Comparison of Factory Methods of Sampling and Testing Brown Crepe Rubber for TC Strain

G. C. IYER

The different sources of variation in sampling a bale are examined and it is shown that sampling at one random location in each of the 5 or 6 biscuits composing a bale leads to the determination of a mean bale strain-value within 5% limits.

The lot mean can be estimated with equal accuracy either by testing the bale samples individually or by testing bulked bale samples on the assumption that the value obtained in the bulked sample test would equal the mean obtainable by testing the corresponding individual bales. Of the two methods of testing, the use of bulked bale samples is preferred for reasons of cost although it gives a less reliable estimate of bale-to-bale variation.

In the factory under consideration, the dry crepe rubber is produced in pads (known as biscuits) weighing about 10 lb and consisting of many layers of approximately 20 cm length. These biscuits are stored on the factory floor. A bale is made up by grouping together 5 or 6 biscuits to weigh 50 lb and then wrapping these in a piece of RSS 3 rubber weighing $8\frac{1}{3}$ lb. There are approximately 38 bales to a ton. The study undertaken was to examine the strain value of the brown crepe rubber used in the bales.

Bales are most easily sampled at the weighing stage, when rubber can be taken from each bale. The sampled bale can then be made up to the required weight with extra rubber. In the experiments described below the sampled bales were systematically selected from a lot. This systematic sampling has the advantage of revealing any trend that may exist in strain value of the biscuits making up the lot.

TC strain is defined as the extension of a standard specimen of rubber produced by a tensile stress of 5 kg per square centimetre of cross-sectional area and is expressed as a percentage of the original length of the specimen. In the laboratory the strain test is known to have an error variation of 3.36 percentage square units (or 2.2% coefficient of variation when the strain level averages 85%). The variation measured by the variance, is defined as the mean value of the squares of the deviations of any set of observations from their true mean and is measured as units square.

SAMPLING WITHIN A BALE

Investigation of Within Bale Variation (Experiment A)

This experiment was carried out on a lot of 50 tons, consisting of 1920 bales. Every hundredth bale was selected systematically at the time of weighing, giving a total of twenty bales in the sample. In each bale, each of the five biscuits was sampled at three random locations. There were thus 300 test pieces, the strain values from which provide the components of variation within a bale as given in *Table 1*.

Choice of a Proper Bale Sample

From *Table 1*, it is evident that the biscuit variation is quite high within a bale, and location variation within biscuits is smaller than test error. Collecting at one location in each of the five biscuits to make a bale sample which if subsequently tested once would give an estimated variation of 3.78 (i.e. $3.36+\frac{2.11}{5}$). This would account for a coefficient of variation of 2.2%, at the mean strain level of 88% in this experiment.

The proper choice of the sampling scheme within a bale so as to determine the bale mean

COMMUNICATION 395

TABLE 1. COMPONENTS OF VARIATION. EXPERIMENT A

Components	TC strain, units sq.		Degrees of freedom of mean squares
1. Bale variation	13.99		19
2. Biscuit variation within bales	6.59		80
3. Location variation within biscuits and test error	5.47		200
(a) Location variation		2.11	
(b) Test error		3.36	
Total variation	26.05		

within 5% of the true value in 95% of the cases, would roughly necessitate that the bale test should have a coefficient of variation below 2.5%. Therefore, as suggested above, sampling at one location in each biscuit and carrying out one test on the bulked sample can be a satisfactory procedure for sampling within bales.

Verification of Suggested Sampling Method (Experiment B)

This experiment was carried out on a lot of 25 tons to examine the efficiency of the method suggested above and also to obtain a better estimate of bale variation. One tenth of the bales were selected, giving 96 bales for sampling. Each of the six biscuits within a bale was sampled at one location; these were bulked to provide one bale sample for testing. To check on the reproducibility of this sampling procedure, two independent samples were collected in an identical manner from each of the chosen 96 bales of the lot. The various components of the analysis of strain test values are given in *Table 2*. The estimate of bale strain determination has a coefficient of variation of 2.7%, compared with the estimate of 2.2% in experiment A.

Efficiency of the Suggested Method of Sampling a Bale

If test error is ignored, sampling error within a bale and bale-to-bale variation would account for a total variation of 17.46 (i.e. 1.77+15.69), which is in excess of the true bale-to-bale variation of 15.69 by 11%. The procedure of testing the entire bulk of sampled bales gives a bale variation of 15.69, when no allowance is made for test error. Thus, in the estimation of the lot mean, the sampling method suggested above entails the use of 11% more bales than the ideal of using the entire bulk of each chosen bale. This assessment of the efficiency slightly underestimates the true efficiency which should provide for test error also.

