Identification of Dispersion Effects in 2^k Factorial Design

Viviane Leite Dias de Mattos
Universidade Católica de Pelotas – UCPel,
Pelotas, RS, Brazil
E-mail: vldm@atlas.ucpel.tche.br

Pedro Alberto Barbetta
Departamento de Informática e Estatística,
Universidade Federal de Santa Catarina – UFSC,
Florianópolis, SC, Brazil
E-mail: barbetta@inf.ufsc.br

Dalton Francisco de Andrade
Departamento de Informática e Estatística,
Universidade Federal de Santa Catarina – UFSC,
Florianópolis, SC, Brazil
E-mail: dandrade@inf.ufsc.br

Robert Wayne Samoryl
Departamento de Engenharia de Produção e Sistemas,
Universidade Federal de Santa Catarina – UFSC,
Florianópolis, SC, Brazil
E-mail: samohyl@deps.ufsc.br

Abstract
The identification of dispersion effects is a very important stage in developing robust products and processes. Several methods to identify dispersion effects are present in statistical and quality engineering literature, especially methods which use 2^K or 2^{K-p} unreplicated factorial designs, such as Box-Meyer, Harvey, Brenemann-Nair and Bergman-Hynén methods. In this paper we considered generalizations of these methods for replicated experiments, and compare them by Monte Carlo simulations, analyzing sensitivity and specificity indicators. We also included joint generalized linear models (joint GLMs) in our comparison. The joint GLMs provides an interesting general framework to fit mean and variance models and it is recommend for this proposal, but it needs specialized software. If the main focus is found only in one or two higher effects, then the Box-Meyer method is an efficient and very simple method. When only one non-null dispersion effect is present, our simulation showed that the Box-Meyer method is the best, even when compared with the joint GLMs. When two non-null dispersion effects are present, the Box-Meyer method is biased, but surprisingly our simulation showed that this method works well.

Keywords: robust design engineering, two-level factorial designs, dispersion effects, Monte Carlo simulations

Introduction

Improvement on quality requires a reduction in the variability of the productive process. In the robust design, popularized by Taguchi, one tries to identify controllable factors in the process, which change the variance of functional characteristics of the product. These factors are, whenever possible, allocated in levels, which minimize the variance, thus allowing the development of products and processes robust to variations in environmental conditions and other noises.

Design and analysis of experiments may be carried out to identify factors which change the expected response (*location effects*) or the variance of the response (*dispersion effects*). The classical process to identify dispersion effects is to carry out replicated designs or experiments with planned variations of noise factors simulated in the laboratory. In each design point, formed by the combination of levels of controllable factors, a dispersion measurement is calculated, such as the sample variance or the signal-to-noise ratio. Then, significant dispersion effects are usually identified by analysis of variance or by normal probability plot of effects (Wu and Hamada, 2000; Montgomery, 2005).

There is a great interest in the industry for two-level fractional factorial designs with K factors and p fractions, usually represented by 2^{K-p}. In these designs, there are $n = 2^{K-p}$ design points, and (n – 1) possible (locations or dispersions) effects can be evaluated.

An experimental study might begin with a great number of possible factors. Taking into consideration the cost of the experiment and supposing that few factors affect the response, it is common to perform an unreplicated 2^k or 2^{k-p} factorial design to make a screening of factors to identify apparently significant effects. Usually, these methods consist of building a location model (or mean model) and the residuals of this model are used to identify dispersion effects (see discuss in the next section). However, a more sophisticated statistical model may be applicable: the joint generalized linear models (joint GLMs). These models address simultaneously the mean and the variance of the response.

This paper considers 2^k factorial designs with replications (less than five) and compares several dispersion effect identification methods by the normal probability plot approach. The comparison study is carried out by Monte Carlo simulations. The simulation design and the analysis of their results have been done applying experimental design methodology. The methods are evaluated by *sensitivity* and *specificity* indicators. *Sensitivity* is the capacity to detect non-null effects; and *specificity* is the likelihood of you not identifying null effects as non-null effects. We hope that the results of our simulations will help the researchers to identify some situations where one method is better than others. Based on these results, we describe some recommendations to identify dispersion effects.

In follow we present several methods for identifying dispersion effects in experiments. Some generalizations for replicated experiments are commented. We developed a simulation study to compare the several methods and illustrated them by an example. Finally, we summarize our study and describe some recommendations to identify dispersion effects.

Methods for Replicated Experiments

Consider the experimental data, y_{ij} (i = 1, 2,..., n and j = 1, 2,..., m), satisfying the following linear model (Equations 1 and 2):

$$y_{ij} = x_i^t \beta + \sigma_i e_{ij} \quad (1)$$

$$g\left(\sigma_i^2\right) = z_i^t \theta \quad (2)$$

where $\beta = (\beta_0, \beta_1, \ldots, \beta_{n-1})^t$ and $\theta = (\theta_0, \theta_1, \ldots, \theta_{n-1})^t$ are unknown parameter vectors; x_j^t and z_i^t are the rows of the design matrices, both with dimensions 1 x n; e_{ij} are the experimental errors; g is a link function for the variance; n is the number of design points, and m is the number of replications. Usually it is assumed that e_{ij} are independent and follow a normal distribution with mean zero and variance one. The first element of vector θ in (Equation 2) is associated with the constant term; and the other ones are the dispersion effects.

When the expectation operator is applied in (Equation 1), the model is called *mean or location model*. Model (Equation 2) is called *variance* or *dispersion model*. The *g* function may be an identity, which leads to a *linear model*, or a logarithmic function, which in turn leads to a *log-linear model*. The lather is more common, since the logarithmic transformation produces positive estimates and stabilizes the variance (Bartlett and Kendall, 1946). In a *log-linear model*, one defines the k^{th} *dispersion effect* as half of the logarithm of the ratio of the level coded as +1 and the level coded as –1.

