JRC 85571

ISSN 1977-5296

ESARDA Bulletin

A Non Original Contribution to the Sampling Problems in Safeguards

A. Rota
CEC, Safeguards Directorate, Luxembourg

Introduction - Purpose of the Paper

1. In recent years some progress has been made in the field of statistical sampling techniques for use in nuclear material verification. Historically (and in very rough and broad terms) we can list three stages for the solution given to this basic safeguards problem:

1. The "attribute" and "variable" sampling approach mentioned in the examples contained in the Subsidiary Arrangements to the NNWS-VA.
2. The introduction of the idea of "variable sampling in attribute mode" and the approach based on attribute sampling, variable in attribute mode sampling and variable sampling.
3. The "new" consideration of the value of attribute sampling involving a critical analysis of the approach used in stage 2 with attention payed to :
   - the quantitative results obtained by the measurements carried out with the attribute tester and
   - the consideration of the false alarms implications.

Synopsis of the symbols in this paper and in the references 2,3, and 4

		This paper	2	3	4
Goal quantity		G	G	M	M
Amount per item		A	A	$\bar{x}$	$\bar{x}$
Attribute tester standard	absolute	δ A	σ	$\delta\bar{x}$	$\delta\bar{x}$
	relative	δ	$\overline{\sigma}$	δ	δ
Size of partial defects	absolute	s δ A	s δ A	$\delta\bar{x}$	$\delta\bar{x}$
	relative	s	s	δ/δ	δ/δ
Critical value for the attribute tester (rel.)		t	q	k/δ	t/δ
Attribute sample size		n_1	n_1	n_a	n_a
Var. sample size		n_2	n_2	n_u	n_u

(Note that when symbols have the same meaning in all papers or when they are used only by one author they are not reported in this synopsis)

2. At present the stage 2 approach is used in practice. Attribute sampling is used to detect "gross defects", variable sampling in attribute mode is used to detect "medium defects" and variable sampling is used for bias defects (and D statistic).

3. IAEA Manual F /1/ suggests using the critical value for measurements of possible defects in attribute mode for the attribute tester at a level of 4σ, σ being the error standard deviation of the attribute tester used. This allows us to avoid the consideration of the false statement risk (see further discussion in section 17).

4. This paper intends to describe and discuss, as simply and informally as possible, the third stage mentioned above. We will move from the critical analysis of stage 2. Stage 1 has been mentioned only for historical reasons and will not be considered further.

Description and Critical Analysis of Stage 2

5. The origin of the recent interest in this matter may be found in the critical analysis of the recipes currently used for sampling techniques, as described in /1/ and as (simplistically) implemented in the field.

6. The analysis and the criticism, in both its destructive and constructive parts, are summarized in the following by means of an example, relevant to one Material Balance stratum. In fact (see section 4.4.1.2 of /1/ and II of /2/) it is given as known that the detection of at least one anomaly, linked to the diversion by gross or medium defects of an amount equal to or larger than G, is assured (with a prefixed risk of non detection β_o) if the sample size for every stratum is calculated as if the quantity G were diverted only from that stratum.

7. It may be useful to recall that, in current safeguards jargon, defects (the difference between what should be in an item or batch, according to records, and what actually is - physically - in it) are classified as gross or medium size, whenever they concern all the material of an item or batch, or a part of it, which may be recognized by a single measurement of the instrument used as attribute tester.

8. Let us assume we have to verify a stratum (quite homogeneous) of N items of very similar size, each containing A units of nuclear material. If the detection goal is of G units and if a β risk of non detection is accepted, it is well known that a sample of size n, given by the approximate equation

$$n = N(1-\beta^{A/G}) \qquad (1)$$

must statistically contain at least one defected item.

9. This "at least" means that the sample will contain, with probability 1-β, at least one defective item if the diversion of G has been operated by emptying completely

$$r(A) = G/A$$

items. In such a way the number of defective items is minimized. If the diversion of G is obtained by partially emptying the necessary number of items, this number of defects is higher so that the probability of inclusion of defective items in the sample increases.

10. In the hypothesis that equal amounts x are taken out from a number of items to accumulate the quantity G, the following elementary relationship between the number of defective items r(x) and x exists :

$$r(x) = G/x > r(A) \quad (x < A)$$

11. Equation (1) may be used to calculate any sample size n(x) which contains, with probability 1-β, at least one defective item, if all defects are of size x. n(x) may be

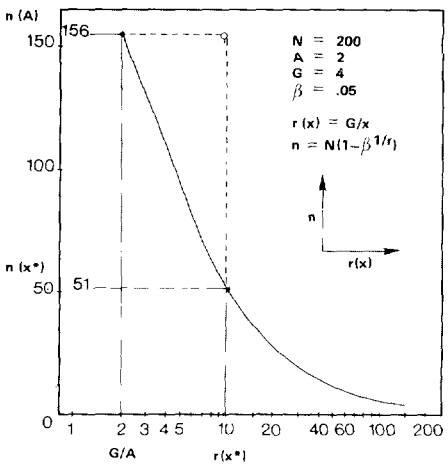

Figure 1 - Attribute sample size n versus r, the number of defected items, for a specific case defined by the parameters N, A, G and β

calculated and plotted for any set of parameters N, A, G and β (provided that NA is greater than G). In general the shape of n versus x is of the type of the plot given in Fig. 1. In view of its practical use, it is important to understand fully the implication of this plot.

12. If an inspector uses an attribute tester capable (but only capable) of assessing with certitude whether or not an observed item is empty, the examination of a sample of size n(A) will allow him to discover an anomaly in the hypothesis, but only in the hypothesis that the diversion occurred by emptying completely the necessary number of items.

13. If the diversion of the amount G has occurred through defects of a specific size x^* and if the inspector examines a sample of size $n(x^*)$ with an instrument capable of 100% assurance ($\alpha = 0$) to detect a defect x^* or larger, then at least one such defective item will be included in the sample $n(x^*)$ with probability $1-\beta$.

14. However,

i) if the diversion of G was obtained through defects of larger size ($x > x^*$) the relevant number of defective items decreases ($r(x) < r(x^*)$) and the risk of non inclusion of at least one defective item in the sample size $n(x^*)$ becomes larger than β;

ii) if the diversion of G was obtained through defects of smaller size ($x < x^*$) the detection probability of at least one item becomes larger than $1-\beta$, provided that the instrument used is capable of recognising such defects. If this is not the case, $1-\beta$ becomes identical to 0.

15. Intuitively one understands immediately that a partial answer to the problem of the larger spectrum of diversion strategies is that of oversampling. Repeating the above reasoning starting from the end, let us see which practical result may be obtained in the hypothesis of the availability of an instrument capable (with $\alpha = 0$) of detecting defects of size x^* or larger. If a sample of n(A) items is examined, it is possible to affirm that in such a sample there is, with probability at least $1-\beta$, at least one defective item, irrespective of the strategy used to divert the amount G, provided that the size of defects is at least x^*.

16. The practical consequence of the above reasoning is found in /1/, sections 4.3.1 and 4.3.2, from where the following guidelines are derived :

i) Define the attribute tester precision and accuracy through its relative standard deviation δ (r.s.d. of the measurement instrument used for the specific stratum).

ii) Define for that stratum, characterized by the parameters N and A, a critical threshold

$$t\delta A \qquad (2)$$

so that, if a discrepancy exceeds this threshold, the relevant item is "labelled as a defect in the attribute tester inspection with probability 1"; if the discrepancy is smaller the anomaly is neglected (if it reflects a real defect it "will be detected with probability zero": $\beta = 1$).

iii) Calculate the sample size $n_a = n(A)$, select it randomly from the population and measure it with the attribute tester.

iv) Calculate the sample $n_v = n(t\delta A)$, select it randomly from the population and measure it with the variable tester.

17. In /1/, section 4.3.2.1, it is suggested to use, for practical implementation, a value of(*)

$$t = 4.$$

This value has the advantage that if an item is observed as anomalous (the discrepancy exceeds the amount $4\delta A$) there is a statistical risk of only 0.003% (practically zero) that the item in question is not defective.