TESTING BALE SAMPLES

The bale samples, composed of bits, one bit from each biscuit in the bale, can be tested for strain in the laboratory in either of two ways:

(i) The bale samples are tested individually, leading to the same number of tests as there are bale samples.

(ii) A fixed number of bale samples are bulked (or blended) and then tested, so that there are fewer tests than bale samples.

Both these testing procedures lead to estimates of lot mean strain value, the degree of accuracy of which can be predetermined. The first method would provide a better estimate of the lot mean and bale variation. In the second method, the lot mean calculation depends on the assumption that the value obtained by the bulked sample test would equal the mean of the individual bale tests; in addition the bale variability estimate, obtained as shown later on, is less reliable. Nevertheless, the choice of method of sampling and testing may depend not only on statistical considerations but on the important question of cost. For such a comparison, the identical permissible

TABLE 2. COMPONENTS OF VARIATION. EXPERIMENT B

Components	TC strain, units sq.		Degrees of freedom of mean squares
1. Bale variation	15.69		95
2. Within bales, sample variation and test error	5.13		96
(a) Sample variation		1.77	
(b) Test error		3.36	
Total variation	20.82		

limits for the lot mean must be stipulated for the two methods; but, as is shown below, the number of sampling and testing operations (and consequently the cost) can be reduced to a relatively low level in method (ii) without deviating from the required standard of accuracy.

(i) *Testing Bale Samples Individually*

When the bale samples are tested individually, there will be one test for each bale sample. By averaging these results the lot mean is obtained. The bale test results provide the bale variation from which the standard deviation of the strain value is calculated. Prediction on the basis of theoretical distribution about the percentage of bales within any stipulated strain limits is easy.

Number of bales sampled. The limits of the true lot mean about the observed lot mean can be evaluated (at any high level of probability) by the following formula:

$$d = \pm t_{n-1} \frac{S}{\sqrt{n}} \sqrt{\left(\frac{N-n}{N-1}\right)}$$

where

N refers to the total number of bales in the lot

n refers to the number of bales sampled

d refers to the acceptable limit (or margin) in the lot mean estimate, in units of TC strain (%)

S denotes the standard deviation among bale strain determinations

and t_{n-1} denotes the critical value for $(n-1)$ degrees of freedom in the Student's 't' distribution for the chosen level of significance.

The above formula can be rewritten as shown below, to provide the number of bales to be sampled for chosen limits for lot mean.

$$n = \frac{(t_{n-1})^2}{d^2} \cdot \frac{S^2}{} \left(\frac{N-n}{N-1}\right) \quad \ldots\ldots \quad (1)$$

Using approximate values of S and d, we can estimate the necessary sample size to be chosen for any lot size. *Table 3* gives the number of bales to be sampled in different lots for a given acceptable margin of 1.7% (i.e. 2% of the mean when the mean level is 85%) for the estimated lot mean and the effect on the acceptable margin of sampling 24 to 50 bales assuming $S^2 = 21$ from experiment B.

(ii) *Testing Bulked Bale Samples*

The lot mean strain is estimated by averaging the test values. The assumption is that the test result on bulked bale sample would equal the mean of the separate bale strain values. Therefore if the same fixed number of bales is involved each time, the average of the tests on the bulked samples would be the estimated lot mean.

Number of bales sampled and bulked. When testing error is controlled, a good estimate is available for this variation. The bale sample variation excluding test error is assumed to be constant in the lots.

The number of bale samples to be bulked for a single test depends on the ratios of the variations of bale samples to the test variation and also on the relative costs of testing and

TABLE 3. EFFECT OF DIFFERENT SAMPLE SIZES ON THE ACCEPTABLE MARGIN FOR TRUE LOT MEAN

Lot size Tons	Lot size Bales	Number of bales for obtaining $d=1.7\%$	Sample size proposed: number of bales	Sample size proposed: sampling fraction	Acceptable margin by which observed lot mean may differ from true lot mean with 95% confidence (i.e. $d\%$)
5	192	27	24	1/8	1.82
10	384	28	25	1/15	1.83
15	576	29	29	1/20	1.70
20	768	29	30	1/25	1.68
25	960	30	48	1/20	1.30
30	1152	30	46	1/25	1.34
35	1344	30	44	1/30	1.37
40	1536	30	51	1/30	1.27
45	1728	30	50	1/35	1.29
50	1920	30	48	1/40	1.32

sampling. The number of samples to be selected from the lot depends on the margin acceptable in the estimated mean strain of the lot. The formulae appropriate for these are given below on the assumption that the sampling fraction is negligible (i.e. n_b is below 10% of the total number of bales in the lot)

$$n_a=\frac{S_b}{S_a}\sqrt{\left(\frac{C_a}{C_b}\right)} \quad \ldots\ldots\ldots \quad (2)$$

$$n_b=\frac{S_b}{\left(\frac{d}{2}\right)^2\sqrt{C_b}}\left(S_b\sqrt{C_b}+S_a\sqrt{C_a}\right) \quad \ldots \quad (3)$$

$$C=n_b\,C_b+\frac{n_b}{n_a}\,C_a \quad \ldots\ldots\ldots \quad (4)$$

where n_a denotes the number of bale samples to be bulked for a single test

n_b denotes the number of bales to be sampled from the lot

d denotes the acceptable limit (or margin) in the lot mean estimate (limits of ± 2 standard errors accepted) in units of TC strain (%)

S_a and S_b denote the standard deviations of testing and bale samples respectively

C_a and C_b denote the cost of a single test and the cost of sampling a bale respectively

and C denotes the total cost involved.

