methods, all uranium determinations will be overestimated in samples containing trace quantities of thorium.
In isotopic correlation methods, if the different laboratories use the same isotopic ratios to determine the Pu/U ratio and if this correlation is perturbed by variations in the irradiation conditions, the effects of these variations will be proportional from one laboratory to another. This was the case with the "Benchmark" exercise organized by the ESARDA RIV group /2/. Another reason for correlated errors in this case was that all the laboratories used the same isotopic analysis results for their calculations.
Figure 3 shows an example of correlation between the errors of laboratories i = 1 and i = 2.
We examined the case of overall error correlation in which the B_{ij} terms can be expressed in the following form :

$$B_{ij} = B_i + C_i \delta_j$$

where C_i is a characteristic factor of laboratory i, and where δ_j is a quantity that varies in random fashion with the measured quantity : its expected value is zero and its

standard deviation is σ_δ.

$$E(\delta_j) = 0$$

$$\text{Var} \ (\delta_j) = \sigma^2_\delta$$

Formula (1) then becomes :

$$x_{ij} = \mu + A_j + B_i + C_i \delta_j + \epsilon_{ij} \qquad (4)$$

The overall errors $(x_{ij} - \mu_j)$ for the different laboratories for quantity j take the following form :

$$x_{ij} - \mu_j = B_i + C_i \delta_j + \epsilon_{ij}$$

The covariance for the errors of laboratories i and i' is equal to :

$$\text{Cov} \ (x_{ij}, x_{i'j}) = C_i C_{i'} \sigma^2_\delta$$

In order to complete the general model (1) it was also assumed that in each laboratory the variance of the random error ϵ_{ij} was constant and designated by σ^2_i.

Choice of Reference Values

Since the actual values of the measured quantities are always unknown, it is necessary to choose a reference value for each quantity. Two types of reference are possible.

External reference

A reference independent of the laboratory circuit results is available in some experimental situations. For example, an expert laboratory may be selected as a reference. In the Benchmark exercise, for example, the reference laboratory may be

the operator's laboratory used to analyse the uranium and plutonium by isotopic dilution.

Moreover, specially prepared quantities are frequently used for the circuit in order to have values given in advance. In chemical analysis, for example, the circuit may involve synthetic solutions; in such cases, the reference value is the desired nominal value.

By convention, the reference value is assumed to be free of any systematic errors. It may, however, be affected by a random error : reference laboratory measurement error or synthetic solution preparation error. The reference value may thus be expressed in the following form :

$$y_j = \mu + A_j + \epsilon_{yj} \qquad (5)$$

where ϵ_{yj} is a random variable with a mean value of zero. It is further assumed that its variance is independent of the quantity measured :

$$Var (\epsilon_{yj}) = \sigma_y^2$$

Inter-laboratory mean reference

If a reference laboratory is unavailable, the mean values of the different results obtained for each quantity measured may be taken as the reference, i.e. the quantities

$$\overline{x}._j = \frac{1}{p} \sum_i x_{ij}$$

For a linear model:

$$x_{ij} = \mu + A_j(1 + C_i) + B_i + \epsilon_{ij}$$

The means $\overline{x}._j$ take the following form:

$$\overline{x}._j = \mu + A_j(1 + \overline{C}) + \overline{B} + \overline{\epsilon}._j$$

To use them as a reference implies that the following condition is met:

$$\Sigma C_i = \Sigma B_i = 0$$

so that the means are free of systematic errors

$$\overline{x}._j = \mu + A_j + \overline{\epsilon}._j$$

Error Estimation in the Additive and Linear Models

For an additive model represented by relation (2), the systematic errors B_i can be estimated by the following relations :
- with an external reference:
$$\widehat{B}_i = \overline{x}_i. - \overline{y}$$
- with reference to the mean:
$$\widehat{B}_i = \overline{x}_i. - \overline{x}$$

where $\overline{x}_i$ is the mean of the q results from laboratory i, and $\overline{x}$ is the overall mean of all the results.

Moreover, if all the variances σ_i^2 of the random errors have the same value, it may be estimated by conventional variance analysis /1/ /3/.

When these variances are different, GRUBBS showed that they can be estimated by using the estimated variances and covariances of the results /4/. This method is discussed in the Appendix.

For a linear model, the variance estimation method differs depending on whether the reference is an external laboratory or the inter-laboratory mean. JEACH described a method for estimating the systematic errors and the variances of the random errors in the case of an external reference /5/. MANDEL examined the linear model with reference to the mean in the special case where all the variances σ_i^2 of the random errors have the same value σ^2 /6/ : the C_i coefficients are estimated by computing the slope of the regression line for x_{ij} versus $x._j$ for each laboratory (Fig. 4).

It can be shown that the expected value for the estimate C_i is equal to C_i if the variance σ^2 of the random errors is very small compared with the variance σ^2_A of the actual values. This condition is always met in practice when a linear model is examined.

MANDEL then showed that a variance analysis leads to the following estimation of σ^2:

$$\hat{\sigma}^2 = \frac{1}{(p-1)(q-2)} [Q_{tot} - Q_{lab} - Q_G - \sum_j (\overline{x}_j - \overline{x})^2 \sum_i \hat{C}_i^2]$$

where Q_{tot}, Q_{lab} and Q_G are the sums of the squares characterizing the total variation, the variation among laboratories, and the variation among the quantities measured.

When the variances σ_i^2 have different values but are small compared with the variance σ^2_A of the actual values, the C_i coefficients can still be estimated by the same regression method. The resulting estimates can be used to modify GRUBBS'
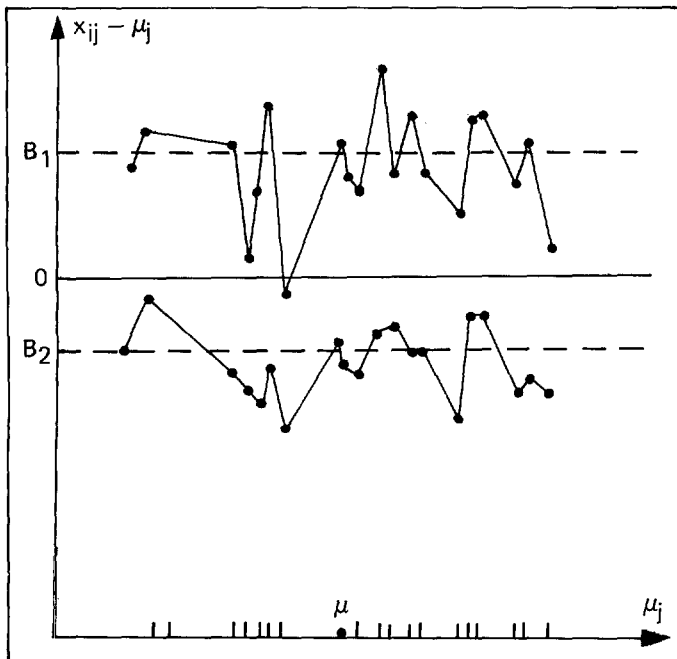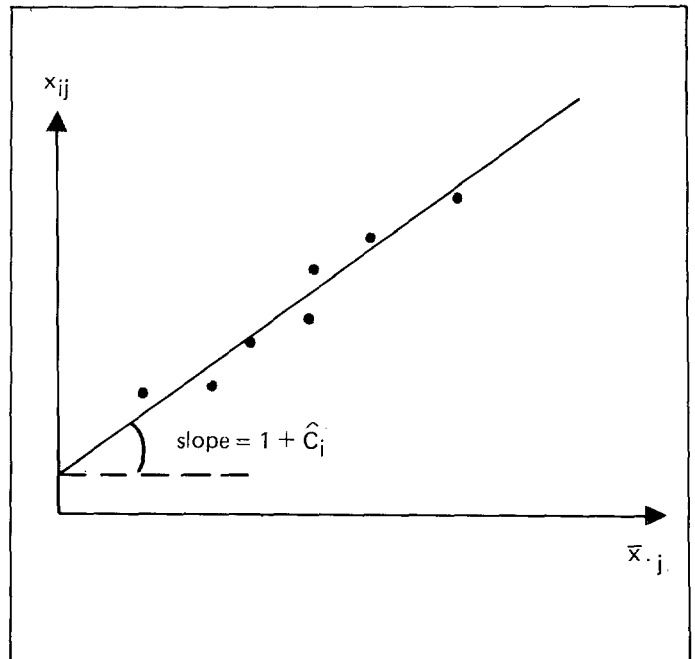

Fig. 3 - General Model for Error Correlation

Fig. 4 - Estimating the C_i Coefficients

formula /1/. The formulas used are indicated in the Appendix.

Error Estimation in the General Model

If the errors are correlated among the different laboratories but independent of the actual value, i.e. if the model takes the general form (4), it is no longer possible to use the inter-laboratory mean as the reference: the C_i coefficients generally have the same sign and their sum cannot be assumed equal to zero. An external reference y_j of the form (5) is therefore necessary.

