ELSEVIER

Computer Physics Communications 117 (1999) 35–43

Computer Physics Communications

Analysis of variance designs for model output

Michiel J.W. Jansen [1]

Centre for Biometry Wageningen, P.O. Box 16, 6700 AA Wageningen, The Netherlands

Received 22 April 1998

Abstract

A scalar model output Y is assumed to depend deterministically on a set of stochastically independent input vectors of different dimensions. The composition of the variance of Y is considered; variance components of particular relevance for uncertainty analysis are identified. Several analysis of variance designs for estimation of these variance components are discussed. Classical normal-model theory can suggest optimal designs. The designs can be implemented with various sampling methods: ordinary random sampling, latin hypercube sampling and scrambled quasi-random sampling. Some combinations of design and sampling method are compared in two small-scale numerical experiments. © 1999 Elsevier Science B.V.

Keywords: Variance-based, regression-free, uncertainty analysis; Experimental design; Latin hypercube sampling; Scrambled quasi-random sampling

1. Introduction

We study a scalar output Y of a deterministic model, $Y = f(X_1 \ldots X_k)$, in which $X_1 \ldots X_k$ are stochastically independent inputs or groups of inputs. The groups X_i may have different sizes. It will be assumed that Y has finite mean and variance.

Under these assumptions, output Y can be decomposed into mean, main effects and interactions up to order k,

$$Y = \mu + \sum_i \epsilon_i + \sum_{i<j} \epsilon_{ij} + \cdots , \tag{1}$$

in which ϵ_i depends on X_i; ϵ_{ij} on X_i and X_j, etcetera [3,16]. The ϵ's have zero mean, and variances σ_i^2, σ_{ij}^2, $\ldots$; they are uncorrelated but need not be independent. The variance is composed as

$$Var[Y] = \sum_i \sigma_i^2 + \sum_{i<j} \sigma_{ij}^2 + \cdots . \tag{2}$$

[1] E-mail: m.j.w.jansen@cpro.dlo.nl

PII S0010-4655(98)00154-4

Formulas (1) and (2) are called the analysis of variance (ANOVA) decomposition of Y.

With respect to input group X_i, two variance components are particularly relevant for uncertainty analysis: the *main-effect variance* of X_i and what will be called the *all-effects variance* of X_i. The main-effect variance of X_i, $\mathrm{MEV}[X_i]$ for short, is defined as the expected variance reduction when X_i would become perfectly known, that is,

$$\mathrm{MEV}[X_i] = \mathrm{Var}[Y] - E[\mathrm{Var}[Y|X_i]] \,. \tag{3}$$

By the well-known relation, $\mathrm{Var}[Y] = E[\mathrm{Var}[Y|X_i]] + \mathrm{Var}[E[Y|X_i]]$, MEV can be written as

$$\mathrm{MEV}[X_i] = \mathrm{Var}[E[Y|X_i]] \,, \tag{4}$$

and under the independence assumptions mentioned, MEV is equal to

$$\mathrm{MEV}[X_i] = \sigma_i^2 \,. \tag{5}$$

The all-effects variance of X_i, $\mathrm{AEV}[X_i]$ for short, is defined as the expected variance remaining as long as X_i stays unknown, whereas all other inputs, which will be collectively denoted by $X_{(i)}$, are perfectly known; that is,

$$\mathrm{AEV}[X_i] = E[\mathrm{Var}[Y|X_{(i)}]] \,. \tag{6}$$

This variance component can also be written as

$$\mathrm{AEV}[X_i] = \mathrm{Var}[Y] - \mathrm{Var}[E[Y|X_{(i)}]] \,. \tag{7}$$

Under the assumed input group independence, AEV is equal to the sum of all variances σ^2 in the right-hand side of (2) in which subscript i is present, or more formally,

$$\mathrm{AEV}[X_i] = \sum_{i \in z} \sigma_z^2 \,. \tag{8}$$

It follows from (3) and (6) that $\mathrm{MEV}[X_i]$ and $\mathrm{AEV}[X_{(i)}]$ are complementary,

$$\mathrm{Var}[Y] = \mathrm{MEV}[X_i] + \mathrm{AEV}[X_{(i)}] \,. \tag{9}$$

Moreover, (5) and (8) lead to the inequality

$$\mathrm{MEV}[X_i] \leq \mathrm{AEV}[X_i] \,. \tag{10}$$

The variance components MEV and AEV are known under many names: importance measures, sensitivity estimates, global sensitivity indices, top and bottom marginal variances; the classical correlation ratio is the main effect variance as fraction of the total variance [16,7,5,8,18].