In replicated experiments (m > 1), it is usual to calculate the sample variance in each design point and apply the logarithm over these variances. In a 2^K or 2^{K-p} factorial design, the least squares method applied in the logarithm of the sample variances results in the estimate D_k^s of θ_k (k = 1, 2,..., n – 1) (Equation 3):

$$D_k^S = \frac{1}{n}\left[\sum_{i(k+)} \log(s_i^2) - \sum_{i(k-)} \log(s_i^2)\right] = \frac{1}{n}\log\left(\frac{\prod_{i(k+)} s_i^2}{\prod_{i(k-)} s_i^2}\right) \quad (3)$$

where: s_i^2 the sample variance in the i^{th} design point (i = 1, 2, ..., n); i(k+) represents runs performed on the level coded as +1 of the effect k; and i(k–) performed on the –1 level.

Another estimator of θ_k is based on the ratio between the arithmetic means of the sample variances (Box and Meyer, 1986). It is denoted by (Equation 4):

$$D_k^R = \frac{1}{2}\log\left(\frac{\sum_{i(k+)} s_i^2}{\sum_{i(k-)} s_i^2}\right) \quad (4)$$

Nair and Pregibon (1988) show that the statistics D_k^s is the maximum likelihood estimators of θ_k if the true model is saturated. A model is said to be saturated if all the elements of the vector θ are non-null. On the other hand, D_k^R is the maximum likelihood

estimator of θ_k if model (2) is restricted (that is, θ_k is the only non-null element of θ). Moreover D^R_k is more efficient than D^s_k in the estimation of θ_k if this is the only non-null dispersion effect, but it is a biased estimator when more than one non-null dispersion effect exists. As Nair and Pregibon (1988), we refer to the statistics D^s_k and D^R_k as the S and R methods respectively.

Methods for the Unreplicated Designs

For unreplicated designs (m = 1), the methods to identify dispersion effects are usually based on residuals, r_i (i = 1, 2, ..., n), of a non-saturated mean model adjusted for the observations. However, these methods have the inconvenience of depending on the adjusted mean model (see Pan, 1999).

Harvey (1976) proposed to model variance taking logarithms of the squared residuals. It has been used to identify dispersion effects in unreplicated experiments (Brenneman and Nair, 2001). In 2^k or 2^{k-p} factorial design, this method consists of substituting the sample variances for squared residuals in Equation 3, which is (Equation 5):

$$D_k^H = \frac{1}{n}\left[\sum_{i(k+)} \log(r_i^2) - \sum_{i(k-)} \log(r_i^2)\right] = \frac{1}{n}\log\left(\frac{\prod_{i(k+)} r_i^2}{\prod_{i(k-)} r_i^2}\right) \quad (5)$$

Box and Meyer (1986) proposed to identify dispersion effects replacing the arithmetic means of the squares residuals in (Equation 4), so (Equation 6):

$$D_k^{BM} = \frac{1}{2}\log\left(\frac{\sum_{i(k+)} r_i^2}{\sum_{i(k-)} r_i^2}\right) \quad (6)$$

To identify the apparently significant dispersion effects, D^H_k or D^{BM}_k are usually placed on a normal probability plot (Montgomery 2005, p. 241). Bergman and Hynén (1997) suggested a modification in Box and Meyer's statistic to produce the sums of squares of the numerator and denominator independent, and consequently the ratio of the mean squares must follow an F distribution. They presented several computational ways to calculate this statistic, in particular Equation 7:

$$D_k^{BH} = \frac{\sum_{\ell} r^2_{(k+)\ell}}{\sum_{\ell} r^2_{(k-)\ell}} \quad (7)$$

where $r_{(k+)\ell}\,(\ell = 1, 2, \ldots, n/2)$ are the residuals of the model built with the observations associated to level +1 of the effect k; and similarly $r_{(k-)\ell}$ are the residuals for the observations associated to level –1 of the effect k. Notice that we need to adjust two

independent regressions for each effect. The dispersion effect may be identified by F tests or placing $\log\left(D_k^{BH}\right)$ on a normal probability plot.

Brenneman and Nair (2001) showed that D_k^{BH} may be biased and does not follow an *F* distribution if more than one dispersion effect is non-null. To reduce this bias, they suggested using Bergman and Hynén´s residuals applied in Harvey's formulation, which is (Equation 8):

$$D_k^{BN} = \frac{1}{n}\log\left(\frac{\prod_{\ell} r_{(k+)\ell}^2}{\prod_{\ell} r_{(k-)\ell}^2}\right) \tag{8}$$

Van de Ven (2008) proved that the Brenneman and Nair estimator and two others proposed estimators (not present here) are only different representations of a single estimator for two-level factorial designs.

Generalizations for eeplicated designs

For replicated designs (m > 1), there are m squared residuals, r_{ij}^2 (j = 1, 2, ..., m) in each design point i (i = 1, 2, ..., n). Then all four formulas of the statistics D_k, as described in the last section, may be rewritten in terms of individual squared residuals, r_{ij}^2, or in terms of average of the squared residuals in each design point, which is (Equation 9):

$$\overline{r_i^2} = \frac{\sum_{j=1}^{m} r_{ij}^2}{m} \quad (i = 1, 2, \ldots, n) \tag{9}$$

In terms of the Box and Meyer's statistic, the two procedures are arithmetically equal, as shown below (Equation 10):

$$\begin{aligned} D_k^{BM(individual)} &= \frac{1}{2}\log\left(\frac{\sum_{i(k+)}\sum_{j=1}^{m} r_{ij}^2}{\sum_{i(k-)}\sum_{j=1}^{m} r_{ij}^2}\right) = \frac{1}{2}\left(\log\sum_{i(k+)}\sum_{j=1}^{m} r_{ij}^2 - \log\sum_{i(k-)}\sum_{j=1}^{m} r_{ij}^2\right) = \\ &= \frac{1}{2}\left(\log\sum_{i(k+)} m.\overline{r_i^2} - \log\sum_{i(k-)} m.\overline{r_i^2}\right) = \frac{1}{2}\log\left(\frac{\sum_{i(k+)}\overline{r_i^2}}{\sum_{i(k-)}\overline{r_i^2}}\right) = D_k^{BM(average)} \end{aligned} \tag{10}$$

The extension of this equivalence for the Bergman and Hynén's statistic is straightforward. However the same kind of equivalence is not true for Harvey's or Brenneman and Nair's statistics. Mattos (2004) justified by Monte Carlo simulations that it is better to use the average of the squared residuals in each design point. If the technique is applied directly on individual squared residuals, some squared residuals may be close to zero merely by chance (*inliers*). In this case, the logarithmic transformation causes

discrepant values. The *inliers* are less probable if we use the average of squared residuals (see Barbetta et al., 2000).