In order to estimate the C_i coefficients, at least three "correlated" laboratories and the reference laboratory are required. As shown in the Appendix, an unbiased estimate of each variance σ_i^2 can be given only if the number q of quantities measured is very large, or if the variance σ_A^2 of the actual values is zero.

This condition can be met by considering the difference for each laboratory with regard to the reference, i.e.

$$z_{ij} = x_{ij} - y_j$$

From (4) and (5), the resulting model is:

$$z_{ij} = B_i + C_i\delta_j + (\epsilon_{ij} - \epsilon_{yj}) \qquad (6)$$

The variances of the correlated errors $C_i\delta_j$ can be estimated, but it is impossible to discriminate between the independent errors of laboratory i and the reference.

Let it be assumed that three laboratories and the reference laboratory are available, and that the objective is to estimate the variance σ_i^2 of the random errors and the variance $C_i^2\sigma_\delta^2$ of the correlated errors of laboratory i = 1.

If: $S_{ii'} = \dfrac{1}{q-1}\sum_j (z_{ij}-\bar{z}_{i.})(z_{i'j}-\bar{z}_{i'.}) =$

$$= \frac{1}{q-1}\left[\sum_j z_{ij}z_{i'j} - \frac{\sum_j z_{ij}.\sum_j z_{i'j}}{q}\right]$$

then, from model (6), with E designating the mathematical expectation:

$$E(S_{11}) = C_1^2\sigma_\delta^2 + (\sigma_1^2 + \sigma_y^2) \qquad (7)$$

$$E(S_{12}) = C_1 C_2 \sigma_\delta^2$$

$$E(S_{13}) = C_1 C_3 \sigma_\delta^2 \qquad (8)$$

$$E(S_{23}) = C_2 C_3 \sigma_\delta^2$$

Solving the three equations (8) for $C_1^2\sigma_\delta^2$ gives the estimate :

$$\frac{S_{12}S_{13}}{S_{23}}$$

If the variable δ_j is normal, the expected value for this quantity is :

$$E\left[\frac{S_{12}S_{13}}{S_{23}}\right] = \frac{q+1}{q-1}C_1^2\sigma^2 + \frac{\sigma_1^2 + \sigma_y^2}{q-1} \qquad (9)$$

Combining relations (7) and (9) gives the unbiased estimates:

$$\hat{\sigma}_1^2 + \hat{\sigma}_y^2 = \frac{q+1}{q}S_{11} - \frac{q-1}{q}\frac{S_{12}S_{13}}{S_{23}} \qquad (10)$$

$$\widehat{C_1^2\sigma_\delta^2} = \frac{q-1}{q}\frac{S_{12}S_{13}}{S_{23}} - \frac{S_{11}}{q} \qquad (11)$$

If p correlated laboratories are available (where p > 3) in addition to the reference, and if once again the errors of laboratory 1 are studied, there are (p-1)(p-2)/2 pairs ii' where i and i' are not 1, and thus the same number of estimates of the form (10) and (11). The mean estimates are computed as follows:

$$A_1 = \frac{2}{(p-1)(p-2)}\sum_i \sum_{i'} \frac{S_{1i}S_{1i'}}{S_{ii'}}$$

Laboratory (i)	1	2	3	4	5
σ_i^2	0.250	0.360	0.090	0.040	0.010
$C_i^2\sigma_\delta^2$	0.111	0.640	0.062	0.160	0
RESULTS (1st batch)	12.048	12.255	12.651	12.358	12.821
	12.672	13.125	13.106	13.145	13.045
	10.762	11.341	11.030	11.030	11.106
	10.996	8.788	10.157	10.018	10.439
	10.500	10.962	11.215	10.866	10.761
	14.433	15.633	14.347	14.552	13.886
	7.559	7,849	8.151	7.585	8.453
	13.868	15.182	13.792	14.059	14.488
	12.399	12.051	13.186	12.715	13.061
	8.315	7.830	7.378	7.493	7.289
	11.410	10.954	11.070	10.690	10.859
	11.411	11.476	11.260	11.614	11.206
	13.717	14.914	14.639	14.528	.14.969
	10.409	10.758	11.145	10.149	11.170
	10.189	10.218	10.882	10.635	10.473
RESULTS (2nd batch)	11.603	11.326	11.021	11.132	10.656
	10.390	10.526	10.618	10.845	10.769
	13.683	14.816	13.735	13.729	13.425
	11.991	12.333	12.314	12.939	12.159
	8.777	9.381	7.880	8.552	7.888
	14.364	16.607	14.999	14.972	14.781
	10.550	10.380	10.655	10.730	10.400
	11.780	11.460	12.032	11.960	13.001
	11.689	11.962	11.282	11.290	10.976
	11.187	10.966	10.247	10.892	10.570
	12.683	11.644	13.286	12.834	12.658
	13.767	13.245	13.757	13.152	13.422
	12.247	12.471	12.534	12.792	13.000
	11.373	11.110	11.415	10.860	11.658
	11.566	11.947	12.263	11.874	12.037

Table 1 - Numerical example

Laboratory (i)	1	2	3	4	5
TOTAL VARIANCE					
Actual value σ_i^2	0.361	1.000	0.152	0.200	0.010
Estimate (Grubbs):					
1st batch	0.285	0.456	0.057	0.009	0.137
2nd batch	0.142	0.563	0.103	0.034	0.219
All results	0.208	0.493	0.087	0.023	0.180
INDEPENDENT ERRORS					
Actual value $\sigma_i^2 + \sigma_y^2$	0.260	0.370	0.100	0.050	
Estimate:					
1st batch	0.310	0.511	0.074	0	
2nd batch	0.179	0.536	0.119	0.077	
All results	0.215	0.479	0.089	0.029	
CORRELATED ERRORS					
Actual value $C_i^2 \sigma_\delta^2$	0.111	0.640	0.062	0.160	
Estimate:					
1st batch	0.121	0.107	0.031	0.212	
2nd batch	0.213	0.366	0.058	0.171	
All results	0.196	0.267	0.047	0.207	

Table 2 - Variance estimates

For example, if p = 4 laboratories:

$$A_1 = \frac{1}{3}\left[\frac{S_{12}S_{13}}{S_{23}} + \frac{S_{12}S_{14}}{S_{24}} + \frac{S_{13}S_{14}}{S_{34}}\right]$$

Relations analogous to (10) and (11) are then obtained:

$$\hat{\sigma}_1^2 + \hat{\sigma}_y^2 = \frac{q+1}{q}S_{11} - \frac{q-1}{q}A_1 \quad (12)$$

$$\widehat{C_1^2 \sigma_\delta^2} = \frac{q-1}{q}A_1 - \frac{S_{11}}{q} \quad (13)$$

If a negative value is obtained for $(\widehat{C_1^2 \sigma_\delta^2})$ then:

$$\hat{\sigma}_1^2 + \hat{\sigma}_y^2 = S_{11}$$

$$\widehat{C_1^2 \sigma_\delta^2} = 0$$

If a negative value is obtained for $(\hat{\sigma}_1^2 + \hat{\sigma}_y^2)$, then

$$\hat{\sigma}_1^2 + \hat{\sigma}_y^2 = 0$$

$$\widehat{C_1^2 \sigma_\delta^2} = S_{11}$$

The formulas for the other laboratories are obtained by circular permutation of the subscripts. For 4 laboratories, for example:

$$A_2 = \frac{1}{3}\left[\frac{S_{12}S_{23}}{S_{13}} + \frac{S_{12}S_{24}}{S_{14}} + \frac{S_{23}S_{24}}{S_{34}}\right]$$

$$A_3 = \frac{1}{3}\left[\frac{S_{13}S_{23}}{S_{12}} + \frac{S_{13}S_{34}}{S_{14}} + \frac{S_{23}S_{34}}{S_{24}}\right]$$

$$A_4 = \frac{1}{3}\left[\frac{S_{14}S_{24}}{S_{12}} + \frac{S_{14}S_{34}}{S_{13}} + \frac{S_{24}S_{34}}{S_{23}}\right]$$

If the variance σ_y^2 of the reference is negligible compared with the variance of the other errors, then the variance of the total error of laboratory 1, for example, is estimated by the sum of quantities (12) and (13), i.e. by S_{11}.

Numerical Example

Table 1 indicates the variances of 5 laboratories. These values were used to simulate 2 sets of 15 results per laboratory.

If GRUBBS' method is applied to the five laboratories, the estimated total variances

given in Table 2 are obtained. It can be seen that these estimates are all underestimated for the first four laboratories which have correlated errors, and significantly overestimated for the fifth laboratory for which the errors are independent of the others.