18. However, the sample size calculated for the variable (in attribute mode) tester, $n(G/4\delta A)$, suggests that the attribute tester is capable of detecting defects of size $4\delta A$, which is not strictly true and confuses the size of the defect with the threshold setting on the measurement instrument.

19. The approach outlined in section 16 above, implies that the performance of the variable tester must be of a quality much higher than that of the attribute tester. In fact, it should be possible to detect with probability $1-\beta$ any anomaly smaller than $4\delta A$, which may indicate an overall diversion of G. In practice the results of this second set of verifications shall be usable either:

- for identification of single defective items (when the single defect is within the detection capability of the tester used) or
- for the detection of a diversion covered by bias defects (defects statistically detectable only through the cumulative observation of the whole measured sample, which are out of the detection capability for single items).

20. Usually testers with such performances are those involving weighings and chemical destructive analyses. They are quite costly and time consuming (both during and after the inspections) and therefore should, as far as possible, be compatible with the attainment of the goals /5/ minimized.

The Passage from Stage 2 to Stage 3

21. The practical impossibility of the extensive use of such variable testers, and the feeling that the potential of presently available measurement instruments was not fully exploited in the function of attribute testing, led some researchers to make a thorough analysis of the sampling approach described in /1/ and currently implemented (as far as possible) by the safeguards inspectorates.

22. The principal idea was the incomplete exploitation of the quantitative results of the attribute testers, the detection capability of which was used only for measured discrepancies larger than $4\delta A$ even if smaller discrepancies may be observed and provide fruitful information.

23. When the defect under observation has a size comparable with the sensitivity of the instrument used, the data analysis must obviously take into account :

- a component of the non detection risk (β) which does not depend only on the (attribute) sample size;
- a risk of false alarm (which - see section 17 above - was lowered down to insignificant values by selecting an appropriate critical threshold for the attribute test).

24. Note that a threshold $t\delta A$ for defects (section 16.) corresponds to a threshold of $A(1-t\delta)$ for the tester: above $A(1-t\delta)$ the measure is accepted, below it is rejected. For a graphical interpretation see the abscissa of Fig. 3.

Very generally speaking, the recent developments in the sampling considerations /2,3,4/ exploit the performances of the available attribute testers in their region of answer higher than $A(1-4\delta)$. An attempt to visualise the situation is made in the following with the help of Figs. 2 to 4.

25. Once the performance of an attribute tester is known (δ is known), it is possible to draft Fig. 2 where the false alarm risk, α, is plotted as a function of the threshold T for acceptance or rejection of the observations. Physically $T = A(1-t\delta)$. If both A and T are expressed in terms of absolute standard deviations ($\delta A = \sigma$), α is given by the expression :

$$\alpha = \Phi(-(A-T)/\delta A) = \Phi(-t) \qquad (3)$$

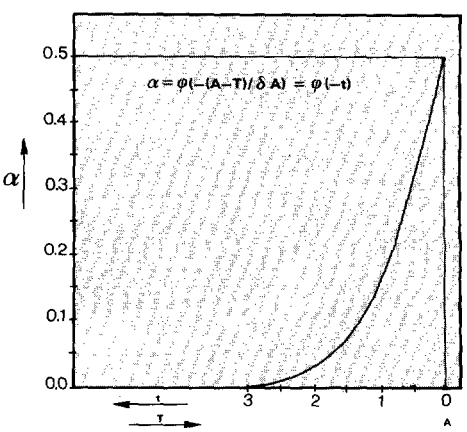

Figure 2 - False alarm rate α versus difect size t (expressed in δ A units) or absolute threshold setting T

(*) For the sake of simplicity 4 is used in the following instead of t. t will be used in the following with the same threshold meaning, but in a more general context.

where Φ is the cumulative normal distribution function.

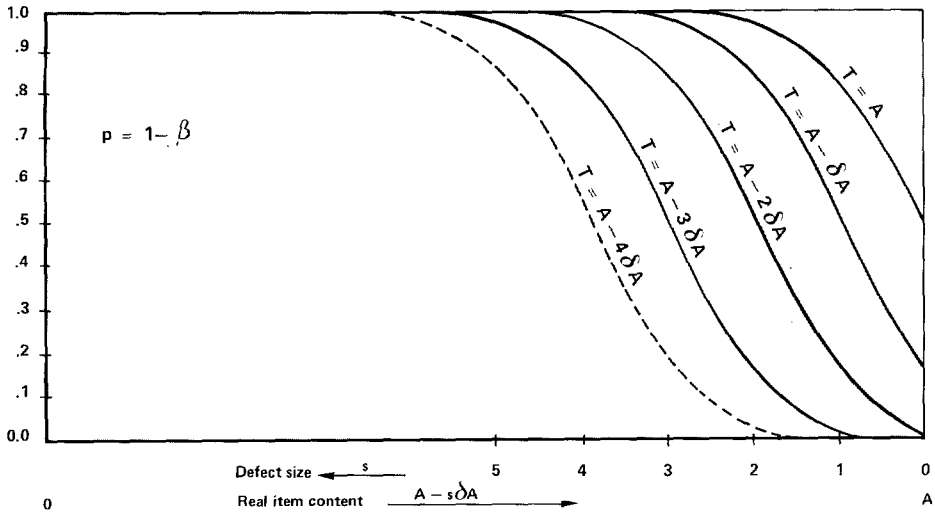

Figure 3 - Detection probability of defects versus real item content for different threshold settings. T is the rejection threshold for defects of real absolute size j δ A (relative size t = j)

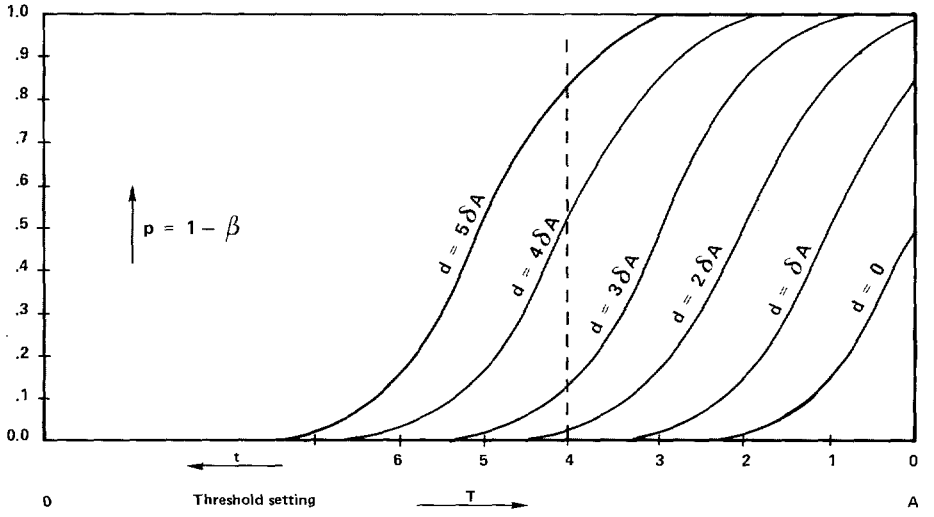

Figure 4 - Detection probability of defects versus threshold setting, expressed in t units or in absolute value T, for different defect sized d = s δ A

26. Note that α (the probability that a non defected item is measured as defective under the threshold T) has obviously no relationship with possible defect sizes.

27. Figure 3 describes the probability p = 1-β to classify as defective an item really defective versus the real item content (on the abscissa) for various choices of the critical T value. If the abscissa is read from A backwards, the size of the concerned defect is directly measurable. For general use the values of T and of the defects are expressed in terms of standard deviation δA of non defected items.

28. The same information may also be plotted (Fig. 4) to describe the probability p = 1-β to classify as defective an item as a function of the threshold setting T for various sizes of the defected items.

29. The plots of Figs. 3 and 4, considered together with those of Fig. 2, illustrate the performance obtainable by the specific tester used (characterized by δ) when a stratum, identified by the item content A, is examined. The dotted lines used in both Figs. 3 and 4 for a t = 4 (equivalent to a threshold setting of A(1-4δ)) illustrate the situation existing when the recommended decision rule, mentioned in section 17 above, is used.