If we use saturated mean models, then the Box-Meyer´s method and the R method are equivalent, that is $D_k^{BM} = D_k^R$. Therefore, if we use the average of squared residuals on each design point, then $D_k^H = D_k^S$.

The joint GLMs

The joint generalized linear models (joint GLMs) for mean and variance provides a general framework for testing and modeling both location and dispersion effects. The model given by (Equation 1 and 2) is a particular case, but an important one of joint GLMs. The main advantages of this approach are:

1) use all information of the data, adjusting models for mean and variance simultaneously;
2) the models are defined for any design, not only for the 2^K or 2^{K-p} factorial design;
3) it is not necessary to transform the response to stabilize variance or to obtain a normal distribution of the error terms, since one can consider any probability distribution in the exponential family in the GLM (see McCullagh and Nelder,1989); and
4) one fit a good model by stepwise methods, as the tests of significance of the coefficients are more appropriate than other methods (see Grego, 1993).

The problem with joint GLMs is that they require more specialized software. In the general sense, the identity link function is used for the mean model with normal error. In this situation, the natural choice of the variance model is the gamma distribution and log link function. The parameters of these models may be estimated by iterative weighted least squares (IWLS), using the squares of the standardized residuals as the response in the variance model, and the weight least square (WLS) for the mean model (weights equal $1/\hat{\sigma}_i^2$), $\hat{\sigma}_i^2$ is the predicted variance for the i^{th} design point of the variance model. Engel and Huele (1996) show that little iteration usually performs better than complete convergence, but there isn't an ideal number of iterations because it depends on the variance heterogeneity and the degrees of freedom in residuals.

Grego (1993) and Pinto and Leon (2006) recommend starting the model with all possible locations and dispersion effects (saturated models) and then, the non-significant effects removed by backward elimination, first the variance model and then the mean model.

In replicated experiments, Grego (1993) proposed applying joint GLMs on the pair $(\bar{x}_i, s_i^2)$, where $\bar{x}_i$ and s_i^2 are respectively the sample mean and sample variance in the design point i (i = 1, 2, ..., n). But Lee and Nelder (2003) suggested considering the individual observations in the joint GLMs. Thereby it is possible to make several model-checking plots, although both procedures result in the same point estimates of the elements of β and θ.

Monte Carlo simulation study

The comparative studies of the methods previously discussed were carried out by Monte Carlo simulations. The study is based on simulated experiments generated by a 2^4 factorial design with two or four replicates on each design point. The coded factors of the simulated experiments are denoted by x_1, x_2, x_3 and x_4. All samples are simulated considering that the location and dispersion effects are described respectively by Equations 1 and 2, where the link function of the variance model is the logarithm function, and normal distribution error in (Equation 1). For example, one set of simulated experimental data is generated by $\beta^t = (2, 2, 2, 0, 1, 1, 0, \ldots, 0)$ and $\theta^t = (0, 0.549, 0, 0.896, 0, \ldots, 0)$, which is (equations 11 and 12):

$$y_{ij} = 2x_{1i} + 2x_{2i} + 2x_{3i} + x_{1i}x_{2i} + x_{1i}x_{3i} + \sigma_i^2 e_{ij} \quad (11)$$

$$\log\left(\sigma_i^2\right) = (0.549)x_{2i} + (0.896)x_{4i} \quad (12)$$

where x_1, x_2, x_3 and x_4 are the -1 and +1 coded factors of the 2^4 design (i = 1, 2, ..., 2^4 and j=1, 2) and e_{ij}: N(0, 1). The coefficients 0,549 and 0,896 are the dispersion effects associated with the factors x_2 and x_4, respectively.

The important features to detect dispersion effects are the number of degrees of freedom for estimating the variance function, the number of the true non-null dispersion effects and the magnitude of the dispersion effects (Engel and Huele, 1996). We considered these features in our project of simulation but introduced others for the exploration task. All the factors that we used are described in Table 1.

In the simulation study, two and five location effects (Factor A of Table 1) and one and two dispersion effects (Factor C) are considered, since location effect is more common than the dispersion effects, so more location effects are considered in the simulation

Table 1 – Description of factors and levels used in simulation project.

Factor	Description	Level –1	Level +1
A	Number of non-null location effects	Two main effects: β_1 and β_2	Three main effects and two interactions: β_1, β_2, β_3, β_{12} e β_{13}
B	Magnitude of location effects.	$\beta_i = 1$ and $\beta_{ij} = 0.5$ for i and j equal to indexes specified in factor A. Other elements of β are equal to zero.	$\beta_i = 2$ and $\beta_{ij} = 1$ for i and j equal to indexes specified in factor A. Other elements of β are equal to zero.
C	Number of dispersion effects	θ_k (one dispersion effect), where k = 1 or k =4 (see factor D)	θ_k and θ_2 (two dispersion effects). θ_2 is fixed equal 0.693.
D	Coincidence between location and dispersion effects	$\theta_k = \theta_1$ (the simulated factor x_1 is associated with location effect β_1, and dispersion effect θ_1)	$\theta_k = \theta_4$ (the simulated factor x_4 is associated with only a dispersion effects, θ_4)
E	Magnitude of dispersion effects	$\theta_k = 0.549$ and $\theta_i = 0$ for $i \neq k$ and $i \neq 2$	$\theta_k = 0.896$ and $\theta_i = 0$ for $i \neq k$ and $i \neq 2$
F	Number of replications	2	4

than the dispersion effects. The Factor E controls the magnitude of dispersion effects. The values 0.549 and 0.896 correspond to an increase in variance with a factor three and six, respectively. It might be seen by Equation 12: if x_{2i} is zero, the relation between $x_{4i} = 1$ and $x_{4i} = -1$ produces ratio of the variances equal to six. The choices of the dispersion effects for simulation were based on the review of the practical studies and the empirical analysis of the power of the procedures to detect significant effects.