Designs that allow the estimation of variances with a reasonable accuracy tend to be large, so there is a need to construct efficient designs, e.g. [8,18]. The efficiency of a design to estimate these variance components depends on the model, the selected output and the distributions of the inputs. By that fact, general conclusions are hard to obtain. In this paper we discuss the case that ϵ_i, $\epsilon_{ij}, \ldots$ are independently normally distributed. This case can be studied with classical ANOVA theory for random effect models, which has a rich literature, e.g. [14]. Although normality will not often arise in the uncertainty analysis of a model, it is hoped that the efficiency properties derived will be robust, and that a normal analysis may be helpful in suggesting promising designs.

Numerical experiments with the designs proposed can be conducted with several sampling methods, ordinary random sampling being the most obvious candidate, while latin hypercube sampling [9,11,17] and quasi-random sampling [10] constitute possibly more efficient alternatives. The purpose of the alternative methods is to obtain more accurate estimates by a more even covering of the sample space. The problem with quasi-random sampling is that it is actually non-random, which precludes statistical assessment of the estimation accuracy. But the recent invention by Owen [12] of randomly scrambled quasi-random sampling may surmount this problem.

Many theoretical points can be discussed when one discerns only two input groups; designs for that situation are discussed in Section 2. Section 3 continues with designs for more than two input groups. Some analytical normal-theory results follow in Section 4. Next, in Section 5, alternatives to ordinary random sampling are briefly mentioned. Section 6 discusses some experimental results for a non-normal case. The paper closes with a discussion in Section 7.

2. Designs for two independent input groups

Denote the inputs under study by U and the complementary inputs by V. For instance, with $k = 5$, $U = \{X_2, X_5\}$ and $V = \{X_1, X_3, X_4\}$. With some abuse of notation, we denote the model output studied by $f(U, V)$.

2.1. Nested design

According to (9), the total variance Var[Y] is equal to the sum of MEV[U] and AEV[V]. These variances can be estimated without bias from a nested design that has an $m \times n$ data matrix,

$$Y_{i,j} = f(U^{(i)}, V^{(i,j)}) \quad (i = 1 \ldots m;\ j = 1 \ldots n). \tag{11}$$

We indicate independent draws from a random input by using different superscripts between round brackets. For instance, $V^{(1,1)}$, $V^{(1,2)}$ and $V^{(2,1)}$ denote independent draws of V. The design can be analyzed with standard ANOVA, e.g. [6]. A similar design, with U and V interchanged, allows estimation of MEV[V] and AEV[U].

2.2. Combi-nested design

Sobol' [16] proposed a combination of two nested designs with $n = 2$ to estimate MEV and AEV of U and V. The resulting design has an $m \times 3$ data matrix defined by

$$\begin{aligned} Y_{i,1} &= f(U^{(i,1)}, V^{(i,1)}), \\ Y_{i,2} &= f(U^{(i,1)}, V^{(i,2)}), \\ Y_{i,3} &= f(U^{(i,2)}, V^{(i,1)}). \end{aligned} \tag{12}$$

The analysis consists of ANOVA of the two constituent nested designs.

2.3. Alternating design

Unbiased estimation of MEV and AEV of U and V is also possible with the following alternating design. When going through its $m \times 2$ data matrix in reading order, one encounters in turn new draws of U and V,

$$Y_{i,1} = f(U^{(i)}, V^{(i)}), \qquad Y_{i,2} = f(U^{(i)}, V^{(i+1)}) \qquad (i = 1 \ldots m). \tag{13}$$

The total variance, Var[Y], can be estimated unbiasedly by the variance of either column. The all-effects variance of V is estimated without bias by the mean of the series $\frac{1}{2}(Y_{i,1} - Y_{i,2})^2$. Consecutive terms of this series are correlated, whereas terms further apart are independent. Thus, the variance of the mean of the series can be estimated in the same way as the variance of the mean of a moving average time-series. Similarly, the all-effects variance of U is estimated by the mean of the series $\frac{1}{2}(Y_{i,2} - Y_{i+1,1})^2$. The main effect variance MEV[U] can be estimated as the complement of AEV[V] with respect to Var[Y].