30. It is immediately evident that the "probability area" (on the right hand side of the dotted lines), which illustrates the measurement system's capability to detect defects smaller than those selected by the adopted decision rule, is not fully exploited. It must nevertheless be noted that in such a zone one must be careful because p < 1 and α > 0 values may occur. The simple rule of t = 4 has been used until now in order to "stay on the safe side" and because it was not yet known how to exploit the above mentioned "probability area".

31. In fact, this exploitation involves mathematical and statistical calculations which are not simple, need a lot of approximations and are not easy to be described in simple terms. However, now some results have become available and these will be useful to inspectors. It is demonstrated that an appropriate treatment of the attribute tester data may be done so that, eventually, the same detection probability may be achieved with a sample to be measured "by variable", significantly smaller than that suggested by the rule of t = 4, used up to now.

The New Approach to Sampling (Stage 3)

32. Intuitively the smaller sample sizes may be justified by the following type of reasoning:

- A full exploitation of the data obtainable from the attribute testers allows, for each considered stratum, the statistical detection of defects of a size smaller than 4δA.
- In such a case the variable testers are used only for the detection of items defected for less than 4δA.
- But, in such a hypothesis, the diversion of a fixed amount G implies the falsification of more than G/4δA items. It follows that samples of smaller size are sufficient to assure, with the required probability 1-β, that at least one defected item is included in the sample to be measured by the variable tester. (The characteristics of such a tester have already been identified above, section 19.)

33. The algorithm for the calculation of the variable in attribute mode sample size is obtained from a process which starts from the formulation of the non detection probability of defected items both for level 1 (attribute) and 2 (variable) sampling. Level 1 sample size is calculated by Eq. (1).

34. If :

- β_1 indicates the probability that no anomalous level 1 measurement is observed;
- β_2 indicates the probability that no anomalous level 2 measurement is observed, and
- β_1 and β_2 are independent,

the product : $\beta_1\beta_2 = \beta$ must satisfy the requirement that the overall β (for all strata) is equal to or lower than the desired prescribed amount (e.g. that fixed in the safeguards goals /5/).

35. In turn β_1 depends on :

- the probability that a defective item is included in the sample by attribute (size = n_1) and on
- the conditional probability that the attribute tester provides for an alarm when the defective item is measured (remember that now an α > 0 is not impossible to be met).

And β_2 depends on n_2, the sample size for level 2 measurement that, in such a way is indirectly defined.

36. Two authors derived approximate formula for the calculation of n_2, satisfying the requirement to attain the detection goals quantified by G and β_o. The formal developments proposed are different, as is the final formula obtained. However, for a number of numerical examples, these different formulae lead to very similar numerical results. The two solutions proposed are outlined in the following. They are due respectively to J.B. Sanborn /2/ and to J.L. Jaech /3/. Jaech's formulation is under consideration by the Agency, for inclusion in a new version of the Manual F. the part of this draft that has been made available to the writer has also been taken into consideration /4/.

37. Going a little further in the description of the development of the algorithms for the calculation of n_2, the nature (systematic or random) of the error inherent to the attribute tester has to be taken into account. This is done by both authors, and both deal with the extreme cases of an "all random" or "all systematic". Both authors agree that the "all systematic" case gives the higher sample size, so that the "all systematic" is conservative in respect of an "all random" hypothesis.

38. It is, however, wise to stress that :
- the numerical results obtained using "all random" and "all systematic" hypotheses are very similar;
- even if the result of a "mixed" case is not yet studied, it seems that the "all systematic" case may give conservative values of n_2;
- (last but not least) the "all systematic" case is formally less complex.

39. The basic data are obviously the same for the two authors. (Here the symbolism of Sanborn is used with a few minor exceptions. A synopsis of the symbols used in different references is given in the Appendix.) These basic data are those listed in section 8 above : N, A and G; the overall accepted risk of non detection β_o and the relative standard deviation δ of the attribute tester (section 16(i) above).

40. For both authors the attribute sample size is calculated by the "usual" Eq. (1).

41. In his approach Sanborn suggests remeasuring with the variable tester those items which
- are classified as defectuous by the attribute tester bur for which
- a risk of false alarm practically larger than zero exists.

42. Note that in the Sanborn approach the critical value t is optimised as function of δ, so that items classified as defected for a certain observed amount may not actually be anomalous but only classified as such for a statistical effect (α error).

43. A similar suggestion does not appear in Jaech's approach, where it seems that a pre-established risk of false alarm rate is accepted as input datum of the problem.

Sanborn Results

44. The total number of items expected to be measured by the variable tester is given by

$$n_2 \text{ total} = n_2 + n_f - n_1 n_f / N \qquad (4)$$

where

$$n_2 \geqslant (N\delta A/G)\, f(\beta_o, t, \delta) \qquad (5)$$

n_2 is the sample size calculated on the basis of the conditions mentioned in section 36 above;

$$n_2 = N(1-\beta_o^{A/G})\, \Phi(-t)$$

is the expected number of items to be remeasured after a doubtful result of the attribute tester (see section 41);

$$n_1 n_f / N$$

is a term which takes into account the probability that the two independent samples n_2 and n_f overlap; and

$$f(\beta_o, t, \delta)$$

is a known (tabulated) function of its three parameters. t (see also (2) in section 16.(ii)) may be defined by the expression

$$t = (A-T) / \delta A$$

and is the number of absolute standard deviations of the tester exceeding which the difference between the declared value (A) and the result of the level 1 measurement is considered to be anomalous. (The definition of T is consistent with that given in section 25. The parameter t is called q in /2/.)

45. Tabulations of f may be found in /2/. f is a very slowly increasing function of δ, but this dependence, at least in the domain of our interest δ = 0.01—0.2, is so small that it may easily be neglected and the conservative value of f for δ = 0.2 may be used for practical implementation.

46. If t has not to be fixed by other considerations (so that f is defined), it may be considered as a free parameter to be used for the minimization of the effort quantified by n_2 total (this is the "appropriate treatment" mentioned in section 31. above).

47. It might have been logically anticipated, as it results in practice, that f is a monotone increasing function of t : n_2 grows with t. But when n_2 increases the number n_f of expected follow-up measurements decreases so that a value of t, and then of n_2, which optimizes (minimizes) the n_2 total expression (4) may be established computationally from the parameters A/G, β_o and δ. The calculation shows that, as a general rule, the dependence of optimal values of t on A/G and β_o is small if not negligible. δ remains then the only parameter influencing the optimization of t and Sandorn suggests the following table as a "reasonable guide" to choose t on the basis of δ:

δ	t
0.15	3
0.015 - 0.04	2.5
0.04 - 0.1	2
0.1	1.5

48. The full procedure of verification by sampling is then the following
i) Calculate n_1 by Eq. (1) and randomly select the sample.
ii) On the basis of δ, select the optimum t value using the indicative table of the previous section.
iii) Measure the sample n_1 by the attribute tester with a threshold setting consistent with the optimum value. If one or more defective items are discovered for which the false alarm risk is negligible, stop the procedure and begin the relevant follow-up actions.
iv) Determine the value of f corresponding to the optimum t, the imposed β_o and the conservative δ = 0.2 values.
v) Calculate the expression at the right number of Eq. (5) and approximate the value to the larger integer number.
vi) Measure with the variable tester the sample of size n_2 randomly selected from the N population items.
vii) Measure additionally with the variable tester those items which
   - have been considered doubtful (in the sense that the relevant anomaly has a non zero probability to be a false alarm) after measurement by the attribute tester and
   - are not included already in n_2.

Jaech Results

49. (The notation used until now continues to be adopted, but, in round brackets, the original Jaech notations /3/ is sometimes reported for easy reference to the original text in the next section - see also the Appendix).

50. The following model is used:

It is assumed that all partial defects are of size $sA\delta$ ($j\bar{x}$). There are r_o (n_o) such defects in the stratum: $r_o = G/sA\delta$. The number of such defects that exists among the n_1 (n_a) items measured by the attribute tester is a random variable. The same is true for the number of defects that exist among the n_2 (n_u) items measured by the variable tester. It is assumed that the n_1 and n_2 samples are independently drawn. Finally the assumption is made that a defect of size sA, if measured by the variable tester, is certain to be classified as a defect.