Note that each crossing levels of the first five factors of Table 1 produces a different specification for mean and variance models. For example, if all factors were put in level +1, the sample will be generated by the mean and variance described respectively by Equations 11 and 12. The levels of the factors of Table 1 were crossed as a 2^{6-1} fractional factorial design. For each one of the $2^{6-1} = 32$ design point we generated N = 2,000 samples (2,000 results of a 2^4 design).

The estimates of dispersion effects θ_k ($k = 1, 2, \ldots, n - 1$) for each simulated sample were obtained by eight different methods, separated by comparable groups. In the first group the residuals are calculated on saturated mean model and we compare the methods: S [expression (3)], R [expression (4)], and the joint GLMs (GLM_s). In the second group the mean model includes terms according to the simulation specification. For example, if the sample was simulated according to expression (11) and (12), the mean model will fit only with x_1, x_2, x_3, x_1x_2 and x_1x_3 terms. In this group we compare H (Harvey), BM (Box and Meyer), BN (Brenemann and Nair) and BH (Bergman and Hynén) methods generalized for replicated design [expressions (5), (6), (8) and (7) adapted for average of the squared residuals in each design point]. Also, we consider the joint GLMs with the mean and variance models specified according to the simulation specification (we denoted it as GLM_t). In all of these methods, the estimation process uses only the results of the simulated samples.

For GLM_s and GLM_t, we adopted the IWLS estimation method. As we discussed before, Engel and Huele (1996) recommend fixing a small number of iterations to prevent the *inlier* problem. But there isn´t an ideal iteration number for all cases, and then we decided to five iterations.

One way to detect significant effect automatically is by Lenth's method, that estimates the (pseudo) standard errors based on robust statistics (see Montgomery, 2005, pp. 234). This method had been used in some cases where the main goal is to detect significant effect from 2_{K-p} unreplicated design, for example in Schoen (2004). But we choose the traditional statistics to sample from normal distribution (the dispersion effects estimates are approximately normal distribution). We identify the effect as apparently significant when its estimate is two standard deviations apart from the average of all effects estimates.

For each sample generated we estimate θ_k by the eight methods discussed before. We denote by D_k one estimate of θ_k. For situations where we need to specify the method we will use a subscript. As we describe before, the k^{th} effect was considered significant when:

$$\left|D_k - \overline{D}\right| > 2S \qquad (13)$$

where $\overline{D}$ is the average of D_k and S is the standard deviation of D_k ($k = 1, 2, \ldots, 2^K - 1$). In both statistics we excluded the two highest absolute values of D_k, because we know by simulation models that they might be truly significant, so they are *outliers* in the set of the effects. We think that the process reproduces automatically what a researcher does when he (or she) analyses a normal probability plot of the effects.

Based on N = 2,000 simulated samples we calculated for each one of the 32 design points and each of the eight methods, the following:

- the proportion of times in which *all* the *non-null* dispersion effects were correctly identified (an estimate of the probability of correct identification, PCI); and
- the proportion of times in which *all* the *null* main or second order dispersion effects were identified correctly; in other words, one minus the proportion of times in which *some null* effects were identified incorrectly as significant (one minus the estimate of the probability of incorrect identification, 1 - PII).

The PCI and 1 – PII may be interpreted as the *sensibility* and the *specificity* of the methods, respectively. The higher these two figures the better the methods.

The Simulation Results

Some caution is necessary to compare the results of the simulated experiments, because in some methods we use information of the simulation models, and in others we don´t use any information. So, we can only compare the R, S and GLM_s methods among themselves, since these methods do not use any information from the simulation models, and H, BM, BN and BH methods among themselves, since these use the same information from the mean model. In GLM_t the true terms of the mean and variance models are used, so the results of this method are not comparable with the others, but for the sake of classification it is placed in the same group with the H, BM, BN and BH methods.

Table 2 shows the averages for the PCI and for the 1 - PII of each method calculated in the 32 design points.

In Table 2, an interesting result is the supremacy of methods which use the ratio of arithmetic (R, BM and BH) means compared to the ones which use geometric means (S,

Table 2 – Estimates of sensibility (PCI) and specificity (1 – PII) of the methods.

Method	PCI	1 - PII
S (Nair and Pregibon)	0.496	0.516
R (Nair and Pregibon)	0.526	0.554
GLM_s	0.491	0.518
H (Harvey)	0.632	0.542
BM (Box and Meyer)	0.636	0.554
BN (Brenemann and Nair)	0.623	0.551
BH (Bergman and Hynén)	0.626	0.571
GLM_t	0.687	0.553

Standard errors of estimates are less than or equal to 0.011.

H and BN). Perhaps this happened because the simulated data were generated with one and two dispersion effects. Nair and Pregibon (1988) and Brenneman and Nair (2001) showed that the R method and their derivates (BM and BH) were appropriate when there is only one dispersion effects, however they produce biased estimates when more than one dispersion effect are present.

On the other hand, the use of modified residuals (as the BH and BN methods), suggested by Bergman and Hynén (1997), and Brennemann and Nair (2001), do not bring relevant gains if the identification of the effects is made by a normal probability plot of effects, as our simulation approach. We note also that the use of GLM with the saturated mean and variance models (GLM_s) performs worse than the simple R method, and approximately equal to the simple S method.

Table 2 also shows that the relative complexity in the calculation of dispersion effects by methods BH and BN is not compensated by gains in sensitivity (PCI) and specificity (1 – PII), even though the use of modified residuals had originally been proposed to allow the identification of dispersion effects by formal statistics tests, which is not being considered in this study. And the poor performance of the GLM_s (Table 2) is justified by the use of the saturated models in the GLM approach as a first part of a more general approach: the non-significant effects being removed hierarchically to produce a model in that all of the effects are significant.

The second group of estimates (H, BM, BN and BH) presents PCI higher than the first group (S, R and GLM_s). This result was expected because in the second group we use only the correct terms in the mean model. But the result suggests us that a good practice is aggregate the non-significant high order interaction in the error term.

Figures 1 to 5 show the more important interactions between the performance of the methods and the factors in Table 1.