If the fifth laboratory is taken as a reference, then the formulas presented in the preceding paragraph give the estimates shown in Table 2. All of them are satisfactory except for $C_2^2 \sigma_\delta^2$, which is understimated by a factor of 2.

Conclusions

A method was developed to interpret the results of an inter-laboratory circuit when the errors are not independent of one another.

The estimated variances obtained using this method were verified by a numerical example based on a simulation, and the results were satisfactory.

Appendix

GRUBBS' Formula (Additive Model)

Consider the relation:

$$S_{ii'} = \frac{1}{q-1}\sum_j (x_{ij} - \bar{x}_{i.})(x_{i'j} - \bar{x}_{i'.}) = \quad (14)$$

$$= \frac{1}{q-1}\left[\sum_j x_{ij}x_{i'j} - \frac{\sum_i x_{ij}\sum_i x_{i'j}}{q}\right]$$

Each quantity S (where $i \neq i'$) is an estimate of σ_A^2.

The variance σ_i^2 of the random errors of laboratory i is estimated by:

$$\hat{\sigma}_1^2 = S_{ii} - \frac{2}{p-1}\sum_{i'} S_{ii'}$$

$$+ \frac{2}{(p-1)(p-2)}\sum_{i'}\sum_{i''} S_{i'i''} \quad (15)$$

Linear Model with Reference to the Mean

The quantities $S_{ii'}$ are defined by (14), and the quantities $V_{ii'}$ by:

$$V_{ii'} = \frac{S_{ii'}}{(1+\hat{C}_i)(1+\hat{C}_{i'})}$$

where the $\hat{C}_i$ and $\hat{C}_{i'}$ coefficients are estimated by regression formulas: $\hat{C}_i$ is the calculated slope of the regression line for x_{ij} versus $\bar{x}_{.j}$

Each quantity $V_{ii'}$ is an estimate of σ_A^2 and the variance σ_i^2 is estimated by a relation similar to (15):

$$\hat{\sigma}_i^2 = \hat{S}_{ii} - (1+\hat{C}_i)^2 \left[\frac{2}{p-1} \sum_{i'} V_{ii'} \right.$$
$$\left. - \frac{2}{(p-1)(p-2)} \sum_{i'} \sum_{i''} V_{i'i''} \right]$$

General Model with Three Laboratories and an External Reference

If the quantities $S_{ij'}$ are defined by relation (14) and S_{iy} by:

$$S_{iy} = \frac{1}{q-1} \sum_j (x_{ij} - \bar{x}_{i.})(y_j - \bar{y})$$

then σ_A^2 can be estimated from a linear combination of the S_{iy} quantities:

$$\sigma_A^2 = \sum_i k_i S_{iy} \qquad (16)$$

where $\sum_i K_i = 1$

This quantity can be used to estimate the variance $C_1^2 \sigma_\delta^2$; thus:

$$E(S_{12}) = \sigma_A^2 + C_1 C_2 \sigma_\delta^2$$

$$E(S_{13}) = \sigma_A^2 + C_1 C_3 \sigma_\delta^2$$

$$E(S_{23}) = \sigma_A^2 + C_2 C_3 \sigma_\delta^2$$

and therefore:

$$\widehat{C_1^2 \sigma_\delta^2} = \frac{(S_{12} - \hat{\sigma}_A^2)(S_{13} - \hat{\sigma}_A^2)}{S_{23} - \hat{\sigma}_A^2}$$

Moreover:

$$E(S_{11}) = \sigma_A^2 + C_1^2 \sigma_\delta^2 + \sigma_1^2$$

hence the estimate:

$$\hat{\sigma}_1^2 = S_{11} - \hat{\sigma}_A^2 - \widehat{C_1^2 \sigma_\delta^2} \qquad (17)$$

Our objective was therefore to determine the k_i coefficients to obtain an unbiased estimate (17). Since laboratory 1 is the subject of the study, laboratories 2 and 3 have the same relation to it, and estimate (16) takes the following form:

$$\hat{\sigma}_A^2 = k(S_{2y} + S_{3y}) + (1-2k)S_{1y} \qquad (18)$$

In this case, if the errors are normal variables, the expected value of σ_1^2 to within the second order with respect to the variances is:

$$E(\hat{\sigma}_1^2) = \frac{q-2}{q-1}\sigma_1^2 - \frac{2C_1^2 \sigma^2}{q-1} - \frac{\sigma_A^2}{q-1} \cdot$$
$$\cdot \left[1 + \frac{(4k-1)C_1^2 + (1-2k)C_1(C_2 + C_3)}{C_2 C_3} \right]$$

Coefficient k may be chosen equal to 0.25 in order to cancel the $C_1^2 \sigma_A^2$ term, then (18) becomes:

$$\hat{\sigma}_A^2 = \frac{1}{4}(S_{2y} + S_{3y}) + \frac{1}{2}S_{1y}$$

Thus:

$$E(\hat{\sigma}_1^2) = \frac{q-2}{q-1}\sigma_1^2 - \frac{2C_1^2 \sigma_\delta^2}{q-1} -$$
$$- \frac{\sigma_A^2}{q-1}\left[1 + \frac{C_1(C_2 + C_3)}{2C_2 C_3} \right]$$

This estimate is correct only if the number of samples q is very large. If this condition is not met, the variances can be correctly estimated only if the variance σ_A^2 is zero.

References

/1/ AFNOR XO6J-DOC 72 — Estimation des erreurs à partir de comparaisons entre laboratoires (Standard in preparation)

/2/ A. GIACOMETTI, The ESARDA Exercise to Test the Performance of ICT Procedures. ESARDA Conference, Venice, May 1984

/3/ ISO Standard 5725. Precision of Test Methods: Determination of Repeatability and Reproducibility for a Standard Test Method by Inter-Laboratory Tests

/4/ F.E. GRUBBS. On Estimating the Precision of Measuring Instruments and Product Variability. J. Am. Stat. Assoc. 43 (1946) p. 243

/5/ J.L. JAECH. Extension of Grubbs' Method when Relative Biases Are not Constant. Nuclear Materials Management J., Spring 1979, p. 76

/6/ J. MANDEL. The Analysis of Structured Data. J. Wiley & Sons, 1964

Near Real Time Material Accountancy
Calculation of the Parameters Used in Pages's Test

Barry J. Jones,
British Nuclear Fuels plc
Risley Warrington (U.K.)

Abstract

A pre-requisite for the application of Pages's Test for the detection of materials diversion is the choice of two parameters, generally referred to as H and K, so that the test has the desired performance in terms of false alarm probability (FAP) and detection capability. Whereas these parameters have previously been derived by simulation, a method is now described whereby H and K can be calculated for a given FAP. The advantages which come with this approach are improved accuracy and an opportunity to readily obtain values of the parameters appropriate to FAPs of less than 5%.

Introduction

Let us denote a series of MUF values as, MUF_1, MUF_2, ..., MUF_i. First calculate Y_1, Y_2, ..., Y_i as the series of SITMUF /1/ values generated by MUF_1 ... MUF_i. Then carry out Page's Test on the values Y_1, Y_2, ... Y_i using the test statistic defined by

$$T_o = 0$$
$$T_i = max(0, T_{i-1} + Y_i - K) \qquad i > 0$$

An alarm is given if $T_i > H$. When no diversion occurs, any MUF series leads to a series of SITMUF values which are independent values from a Gaussian distribution with mean zero and variance 1. Figure 1 illustrates the process for a single campaign, using H = 2 and K = 0, and shows the random fluctuation of the test statistic until it exceeds the alarm level at period 21. Also shown at periods 7, 8. 9, 14, 18 and 19 is the resetting of the test statistic (from a negative value) to zero before proceeding to the next period.

Simulation of H K Values

If the number of periods comprising a campaign is fixed, then the number of campaigns generating a false alarm will depend on the choice of the parameters H and K. For 10,000 simulated campaigns, Table I shows the numbers generating false alarms for a range of values of H and K and a campaign length of 21 periods. (Throughout this paper a campaign length of 21 periods has been used for illustrative purposes and corresponds to the implementation of weekly balance periods in a plant operating with two campaigns per year separated by maintenance periods of 5 weeks.) The false alarm probability is defined as the probability that a false alarm will be generated at any time within the length of this campaign. For purposes of this mathematical analysis, the occurrence of an alarm ends the campaign concerned.

The behaviour of the test statistic T was described earlier. At the beginning of a campaign T_o has the value zero and once T_i exceeds H an alarm is registered and the campaign is ended. At periods during the campaign the value of T will, on average, tend to increase. The reason for this is that whenever Y_i is positive T increases by the full amount of Y_i but, if Y_i is negative, T may not decrease by the full extent of Y_i since T is not allowed to fall below zero. Therefore, the likelihood of a given campaign ending increases as that campaign continues. Furthermore, the number of campaigns still running decreases as the number of periods increases. If these two factors are taken into account, the expected observation would be that the number of alarms per period would increase as the campaigns progressed from start-up, pass through a maximum rate, and then decline as the falling number of campaigns still running becomes the controlling factor.