51. The calculation of n_2 is obtained through an algorithm which makes also use of the following parameters :

- α probability that a non defected item is classified as a defect by the attribute tester;
- q probability that a defect of size sA is detected by the attribute tester;
- t critical value for the attribute tester (same definition as in sections 6 and 44 above - called k in /3/).

52. The value of n_2 to be used is obtained by means of a "trial and error" process, which allows the definition of the optimum (maximum) sample size to be used.

53. The process follows the following steps :

i) Choose a value for q

ii) From the prescribed α value (considered as one of the input data of the problem) and the characteristic δ of the attribute tester, calculate the critical value t. t is defined by the relationship

$$\alpha = \Phi(-t) \qquad (6)$$

This function is equivalent to (3) and is graphically represented in Fig. 2.

iii) Once t and δ are fixed, the relationship

$$q = \Phi(s-t) \qquad (7)$$

links the parameters q and t. It is then possible to calculate the s corresponding to the value of q arbitrarily chosen at step (i).

iv) Fixed s, calculate r_o:

$$r_o = G/s\delta A$$

the number of detected items present in the population, according to the model chosen.

v) Calculate n_2 as follows:

$$n_2 = \ln(\beta_o/1-q)/\ln(1-r_o/N)$$

54. One "trial" is now available.

vi) The solution is found by assigning different values to q and by finding, by trial and error, the value which maximizes n_2 and which is then the sample size to be used for measurements with the variable tester.

55. This algorithm represents the approximate solution offered in /3/ for the "all systematic" case (see sections 37 and 38). It seems to the writer that this is the most simple and workable of all the more precise and complex cases treated by this author. This solution is included in /4/, where tables of s as a function of q and α, which may be very helpful for the calculations involved by the algorithm, are mentioned, but unfortunately not yet attached to the presently available draft copy. Actually they are not difficult to prepare when tabulation of the function Φ is available.

56. A superficial reading of /3/ may give the misleading impresssion that the sampling is designed only for defects of a specific size ($s\delta A$). But the measurements made by the attribute tester are in fact such that all diversion strategies aimed to divert an amount of G by means of defects greater than or equal to the minimum specified size, are eventually covered.

57. The problem of the false alarms is, instead, not very clearly discussed. The reading of references /3/ and /4/ suggests that α is an "input" parameter for the sampling exercise. If this is simply the case, Jaech seems to accept the specified α risk so that any item measured as defective (observed value under the critical threshold T) has to be treated as such without inquiring further if this anomaly is real or only apparent.

Further Considerations

58. A number of questions inherent to the practical use of the suggested sampling algorithms exist, first of all the approximation of strata with a group of items ideally with equal content A. This difficulty is rather linked to the calculation of n_1 : this problem has already been addressed and a number of solutions are suggested (substratification, conservative approach, use of mean values, which may be chosen to suit practical cases). The sole requirement for the calculation of n_2 is that the absolute standard deviation ($sA\delta$) of the attribute tester may be considered practically constant for all the items of the considered stratum.

59. The description of the sampling technique described in sections 53 and 54 (Jaech) is not as complete as the description of the technique given in section 48 (Sanborn). However it must be clear that the first case also includes an identical step relevant to the attribute tester measurements (for which the t value corresponding to the maximum n_2 is used).

60. What remained implicit for both sampling techniques is the fact that the suggested procedures must be applied in the same way for all the strata relevant to the specific inspection (e.g. PIV) or to the group of inspections (e.g. for IC verifications) or to a full Material Balance.

61. Because of the basic principle mentioned in section 6, one must be conscious that the discovery of one anomaly in one stratum is not an indication that something is wrong in that stratum only, but an indication that something may be wrong in the whole strata considered together. This is the consequence of the fact that, for sampling calculations, the whole goal quantity G has been used for each stratum: the discovery of one anomaly must involve the beginning of an inquiry covering all the strata.

62. Jaech /4/ describes explicitly the "model" he uses for the development of his technique while Sanborn lets the reader deduce such a piece of information (something which is not always trivial). However, it seems that the Jaech conclusions, according to his hypotheses, are limited to the consideration of defects greater than or equal to sSA; this lower limit is not explicitly deducible from Sanborn's approach, giving the impression that the lower limit he may considers is that of the measurement capability of the variable tester.

63. Another difference between the two techniques is that of the false alarm treatment for the attribute tester. The remeasurement of "doubtful" items (see section 41) allows Sanborn to neglect the problem. Jaech's attitude seems to be that of accepting to run a certain risk because α is considered simply an input parameter for his sample size calculations.

References

/1/ IAEA Safeguards Technical Manual Part F, Vol. 3, Edition 1982 (IAEA-TECDOC-261) Ch. 4: Design of Inspection Plan

/2/ Jonathan B. SANBORN. "Attributes Mode Sampling Schemes for International Material Accountancy Verification". In the INMM proceedings of the Winter 1982 Symposium, p. 34

/3/ John L. JAECH. "Sample Size Determination for the Variables Tester in Attribute Mode". In INMM Summer 1983 Bulletin, p. 21

/4/ IAEA Draft for a new edition of Manual F/111, pages 55 to 71, private communication, 1984

/5/ W. GMELIN, B. MATH, B.W. SHARPE, W. STANNERS, G.V. LANDRESSE and B. LOVE. "Notes on Inspection Goals". 6th ESARDA Annual Symposium 1984 (ESARDA-17), p. 19

Publication of a Book on

Statistics Applied to the Interpretation of Measurement Data

CETAMA
Centre d'Etudes Nucléaires de Fontenay-aux-Roses
Boite postale 6
92265 Fontenay-aux-Roses (Cedex) France

In 1978, the CEA's CETAMA* working group No. 11 ("Statistics") published a work entitled **Statistics Applied to the Interpretation of Measurement Data.** As this book is now out of print, a new and fully revised edition is scheduled for publication at the end of 1985**. This article discusses its contents and the problems it may help to solve.

The book is divided into six parts, and is completed by a section of tables and charts:

I. Definitions
II. Analysis of an Observed Distribution
III. Evaluation of Measurement Methods
IV. Problems of Comparison
V. Relating Two Variables
VI Interlaboratory Circuits

The work is illustrated by 78 numerical examples to provide more tangible applications for the otherwise abstract discussions due to the inevitable use of mathematical symbology. The examples are generally taken from analytical chemistry, a field in which it is especially difficult to determine all of the factors affecting a measurement result for which analysis laboratories must provide the user with a maximum of information at the lowest cost.

I. Definitions

Part I begins with a discussion of random variables, as statistical applications are conceivable only where at least one of the variables studied is random in nature. Statistical interpretation of measurement results thus requires that the experimental findings be assimilable to particular values of a random variable, and that the occurrence frequency of each value (i.e. its probability) be given by a mathematical law governing the probability of random variables.

This discussion is followed by a review of some basic definitions as well as the characteristics of the most common probability laws.

II. Analysis of an Observed Distribution

Part II includes three chapters covering the operations used to determine and characterize usable observations.

II.1 Determining the Distribution Law

The first section of chapter II.1 discusses the ways for verifying that the series of experimental results is actually random, i.e., that the value of any given result is independent of the values of the previous results. This random character is the basis of statistics, but it cannot always be taken for granted and must often be verified. For example, successive spectrophotometric measurements on a solution with a coloration likely to fade in time may produce results that are not random because the values gradually decrease.

In practice, this verification of the randomness of a series of successive observations is not only a preliminary step for many statistical procedures, but also an effective method for directly solving much common problems as the following:

- Is the response of a measurement instrument subject to drift ?
- Has a measurement instrument reached stable operation after startup ?
- Does the impurity content of a metal ingot vary systematically from top to bottom ?
- Does a given factor (e.g. the temperature) affect the result ?
- Does a qualitative factor (e.g. the identity of the operator) affect the result ?

Once it has been determined that the observation data may be considered as particular values of a random variable, the following chapters provide a way of evaluating the mathematical form of the relevant probability law.