2.4. Crossed design

Obviously, unbiased estimation of MEV and AEV of U and V is possible via ANOVA of a crossed design [6] that has the data matrix

$$Y_{i,j} = f(U^{(i)}, V^{(j)}) \quad (i = 1 \ldots m;\ j = 1 \ldots n)\,. \tag{14}$$

For large m and n the crossed design soon becomes inefficient, because the interaction variance σ^2_{UV} is estimated with much better accuracy than the main effect variances σ^2_U and σ^2_V. One should use replicated crossed designs instead.

3. Designs for more than two independent input groups

With increasing number of input groups, replicated crossed designs, even at two levels per input, soon become inefficient. Combi-nested designs for each input group are more promising [16,18].

3.1. Winding stairs design.

A winding stairs design [5] allows estimation of MEV and AEV of all input groups considered. When going through the matrix of this design in reading order, one encounters cyclically new draws of $X_1, X_2 \ldots X_k$,

$$Y_{i,j} = f(X_1^{(i+\theta(j-1))}, X_2^{(i+\theta(j-2))}, \ldots, X_k^{(i+\theta(j-k))}) \tag{15}$$

($i = 1 \ldots m$; $j = 1 \ldots k$). The function θ is the unit jump function: $\theta(s) = 0$ if $s < 0$, and $\theta(s) = 1$ otherwise. With three input groups, for instance, the first few rows of the data matrix have the form

$$\begin{array}{lll}
f(X_1^{(1)}, X_2^{(1)}, X_3^{(1)})\,, & f(X_1^{(1)}, X_2^{(2)}, X_3^{(1)})\,, & f(X_1^{(1)}, X_2^{(2)}, X_3^{(2)})\,, \\
f(X_1^{(2)}, X_2^{(2)}, X_3^{(2)})\,, & f(X_1^{(2)}, X_2^{(3)}, X_3^{(2)})\,, & f(X_1^{(2)}, X_2^{(3)}, X_3^{(3)})\,, \\
f(X_1^{(3)}, X_2^{(3)}, X_3^{(3)})\,, & f(X_1^{(3)}, X_2^{(4)}, X_3^{(3)})\,, & f(X_1^{(3)}, X_2^{(4)}, X_3^{(4)})\,, \\
f(X_1^{(4)}, X_2^{(4)}, X_3^{(4)})\,, & f(X_1^{(4)}, X_2^{(5)}, X_3^{(4)})\,, & f(X_1^{(4)}, X_2^{(5)}, X_3^{(5)})\,.
\end{array}$$

Two columns of a winding stairs sample constitute an alternating design. Column 3 and 5, for instance, of a winding stairs sample with $k > 6$ sources constitute an alternating design for $U = \{X_4, X_5\}$ and $V = \{X_6 \ldots X_k, X_1 \ldots X_3\}$. Thus, a winding stairs design allows estimation of MEV and AEV of all input groups X_i, and also of pools of groups that are adjacent in the cycle of new draws.

4. Some analytical results

This section treats some efficiency results about nested and alternating designs, under the assumption that the ϵ's in the ANOVA-decomposition (1) are normally distributed.

4.1. Nested design

In the nested design (11), $Y_{i,j}$ can, according to (1) and (2), be considered as

$$Y_{i,j} = \mu + \eta_i + \theta_{i,j} ,$$

in which the η's and θ's are uncorrelated (but in general not independent) with means 0 and variances $\tau^2 \equiv \sigma_u^2$ and $\phi^2 \equiv \sigma_v^2 + \sigma_{uv}^2$ respectively. Each element $Y_{i,j}$ has variance $\tau^2 + \phi^2$, the covariance of two elements in the same row equals τ^2.

