Appl. Statist. (1996)
45, *No.* 1, *pp.* 47–61

Dispersion Measures and Analysis for Factorial Directional Data with Replicates

By CHRISTINE M. ANDERSON†

University of Western Ontario, London, Canada

and C. F. JEFF WU

University of Michigan, Ann Arbor, USA

[Received June 1994. Final revision April 1995]

SUMMARY

Determining the nature of the connection between a number of factors and the spread of a directional response in industrial experiments has not been considered much in the literature. Several dispersion measures are explored and their relationships described. The circular variance is a good dispersion measure that transforms the angular dispersion into a statistic measured on a linear scale. Once this transformation has been performed, established techniques for analysis can be employed for analysing factor influences on the directional dispersion. The method proposed is used to analyse data from an experiment involving the balancing of automotive flywheels.

Keywords: Circular data; Directional data; Dispersion analysis; Variation reduction

1. Introduction

Studying factor effects on dispersion and improving quality through the reduction of variation are the main ideas in robust parameter design that were popularized by Taguchi (1986). In some applications, reducing the spread of the data by selecting an optimal combination of factors is the primary goal of experimentation. In this paper we examine techniques for analysing the influence of experimental factors on the dispersion of a *directional* response located on a unit circle. How the treatment of control and noise factors differs for the analysis is also explored.

For example, in the automotive industry, a number of rotating parts (such as brake rotors, flywheels, crankshafts and tyres) need to be precisely balanced to prevent excessive vibration. We can measure imperfection in the part by identifying the direction, which is disproportionately heavy (or light), and the magnitude of the imbalance. This paper considers analysing the spread of the directional component only. If a combination of factors is found that locates all the imbalances close together, then one of several strategies can reduce production costs. In some cases, a global corrective action could be taken to adjust the process to reduce the number of unbalanced parts. In other cases, the parts still may need to be individually corrected,

†*Address for correspondence*: Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, N6A 5B7, Canada.
E-mail: cmanders@stats.uwo.ca

 0035–9254/96/45047

but the related cost can be lowered simply by having the imbalances all located close to one another on the part.

We now present some of the issues that arose from a real industrial experiment at an automotive production plant involving the balancing of engine flywheels. The response obtained from each flywheel was a location on the circumference of the part where a corrective adjustment would be required to balance the part. In this case, the particular location on the part where the corrections will be made is not critical, as long as all the corrections can be made close to one identifiable point. Hence, this is clearly a case where dispersion analysis alone is of primary interest.

The process of determining the location of the imbalance is quite precise and uniquely determines a single point where the corrective action should be taken. A 2^4 full factorial experiment was run with 10 observations at each set of factor combinations. The four factors thought to influence the dispersion of the imbalance are as follows:

(a) the location of a butt weld to the flywheel (factor A), either fixed (coded F) or random (coded R)—in current production, the selection of the location for joining these two critical pieces together was determined randomly;
(b) flywheel radius grade (factor B), either low (coded L) or high (coded H)—current levels used the lower grade flywheels where a larger difference in radius (larger than 0.005 units) was observed;
(c) flywheel thickness grade (factor C), either L or H—currently, the low grade thickness was used with a larger difference in flywheel thickness deemed acceptable; the new high grade requirement had less than a 0.001-unit difference in thickness;
(d) the size of the counter-weight attached at 0° (factor D), either L or H. The size of the counter-weight in production was at present at the low level.

In addition, a fifth factor E (ring gear imbalance) which is very difficult and expensive to control was observed and a number of each of the three levels (L, M, H) were used in each of the 16 factor combination groups. Typically, each of the groups consisted of 1–3 low, 5–7 medium and 1–3 high observations. The data are provided in angular form (measured in degrees) in Table 1 with the level of factor E in parentheses below.

As a preliminary test of group dispersion differences, Bartlett's test for homogeneity of von Mises concentration parameters as described in Stephens (1982) can be used. Further details of the test are provided in Section 3. The test disregards the structure of the factorial experiment and considers each combination of factors as a different group. For the flywheel data, the significance level of this test is approximately 0.0001. Hence, we conclude that there are real differences between the dispersions of 16 factor combinations. Several questions arise from this conclusion, which will be studied in the remainder of the paper.

(a) Can the relative importance of the four controllable factors be assessed to determine how improvements to the process should be approached?
(b) Since factor C is actually expensive to control (and would greatly increase production costs), can a combination of the other factors be found that is robust to the different levels of this factor?
(c) How can the information about the noise factor E be incorporated to give greater insight into the working of the process?