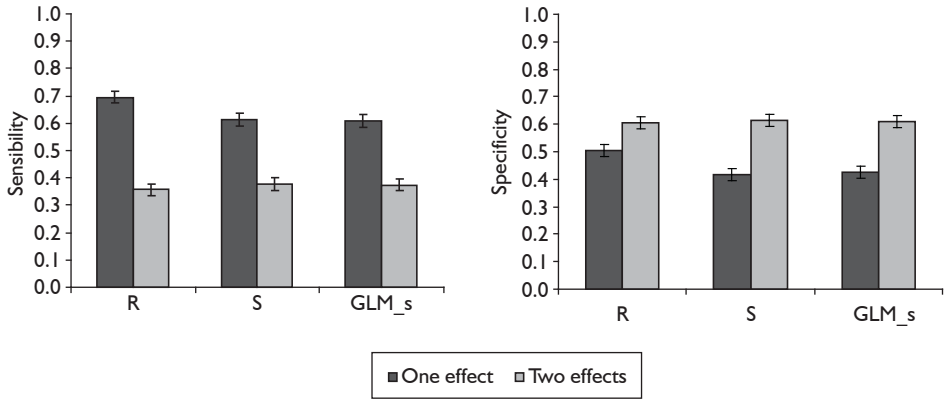

Figure 1 – Comparison of methods of the saturated models by the number of non-null dispersion effects. Note: The bars represent the average of the PCI or the 1 – PII, and the markings at the top of the bars represent 95% confidence intervals.

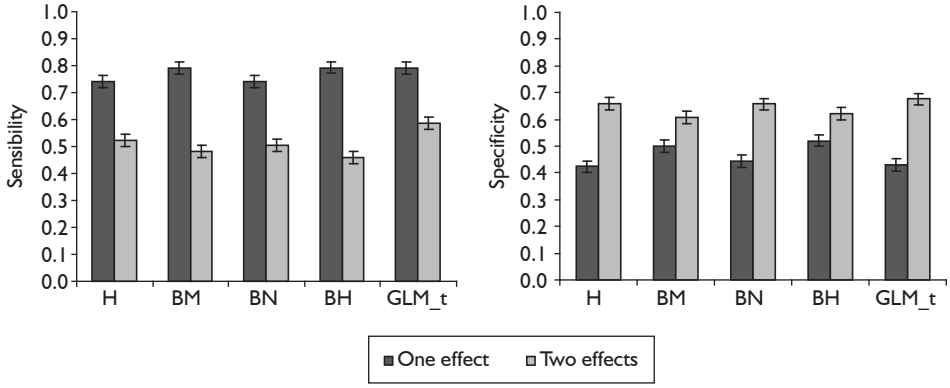

Figure 2 – Methods where we assume the true terms in the mean model by different number of non-null dispersion effects.

In Figure 1, we see that the R method is better than the S or GLM_s methods when we have only one dispersion effect. In this condition, methods BM and BH are also better than methods H and BN (Figure 2). In other words, we could say that the generalized Box-Meyer method performs better than the generalized Harvey method when we have only one dispersion effect. The term "generalized" is because we adapted the methods for replicated experiments.

With two dispersion effects, the methods showed slight differences. For the saturated models (Figure 1), the differences are not significant. Therefore, when we use the correct terms of the mean model (Figure 2), the H and BH methods are slightly better than the BM and BH methods.

In previous studies, this interaction had already been partially explored. Nair and Pregibon (1988) showed that the R method was more efficient in the presence of only one dispersion effect, but not adequate to identify more than one dispersion effect. In unreplicated experiments with two or more dispersion effects, methods BM and BH have structural bias, in other words, the calculation of effects, even in theory (with model parameters), does not make it possible to isolate the real effects, thus leading to an erroneous identification of inexistent effects (see Brenneman and Nair, 2001). However, in our simulation (for experiments with two and four replications), we note that these methods still perform well in the presence of two dispersion effects.

The interaction plots between methods and number of replications are shown in Figures 3 and 4. We note that the R method performs better than the S and GLM_s when we have only two replicates (Figure 3). Similarly, the BM and the BH methods perform better with two replicates (Figure 4). When we use four replications, the differences between the methods are very small. Thus, the generalized Box-Meyer method performs better than that of the generalized Harvey method when we have fewer replications.

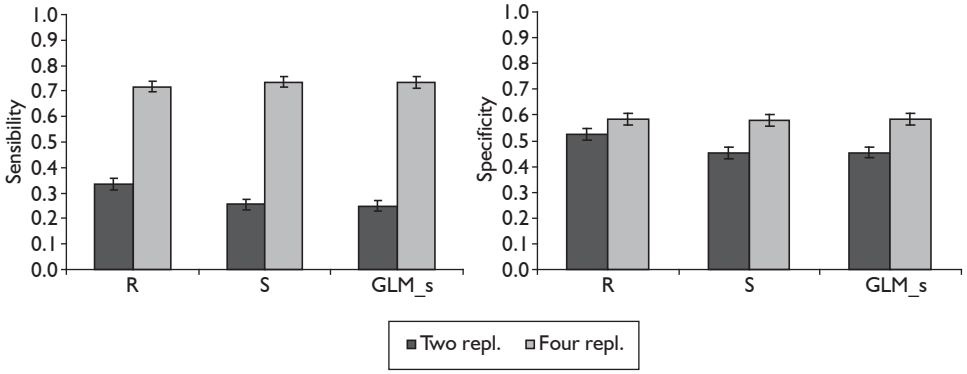

Figure 3 – Comparison of methods of the saturated models by the number of replications.

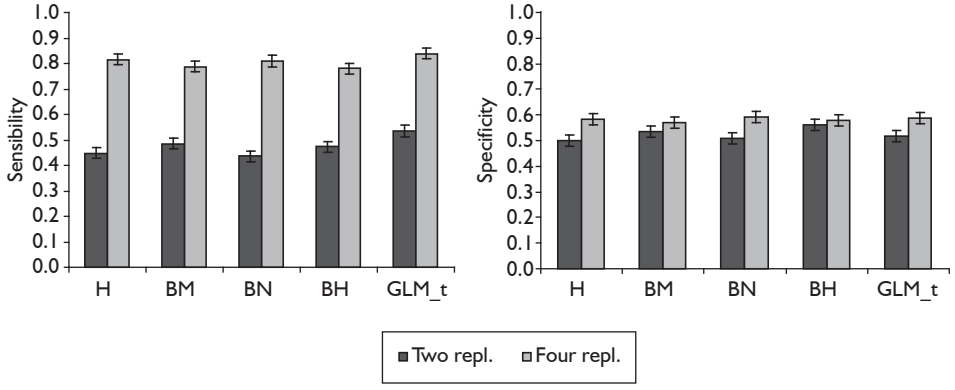

Figure 4 – Methods where we assume the true terms in the mean model for different number of replications.