II.2 Estimating the Variance and the Mean

A probability law is characterized by different parameters, the most important of which are the mean (position parameter) and the variance (dispersion parameter). These two parameters alone are sufficient to define the probability law for a normal distribution.

In general, however, the mean and variance of the law, which represents an infinite number of observations, are unknown and must be estimated from a finite number of observations.

These estimates must be completed by an estimate of the possible error, i.e. the maximum probable deviation between the true unknown value and the estimated value. Chapter II.2 discusses the procedures for estimating the variance, the mean and their error ranges, depending on whether or not the mathematical form of the distribution law is known.

II.3 Outlier values

The presence of values that deviate considerably from the mean in a set of observations suggests that they do not obey the same probability law as the others, i.e. that they are outlier values that must be eliminated before statistical treatment. However, since these deviations may also be extreme values for the population, they must only be rejected on the basis of objective criteria.

Chapter II.3 provides statistical ways for detecting outlier values with a normal distribution law.

III. Measurement Results

Two steps are involved in estimating the value of a quantity measured directly or indirectly :

- the actual estimate based on the experimental results;
- the definition of a confidence interval with a specified probability of containing the true value.

Part III addresses the following points.

III.1 Systematic and Random Errors

It is shown in this chapter that the distinction between random and systematic errors depends essentially on the measurement conditions.

III.2 Characteristics of a Measurement Method

A statistical definition is given of the principal characteristics of a measurement method : repeatability, reproducibility, accuracy, sensitivity and detection limit.

III.3 Error Propagation

This chapter discusses the error

* CETAMA (Commission d'Etablissement des Méthodes d'Analyse)
CEA (Commissariat à l'Energie Atomique)
Centre d'Etudes Nucléaires de Fontenay-aux-Roses

** Editions Masson, 120 Bd St. Germain. 75280 Paris Cedex 06

calculation specifying the confidence interval on the value of a quantity depending on whether it is measured directly or indirectly as a function of one or more different physical quantities.

III.4 Expressing the Result

The final chapter in Part III summarizes the practical rules for expressing the measurement result.

IV. Comparisons

Measurement results are often used to compare the quantities estimated from them. In practice, this involves comparing certain characteristic parameters (means, variances, proportions) of the infinite statistical populations as represented by a finite number of experimental observations. The following points are examined.

IV.1 Basic Concepts

This chapter reviews the principles of statistical tests, by which experimental observations are used to determine the nature of a population, the form of a distribution law, the parameters of this law, etc.

IV.2 Comparing Mean Values

The mean value of the observed distribution may be compared with a reference value or with the mean of another observed distribution. The comparison between two means is improved when both variables can be matched. The following are only a few examples of the applications of this concept:

- Checking the accuracy of a measurement method
- Checking the composition of a synthetic compound
- Checking product specification compliance
- Checking the theoretical value of a quantity
- Checking the agreement between two measurement methods, two instruments or two laboratories.

IV.3 Comparing Variances

Comparing a variance with a reference value or comparing two variances generally involves evaluating the reproducibility of measurement methods, instruments or laboratories, for example to choose between two procedures or to verify the stability and reproducibility of measurements over time.

IV.4 Comparing Proportions and Acceptance Inspections

This chapter deals primarily with inspection by attributes in which the items submitted to examination are designated as satisfactory or defective.

IV.5 Analysis of Variance

Variance analysis is a statistical technique used to detect the possible influence of one or more factors on a measurement result when the result can be considered as a random variable. This type of analysis can show the existence of a link among several quantities, but cannot represent this link by a mathematical relation.

Variance analysis is a particularly economical and advantageous way of studying the effects of several factors on a phenomenon and has innumerable applications, two of which are examined in the next two chapters.

IV.6 Investigating the Causes of Error

This chapter describes the application of variance analysis to two special problems:

- Evaluating the contribution of one or more phases of the operating procedure to the overall uncertainty in the measurement result;
- Estimating the mean value of a large number of results obtained under different conditions using the same method.

IV.7 Homogeneity of a Material or of a Set of Parts

Possibilities of application to homogeneity analysis are presented for various statistical techniques already discussed, and for variance analysis in particular.

V. Relating Two Variables

Measurement results may be used to find a relationship between two quantities. The application of the results may also depend on the existence of such a relationship, either because the measurement method requires calibration or because the quantity to be evaluated is dependent on other measured quantities X, Y, Z, etc. It is important to note that the confidence interval for the final result is not the same depending on whether X, Y, Z are independent or not.

The first two chapters in Part V deal with regression analysis.

V.1 Linear Regression

The regression method is used to express the relationship between two quantities when one of them is subject to nonrandom variations. The regression is termed linear if the relationship can be represented by a straight line. An important application to measurement techniques is plotting the calibration line representing the measured quantity directly as a function of the desired quantity. These quantities are generally different in nature: for example, the first may be an optical density and the second a concentration.

V.2 Extended Linear Regression

Nonlinear regressions are handled either by changing the variables to obtain a linear regression or by using the multiple regression technique. This method can also be used to estimate a linear relation between a random variable Z and several nonrandom variables X_1, X_2, ...

These two chapters cover the main problems related to calibration:

- Computing a calibration curve
- Determining the confidence interval for a result obtained from a calibration curve
- Checking that the slope of the calibration line corresponds to the expected theoretical value
- Checking whether the calibration line intersects a specified point (e.g. the origin)
- Checking whether a new measurement of a reference quantity is compatible with the previous calibration
- Comparing two quantities estimated using the same calibration line or two different calibration lines
- Comparing calibration lines that may differ by the moment they were plotted, the extent of the calibration range, the presence or absence of an element in the reference products, the origin of the reference quantities, etc.

V.3 Correlation

The correlation method may be used to detect a certain correspondence between the numerical values of two random variables without any notion of a cause-effect relationship.

V.4 Relating Two Variables

This chapter summarizes the statistical methods that may be used to detect a relationship between two variables.

VI. Interlaboratory Circuits

Variance and regression analysis may be applied to interlaboratory circuits in order to evaluate two special problems :

- The existence of systematic discrepancies between the results provided by several laboratories
- The reproducibility of the results from each laboratory.

In conclusion, we feel that this work provides an overview of the many possibilities offered by statistical analysis. In particular, statistical treatment allows more thorough implementation of measurement results, often giving greater productivity either by providing better data for a given number of observations or by ensuring all of the information required despite a smaller number of measurements.

Nevertheless, it must not be forgotten that statistical calculations are based on explicit or implicit hypotheses that may not hold. It is indispensable to understand these hypotheses fully, to keep them always in mind and, before adopting the conclusion furnished by the calculations, to check scrupulously that they are acceptable given the conditions under which the experimental observations were made.

Activities of the ESARDA Working Groups

1. Low Enriched Uranium


P. Boermans, FBFC, Dessel, Belgium
Convenor ESARDA LEU WG

V. Verdingh, CBNM, Geel
Secretary ESARDA LEU WG

Introduction

The ESARDA Working Group on LEU was constituted in 1978. The basic idea was the creation of a plant specific working group, for which the prime input should come from plant operators.

The terms of reference of the group are as follows:

a. The ESARDA Working Group for LEU Conversion and Fuel Fabrication Plants is composed of representatives from plant operators, safeguards authorities and research establishments.
b. The group acts as a free podium for discussion, to exchange and demonstrate experience gained in Nuclear Materials Management by members of the Working Group.
c. The Working Group endeavours throughout its work, to ensure that safeguards procedures can draw largely, if not entirely, upon comprehensive Nuclear Materials Management Systems.
d. The Working Group will restrict actions to the pragmatic development of current Nuclear Materials Management technology taking into account the requirements of both plant operations and safeguards. The work commitment of the plant operators will be limited to the transfer of knowledge and assisting each other to implement improvements with the least average cost for all members.

The role of the Research and Inspection Organisations will be to advise upon the developments in Nuclear Materials Management Systems and to support R & D work if required.

