bulk solids handling Volume 3, Number 1, March 1983

Sampling Procedures for Bulk Solids

R. Mathar and B. Rauhut, Germany

Summary

Judging the state of bulk materials from observations, necessitates the use of statistical sampling procedures. In this paper we employ a certain stochastic model to determine the portion of material which must be included in a sample. A numerical example is calculated in detail.

1. Introduction

The content of a particular characteristic in raw material, e.g., the content of metal in iron ore, the content of sulphur in lime, or the content of ashes in coal, cannot be assumed to be constant but exhibits naturally some kind of variability. Therefore the amount of the characteristic of interest (with respect to a fixed unit of measurement) is usually interpreted as a random variable, which can be characterized sufficiently precisely for practical purposes by its expected value μ and variance σ^2.

These, usually unknown values must be estimated by observations. In the case of bulk solids one sees that only a very small proportion of the material can be examined with respect to the characteristic. Therefore it is inevitable to employ sampling procedures to conclude the state of all of the material from the observations.

Since every mined raw material has to be transported at some time the bulk, which has to be judged, can be imagined as a stream (e.g., a moving belt conveyor with a constant load or a sequence of containers) the properties of which change in time. In using sampling procedures the following questions arise:

(i) How much material has to be taken in each sample?
(ii) At what distances should the samples be taken?
(iii) How must the sampled material be processed further (technically and statistically)?

Usually two conflicting points of view occur. With costs in mind, as few samples as possible should be taken, on the other hand, a high statistical accuracy can only be achieved with many samples. Often a compromise is found by prescribing the statistical accuracy and then deciding whether the number of samples required and the expense of their processing is economically justifiable.

Given this, we proceed as follows, in the course of which we answer the first question asked above. To be more precise: Having a rough knowledge of μ and σ^2 from the past we determine the size of each sample to be taken from the material in flow. This size must be chosen large enough to reduce the short-duration, irregular variation of the considered characteristic to a prescribed degree.

2. The Stochastic Model

Our aim now is to determine the content of some contaminating substance in a flow of raw material, e.g., sulphur in limestone.

This content (expressed as percentage of weight) is interpreted as a random variable with expectation μ and variance σ^2 where the last quantity depends on the amount of reference material. This means that, for example, 100 g-samples exhibit a much higher variability in the amount of the considered characteristic present, than samples of 100 kg mass which form the mean of 1,000 small portions.

To describe the inhomogeneity of the raw material in a model and to make conclusions concerning the size of a single sample we assume the following:

(i) The content of the contaminating substance can be interpreted for a certain amount of reference material B (say 100 kg $\leq B \leq$ 1,000 kg) as random variable Y_t with fixed expectation μ and fixed variance σ^2.
(ii) The sequence of these random variables in the material flow forms a weak stationary stochastic process.

The values of μ and σ^2 characterize the stochastic behaviour of an amount B of reference material but in no way describe the state of smaller portions of material, say laboratory samples of only a few grams weight. Under the (worst case) assumption that in such small samples the proportion of contaminating substance may vary between 0 % and 100 % the following model seems to be adequate.

The content, X_t, of the contaminating substance related to 1 g mass of material drawn in a sample at time t forms a random variable with

$$X_t = Y_t \cdot Z_t, \tag{1}$$

where Z_t is a multiplicative random disturbance, which is stochastically independent of Y_t and Z_τ for $\tau \neq t$. Its expectation is assumed to be $EZ_t = 1$ and its variability is "much higher" than σ^2, i.e., Var $Z_t = \sigma_1^2 \gg \sigma^2$. Note that $EZ_t = 1$ implies $EX_t = E(Y_t \cdot Z_t) = EY_t \cdot EY_t = \mu$.

Under these assumptions as a first step an upper boundary for σ_1^2 can be calculated:

A good approximation to the support of the probability distribution of Z_t may be assumed to be the interval $[0, \frac{100}{\mu}]$. ($Z_t = 0$ yields material without any contaminations, $Z_t = \frac{100}{\mu}$ material consisting only of contaminating substance).