Finally, the simulated experiment showed that the sensitivity of the simple H and BM methods are worse when there are coincidences between the locations and the dispersions effects. This does not happen with more elaborate methods (BN, BH and GLM_t) (Figure 5). However, this interaction was not significant with the methods based on saturated models or in the analysis of specificity.

Application Example

We use an experiment of the compression strength of concrete described in Mattos (2004) to illustrate the dispersion effects identification methods. The original purpose of this experiment was to study the compressive strength of concrete containing the addition of rice husk ash, but here we only use it to discuss the application of the methods.

The concrete is obtained from a designed mix of cement, water, aggregate and, if necessary, add mixtures and other additions. The type of structure being built as well

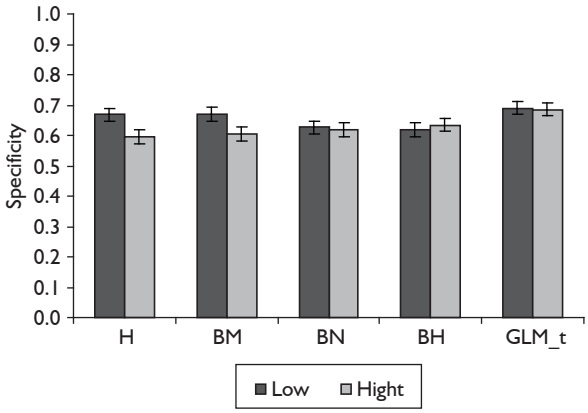

Figure 5 – Interaction plot between the methods and the level of coincidence of location and dispersion effects.

as the method of construction determines how the concrete is placed and therefore the composition of the concrete mix. The concrete quality is appraised for engineering properties of fully cured concrete and how it performs as a structural material.

Table 3 presents the levels of the five factors, which were crossed to perform a 2^5 factorial design. For each design point we observed the compressive strength of three samples of concrete. Note that we didn´t have three genuine replications, because we didn´t repeat all the process. So, in this experiment the variance of replications can be lesser than the true variance of the experimental error. Table 4 presents the responses.

The least square estimates of the location effects results in several terms highly significant as the main effects: A, D and E, and the interactions: AE, BCD and ABCD. The Figure 6b presents the location effects in a normal probability plot. The presence of the high-order interaction produces a quite complex mean model, and the use of transformations, as a Box-Cox approach (Montgomery, 2005, pp. 32), did not improve the results satisfactorily.

We estimated the dispersion effects based on the variance in the replications, or equivalently, based on the residuals of the saturated model. The S and R methods are shown in Figures 6b and 6c. Both methods suggest only the main effect A as apparently

Table 3 - Description of the factors and levels used in the experiment.

Factors	**Description**	**Coded level**	
		-1	**+1**
A	Water-blinder ratio	0.35	0.60
B	Addition (%)	0	10
C	Time of mixture (seconds)	100	300
D	Compaction	mechanical	by hand
E	Age (days)	3	28

Table 4 – The design and data of the concrete experiment.

Run	A	B	C	D	E	Y1	Y2	Y3	Run	A	B	C	D	E	Y1	Y2	Y3
1	–1	–1	–1	–1	–1	57.30	55.90	58.94	17	1	–1	–1	–1	–1	27.50	32.47	34.82
2	–1	–1	–1	–1	1	65.57	60.10	54.11	18	1	–1	–1	–1	1	38.20	38.20	37.56
3	–1	–1	–1	1	–1	48.38	50.93	54.75	19	1	–1	–1	1	–1	28.65	26.36	26.10
4	–1	–1	–1	1	1	52.33	54.37	40.74	20	1	–1	–1	1	1	33.23	35.01	35.65
5	–1	–1	1	–1	–1	53.35	49.85	45.45	21	1	–1	1	–1	–1	32.78	27.88	29.92
6	–1	–1	1	–1	1	62.71	38.71	59.21	22	1	–1	1	–1	1	35.90	41.51	36.16
7	–1	–1	1	1	–1	45.20	45.58	56.34	23	1	–1	1	1	–1	26.10	29.28	26.35
8	–1	–1	1	1	1	60.73	48.06	54.75	24	1	–1	1	1	1	35.01	37.30	37.69
9	–1	1	–1	–1	–1	56.79	56.53	53.35	25	1	1	–1	–1	–1	28.52	31.07	26.10
10	–1	1	–1	–1	1	49.53	55.39	46.47	26	1	1	–1	–1	1	37.56	32.47	35.20
11	–1	1	–1	1	–1	52.71	48.38	51.95	27	1	1	–1	1	–1	23.94	22.92	26.74
12	–1	1	–1	1	1	57.61	40.87	60.35	28	1	1	–1	1	1	27.69	31.83	25.21
13	–1	1	1	–1	–1	60.03	57.04	56.66	29	1	1	1	–1	–1	27.76	32.72	32.59
14	–1	1	1	–1	1	66.84	60.81	62.39	30	1	1	1	–1	1	34.12	34.06	39.98
15	–1	1	1	1	–1	46.60	44.56	53.99	31	1	1	1	1	–1	31.19	28.65	26.86
16	–1	1	1	1	1	40.74	43.42	43.29	32	1	1	1	1	1	36.61	35.65	33.74

significant. The generalized Harvey and Box-Meyer methods produce normal plots quite similar (not presented here). The last two methods are built based on the residuals of the mean model with only apparently significant terms. Figure 6d presents the dispersion effects estimated from joint GLMs approach and the result looks quite similar. In joint GLMs we also used the saturated mean model because it is the common practice to detect dispersion effects.