The organisations represented in the Working Group are the following:

Plant-operators

BNFL	Springfields, U.K.
EXXON	Lingen, FRG
FBFC	Dessel, Belgium
FBFC	Romans, France
FN	Boscomarengo, Italy
RBU	Hanau, FRG

Research and Inspectorate

CEN/SCK	Mol, Belgium
ECN	Petten, The Netherlands
KFK	Karlsruhe, FRG
UKAEA	Harwell, U.K.
CEC	Safeguards Directorate, Luxembourg
CEC	Joint Research Centre, Ispra and Geel

The Working Group meets at regular intervals to evaluate actions going on and to decide upon new actions. Up to this date the Working Group met 11 times. The last meeting was hosted by RBU at Hanau on the 19th and the 20th of June 1984. An informal meeting was organized on the occasion of the ESARDA Symposium in Liège, on the 22nd of May 1985.

The activities of the Working Group can be grouped under all of the three following items :

- system studies
- DA measurement techniques
- NDA measurement techniques,

each of the mentioned topics specificially related to LEU facilities.

The work of the Group has been reported on different occasions : on the INMM Symposium in Albuquerque in 1979 /1/; on the ESARDA Symposium in Edinburgh /2/ in 1980, and on the ESARDA Symposium in Versailles in 1983 /3/.

This document reports on the major achievements of the Group and on some perspectives for the future; it is not intended to be a detailed report on each of the different topics.

System Studies

The Working Group considered the following subjects :

- verification schemes
- computerized nuclear material management systems
- applications of NUMSAS to plant data.

Verification schemes

The Working Group performed studies to evaluate the consequences of the verification effort in a reference plant with a flow of about 600 t $U.y^{-1}$ and an average enrichment of 3%. Different verification schemes were considered.

A first scheme was based on the verification of input and output and on the verification of the physical inventory once a year. This scheme has the advantage that the details of the plant operation are not imperative. The verification of the material flow (input and output) has been evaluated considering the LEU conversion/fabrication plant as a part of a fuel cycle subjected to an integrated safeguards system.

The verification problems are specific for different materials involved. The materials can thus be divided into two broad categories :

- materials which are transferred in standardized containers of small dimensions (UO_2 powders, UO_2 pellets) and that can be easily sampled for the determination of fissile material content by DA or by NDA measurement.
- materials which are transferred in items of large geometries (UF_6 cylinders, UO_2 fuel assemblies) so that a verification of the fissile material content requires special procedures : NDA measurements, containment and sealing techniques.

In the transfer of the first category of materials it seems important, in order to reduce the effort, to make only one verification either at the outlet of a plant or at the entrance of the following plant. In principle it seems preferable to perform the verification at the outlet of the plant and to conserve the information through the use of containers which can be sealed and verified for integrity.

The second verification scheme is based on an optimized utilization of a Fast Response Computerized Accountancy System, together with statistically based sample verification. This scheme, in contrast with the

first, requires a detailed knowledge of the plant operations and will lead to an intrusive control system. The members of the Working Group considered that, from the operators point of view, one of the advantages of a Fast Response Computerized Accountancy System may be the possible reduction of the frequency or totality of physical inventory taking. With this approach the detection time may also be reduced to a significant extent, which is an attractive safeguards feature.

A third approach to safeguarding LEU plants, largely based on containment, can be considered. An extended use of containment to overcome the limitations in the use of accountancy is a point of increasing interest in safeguards. This approach has been mentioned in the Working Group but no specific action was started in this area. The system seems more attractive for future plants (when foreseen in the conception of the plant) than for existing facilities.

Computerized nuclear material management systems

All Working Group members are using computerized accountancy systems. The degree of implementation of the computer systems is different from plant to plant, but in general it can be said, that significant progress in this field has been made in the last two years. Some Working Group meetings were specially dedicated to this topic.

Some of the plants are fully computerized, in the sense of using a quasi real time accounting system.

In particular, a near real time nuclear materials management system has been developed and implemented at FBFC-Dessel, and presented to the Working Group.

The system permits an interrogation of stocks and movements following different parameters. Hard copies following different layouts exist, for instance the I.C.R., the L.I.I. and the P.I.L., on listing and on magnetic disk.

Attention has to be paid to the fact that the concerned fissile material data are used for Safeguards control as well as for other computerized applications as :

- safety reports
- fissile material balances for customers
- financial reports (insurance, etc.)
- manufacture reports (production reports)
- custom reports
- ...

Thus the system has to be considered as an "integral" material follow-up application.

Another near real time nuclear materials management system has been developed and implemented at BNFL-Springfields under a Joint Collaboration Contract between BNFL and JRC-Ispra and has been reported at the ESARDA Symposium on Nuclear Materials Management /4/.

The project was made possible through the LEU WG and incorporated safeguards objectives within its framework from its very conception.

Through this work :

- the facilities have been able - assisting each other - to introduce a more comprehensive system for LEU nuclear materials accountancy and control than otherwise would have been possible;
- JRC-Ispra has gained experience in the field and influenced the design of a computerized materials management system for a large plant (BNFL);
- systems can be easier to audit by the inspectorate since consideration was given to simplicity and "transparency" during the design stages with audit requirements in mind;
- the system can produce the Euratom reports (I.C.R., P.I.L., M.B.R., L.I.I.) directly.

Applications of NUMSAS to plant data

The NUMSAS package (NUclear Material Statistical Accountancy System) developed by JRC-Ispra is a statistical tool for the analysis of materials balance information /5/. The Group expressed interest in this package since the Euratom Safeguards Directorate had indicated that it intended to use the package on a routine basis. Following initial discussions, an exercise was arranged through the Group to apply NUMSAS to plant data from FBFC-Dessel. (Actually the code is applied by the Inspectorate on a routine base for this plant.)

The Group had noted that the analysis of material balance data at such LEU plants would be difficult without the use of computerized tools because of the volume of data involved. The code is designed to give MUF, the standard deviation of the MUF, for each primary error source component, the contribution to the overall variance of the material balance. The Group noted the following main points arising from the experiment :

- the main effort required to implement the code is concerned with collecting the measurement uncertainty data for the primary error sources and to allocate the error paths;
- the definition of systematic error in some cases was necessarily rather arbitrary and the treatment of systematic errors was problematic;
- in the case considered, with 10,000 entry lines, there were 65 distinct error paths comprising 36 primary error source components;
- the code is designed to eliminate all correlated batches from the calculation of MUF variance. In this case, it was only possible to eliminate those batches which appeared with an unchanged batchname on both concerned Physical Inventory Listings and I.C.R. reports for the intermediate period;
- the results of the run prompted a review of the error table. The merit of the code and of this exercise was that it highlighted those factors which made the greatest contribution to uncertainty in the material balance.

DA Measurement Techniques

These activities are considered to be part of the planned scheme for plant specific LEU inter-laboratory measurement evaluation programmes.

Determination of U-content by gravimetry

Members of the Group have taken part in an inter-laboratory comparison exercise for the determination of "U" in "UO_2". JRC-Geel produced a certified reference batch of 10 g UO_2 pellets for this exercise /6/ and used direct uranium measurements by potentiometry and controlled potential coulometry for the certification.

For verification purposes the uranium content was checked by the determination of the total impurity content and substracting it from a 100%, taking into account the O/M ratio.

The exercise involved analytical laboratories from all of the European LEU fuel fabricators and from certain Research Centres. The exercise served to confirm the excellence of "U" in "UO_2" determination for pure pellets using the gravimetric method at analytical laboratories throughout Europe /7/.

Determination of U-content by potentiometric titration

A similar inter-laboratory comparison exercise as on the gravimetric determination of U in UO_2 pellets has already been set up for the determination of U by potentiometric titration. The exercise will allow the calculation of the accuracy and precision in terms of reproducibility and repeatability of laboratory procedures for routine uranium determinations by the potentiometric titration method. It will also enable the calculation of the systematic deviation in these procedures in terms of deviation from certified values provided by CBNM.

The experiment will be conducted in two phases. In a first phase pure uranylnitrate solutions will be circulated followed in a second phase by similar uranium solutions doped with impurities. It will be requested to perform routine potentiometric analyses. The experiment will be performed in the period 1985-1988. Similar experiments will be organized in well specified and justified cases.

Sampling in LEU plants

An enquiry on the sampling procedures of different types of materials in LEU plants was undertaken. The enquiry took into consideration, UF_6, UO_2 powder, UO_2 pellets, residues, waste and other materials.