The nested design can be used to estimate MEV$[U] = \tau^2$ and AEV$[V] = \phi^2$ with standard ANOVA, e.g. [6]. The mean squares for row and row.column are independently distributed proportional to two χ^2-distributions,

$$\mathrm{MS}_{\mathrm{row}} \sim (n\tau^2 + \phi^2)\chi^2_{m-1}/(m-1) ,$$
$$\mathrm{MS}_{\mathrm{row.col}} \sim \phi^2 \chi^2_{m(n-1)}/(m(n-1)) ,$$

implying for the expected mean squares and degrees of freedom that

$$E[\mathrm{MS}_{\mathrm{row}}] = n\tau^2 + \phi^2 , \qquad \mathrm{DF}_{\mathrm{row}} = m - 1 ,$$
$$E[\mathrm{MS}_{\mathrm{row.col}}] = \phi^2 , \qquad \mathrm{DF}_{\mathrm{row.col}} = m(n-1) .$$

Accordingly, we have the following estimates:

$$\hat{\phi}^2 = \mathrm{MS}_{\mathrm{row.col}} ,$$
$$\hat{\tau}^2 = (\mathrm{MS}_{\mathrm{row}} - \mathrm{MS}_{\mathrm{row.col}})/n .$$

The total mean square underestimates the total variance $\tau^2 + \phi^2$:

$$E[\mathrm{MS}_{\mathrm{tot}}] = \tau^2 + \phi^2 - [(n-1)/(mn-1)]\tau^2 ;$$

but the bias tends to zero when the number of rows, m, tends to infinity, at a fixed number of columns, n. It may happen that $\hat{\tau}^2$, as calculated above, assumes a negative value. In that case the original unbiased estimates may be replaced by $\hat{\tau}^2 = 0$ and $\hat{\phi}^2 = \mathrm{MS}_{\mathrm{tot}}/\mathrm{DF}_{\mathrm{tot}}$. Uncertainty analysis via ANOVA of a nested design is also discussed in [8], with a slightly different solution for the problem of possibly negative variance estimates.

The efficiency of the nested design has been treated by Robertson [13]. The author treats the estimation of the correlation ratio or fraction of variance accounted for by rows, that is, the quantity

$$r = \tau^2/(\tau^2 + \phi^2) ,$$

which is estimated by substituting the estimates $\hat{\tau}^2$ and $\hat{\phi}^2$. The resulting estimate for r is distributed as $\hat{r} \sim (\alpha z_1 + \beta z_2)/(\gamma z_1 + \delta z_2)$, in which z_1 and z_2 are independently χ^2-distributed, while the degrees of freedom and the coefficients $\alpha \ldots \delta$ depend on the values of ϕ^2 and τ^2, and on the design parameters m and n according to the previous formulas. The variance of $\hat{r}$ may be shown to be approximately equal to

$$\mathrm{Var}(\hat{r}) \approx 2[1 + (n-1)r]^2(1-r)^2/[n(n-1)(m-1)] .$$

Robertson [13] derived that, with a fixed number of function evaluations mn, this approximation of the variance of $\hat{r}$ is minimal when the number of columns is equal to $n^* = 1/r$. Or rather, an integer number close to n^* is optimal. Obviously, one should have some advance knowledge of the magnitude of r in order to be able to derive a design that is not too far from optimal.

4.2. Alternating design

The all-effects variance of V is estimated without bias by the mean of the series $d_i \equiv \frac{1}{2}(Y_{i,1} - Y_{i,2})^2$ from (13). Consecutive terms of this series are correlated, whereas terms further apart are independent. Thus, the correlation structure of the series is as of a first-order moving average time-series. Each term d_i is distributed as $(\sigma_V^2 + \sigma_{UV}^2)\chi_1^2$. Consecutive terms have correlation $\rho = \frac{1}{4}(\sigma_V^2/(\sigma_V^2 + \sigma_{UV}^2))^2$, whereas terms further apart are independent. Thus, the mean, M_{alt} say, of $d_1 \ldots d_m$ has variance

$$\begin{aligned}\text{Var}[M_{\text{alt}}] &= [(\sigma_V^2 + \sigma_{UV}^2)^2/m][1 + 2\rho(m-1)/m] \\ &= [(\sigma_V^2 + \sigma_{UV}^2)^2 + \sigma_V^2(m-1)/(2m)]/m \\ &\approx (\text{AEV}[V]^2 + \tfrac{1}{2}\text{MEV}[V]^2)/m. \end{aligned} \tag{16}$$