In practice, it is interesting to have different factors with location and dispersion effects, because in this situation you can make independent adjustment of the mean and the variance of the process – the *separability* principle; see Box (1988) and Lee and Nelder (2003). But the factor A has both location and dispersion effects. The negative dispersion effect indicates that water-blinder ratio at 0.60 (coded level +1 of factor A) shrinks the variance of the compressive strength. But the negative location effect of this factor indicates that water-blinder ratio at 0.60 produces a lower mean of the compressive strength than at 0.35.

The main issue of our article is the dispersion effects' identification problem, but usually the researchers intend to build means and variance model. For this purpose the joint GLMs approach provides an interesting general framework. Following Grego (1993), Lesperance and Park (2003), and Pinto and Leon (2006) we start with a saturated model for both mean and variance, obtaining an appropriate model for the variance through backward procedures, then obtaining an appropriate model for the mean, again through backward procedures. The saturated model of the mean extracts all mean influence from the residuals. Therefore a model of variance with only high significant terms is better.

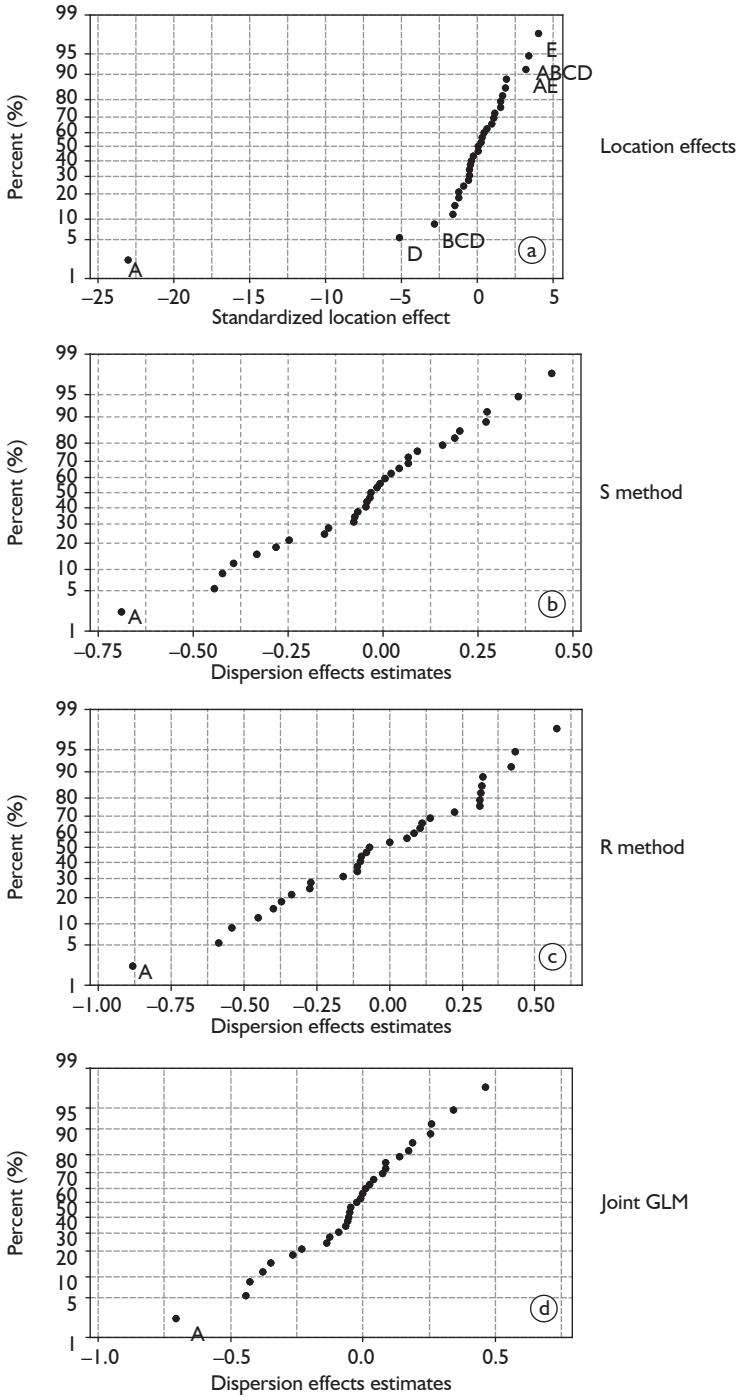

Figure 6 – Normal probability plot of location and dispersion effects.

Iterating with estimated weights is only efficient for IWLS estimates under large-variance heterogeneity with weights that are estimated accurately (Engel and Huele, 1996).

First, we verified that all three or higher interactions were not significant in variance model through joint GLMs approach (by the difference of deviance test). Then we identify the significant effects: A ($p < 0.0001$), E ($p < 0.05$) and the interaction A*E ($p < 0.05$). Notice that the highest significant effect is the same as what we observe in simple R and S methods. Using IWLS with weights given by variance model, we built a mean model. Again we test if all three or higher interactions may be excluded. Then we test each term hierarchically, obtaining the significant terms ($p < 0.05$): A, B, D, E, A*C, B*E, B*C. For the *marginality rule*, namely when an interaction term is significant all related lower-order interaction and main effects should be included in the model (Lee and Nelder, 2003), then we included the main effect C in the mean model. Notice that the joint GLMs approach produces a simpler and more interpretable model (without high-order interactions) than the least square approach.

Conclusions

Nair and Pregibon (1988) and Brenneman and Nair (2001) showed that the R method (or Box and Meyer´s method for unreplicated experiments) produces a great bias in the presence of more than one dispersion effect, but our simulation study showed that this method, for replicated experiments, has the highest sensibility and the highest specificity when the process has only one non-null dispersion effect. Therefore, we showed that the generalized Box-Meyer method (as we call it) performs as well as the generalized Harvey method, or the joint GLMs with saturated models, when the processes have two non-null dispersion effects. Therefore, when we suspect only one or two dispersion effects, we suggest considering primarily the generalized Box-Meyer method. Also, it is apparently better to fit mean model with the non-significant high-order interaction aggregated in the error term instead of considering a saturated mean model.