The collected information was normalized and communicated to the members. The document was examined and discussed during a Working Group meeting.

Experiments for the determination of sampling errors for pure UO_2 powders in drums have been carried out at FN. The results of the experiments have shown that on standard free flowing UO_2 powder ex ammoniumuranylcarbonate no significant sampling error exists. The determination of the sampling errors in UO_2 powder from other conversion processes and in UO_2 impure powder has also been considered as important, the last one mainly in connection of plant clean-up for inventory taking.

A study of the water pick-up of uranium-oxide powders (UO_2 and U_3O_8) was executed by JRC-Geel. The results were communicated to the Group /8/.

NDA Measurement Techniques

Weighscales intercomparison exercise

The Group discussed weighing and exchanged information on measurement control procedures for weighscales. During the first round of discussions, the Group concluded that the precision and accuracy of weight measurements were not sufficiently well understood. As a result the Group set up an exercise to gather comparative data from a number of weighscales and JRC-Ispra initiated a research programme to define a method of analysis for such data.

The weighscale comparison exercise involved four LEU plant operators in EEC countries, and the JRC-Ispra. JRC manufactured the set of seven standard weights for the exercise. These weights were then sent to each operator in turn, who carried out a sequence of 127 weighings of different combinations of the weights on one or more weighscales. The 127 combinations gave a sequence of weights evenly distributed through the range 0-90 kg.

JRC developed a new methodology for the analysis of random and systematic errors in weighing. An initial assessment of the data showed that the error structure is more complex than had previously been believed since bias is not constant. To overcome this problem JRC developed a second method which assumed a model with linear bias /9/.

The algorithm for analysis of weighing data is now available for use by the operators in computer tape form. The statistical tools behind the methodology are not always easily understandable by non-specialists in this field. Therefore a "school example" is being prepared by the ECN-Petten and the JRC-Ispra for the next Working Group meeting.

The comparison exercise is viewed as a success by the Group since it has led to an improved understanding of weighing errors and has allowed operators to assess their position relative to other operators. Further, the exercise was an example of productive cooperation between the operators and a research centre. Members of the Group have, therefore, decided to set up a second round of the weighscale comparison exercise using the same set of standard weights.

This second round is being executed at this moment and will be finished at the end of the year. In this exercise attention was specially drawn to the higher weight range of balances, used for weighing finished assemblies, and to limited well defined small parts of the ranges of balances, because of their use for standardized containers with slight different weights.

Use of the Neutron Collar in Leu facilities

The Neutron Collar is a special design of neutron coincidence counter developed at Los Alamos laboratories for the verification of fresh fuel assemblies. In function of the discussions on verification schemes the Group discussed ways of safeguarding the assembly area of LEU plants and recognized the potential value of a fuel assembly verification measurement device. The instrument has been tested in the presence of Euratom safeguards inspectors at FBFC Dessel as a part of the Belgian support programme to the IAEA /10/. The Neutron Collar has also been used by Euratom safeguards inspectors at a number of reactor sites and at LEU fabrication plants in the Community.

The experience gained from using the instrument was discussed by the Group and, in particular, the practical problems of use for safeguards verification in an operating plant were addressed.

The main parts noted were as follows :

- to some extent the use of the instrument will be fuel and plant specific in terms of calibration and in terms of the physical constraints imposed by the design of the fuel and the fuel store;
- agreements will be required to comply with site licences and transport regulations (use of neutron source) and covering liability in the event of damage to finished fuel;
- there may be limitations of use with multi-enrichment fuel or with fuel containing burnable poisons;
- the instrument proved to be reliable and robust under plant operating conditions : no noise pick-up was observed even when differest types of electrical devices were in operation in the immediate surroundings. The influence of neutron backgrounds was very small for the active assay, but important for the passive mode.
- after measurement of an assembly, sealing should be possible. Adequate techniques for sealing assemblies are only in the development stage at this moment.

Anyway, the Group concluded that the active Neutron Collar promises to be a useful measurement technique for flow and inventory verification. A new measurement campaign was executed at FBFC in the month of June 1985. The results will be presented on the occasion of a next Working Group meeting.

Use of the rod-scanner for safeguards purposes

Most of the European LEU fuel fabricators have, or soon will have, a rod-scanner in their plant for Quality Control purposes.

It has been noticed that it is possible for safeguards inspectors to make use of the scanners to verify operators' data for finished fuel rods, as it is already the case on a routine base in some facilities.

This scheme was discussed by the Group and a sub-group of expert representatives of plant operators was set up to examine such a possibility.

The initial findings of the sub-group were reported to the 9th meeting of the Group in June 1983. These were :

- use of rod-scanners could present different problems in different plants. For example, the rod-scanner is an integral part of the production line in certain plants, hence use for safeguards purposes could, under certain circumstances, interfere with production;
- preparation of standards may present a problem to the safeguards inspectorate, although some standard rods have already been prepared and are in use at certain plants. Given adequate standards, rod-scanners provide a sufficiently accurate total U235 measurement for verification purposes for standard fuel;
- fuel containing Gadolinium (a burnable poison) present special problems for neutron-activation rod-scanner systems. For these fuels a passive high resolution detection system may be required.

Actually one of the members of the Working Group is preparing a draft for a possible intercomparison exercise to get a better idea of the accuracies and precisions relied to this verification tool.

PHONID-3

JRC-Ispra performed some measurements on LEU material at the UKAEA's Springfields laboratory. The results were promising since the instrument remained stable over a relatively long period, and accuracy for measurements of LEU appeared to be better than that achieved with HEU samples.

During the month of October 1984, a campaign of measurements on U-waste was carried out at the CEN-Mol by means of PHONID-3 device in collaboration with FBFC-Dessel /11/.

The scope of this work was a preliminary check of the feasibility of waste assessment with an active interrogation technique, as proposed by FBFC-Dessel. FBFC prepared 29 sealed plastic bags containing a well known quantity of UO_2 powder and resembling to real waste, to make a calibration curve. It came out that the efficiency

of the system was much better by using thermalized neutrons instead of epithermal neutrons. It must be pointed out that in such conditions the device can easily detect tenths of a gram of fissile material with reasonable counting times. The complete results of this experiment will be reported on the occasion of the next Working Group meeting.

The results of the experiments show that the application of PHONID devices in such field is promising. Yet they have demonstrated the feasibility of this kind of measurements, they also suggest the need of a deeper investigation and refinement.

Conclusions and Perspectives

The Group has provided an excellent forum for an interchange of views between plant operators, research centres and the Euratom Safeguards Directorate. Common problems were identified in view to arrive at common solutions and at a maximum of communality in the safeguards system within the European Community.

In the future the Group will focus on following themes and topics :

- Prior to start new projects, increased thought will be given to the potential cost-benefit factor of the project for all the parties involved : operators, safeguard authorities.
- The exchange of information with other ESARDA Working Groups will be promoted.
- Discussion should be started on safeguarding recycled material, and on mixtures of fresh and recycled materials; problems expected in following fields are related to :
  - neutron measurements
  - gamma measurements.
- there should be further looks for methods for the measurements of low-level waste (dry and in solution).
- Discussions on practical problems related to the introduction of material flow and material inventory data in computerized systems are desired.