If Table II is examined, the shortcomings of deriving H and K by simulation become apparent. Whilst the behaviour described above is followed for the first few periods, the number of alarms per period certainly do not **steadily** increase and later decrease as the campaigns grow longer. This sort of imprecision due to sampling error is typical of results obtained by simulation. The quality of the results will be worse as the number of observed alarms become fewer. The consequence is that poorer results will be obtained as H or K is increased in order to move to lower FAP. This can only be remedied if the number of simulations is increased to compensate. This is not an attractive proposition since already excessively long simulation runs are necessary even to prepare the H K values for 5% FAP and the number of runs required escalates rapidly as the FAP falls.

Further difficulties arise on account of the sequence of random ITMUFs (mean zero, variance 1), Y_i, used in the test algorithm. Each column of figures in Table II required the generation of approximately 200,000 random numbers which were subsequently translated into ITMUFs falling in the range -3 sigma to +3 sigma. The random number

H	K																				
	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
0														2024	1575	1243	908	674	482	348	253
0.5											1257	919	680	491	351	256					
1.0								736	518	369	266										
1.5				727	497	331	229	139	94												

H	K																				
	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
2.0										1078	715	454	293	179	117	70					
2.5									1092	694	427	247	145	87							
3.0						1484	901	514	292	161	86										
3.5					1446	848	455	254	133	62											
4.0				1562	891	478	235	119	54	.											
4.5				1046	561	262	125	59													
5.0			1349	711	350	155	69														
5.5			952	460	208	96															
6.0		1364	667	315	127	55															

To Illustrate how Table 1 is compiled, Table 2 shows the false alarms simulated, period by period, for H= 2 and a range of values of K.

Table 1   Simulated false alarms (per 10,000 campaigns of 21 periods) for various H and K
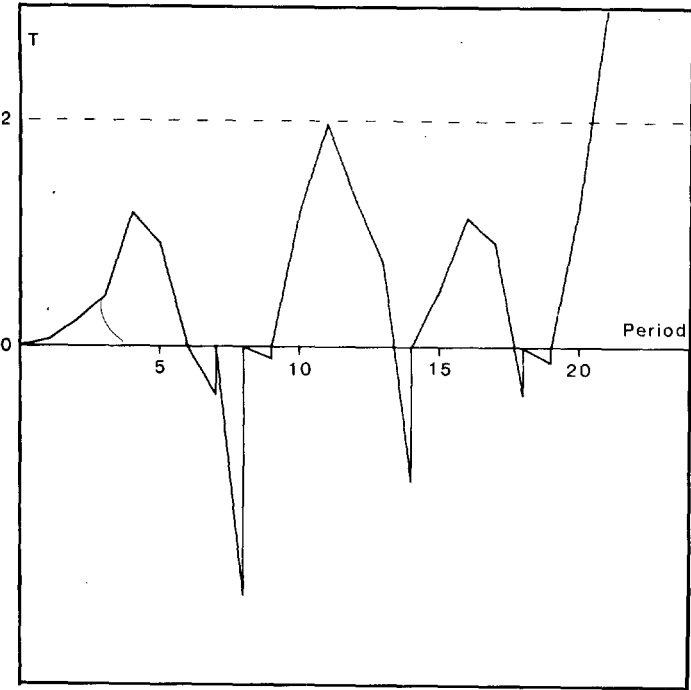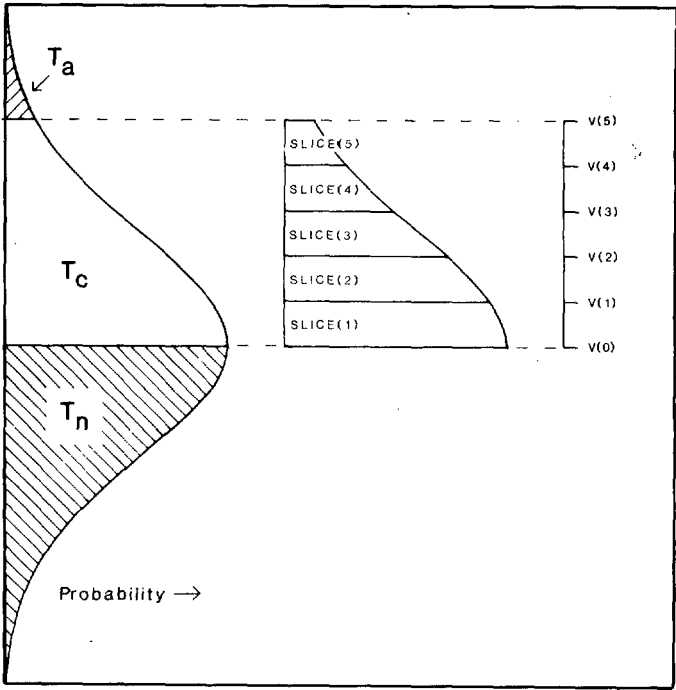
Figure 1 - Behaviour of the test statistic 'T'

Figure 2 - Behaviour of the test statistic 'T'

Period	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
1	12	8	6	4	3	2	1	1	1			
2	39	30	20	13	9	8	6	3	3			
3	59	36	26	17	8	4	2	1	1	1		
4	46	30	21	14	7	6	4	2	2	1		
5	53	30	17	10	3	1						
6	58	47	29	12	10	6	4	2				
7	58	37	24	18	8	4	3	2	1			
8	54	36	24	15	12	9	7	3	3			
9	46	33	22	12	10	7	4	1		1		
10	60	34	23	18	7	3	1		1			
11	58	39	26	14	8	2	1	1	1			
12	52	37	19	17	14	8	4	2	1			
13	58	25	11	6	2	2	2	1				
14	44	26	18	15	9	6	3	1	1			
15	54	34	22	8	9	7	3	1	1			
16	51	31	19	15	7	6	3	1	1			
17	56	39	20	17	10	10	8	6	6	1	1	1
18	56	39	28	20	8	3	2	2	2			
19	56	40	22	14	11	8	5	3	1	1	1	1
20	50	39	24	14	8	5	1					
21	58	45	33	20	16	10	6	2	2			
Total	**1078**	**715**	**454**	**293**	**179**	**117**	**70**	**35**	**28**	**5**	**2**	**2**

Table II Simulated false alarms (per 10,000 campaigns of 21 periods) for H = 2 and various K

| Period | Determination | | | | | | | | | | | Std |
	1	2	3	4	5	6	7	8	9	10	Mean	Dev
1	6	8	10	9	14	9	5	12	7	13	9.3	3.0
2	20	27	18	13	26	15	22	19	18	22	20.0	4.4
3	26	28	26	29	18	24	25	22	26	24	24.8	3.1
4	21	28	24	32	19	24	29	26	33	21	25.7	4.8
5	17	29	28	25	23	24	19	18	23	28	23.4	4.3
6	29	21	19	18	22	22	25	27	38	23	24.4	5.9
7	24	20	31	28	23	22	21	26	25	22	24.2	3.4
8	24	27	15	21	29	32	30	25	22	22	24.7	5.0
9	22	27	14	25	21	15	23	29	18	21	21.5	4.9
10	23	29	24	25	28	27	26	28	25	26	26.1	1.9
11	26	26	18	16	38	24	25	19	26	20	23.8	6.2
12	19	23	27	29	24	22	23	22	21	25	23.5	2.9
13	11	14	20	16	20	22	33	25	23	29	21.3	6.7
14	18	30	25	20	26	26	26	22	14	31	23.8	5.3
15	22	20	28	25	17	23	19	23	30	15	22.2	4.7
16	19	29	25	19	26	25	27	20	25	19	23.4	3.8
17	20	28	13	20	15	25	28	24	27	27	22.7	5.5
18	28	30	21	26	22	27	27	35	22	33	27.1	4.7
19	22	16	21	16	28	33	19	22	27	25	22.9	5.4
20	24	21	18	24	33	26	38	27	33	21	26.5	6.3
21	33	20	29	21	15	13	21	22	29	28	23.1	6.5
Totals	**454**	**501**	**454**	**457**	**487**	**480**	**511**	**493**	**512**	**495**	**484.4**	**22.5**

Table III Simulated false alarms (per 10,000 campaigns of 21 periods) for H = 2 and K = 1.1
10 determinations using different random number seeds

generator has been checked and found to be reliable but, even so, the data generated show pronounced dependence on the random number seed used to initiate the sequence of random numbers which in their turn control the random ITMUFs used. Table III illustrates the variety of results obtained if the parameters H and K are fixed, and the false alarms are generated ten times using a different random number seed each time. Note that the first determination is that already shown in Table II.