4.3. Relative efficiency: combi-nested and alternating design

The mean, M_{cn} say, of the series $d_i \equiv \frac{1}{2}(Y_{i,1} - Y_{i,2})^2$ of a combi-nested design (11) has variance

$$\text{Var}(M_{\text{cn}}) = (\sigma_V^2 + \sigma_{UV}^2)^2/m = \text{AEV}[V]^2/m. \tag{17}$$

Thus, if one would only want to estimate AEV[V], and if one would restrict the combi-nested design to its first two columns, the relative efficiency of alternating versus combi-nested satisfies, because of (16) and (17),

$$\text{eff} = \text{AEV}[V]^2/(\text{AEV}[V]^2 + \tfrac{1}{2}\text{MEV}[V]^2) \geq 2/3,$$

with equality only if AEV[V] = MEV[V], i.e. if there is no interaction (cf. Eqs. (5), (8), (10)). The same applies, with interchange of U and V, for estimation of AEV[U].

However, if one wants to estimate both AEV[V] and AEV[U], the story becomes different, because combi-nested will then need a third column, which reduces the efficiency of combi-nested by a factor 2/3. Then, an alternating design is at least as efficient as a combi-nested one; with equal efficiency if and only if there is no interaction. Thus, under normality, an alternating design is at least as efficient as a combi-nested design.

5. Alternatives to ordinary random sampling

Up to now, it was assumed that the successive input vectors $X_i^{(j)}$ are independent, which corresponds to ordinary random sampling, OR for short. (Remember that $i = 1 \ldots k$ indicates the input vector, while $j = 1, 2 \ldots$ indicates successive draws). Quite often, the accuracy of the estimations may be improved by alternative sampling methods such as latin hypercube sampling and quasi-random sampling, which have been devised to obtain a more even covering of the sample space. These alternative sampling methods can be applied if for each X_i a random draw can be realized by a random draw from a homogeneous distribution on a unit-cube of the same dimension as X_i, followed by some transformation. Because of the independence of the X's, this implies that a full set, $X_1 \ldots X_k$, of input vectors can also be realized by a draw from a multidimensional unit-cube followed by a transformation. The alternative sampling methods deal with sampling from this latter unit-cube. Some algorithms for quasi-random sampling impose restrictions on the dimension of the unit-cube, e.g. [1], which may hamper the uncertainty analysis of models with high-dimensional input.

Latin hypercube sampling (LH), a form of restricted random sampling, is well known in uncertainty analysis [9,17,11]. In this paper we will use LH in conjunction with the method of Iman and Conover [4] to suppress spurious correlations. The repeatability of LH estimates of variance components can be assessed in the statistical way by replication of sampling and estimation procedure; but systematic error, bias, is not detected in this way.

Quasi-random sampling (QR) yields a sample of points in the unit-cube that are, in a sense, maximally avoiding each other, and thus fill the cube very evenly, e.g [10]. The problem with QR sampling is that it is actually non-random, which impedes the statistical assessment of estimation accuracy. But the recent invention by Owen [12] of randomly scrambled QR samples may overcome this problem. When it is scrambled, a QR sample retains its desired equidistribution properties. In this paper we use scrambled Sobol' (SS) samples [15]. The algorithm [1] used for Sobol' samples is restricted to dimensions up to 40.

Since the approach is rather new, we will briefly sketch how a Sobol' sample is scrambled; see [12] for more details. The individual coordinates of the sample take values in the interval $[0, 1)$. The values of each coordinate are independently scrambled in the following way. First the interval $[0, 1)$ is divided into two slices $[0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$; their contents are interchanged or not with probability $\frac{1}{2}$. Subsequently, each half is divided into two quarts, which are interchanged, *within their half*, with probability $\frac{1}{2}$; independently for each half. Then the quarts are splitted into two parts which are independently interchanged *within the quarts*; and so on. Since the scrambled samples have the same equidistribution properties as the original samples, it is possible to reach the superior accuracy of original Sobol' samples, while the repeatability can be assessed by random replication of the scrambling. As with LH sampling, bias is not detected in this way.