Estimates of the above quantities for S_a as 1.833 (i.e. $\sqrt{3.36}$) and S_b as 4.178 (i.e. $\sqrt{17.46}$) are available from experiment B. If the cost of testing is assumed to be 3 to 7 times the cost of sampling bales, the use of formula (2) leads to the conclusion that 4 to 6 bale samples should be bulked for one test.

If testing is assumed to cost 5 units of currency and sampling the bale 1 unit, formula (2) shows that 5 bale samples should be bulked. Formula (3) gives the number of bales to be chosen as

$$n_b=\frac{138.32}{d^2} \quad \ldots\ldots\ldots \quad (5)$$

Bale sample variation. The bale sample variation (consisting of bale-to-bale variation and testing error) may be needed in case one wishes to estimate the percentage of bales in the lot that are within two fixed levels of strain. It is readily available in cases where bale samples are tested individually. It is possible as given below to obtain a similar estimate of individual bale sample variation, to correspond with a single test based on an individual bale.

Knowing S_a^2 from previous work in the laboratory for test errors, the value of S_b^2 can be estimated from the following equation:

$$S_o^2 = \frac{S_b^2}{n_a} + S_a^2 \qquad \ldots\ldots\ldots \quad (6)$$

Where S_o^2 represents the variance of the test results on bulked samples each consisting of n_a bales.

After estimating S_a^2 and S_b^2 it is easy to obtain the variation of test results on individual bales (denoted as S^2) as

$$S^2 = S_a^2 + S_b^2 \qquad \ldots\ldots\ldots \quad (7)$$

Comparison of the Costs of the Two Methods of Testing

The testing of individual bale samples or bulked bale samples provides estimates of the lot mean. By assigning with a given confidence (say 95%) the same permissible limit for the true lot mean to vary from the estimated lot mean for both procedures, the cost of sampling and testing can be compared.

Equation (1) can be approximated for large lots by ignoring the influence of sampling fraction and making use of the value $t=2$ as given below:

$$n = \frac{4S^2}{d^2} \qquad \ldots\ldots\ldots \quad (8)$$

The above equation gives the number of bales to be sampled and in turn the number of tests to be carried out as given below:

$$n = \frac{4}{d^2} S^2$$

$$= \frac{4}{d^2}\left(S_a^2 + S_b^2\right) \qquad \ldots\ldots\ldots \quad (9)$$

From equations (2), (3), (4) and (9), it is possible to evaluate the following three aspects of the second method relative to the first method for a given accuracy of lot mean estimate (d).

The number of extra bales needed (i.e. excess of n_b over n) $= \frac{4}{d^2} S_a^2 (n_a - 1)$ .... (10)

The reduction in number of tests (i.e. reduction of $\frac{n_b}{n_a}$ relative to n)

$$= \frac{4S_b^2}{d^2}\left(1 - \frac{1}{n_a}\right) \qquad \ldots\ldots\ldots \quad (11)$$

The gain in cost units
(i.e. excess of $n\,(C_a + C_b)$ over C)

$$= \frac{4}{d^2}(n_a - 1)^2\, C_b S_a^2 \qquad \ldots\ldots\ldots \quad (12)$$

As an example, assume that the lot mean is to be estimated within 2% limit with 95% confidence when the true strain averages 85%. This gives 1.7% as the expected value for d. Let the cost of testing be 5 units ($C_a=5$) relative to the cost of sampling bales ($C_b=1$). Assuming the results $S_a^2=3.36$ and $S_b^2=17.46$, the first method of testing individual bales requires at least *29* bale samples (*Table 3*) leading to 174 cost units. The alternative method of testing bulk bale samples needs only 10 tests to be carried out on 48 bales sampled from the lot, leading to a total of 98 cost units only.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the valuable help and suggestions of Mr C.A.K. Salgado of the Chemical Division and appreciates the co-operation of the Lee Rubber Co. factory which provided facilities for sampling and of the staff of the Chemical Division who carried out the experiments in the laboratory.

Statistics Division
Rubber Research Institute of Malaya
Kuala Lumpur *June 1965*