Examination of the Page's Test Algorithm

The Page's Test algorithm is used to examine a single stream of data, in this instance a stream of random ITMUFs represented by Y_j. Only when a large number of streams have been examined can a clear picture of the behaviour of the test statistic appear. Since the ITMUFs Y_j are generated according to a known rule (i.e. Gaussian, mean 0, and variance 1) it is possible, in principle, to calculate how the test statistic will behave from period to

period. If this can be done, in practice, it will be possible to calculate the likelihood of an alarm at any chosen period and, by summation, over any chosen interval.

Development of Approach

When a campaign is about to begin, T_0 is set at zero. After one period, the most likely value (for $K = 0$) of T_1 is still zero but T_1 can assume any value in the range + infinity to - infinity. The probability that T takes a value in the interval X to X + dX is given by the normal density function:

$$P(X) = K \times \frac{\exp(-X^2/2)}{(2 \times 3.14159)^{1/2}} dX \quad (1)$$

where K is a constant.

The probability that T will fall within certain limits can be calculated using the integral of equation (1) which, for convenience can be referred to as Integral 1. Then the probability, T_a, that T will exceed the alarm level, H, is the value of Integral 1 between H and + infinity. The probability, T_n, that T will be negative, and therefore reset to zero before the next period, is the value of Integral 1 between - infinity and zero. The probability, T_c, that T will continue into the next period is the value of Integral 1 between zero and H. T_a, T_n and T_c can be easily calculated because the starting value of T, is precisely defined as zero, with probability unity. Behaviour of T in the first period is illustrated in Figure 2.

The behaviour of T in the second period is more difficult to calculate because, whilst the value of T_c is accurately known, it is distributed over the region zero to H. The convolution of two complete normal distributions with the same variance simply gives another normal distribution with twice that variance, but to obtain the convolution of a truncated normal distribution with a complete normal distribution is not so straightforward. To overcome this difficulty, a method was developed in order to allow the value and position of T_c to be represented by a vector.

Suppose T_c is divided into a number of slices and then each slice (the area of which represents the probability of the test statistic residing within the bounds of the slice) is replaced by a probability of the test statistic residing at each bound. Figure 2 shows T_c divided into five slices, SLICE(1) through SLICE(5). Figure 2 shows the vector, V(0) through V(5) which is to be used to represent T_c. SLICE(1) contributes to V(0) and V(1) or, in general, SLICE(n) contributes to V(n-1) and V(n) and so on. Finally the value of T_n is added to V(0) in order to parallel the resetting of the negative test statistic to zero. In principle, the test statistic going into period 2 can be represented by the vector V(0)

through V(5). For the purposes of illustration in Figure 2, a vector of length 6 was used. In practice the length of vector used will affect the quality of the results; this aspect has been the subject of particular study and is discussed later. Furthermore, the way in which each slice is apportioned between the elements of the vector is important. This latter aspect involved the optimisation of the apportionment function and is not described here in detail.

Calculation of the behaviour of the test statistic in period 2 and successive periods will be analogous to that in period 1 except that the starting position will be a vector rather than a single value, T_0 (unity). It follows that a set of slices will be generated from each element of the starting vector and each slice must be apportioned into the appropriate elements of the vector which is to go forward to the next period. Since each element contributes to each slice, it follows that the size of the operation from the computational aspects grows with the square of the number of slices. On the other hand, increasing the number of slices used will increase the quality of the method. Not only did the method of the approach need to be validated but a sensible slice width needed to be determined in order to obtain adequate accuracy consistent with reasonable computing time.

Validation of the Approach

In order to demonstrate its validity, the new approach was used to obtain the convolution of four normal distributions. The latter should also be a normal distribution but with four times the variance of the former. Using the methodology described earlier, a normal distribution was represented by a vector of 181 elements extending over the range + 9 to - 9 sigma in steps of 0.1 sigma. Every tenth element of this vector between + 7 and - 7 sigma is shown in the second column of Table IV. The corresponding vector for the convolution of four normal distributions is shown in the third column. The fourth column indicates the difference between the two vectors. Because the errors incurred are several orders of magnitude smaller than the values of the elements of the vectors, the soundness of the approach is confirmed.

Application of the Approach to Calculation of H K Contours

The methodology was developed further to allow the parameter K to be accommodated. Subsequently a data base of false alarm probabilities was generated for:

- campaigns of 1 to 100 periods
- H in the range 0 to 4 (0.1 steps)
- K from 0 to a maximum of 4 (0.1 steps).

It is now possible to specify values for H, K and the campaign period and to obtain, almost instantaneously, values for the period and cumulative false alarm probabilities by interrogation of the data base. By specifying a range of values for H and K and then interpolating the results, a list of pairs of H and K values corresponding to a chosen FAP can be prepared. This list of values is the basis of the H K Countour and can be prepared, using the data base, in a matter of seconds. The H K Contours for a campaign length of 21 periods and FAPs of 5%, 1% and 0.1% are given in Figure 3. The particular benefit of this new approach, other than the improvement in accuracy, is that, once the investment of setting up the data base has been made, an H K Contour for any FAP and any campaign legnth (up to 100 periods at present) can be generated in seconds.

During the generation of the data base, it was observed that, for each pair of H and K, a period, n, was reached after which

$$FAP_{n+j} = \text{constant} \times FAP_{n+j-1} \quad (j > 0) \quad (2)$$

It was also observed that the vectors defining the test statistic, after period n, were a scalar multiple of the corresponding vector for the previous period. The value of n was found, by inspection, to be approximately ten times the value of H, and to show some dependence on the value of K. The computer program checked for the period n on a case by case basis.

These observations allow a particularly useful innovation to be made in generating the data base. Once period n has been reached, subsequent alarms can be calculated using equation (2) with a reduction in computing time of several orders of magnitude. Furthermore false alarm rates for any period and FAPs for any campaign length can be readily calculated.

Comparison of Results with Simulation Method

Using the new methodology, and accessing the data base, Table V was generated for comparison with Table II prepared using the simulation approach.

Examination of Table V shows that, for a given value of K, the number of alarms per period increases as the campaign progresses from start-up, passes through a maximum rate, and then declines. For convenience of comparison with Table II, the data in Table V has been displayed to at most three significant figures. The data is held to a very much higher level of precision.
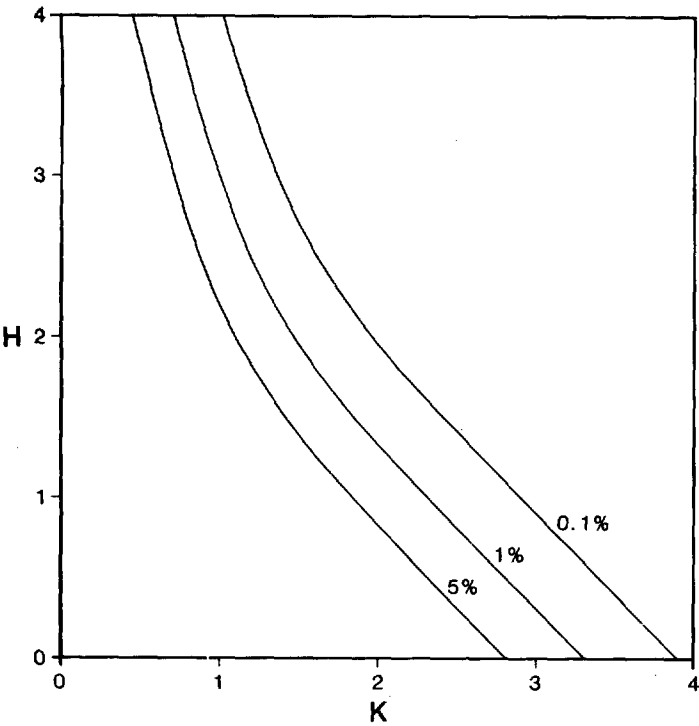
Figure 3 - H K Contours for 5%, 1% and 0.1% FAP

Sigma	P_1	P_4	$P_1 - P_4$
—7	9.50551E-13	9.50653E-13	—1.02455E-16
—6	6.25494E-10	6.25528E-10	—3.47997E-14
—5	1.51665E-07	1.51669E-07	—3.73132E-12
—4	1.35509E-05	1.35510E-05	—1.13782E-10
—3	4.46143E-04	4.46144E-04	—6.87900E-10
—2	5.41259E-03	5.41259E-03	1.26000E-09
—1	2.41971E-02	2.41971E-02	2.63000E-09
0	3.98610E-02	3.98610E-02	—6.20000E-09
+1	2.41971E-02	2.41971E-02	2.63000E-09
+2	5.41259E-02	5.41259E-03	1.26000E-09
+3	4.46143E-03	4.46144E-04	—6.87900E-10
+4	1.35509E-05	1.35510E-05	—1.13782E-10
+5	1.51665E-07	1.51669E-07	—3.73132E-12
+6	6.25494E-10	6.25528E-10	—3.47997E-14
+7	9.50551E-13	9.50653E-13	—1.02455E-16

Legend: For each value of sigma:

P_1 = Populations residing at discrete values of sigma to represent a standard (N(0,1)) distribution.