6. Some experimental results

The purpose of this section is to investigate experimentally, on a small scale, the robustness of the efficiency results that were derived in Section 4 under the assumption that model output and its ANOVA decomposition (1) are normally distributed. Instead of a complex model, we study a simple test function that was chosen to be lognormal rather than normal,

$$f(U, V) = e^U e^V .$$

The inputs U and V are independent normal, with mean 0 and variances log(2) and log(2.5), respectively. Thus, e^U and e^V are lognormal, as well as their product $f(U, V)$. From well-known formulas for the mean and variance of the lognormal distribution, it can be derived that the variance of $f(U, V)$ can be decomposed into $\sigma_u^2 = 5$, $\sigma_v^2 = 7.5$ and $\sigma_{uv}^2 = 7.5$.

The first experiment compares the efficiency of a combi-nested design (12) and an alternating design (13). According to the (normal) theory of Section 4, the alternating design should be at least as efficient as the combi-nested.

The second experiment was performed to compare the efficiency of a two-column and a four-column nested design (11). According to Section 4, the optimal number of columns for estimating the correlation ratio r would be four, since $r = \sigma_u^2/(\sigma_u^2 + \sigma_v^2 + \sigma_{uv}^2) = \frac{1}{4}$. So one would expect that the two-column design is less efficient than the four-column design.

Each design was implemented with three sampling methods: scrambled Sobol' (SS), latin hypercube (LH) and ordinary random (OR); which leads to 6 types of analysis when two designs are compared.

6.1. Combi-nested versus alternating design

The number of rows was equal to 1024 in both designs, so the number of 'model runs' per analysis is 3×1024 for the combi-nested design and 2×1024 for the alternating design. Each of the 6 types of analysis was performed in 30 batches of 1000 analyses. MEV and AEV were estimated per analysis. Per batch of 1000, the means and variances of the estimates were kept. The means were analyzed in order to detect possible biases, while the variances were analyzed in order to compare efficiencies. We corrected for difference in number of function evaluations by multiplying the variances of the alternating design with 2/3. It would have been more sophisticated to compare samples of the same size, but that was precluded by the fact that the software used for scrambled Sobol' sampling forced us to use an integer power of 2 as number of rows. The significance tests performed were: ANOVA on ranks with sampling method and design as treatments; and the Wilcoxon rank sum test to compare different combinations of design and sampling method. Table 1 gives the estimated coefficient of variation, between random replications of the individual analyses, of the estimates of MEV$[U]$.

Table 1
Coefficient of variation of $\widehat{\text{MEV}}[U]$

	SS	LH	OR
combi-nested	0.29	0.49	0.55
alternating	0.28	0.38	0.41

Table 2
Coefficient of variation of $\hat{r}$

	SS	LH	OR
2 columns	0.20	0.29	0.30
4 columns	0.15	0.21	0.23

The further analysis of this particular case leads to the following conclusions:

(a) SS performs much better than LH and OR, despite some small but statistically significant bias. The difference between LH and OR is smaller.

(b) The bias in SS estimates is an order smaller than the random variation. For instance, MEV$[U]$ has an estimated relative bias of -0.04, whereas the coefficient of variation is estimated as 0.29 for combi-nested and 0.28 for alternating.

(c) The alternating design is more efficient than the combi-nested design with LH and OR sampling. With SS, however, the advantage is not significant.

(d) The coefficient of variation of the estimates is large anyhow: for the estimate of MEV$[U]$, for instance, it ranged from 28% to 55%.

6.2. *Two versus four columns in a nested design*

The number of rows was 1024 in the two-column analyses and 512 in the four-column analyses, so that the number of 'model runs' was equal. Each of the 6 types of analysis was performed in 40 batches of 1000 analyses. The correlation ratio was estimated per analysis. Per batch of 1000, the means and variances of the estimates were kept. The method of analysis was the same as in the previous experiment. Table 2 gives the estimated coefficient of variation, between random replications of the individual analyses, of the estimates of r.

The analysis lead to similar conclusions as in the previous experiment:

(a) In this particular case too, SS performs much better than LH and OR. The difference between LH and OR is less pronounced, but statistically significant.

(b) The bias in SS estimates is much smaller than the random variation. For instance, with the two-columns design the estimate of r has an estimated relative bias of $+0.04$, whereas its coefficient of variation is estimated as 0.20.

(c) As expected, the four-column design is more efficient than the two-column design.

(d) The coefficient of variation of the estimate of the correlation ratio r ranges from 15% to 30%.