To deduce an upper boundary for σ_1^2, Z_t is replaced by an extreme variable $\hat{Z}_t$ with $P(\hat{Z}_t = 0) = p$, $P(\hat{Z}_t = \frac{100}{\mu}) = 1-p$, $0 < p < 1$, where $P(E)$ denotes the probability of the event E. The condition $EZ_t = 1$ then gives $p = 1 - \frac{\mu}{100}$ and it follows that

$$\text{Var}\,\hat{Z}_t = \left(1 - \frac{\mu}{100}\right) \cdot 1 + \left(1 - \frac{100}{\mu}\right)^2 \cdot \frac{\mu}{100} = \frac{100}{\mu} - 1 \quad (2)$$

By construction, $\text{Var}\,Z_t \leq \text{Var}\,\hat{Z}_t$ holds. A rough idea of the value of μ then results in an upper boundary of σ_1^2. Let, for example, $\mu = 0.02$ (i.e., B kg of material contain an average of 0.02 % of contaminating material), then it follows that

$$\text{Var}\,Z_t \leq \frac{100}{0.02} - 1 \approx 5{,}000. \quad (3)$$

3. Derivation of Results

Drawing a sample means to take a cluster of N consecutive 1-g-portions out of the material stream. The correlations of the constituent Y_t, Y_s of (1) are nearly 1, i.e.,

$$\text{corr}\,(Y_t, Y_s) \approx 1.$$

The content of the contaminating substance in the sample then can be described by the random variable

$$\frac{1}{N} \sum_{t=1}^{N} X_t.$$

Our aim now is to determine the number N such that the short duration irregular disturbances Z_t cause a negligible effect on the variability of the characteristic in the sample. To do so we proceed by

$$\begin{aligned} \text{Var}\,X_t &= E\,Y_t^2 Z_t^2 - (EY_t\,EZ_t)^2 \quad (4)\\ &= (\sigma^2 + \mu^2)(\sigma_1^2 + 1) - \mu^2\\ &= \sigma^2(\sigma_1^2 + 1) + \mu^2\sigma_1^2. \end{aligned}$$

$$\begin{aligned} \text{Cov}\,(X_t, X_s) &= E(Y_t Y_s)\,E(Z_t Z_s) - \mu^2, \quad (5)\\ \text{Cov}\,(Y_t, Y_s) &\approx \sigma^2. \end{aligned}$$

(4) and (5) yield the variance of the contaminating substance in the sample

$$\text{Var}\left(\frac{1}{N}\sum_{t=1}^{N} X_t\right) = \frac{1}{N^2}\left(\sum_{t=1}^{N} \text{Var}\,X_t + \sum_{t<s} \text{Cov}(X_t, X_s)\right)$$

$$= \frac{1}{N}\left(\sigma^2\sigma_1^2 + \mu^2\sigma_1^2 + \sigma^2\right) + \frac{N-1}{N}\sigma^2$$

$$= \sigma^2\left(1 + \frac{\sigma_1^2(\sigma^2 + \mu^2)}{N\sigma^2}\right).$$

The assumption that the short duration disturbance Z_t may result in an increase in variance of at most $\alpha \cdot 100\,\%$ ($\alpha > 0$) leads to the following inequality

$$\frac{\sigma_1^2(\sigma^2 + \mu^2)}{N\sigma^2} \leq \alpha \text{ or equivalently}$$

$$N \geq \frac{\sigma_1^2}{\alpha}\left(1 + \frac{\mu^2}{\sigma^2}\right).$$

For the above example with $\mu = 0.02$ and variance $\sigma^2 = 5 \cdot 10^{-5}$ together with the upper boundary of σ_1^2 in (3) this results in

$$N \geq 5 \cdot 10^4\left(1 + \frac{4 \cdot 10^{-4}}{5 \cdot 10^{-5}}\right) = 45 \cdot 10^4\,\text{g}.$$

This means that each sample has to consist of at least 450 kg of material.

For other values of μ and σ^2 the results can be obtained in an analogous way.

4. Remarks

Questions (ii) and (iii) of Section 1 can be answered by standard procedures of statistical quality control (see [1], [3], [4]).

The derivation of a subsampling procedure with minimum costs and prescribed accuracy will be treated in a forthcoming paper [2].

References

[1] Cochran, W.G., "Stichprobenverfahren", W. de Gruyter, Berlin, New York (1972)

[2] Mathar, R., "An optimal subsampling procedure for a class of stationary processes", Submitted manuscript.

[3] Stange, K., "Stichprobenverfahren bei der Beurteilung von Massengütern, insbesondere von Kohle", Mitteilungsblatt für Math. Statistik, 6, pp. 204—220 (1954).

[4] Stange, K., "Über Probenahme bei Massengütern", Qualitätskontrolle, 4, pp. 25—28 (1953).

Acknowledgements

This analysis was carried out at the Institut für Statistik und Wirtschaftsmathematik an der RWTH Aachen (Department of Statistics and Economical Mathematics at the Technical University Aachen), Federal Republic of Germany.