P_4 = Populations residing at discrete values of sigma for the distribution resulting from the convolution of four normal distributions using the method developed.

Table IV   Convolution of four normal distribution

Period	'K'											
	0.9	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0
1	18.7	13.5	9.7	6.9	4.8	3.37	2.33	1.59	1.08	0.72	0.48	0.32
2	44.5	30.4	20.5	13.7	9.1	5.93	3.85	2.48	1.59	1.01	0.64	0.40
3	54.2	36.3	24.0	15.7	10.2	6.52	4.16	2.63	1.66	1.04	0.65	0.41
4	57.0	37.9	24.9	16.2	10.4	6.63	4.21	2.66	1.67	1.05	0.66	0.41
5	57.5	38.2	25.1	16.2	10.4	6.65	4.21	2.66	1.67	1.05	0.66	0.41
6	57.5	38.2	25.1	16.2	10.4	6.65	4.21	2.66	1.67	1.05	0.66	0.41
7	57.2	38.1	25.0	16.2	10.4	6.64	4.21	2.66	1.67	1.05	0.66	0.41
8	56.9	38.0	25.0	16.2	10.4	6.64	4.21	2.66	1.67	1.05	0.66	0.41
9	56.5	37.8	24.9	16.2	10.4	6.63	4.21	2.66	1.67	1.05	0.66	0.41
10	56.2	37.7	24.8	16.1	10.4	6.63	4.21	2.66	1.67	1.05	0.66	0.41
11	55.9	37.5	24.8	16.1	10.4	6.63	4.20	2.66	1.67	1.05	0.66	0.41
12	55.6	37.4	24.7	16.1	10.4	6.62	4.20	2.65	1.67	1.05	0.66	0.41
13	55.2	37.2	24.6	16.1	10.4	6.62	4.20	2.65	1.67	1.05	0.66	0.41
14	54.9	37.1	24.6	16.0	10.3	6.61	4.20	2.65	1.67	1.05	0.66	0.41
15	54.6	37.0	24.5	16.0	10.3	6.61	4.20	2.65	1.67	1.05	0.66	0.41
16	54.3	36.8	24.5	16.0	10.3	6.60	4.20	2.65	1.67	1.05	0.66	0.41
17	53.9	36.7	24.4	16.0	10.3	6.60	4.19	2.65	1.67	1.05	0.66	0.41
18	53.6	36.5	24.3	15.9	10.3	6.59	4.19	2.65	1.67	1.05	0.66	0.41
19	53.3	36.4	24.3	15.9	10.3	6.59	4.19	2.65	1.67	1.05	0.66	0.41
20	53.0	36.2	24.2	15.9	10.3	6.59	4.19	2.65	1.67	1.05	0.66	0.41
21	52.7	36.1	24.1	15.9	10.3	6.58	4.19	2.65	1.67	1.05	0.66	0.41
Total	1113	751	498	326	211	135	85.9	54.5	34.4	21.6	13.6	8.49

Table V   Calculated false alarms (per 10,000 campaigns of 21 periods) for H = 2 and various K

Summary and Conclusions

The new methodology described in this paper enables false alarm probabilities to be quickly and accurately calculated for chosen values of the Page's Test parameters, H and K, and for any campaign length. Conversely, if the campaign length and false alarm probability are specified, then the appropriate H K Contour can be quickly and accurately calculated.

Reference

/1/ A.J. WOODS, D.J. PIKE — Use of the Standardised ITMUF for Choosing Parameters of Sequential Tests for Protracted and Abrupt Diversions — Proc. of the Fifth Annual ESARDA Symp. on Safeguards and Nuclear Material Management, 1983, pp. 369-374

Acknowledgements

Calorimetry for Safeguards and Nuclear Material Management

R. Carchon,
SCK/ÇEN, Mol

S. Guardini,
JRC-Ispra

Introduction

Calorimetry is the technique for measuring the heat produced by some substances. It is a well-established and very accurate technique and the method has been applied to a wide variety of heat producing processes, such as chemical reactions and radioactivity /1,2/. Its interest in safeguards is steadily increasing. Plutonium, which develops heat as a consequence of α and β decay, can thus be measured in this way.

The calorimeter can be used to determine or verify the Pu content of a sample, on the condition that its isotopic composition is known. It is calculated following the formula:

$$M = W / P_{eff} \tag{1}$$

where M is the Pu mass, W the power output of the sample and P_{eff} the effective specific power: the latter can be obtained by calculation of the contribution of the different isotopes, according to the formula:

$$P_{eff} = \Sigma_i R_i P_i \tag{2}$$

It is clear that all contributions to the heat production should be taken into account.

The specific heat is increasing with burnup, as shown in Table I, where different burnup values for light water reactors are considered with their isotopic compositions /3/. The specific power values are taken from /4/. The ^{241}Am contribution, which is, besides ^{238}Pu the most important
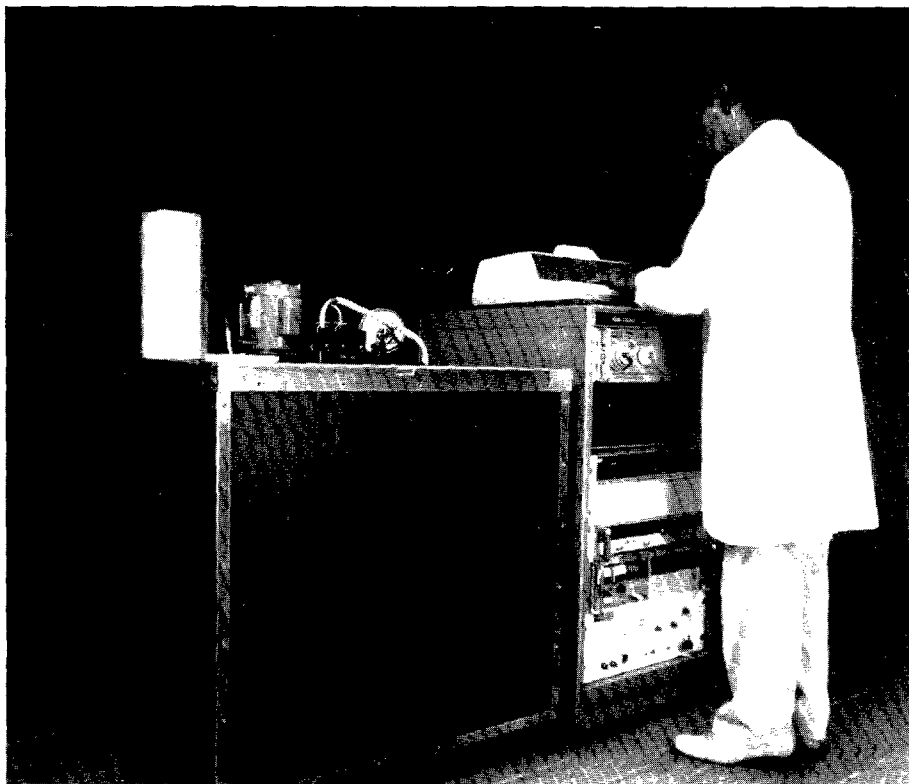

Fig. 1. General view of the calorimeter

contributor to the heat production, is not taken into account; if any Americium is present, the effect of increase in the specific power is still more pronounced. The low specific powers for the U-isotopes (5.72x10⁻⁵mW/g for ^{235}U and 8.45x10⁻⁶ mW/g for ^{238}U) make it clear that U is difficult to measure in reasonable amounts and that in MOX fuel only the Pu is contributing to the calorimetric meas-

Fuel number	Burnup 1000 MWd/t	^{238}Pu (wt %)	^{239}Pu (wt %)	^{240}Pu (wt %)	^{241}Pu (wt %)	^{242}Pu (wt %)	^{241}Am (wt %)	Specific Power (mW/g) without Am-contribution
I	Low	0.01	93.	6.4	0.55	0.04		2.3239
II	8 – 10	0.06	85.68	12.11	1.94	0.22	0.16	2.9189
III	16 – 18	0.24	75.65	18.42	4.59	1.11	0.74	4.2850
IV	25 – 27	1.451	58.945	24.822	10.549	4.233	2.53	11.4910
V	38 – 40	2.0	45.	27.	15.	11.		14.6490
Spec. Power (mW/g) per Pu isotope (*)		567.16 ± 0.57 (0.10 %)	1.9293 ± 0.0053 (0.27 %)	7.098 ± 0.015 (0.2 %)	3.390 ± 0.002 (0.06 %)	0.1146	114.23 ± 0.16 (0.14 %)	

(*) Updated values are expected (ANSI - N 15 - 22 - 1983) for the specific power values and uncertainties, which slightly affect the values but do not change the conclusions.