References

/1/ R. STEWART et al. — The ESARDA Approach to Facility Oriented Safeguards Problems. Proceedings of the 20th Annual Meeting of the INMM, Albuquerque (1979) pp 37-47

/2/ S. SAIGER (RBU Hanau) — Activity of ESARDA Working Group on Safeguarding Low Enriched Uranium Conversion/Fabrication Plants. Proceedings of the 2nd ESARDA Annual Symp., Edinburgh (1980) pp. 27-34

/3/ R. STEWART, A.A. MUSTO (BNFL Springfields) — A Survey of the Interaction between Euratom, the European Research Centres and Plant Operators in the Field of LEU Conversion and Fuel Fabrication. Proceedings of the 5th ESARDA Annual Symp., Vesailles (1983) pp. 339-341

/4/ F. ARGENTESI, L. COSTANTINI, T. CASILLI, B. BURROWS, S.M. FENSOM — NUMSIS - A Real Time System for Nuclear Material Management. Proceedings of the 5th ESARDA Annual Symp., Versailles (1983) pp. 443-446

/5/ F. ARGENTESI, T. CASILLI, M. FRANKLIN — NUMSAS - A Statistical Nuclear Materials Accountancy System in the Euratom Framework. Proceedigs of the 1st ESARDA Annual Symp., Brussels (1979) pp. 388-390

/6/ Y. LE DUIGOU, V. VERDINGH - Certification of a Uranium Dioxide Reference Material for In-Plant Uranium Determinations. COM 4146 (1984)

/7/ V. VERDINGH, W.L. ZIJP — Repeatability and Reproducibility of Gravimetric Uranium Determinations in UO_2 Pellets. COM 4182 (1985)

/8/ V. VERDINGH, B. SLOWIKOWSKI — Water Pickup of Uraniumoxide Powders. Presentation to the LEU-WG, GE/R/AS/06/84

/9/ F. ARGENTESI, H. MUNTAU, S. CANALI — A Standard Method for the Estimation of Scale Accuracy and Precision : A Case Study for the Evaluation of Bias Time Evolution. Proceedings of the 3rd ESARDA Annual Symp., Petten (1982) pp. 113-114

/10/ Optimisation of NDA Measurements in Field Conditions for Safeguards Purposes. CEN, IAEA, FBFC. LASL; BLG-543, February 1981

/11/ P. BOERMANS, R. CARCHON, P. DELL'ORO, G. SMAERS — U-Waste Measurements with PHONID-3 Device. Internal report FMM 104, July 1985

2. Containment and Surveillance

Chr. Brückner, KfK, PAE
Past-Convenor of the ESARDA Working Group on Containment and Surveillance

The activities of the ESARDA Working Group on Containment and Surveillance (C/S) are based on the following terms of reference :

"Objective : To evaluate and recommend procedures for containment and surveillance methods for use by safeguarding authorities to reduce the inspection effort at nuclear facilities".

This report relates to the two years 1984 and 1985, during which the author was acting as convenor of the working group. Four meetings were performed in this time.

There are 20 members of the group, from ENEA-Casaccia, ENEL-Rome and Fabbricazioni Nucleari-Bosco Marengo, Italy, from Harwell and BNFL-Risley, United Kingdom, from SCK/CEN Mol, Belgium, and ECN-Petten, Netherlands, from JRC-Ispra and Euratom Luxembourg, from CEA, France, and from VDEW-Frankfurt, DWK-Hannover, KFA-Jülich and KfK-Karlsruhe, Germany. The IAEA and Sandia National Laboratories, Albuquerque, U.S.A., are represented as observers. this means, that facility operators are represented in the group as well as R & D facilities and Safeguards Inspection Authorities.

In the beginning of the two years period the group started with a broad discussion about a future programme. From the side of the convenor four points were proposed to describe a more detailed programme under the heading of the terms of references :

1. to determine C/S instruments and devices currently applied in the field by Euratom and the Agency
2. to identify where improvements to currently used techniques and devices were required and to identify how these improvements may be effected
3. to establish what techniques and devices were under development, by whom and to what end
4. to understand the requirements of the safeguards authorities for containment and surveillance hardware.

It was agreed that there is a need for the Working Group to continue primarily as a means of interchanging accumulated experience. For discussion of R & D activities the starting point should be the specification of need. In this connection the input from operators was indicated as crucial in deviding on the practical circumstances in which the C/S instruments and devices have to operate and providing an in-depth investigation of the performance of subsequent hardware.

In general the view of the committee was that the Group provided a good forum for the exchange of opinion and information. The members of the Working Group considered that the Group should be retained and should continue to take an active interest in containment and surveillance matters following the terms of references and the points proposed by the convenor. It was, however, indicated as essential that information should be available mainly in written form in advance of meetings so that informed discussions could take place at working group meetings.

In agreement with this line working group meetings were performed with the following structure of the agenda :

1. short oral report of each member about
   - ongoing R & D activities, results, new ideas and so on in the case of R & D people
   - experience with applied hardware, probable needs for improvements or

developments, in care of operators and inspection authorities

2. oral reports about meetings of other committees dealing with containment and surveillance, for example IAEA advisory group meetings, if such meetings were held
3. presentation and discussion of papers about special subjects, worked out and circulated before a meeting. The subjects were indicated in the previous meeting, mainly as a result of the short oral reports of the members.

Presentation and discussion of such papers from the main part of the present working group meetings. Papers about the following subjects were presented and discussed :

- C/S instruments currently used by Euratom, with description of instrument characteristics and application (paper written by Mr. Chare, Euratom Luxembourg)
- VACOSS III seals for safeguards use, describing possible applications (Mr. Haas, Euratom-Luxembourg)
- Adhesive paper seals, describing characteristics of such seals and the status of development (Mr. Walford, AERE-Harwell and Mr. Richter, KFA-Jülich)
- Fuel assembly sealing systems and their tests in the Kahl boiling water power reactor facility (fuel assembly seals demonstration experiment) describing status of development and results of the actual tests in the Kahl reactor (B.C. d'Agraives, JRC-Ispra, J. McKenzie, Sandia Nat. Labs., Chr. Brückner, KfK)
- A new developed electronic sealing system (Mr. Musyck, Mol)
- A film produced for Agency Inspectors training entitled : "Fuel handling in an LWR power plant and camera surveillance for safeguards" which also shows the capability of a film camera system developed in KfK (Chr. Brückner, KfK).

In general it can be summarized that the Working Group forms a useful forum for collection of information and experience for all parties involved, including R & D people to look for the best way with minimum burden for operators for a sufficient performance of safeguards with use of C/S measures.

3. Destructive Analysis

P. De Bièvre, CBNM, Geel
Convenor of the ESARDA Working Group on Destructive Analysis

At its annual meeting in Saluggia (Italy) on 1 and 2 October 1985, the Working Group on Techniques and Standards for Destructive Analysis of the European Safeguards Research and Development Association (ESARDA),

- consisting of individuals, responsible for U and Pu measurements, from about 35 analytical laboratories in the European Community and of 5 observers from laboratories outside the European Community,
- in view of the growing pressure trying to prescribe measurement methods and procedures, with the implied idea that their use automatically ensures a given level of precision and accuracy or reliability,

discussed questions of the responsibility for U and Pu analytical measurements.

The WG expresses its opinion that certain authorities, by their very nature and task, have the responsibility to decide **what** should be measured, e.g. fissionable isotope and/or fissionable element content of a given material, and possibly to what accuracy.

However, the Group also expresses its unanimous opinion that the full responsibility to decide **how** such measurements can best be made, must remain with measurement laboratories and that such a responsibility automatically includes the freedom of the choice of appropriate measurement methods. The latter entails the responsibility of the laboratories to demonstrate the quality of the measurements.

The Group stresses that the quality of a result should be the criterion on which to judge measurements, and not the choice of any given particular method or procedure. Such "quality" or "level of performance" or "state-of-the-practice" (as opposed to "state-of-the-art") can be derived from internal measurement control programmes combined with well conducted interlaboratory measurement evaluation programmes.

ESARDA News

8th ANNUAL ESARDA MEETING (Restricted participation)

Copenhagen, 13-15 May 1986

The eighth Annual Meeting will be held in Copenhagen in the Eigtveds Pakhus.
The attendance will be limited to the ESARDA Steering Committee members, coordinators, working group members and observers.

The title of this internal meeting of ESARDA is :

Capabilities and Objectives of the Use of NDA-DA-C/S Measures in Safeguards

The objectives of the meeting are twofold :

- give the managerial levels of the Association the opportunity to make an assessment of the work performed in various fields through the information given by the working groups and
- promote a mutual technical information of all those who are working in the Association framework.

9th ANNUAL ESARDA MEETING

ESARDA announces that the 9th meeting will be a general **Symposium on Safeguards and Nuclear Material Management** and will be held in London on 12-14 May 1987.

It will be held in the new Queen Elizabeth II Conference Centre near Westminster Abbey.

A call for paper will be circulated soon.