The joint generalized linear models for mean and variance provides a general framework for testing and modeling both location and dispersion effects. The statistical properties of this method suggest that it should be better than other methods previously discussed. However, the generalized Box-Meyer or Harvey methods can be made with simple spreadsheet software, while the joint GLM needs more specialized computation software. Therefore, when we use joint GLM with saturated models as the first step of the backward regression, our simulation showed that it does not always performs better than the R or S method. Then, even in the GLM approach, the normal probability plot with effects estimated by generalized Box-Meyer or R method can be helpful.

The main computational programs in this work were made in IML/SAS and can be downloaded at www.inf.ufsc.br/~barbetta/DispersionEffects.htm.

Acknowledgements

We are grateful to the referrers and the editor. Their comments and suggestions have substantially improved this article.

References

Barbetta, P.A. et al. (2000) Variance regression models in experiments with few replications. Quality and Reliability Engineering International, Vol. 16, No 5, pp.397- 404.

Bartlett, M.S. and Kendall, D.J. (1946) The statistical analysis of variance-heterogeneity and the logarithmic transformation. Journal of the Royal Statistical Society, Serie B, Vol. 8, No 1, pp.128-138.

Bergman B. and Hynén, A. (1997) Dispersion effects from unreplicated designs in 2k-p series. Technometrics, Vol. 39, No 4, pp.191-198.

Box, G.E.P. (1988) Signal-to-noise ratios, performance criteria and transformations. Technometrics, Vol. 30, No.1, pp.1-40.

Box, G.E.P. and Meyer, R.D. (1986) Dispersion effects from fractional design. Technometrics, Vol. 28, No 1, pp.19-27.

Brenneman, W.A. and Nair, V.N. (2001) Methods for identifying dispersion effects in unreplicated factorial experiments: A critical analysis and proposed strategies. Technometrics, Vol. 43, No. 4, pp. 388 - 405.

Engel, J. and Huele, A.F. (1996) A generalized linear modeling approach to robust design. Technometrics, Vol. 38, No. 4, pp. 365 - 373.

Grego, J.M. (1993) Generalized linear models and process variation. Journal of Quality Technology, Vol. 25, No 4, pp. 228-295.

Harvey, A.C. (1976) Estimating regression models with multiplicative heteroscedasticity. Econometrica, Vol. 44, No 3, pp. 461- 465.

Lee Y. and Nelder, J.A. (2003) Robust design via generalized linear models. Journal of Quality Technology, Vol. 35, No 1, pp. 2-12.

Lesperance and Park (2003) GLMs for the analysis of robust designs with dynamic characteristics. Journal of Quality Technology, Vol. 35, No 3, pp. 253-263.

Mattos, V.L. (2004) Identification of dispersion effects in experiments with few replications. (Thesis, Industrial Engineer), Federal University of Santa Catarina, Santa Catarina, Brazil (in Portuguese). Available: http://www.tede.ufsc.br/teses/PEPS4356.pdf. Access: 31th January, 2008.

McCullagh, P. and Nelder, J. (1989) Generalized linear models. Chapman and Hall, London.

Montgomery, D.C. (2005) Design and analysis of experiments. John Wiley, New York.

Nair, V.N. and Pregibon, D. (1988) Analysing dispersion effects from replicated factorial experiments. Technometrics, Vol. 30, No 3, pp. 247-257.

Pan, G. (1999) The impact of unindentified location effects on dispersion-effects identification from unreplicated factorial designs. Technometrics, Vol. 41, No 4, pp. 313-326.

Pinto, E.R. and Leon, A.C.M.P. (2006) Nelder and Lee´s joint modeling of mean and dispersion as an alternative to the Taguchi´s methods. Pesquisa Operacional, Vol. 26, No 2, pp. 203-224 (in Portuguese).

Schoen, E.D. (2004) Dispersion-effects detection after screening for location effects in unreplicated two-level experiments. Journal of Statistical Planning and Inference, Vol. 126, No 1, pp. 289 – 304.

Van de Ven, P.M. (2008) On the equivalence of three estimators for dispersion effects in unreplicated two-level factorial designs. Journal of Statistical Planning and Inference, Vol. 138, No 1, pp. 18-29.

Wu, C.F.J. and Hamada, M. (2000) Experiments: planning, analysis, and parameter design optimization. John Wiley, New York..

Biography

Viviane Leite Dias de Mattos is an associate professor of the Education School at the Catholic University of Pelotas – Brazil. She received the degree of Doctor of Industrial Engineer at the Federal University of Santa Catarina – Brazil, in 2004. Professor Mattos is member of the Brazilian Association of Industrial Engineers and hers areas of interest are Design of Experiments and Statistics Process Control.

Pedro Alberto Barbetta is an associate professor of the Department of Informatics and Statistics at the Federal University of Santa Catarina – Brazil. He received a Ph. D. in Industrial Engineer, in 1998, from the same university. He is the author of two textbook in statistics education. Professor Barbetta is a member of the Brazilian Statistical Association and the Brazilian Association of Industrial Engineers. His areas of interest are Design of Experiments and Multivariate Data Analysis.

Dalton Francisco Andrade is a full professor of the Department of Informatics and Statistics at the Federal University of Santa Catarina – Brazil. He received a Ph.D. degree in Biostatistics, in 1984, from University of North Carolina at Chapel Hill. Professor Andrade is a member of the Brazilian Statistical Association, the Brazilian Section of the International Biometric Society and the American Educational Research Association. His areas of interest are Design of Experiments, Sampling, Longitudinal Data Analysis, Item Response Theory and Hierarchical Models.

Robert Wayne Samohyl, Ph. D. Professor Bob is from Houston, Texas, Ph.D. in Economics from Rice University, and has lived in Brazil permanently since 1978. He is full professor

and recent Chairman of the Industrial Engineering Department, Federal University of Santa Catarina, Florianopolis, Southern Brazil, a region known for a thriving software industry and wonderful beaches. He is recent Vice President of the Brazilian Operations Research Society. He has developed courses for the internet in Statistics, Econometrics and Forecasting, and taught by video-conferencing. He is presently coordinating a government project for implementing forecasting techniques in electrical energy companies.

Submitted: 13th February, 2008.
Accepted: 14th April, 2009.