Table 1 Pu-isotopic compositions for different burnups

urement.

It becomes clear that the specific power can characterize a certain plutonium batch, but that it only becomes sensitive for burnup values above 20,000 MWd/t.

We have to repeat that in the case considered, Americium is not taken into account, that would have caused a more pronounced effect.

This paper will deal with one type of calorimeter only, the so-called isothermal calorimeter, which is most frequently used in safeguards. Details about other calorimeter types can be found in the literature /1,2/. The experiment aimed at the testing of the applicability of calorimetry for the measurement of plutonium. A set of high burnup (25,000 MWd/t) samples, ranging between 60 g and 2.5 kg Pu, were measured by the Mound # 150 calorimeter (Fig.

1). Details are given elsewhere /5,6/. Therefore, some claim will be laid here on the calorimeter accuracy considerations and on the general application of calorimetry in safeguards.

From (1) one can see that for the determination of the Pu-mass of an unknown sample, it is necessary to know the power output W of the sample and its specific power P_{eff}.

In order to be able to determine the sample power output W from the calorimeter response, the calorimeter has to be calibrated against a well known heat source; electrical heating by calibrated thermal resistances can be used or well characterized heat sources such as ^{238}Pu. Details can be found in the literature /4,5/.

. According to (2), the accurate knowledge of the Pu-isotopic composition is required if

the "computational method" is applied.

Thus P_{eff} is derived from the evaluated isotopic abundances R_i and the specific powers P_i of the different isotopes.

In the "empirical method" Pu samples of the batch to be monitored are taken, the Pu content determined and the P_{eff} derived from the previous data /4/.

Analysis of Calorimetry Uncertainty

An analysis of the uncertainty connected with Pu calorimetric measurements is presented in this section and the results are given in Fig. 2 and Table III.

The analytical expression used for computing the relative standard deviation RSD (here assumed as uncertainty) of P_{eff} is taken from /4/ :

$$\epsilon(P_{eff}) = \{ \Sigma_i (R_i P_i / P_{eff})^2 [\epsilon^2 (R_i) +$$
$$+ \epsilon^2(P_i)] \}^{1/2} \qquad (3)$$

where: $\epsilon(P_{eff})$ is the overall RSD on P_{eff} in percent; $\epsilon(R_i)$ is the RSD on R_i in percent; R_i is the isotopic abundance of isotope i; $\epsilon(P_i)$ is the RSD on P_i in percent; P_i is the specific power (W/g) of isotope i.

Correlations are not taken into account here; modifications arise if one would take care of them /8/, resulting in a reduction of the errors, without affecting, however, the global idea, expressed in this paper.

The above expression (3) has been applied to three different isotopic compositions (fuel II, III, IV from Table I) with the error pattern as given in Table II.

In Table III, a summary of the calculations is given.

In the case of fuel II (8-10,000 MWd/t) with an $\epsilon(^{239}Pu)$ settled at .3%, the uncertainty on ^{239}Pu itself is already predominant, and becomes comparable with $\epsilon(^{238}Pu)$ set at 2%.

To reduce strongly $\epsilon(^{239}Pu)$ below .2% is useless if the uncertainty in the knowledge of its specific power is assumed to be .27% as indicated in /4/. It appears clearly that it is an achievable target to keep the P_{eff} uncertainty lower than .5%, if the assumptions of a negligible influence from the power measurement, the heat distribution, and the calorimeter calibration are achieved, and if the P_{eff} uncertainty may be assumed to represent the overall calorimetry uncertainty; otherwise other error sources must be identified and quantified.

In the same table the P_{eff} uncertainty for fuel III is given as well, ranging from 0.35 to 0.64%. $\epsilon(^{238}Pu)$ is now the dominant error source, in competition with the $\epsilon(^{241}Am)$ that can influence the overall error in certain error patterns.

Concerning fuel IV, the $\epsilon(^{238}Pu)$ is definitely the only important error source in any error pattern, even with that high ^{241}Am content. The uncertainty achieved with this

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
$\epsilon(^{238}Pu)$	2.0 %	1.0 %	1.0 %	1.0 %
$\epsilon(^{239}Pu)$	0.3 %	0.3 %	0.3 %	0.1 %
$\epsilon(^{240}Pu)$	0.5 %	0.5 %	0.5 %	0.5 %
$\epsilon(^{241}Am)$	2.0 %	2.0 %	1.0 %	2.0 %

Table 2   Error patterns $\epsilon(R_i)$

Fuel number	burnup (1000 MWd/t)	$\epsilon(P'_{eff}) = \{ \Sigma (R_i P_i / P_{eff})^2 [\epsilon^2(R_i) + \epsilon^2(P_i)] \}^{1/2}$			
		Error Pattern 1	Error Pattern 2	Error Pattern 3	Error Pattern 4
II	8 – 10	0.36	0.31	0.29	0.27
III	16 – 18	0.64	0.45	0.35	
IV	25 – 27	1.21	0.70	0.61	

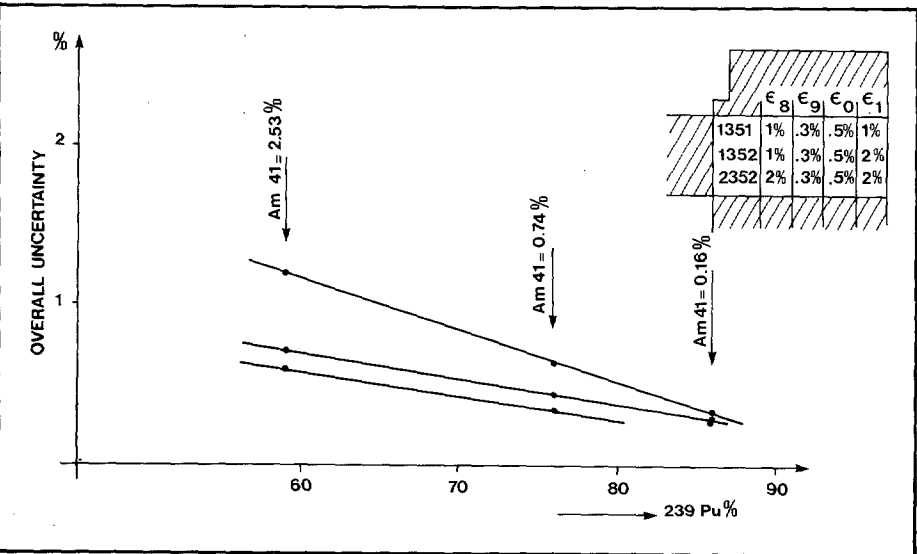Table 3   Overall error on the P_{eff} (light water reactor fuel)

Fig. 2. Overall P_{eff} uncertainty evaluated at different burnups

burnup is around 1%, even with optimistic error patterns and considering other error sources as being negligible.

Application of Calorimetry in Safeguards

Calorimetry is applied in safeguards mainly for the measurement of plutonium. The technique gives the total quantity of plutonium in unknown samples provided that their isotopic composition, or more generally P_{eff}, is known.

Before discussing the field of application of calorimetry in safeguards, it is worthwhile to remember some characteristics of the technique, which can determine the modes of application.

Calibration and biases

Calorimetry does not present major risks of bias coming from the physico-geometrical packaging of the samples; in fact, a different packaging of the same Pu quantity may only influence the equilibrium time and not the result. The technique is, moreover, practically insentitive to other parameters that introduce biases in other measurement systems such as sample weight, neutron self-absorption, neutron multiplication, dead-time, overlapping fissions, etc. None of them influences calorimetry in broad utilization ranges.

Otherwise a "sensitive" bias may be introduced by calorimeter calibration. An incorrect electrical calibration or a bias in the calibration standard (e.g. ^{238}Pu) may, in fact, introduce a systematic uncertainty in the results. But again this source of bias is rather controllable in an easy way, because the "standards" are quite simple and refer to the calorimeter, rather than to the measured sample family, like in almost all other NDA techniques, which necessitate calibration curves or calibration standards to which normalize the results.

Accuracy

It has been said in the previous section that the technique has high precision and high accuracy. In fact, as the heat measurement is a very precise technique, the overall accuracy is normally determined by the uncertainty on P_{eff} and as a consequence by the errors on the isotopic composition.

As the uncertainty on the isotopic abundances varies considerably with the isotope, the overall uncertainty of the result is heavily depending on isotopic composition, as it was shown in the previous section. Namely, at high burnup the ^{238}Pu uncertainty is almost predominant in determining the overall error. But the situation of the neutron coincidence counting (NCC) technique is, as far as this aspect is concerned, almost the same: in fact, the NCC determines Pu content through the measurement of spon-

taneous fission neutrons from even Pu nuclei and principally from ^{240}Pu which is determined with high uncertainty, as compared with odd nuclei, at least if γ spectrometry is used for its determination.