TABLE 1
Automotive flywheel data

Group	*A*	*B*	*C*	*D*	*Data (deg) (with levels of E)*									
1	R	L	L	L	133 (L)	175 (M)	178 (M)	178 (M)	153 (M)	190 (M)	221 (M)	177 (M)	281 (H)	190 (H)
2	R	L	L	H	139 (L)	61 (L)	109 (M)	187 (M)	74 (M)	351 (M)	309 (M)	236 (M)	69 (H)	320 (H)
3	R	L	H	L	111 (L)	122 (L)	105 (M)	49 (M)	189 (M)	188 (M)	177 (M)	151 (M)	62 (H)	329 (H)
4	R	L	H	H	170 (L)	19 (M)	337 (M)	171 (M)	114 (M)	341 (M)	10 (M)	266 (H)	201 (H)	162 (H)
5	R	H	L	L	127 (L)	215 (L)	125 (M)	188 (M)	187 (M)	175 (M)	162 (M)	172 (M)	169 (H)	82 (H)
6	R	H	L	H	150 (L)	84 (L)	113 (M)	318 (M)	84 (M)	353 (M)	301 (M)	12 (M)	82 (H)	351 (H)
7	R	H	H	L	152 (L)	164 (L)	180 (M)	187 (M)	159 (M)	149 (M)	127 (M)	148 (M)	175 (H)	201 (H)
8	R	H	H	H	184 (L)	128 (L)	177 (L)	186 (M)	163 (M)	178 (M)	196 (M)	155 (M)	150 (H)	120 (H)
9	F	L	L	L	154 (L)	200 (L)	147 (M)	133 (M)	171 (M)	318 (M)	100 (M)	108 (M)	86 (H)	73 (H)
10	F	L	L	H	165 (L)	31 (M)	51 (M)	314 (M)	84 (M)	267 (M)	135 (M)	318 (H)	14 (H)	198 (H)
11	F	L	H	L	345 (L)	43 (L)	4 (M)	295 (M)	75 (M)	138 (M)	149 (M)	141 (M)	198 (H)	175 (H)
12	F	L	H	H	153 (L)	194 (L)	136 (M)	144 (M)	206 (M)	151 (M)	202 (M)	104 (H)	188 (H)	207 (H)
13	F	H	L	L	140 (L)	170 (M)	62 (M)	109 (M)	127 (M)	132 (M)	116 (M)	94 (M)	183 (M)	134 (H)
14	F	H	L	H	340 (L)	111 (L)	128 (M)	327 (M)	81 (M)	301 (M)	3 (M)	335 (M)	215 (H)	334 (H)
15	F	H	H	L	160 (L)	152 (M)	187 (M)	158 (M)	143 (M)	91 (M)	200 (M)	143 (M)	84 (H)	191 (H)
16	F	H	H	H	171 (L)	156 (L)	171 (M)	195 (M)	159 (M)	153 (M)	188 (M)	125 (M)	107 (H)	98 (H)

Before addressing these experiment-specific questions, some more fundamental issues need to be discussed. We begin with a brief review of notation and basic quantitites for directional data. Consider the simplest situation of a single population of directional data responses located on the circumference of the circle. From this population we obtain a sample, $(\theta_1, \ldots, \theta_n)$. The θ_i can also be identified as vectors of unit length starting at the origin and pointing in the direction of their angle, with the usual convention that 0° points horizontally to the right, with positive angles rotating counter-clockwise. The vector that corresponds to angle θ_i is called u_i, and from this vector representation we can calculate the resultant vector by summing the vectors. The overall resultant vector $u_.$ of the sample has length $R = (c^2 + s^2)^{1/2}$, where $c = \Sigma_i \cos\theta_i$ and $s = \Sigma_i \sin\theta_i$. Another quantity of interest is $\bar{R} = R/n$, the standardized length of the resultant vector. The average direction for the sample, frequently called the mean direction, is defined to be the angle of the resultant vector. See Fisher (1993), p. 31, for details.

A common choice for a distribution on the circle is the von Mises distribution with mean direction μ and concentration parameter k, denoted VM(μ, k). It has many of the desirable properties associated with the normal distribution for linear data (Mardia (1972), pages 55–58) and has the probability density

$$f(\theta) = \frac{1}{2\pi\, I_0(k)} \exp\{k \cos(\theta - \mu)\},$$

where $I_0(k)$ is the modified Bessel function with $k \geqslant 0$ and $\theta \in (-\pi, \pi)$.