7. Discussion

Analytical and, admittedly anecdotal, experimental findings indicate that worthwhile efficiency improvements may in principle be obtained by judicious choice of analysis of variance design and sampling method. In general, scrambled quasi-random sampling, rather than latin hypercube or ordinary random sampling, may have a beneficial effect. A disadvantage of scrambled quasi-random and latin hypercube sampling is that accuracy has to be assessed by replication, whereas with ordinary random sampling the accuracy can be assessed from a single experiment.

The Fourier Amplitude Sensitivity Test has not been treated, although it permits estimation of the same variance components [2,18]. This design is of an altogether different type; moreover, it is non-random, which demands other, non-statistical, methods to assess accuracy.

Under the assumption of normality, the accuracy of the estimators can be calculated analytically as function of the variances to be estimated. The problem is that these variances are not known in advance. Thus, the possibility to extend designs according to results of intermediate analyses constitutes an ingredient of efficiency. Most designs mentioned can be extended in width and in length. In width, by starting with a sample of pooled sources, and splitting only those pools that appeared to be important. In length, by adding rows to the data matrix if it appeared that some estimates are not yet sufficiently accurate.

Efficiency is a clear concept when there is only one estimand of interest, like in the case of the nested design, where it was possible to pinpoint the optimal design. When the number of estimands is large, however, it is not easy to formulate an adequate optimality criterion. By this fact, it is hard to make general statements about optimality.

References

[1] P. Bratley, B.L. Fox, Algorithm 659: implementing Sobol's quasirandom sequence generator, ACM Trans. Math. Software 14 (1988) 88–100.
[2] R.I. Cukier, H.B. Levine, K.E. Shuler, Nonlinear sensitivity analysis of multiparameter model systems., J. Comput. Phys. 26 (1978) 1–42.
[3] B. Efron, C. Stein, The jackknife estimate of variance, Ann. Stat. 9 (1981) 586–596.
[4] R.L. Iman, W.J. Conover, A distribution-free approach to inducing rank correlation among input variables, Commun. Statist. Simul. Comput. 11 (1982) 311–334.
[5] M.J.W. Jansen, W.A.H. Rossing, R.A. Daamen, Monte Carlo estimation of uncertainty contributions from several independent multivariate sources, in: J. Grasman, G. Van Straten, eds., Predictability and Nonlinear Modelling in Natural Sciences and Economics (Kluwer, Dordrecht, 1994) pp. 334–343.
[6] P.W.M. John, Statistical Design and Analysis of Experiments (MacMillan, New York, 1971).
[7] B. Krzykacz, SAMOS: a computer program for the derivation of empirical sensitivity measures of results from large computer models, GRS-A-1700, Gesellschaft für Reaktorsicherheit, Garching (1990).
[8] M.D. McKay, Variance-based methods for assessing uncertainty importance in NUREG-1150 analyses, Los Alamos National Laboratory, LA-UR-96-2695, Los Alamos (1996).
[9] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (1979) 239–245.
[10] H. Niederreiter, Random Number Generation and Quasi-Monte Carlo Methods (SIAM, Philadelphia, PA, 1992).
[11] A.B. Owen, A central limit theorem for latin hypercube sampling, J.R. Statist. Soc. B 54 (1992) 541–551.
[12] A.B. Owen, Scrambled net variance for integrals of smooth functions, Ann. Statist. 25 (1997) 1541–1562.
[13] A. Robertson, Experimental design in the evaluation of genetic parameters, Biometrics 15 (1959) 219–226.
[14] S.R. Searle, G. Casella, C.E. McCulloch, Variance Components (Wiley, New York, 1992).
[15] I.M. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals, USSR Comput. Math. Phys. 7 (1967) 86–112.
[16] I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, Matematicheskoe Modelirovanie 2 (1990) 112–118 [in Russian]; Translated in: Math. Model. & Comput. Exp. 1 (1993) 407–414.
[17] M. Stein, Large sample properties of simulations using latin hypercube sampling, Technometrics 29 (1987) 143–151.
[18] S. Tarantola, A. Saltelli, K.P.S. Chan, Quantitative methods for global sensitivity analysis, in: R.M. Cooke, ed., Report on Benchmark Workshop Uncertainty/Sensitivity Analysis Codes, European Safety and Reliability Association, Technical Committee on Uncertainty Modelling (1997).