Measurement control and procedure

The state of the art of calorimetry applied to safeguards as far as control of measurement, measurement procedures, etc. are concerned, may be summarized in saying that very accurate and reliable instruments exist for different sample types (bulk, rods, small samples). The measuring procedures are also quite well developed so that they give reliable performances. The technique is essentially easy. It may be stated in a phase of transition from laboratory stage to field of inspection.

Measuring time

The measuring time for bulk calorimeters may range from 1 hour to 6-8 hours per sample; for long (pin) calorimeters the measuring time is of 1-2 hours /7/ while for the small sample calorimeter measuring times of the order of 1 hour are achievable. The applications of suitable equilibrium prevision techniques has the tendency of decreasing these measuring times (but not substantially in the next future). By taking certain precautions, however, these times can be reduced considerably.

It is often stated that the measuring times of calorimetry are very long as compared with other NDA techniques. This is only partially true:

a) other NDA techniques too have long measuring times (hours): e.g. accurate enrichment measurement through γ-spectrometry or NCC measurements with determination of higher order multiplets;

b) as it has already been pointed out, one of the characteristics of the calorimetry (to be put in more evidence) is that the technique is free from any standards specific to the population to be measured (the encapsulation and the mechanical characteristics does not influence the result). This characteristic, rather than the level of uncertainty, assimilates calorimetry to DA. For this reason calorimetry should in certain cases be compared rather with DA.

This fact has the consequence that the in situ inspection activity may become more independent from DA laboratories if one identifies inspection strategies (which is not obviously the scope of this paper) that make more use of calorimetry.

Preparation of standards

Calorimetry may be used for the preparation of accurate plant specific standards,

complementing in an essential way the mass spectrometry in characterizing the items which are not dismantled. The preparation of secondary (working) standards from primary ones as well may be carried out using calorimetry and substituting essentially other techniques which are either precise or more time-consuming ones.

Calibration of instruments

In a plant where many plutonium monitoring instruments are operating the presence of a calorimeter which can be used to calibrate those instruments may be very useful. In fact, provided that the isotopic composition of a sample is known (e.g. with precise γ-spectrometry measurements), the response of another neutron counting instrument (of Shift Register type for instance) can be normalized in terms of Pu total content, to the calorimeter results.

Reference point

A general aspect, that includes also the two points mentioned before, refers to the importance of having in a plutonium handling facility or laboratory always a calorimeter (or more of different applications) installed for reference purposes even not directly connected with inspection practices.

Direct inspection or inventory activities

Very important applications of calorimetry can be found in the field of direct inspection activities, when the required detection time is not limited to minutes.

Actually almost no other known technique might be more reliable and successful than calorimetry when applied to bulk quantities.

MOX pins may be measured with calorimetry, when high precision absolutewise measurements are required. For consistency measurements of pin batches or when fast inventory results are required, NCC techniques are better employed because pin measurements are less influenced by perturbation of the neutron signal.

Small samples as well are better (more quickly) measured with NCC or active interrogation techniques, because generally neutron multiplication or self shielding or dead time are well controlled parameters in such samples.

Fuel elements, wastes and discards might possibly be measured by calorimetry, but the cost-benefit ratio is probably too high to be reasonable.

Conclusions

Calorimetry has been discussed as a Pu-assay instrument. For that purpose, the heat emitted by the sample and its isotopic composition have to be known.

The contribution of every Pu-isotope to

the uncertainty on the effective specific power has been studied and turned out to be seriously dependent on the burnup of the Pu-sample under study.

Possible applications of calorimetry on laboratory scale and in-field conditions have been presented.

References

/1/ S.R. GUNN - Radiometric Calorimetry: A Review - Nucl. Instruments and Methods, Vol. 29 (1964) 1 and Vol. 85 (1970) 285

/2/ J.A. MASON - The Use of Calorimetry for Plutonium Assay. SRDP-R100 (1982)

/3/ M.E. ANDERSON, J.F. LEMMING - Selected Measurement Data for Plutonium and Uranium. MLM-3009 (1982)

/4/ Calibration Techniques for the Calorimetric Assay of Plutonium-Bearing Solids Applied to Nuclear Materials Control. ANSI N15-22 (1975)

/5/ C. BEETS, R. CARCHON, P. FETTWEIS, M. CORBELLINI, D. D'ADAMO, S. GUARDINI, W.W. RODENBURG, W.W. STROHM, S. FIARMAN, A. KEDDAR - Test of Calorimetry for High Burnup Plutonium. BLG 559 - EUR 9181.EN (1984)

/6/ C. BEETS, R. CARCHON. P. FETTWEIS, M. CORBELLINI, D. D'ADAMO, S. GUARDINI, W.W. RODENBURG, W.W. STROHM, S. FAIRMAN, A. KEDDAR - Pu-Content Verification by Calorimetry - Proc. of the 6th Annual Symp. on Safeguards and Nuclear Material Management, Venice 1984

/7/ R.B. PERRY, A. KEDDAR - Field Test of the One-Meter Fuel Rod Calorimeter. ANL-NDA-10 (1982)

/8/ T.E. SAMPSON, R. GUNNINK - The Propagation of Errors in the Measurement of Plutonium Isotpic Composition by Gamma-Ray Spectroscopy. LA-UR 83-1520 (1983).

ESARDA / INMM Joint Specialist Meeting on NDA Statistical Problems

Some of the participants at the meeting

A specialist meeting on the statistical problems encountered in the evaluation of NDA measurements was held at Ispra on 12th - 14th September 1984. The meeting was jointly sponsored by ESARDA and the INMM and was coordinated by M. Franklin (Ispra) and J. Markin (Los Alamos). Thirty specialists in NDA and/or statistics representing the European Community countries, the Commission of the EC, the U.S.A., Israel and the IAEA participated in the meeting.

Fourteen papers were presented in a workshop format which was designed to allow ample time for discussion. The papers illustrated the growing trend in NDA to look for instrument error modelling which can provide an integrated rationale for calibration, measurement control programs and operator inspector comparisons.

The proceedings will be distributed to participants and authors at the end of 1984. A limited number of copies are available for interested specialists on request.

Preview for Liège, Belgium

7th ESARDA Symposium

R. Carchon
CEN/SCK Mol
L. Stanchi
JRC-Ispra

The city of Liège was founded in 705 after the death of St. Lambert. A chapel was built which soon became a place of pilgrimage. In the 8th century St. Hubert moved his residence as bishop from Maastricht to Liège. This city has thus a clearly religious origin and character.

At the end of the 10th century Bishop Notger created an independent state ruled by a prince-bishop. This state experienced repeated battles, sacks and destructions, but the churches with their artistic treasures were always preserved. Its independence came to an end after the French revolution and Liège became part of the modern Belgian state in 1830.

Liège lies at the confluence of the rivers Meuse and Ourthe surrounded by hills with many suburbs and is a good example of an industrial and economic centre. It is the third town of Belgium in terms of populations.

The town of Liège is first of all a town where it is pleasant to live, a town on the human scale. Liège's keynotes are "charm and dynamism".

There is always some picturesque corner to discover during strolls through the small streets, by stairs or court yards in quiet places near the centre.

Liège is a town where museums, Romanesque and Gothic churches and theaters are well distributed, and where the cultural life is intense. The artistic development of the towns around the Meuse started in the middle age. Churches and museums contain splendid art objects: ivory carvings, gold and silver. During the renaissance several painters and sculptors contributed to render Liège famous; among them Lambert Lombard and Jean Delcour. Celebrated composers and writers were also born in Liège. We may mention Cesar Franck for the former and Georges Simenon (the father of Inspector Maigret) for the latter.

The economic development of Liège started in the 10th century with the discovery of coal in the vicinity. This gave rise to the development of forges and the manufacture of weapons. In the 19th century, Liège experienced enormous economic development helped by the Meuse which was a very useful transport route. The two world wars of the 20th century had severe effects on the town and its industrial life. After world war II there has been a new industrial development encouraged by the creation of a new artificial transport route, the Albert canal connecting the rivers Meuse and Schelde.

Last but not least, we wish to mention the glass industry with the famous Val St. Lambert crystal.

The ESARDA Symposium will be held in the Congress Hall which is situated on an island in the river Meuse. It is one of the best equipped congress centres in Europe (see top figure).

P.S. Please note in the following call for papers that the Belgian secretary of the symposium is now Mr. R. Carchon, replacing Mr. Franssen who has moved to the IAEA.

We wish also to point out that being next year the year of the IAEA Symposium, as is customary, the annual meeting of ESARDA will not be a general symposium but a specialist meeting like that in Petten in 1982. This meeting will be held in Copenhagen. The topic of the meeting will be communicated later.