The remainder of the paper seeks to develop a method for studying dispersion for a directional response and to provide answers to the specific questions for this industrial problem. Section 2 discusses several dispersion statistics and outlines their relative strengths and some distributional results. Section 3 outlines a strategy for modelling the dispersion for a factorial experiment with replication and discusses two models with desirable intuitive interpretations that are often suggested by the strategy. Section 4 provides details about how to incorporate noise factor effects into the model. Finally, Section 5 illustrates the technique by analysing the flywheel experiment data.

2. Measures of Dispersion

In this section we consider some dispersion statistics which might be suitable for studying the spread of the observed directional data. It is desirable that the statistic is a simple, intuitively pleasing and a computationally convenient measure of the spread of the data, regardless of the shape or distribution of the original data. In addition, it would be advantageous if the measure has known and manageable distributional properties under more restrictive assumptions about the original data.

The standardized length of the resultant vector $\bar{R}$ is the natural measure of dispersion for directional data. Hence, functions of $\bar{R}$ are considered. The circular variance $S_0 = 1 - \bar{R}$ is a common dispersion statistic used to quantify directional variability. It ranges in value from 0 to 1. A value of 0 corresponds to no variation in the data, whereas $S_0 = 1$ means that the data are uniformly distributed on the circumference of the circle. It follows the usual convention that a small value for the variance means that the data are concentrated near the average, but unlike the variance obtained for linear data it has a finite maximum. As well, it is more distributionally robust than the von Mises concentration parameter k suggested by Fisher and Lee (1992), which does not describe a general characteristic of the data without the von Mises assumption.

Calculation of $\bar{R}$, and hence the variance, is straightforward and well defined for all data sets. Rivest (1982) noted that for highly concentrated samples the circular variance properly normalized has approximately the same distribution as the linear variance of the angles measured on a $(0, 2\pi)$ scale. Mardia (1972), p. 113, summarized results about the expectation and variance of S_0 under the von Mises assumption, whereas Watson and Williams (1956) showed that

$$2k(n - R) = 2nk(1 - \bar{R}) \sim \chi^2_{n-1}$$

for concentrated von Mises data. To improve this approximation for small concentrations and sample sizes which are typical of industrial data, the method of matching the first two moments of the circular variance is used. We assume that

$\gamma(1-\bar{R}) \sim \chi^2_f$, for some multiplicative coefficient γ and approximate degrees of freedom f. Solving for the optimal values of these parameters yields

$$\gamma = \frac{2nk(1-A)-1}{k(1-A^2)-A-1/4nk} \tag{1}$$

and

$$f = \gamma\left(1 - A - \frac{1}{2nk}\right) \tag{2}$$

where $A = A(\hat{k}) = I_1(\hat{k})/I_0(\hat{k})$, the quotient of modified Bessel functions.

It is easy to show that $\gamma/2nk \to 1$ and $f/(n-1) \to 1$ as $n \to \infty$ and $k \to \infty$. So equations (1) and (2) can be viewed as an improvement and extension to the χ^2-approximation of Watson and Williams (1956). For $k \geqslant 2$ and $n \geqslant 10$, these two ratios are close to 1 as shown in a numerical study reported in Anderson and Wu (1994). Therefore, in these situations, the χ^2_{n-1}-approximation is adequate. For $k < 2$ or $n < 10$, it is advisable to use the approximation $1 - \bar{R} \sim \chi^2_f/\gamma$.

The circular variance is a good dispersion statistic for a large number of distributions. It is simple and has manageable distributional properties for data originating from a von Mises distribution. Hence, the circular variance or its monotonic transformation satisfies the criteria established earlier.

The circular standard deviation is another alternative statistic to the circular variance. Unlike the linear case where the standard deviation is simply the square root of the variance, for directional data, the form of this new statistic is $s_0 = \{-2\log(1-S_0)\}^{1/2}$, where S_0 is the circular variance. Hence it can be simplified to $(-2\log\bar{R})^{1/2}$.

The circular standard deviation is a non-negative statistic ranging from 0 for no dispersion in the data to ∞ for the data uniformly distributed around the circle. Less is known about the distributional properties of this statistic than the circular variance even when the data come from a von Mises distribution, and hence Mardia (1972), p. 24, commented that the circular variance is 'more useful than s_0 for theoretical investigations'.

To estimate the distribution of s_0, we exploit its relationship to the circular variance. Using Taylor expansions, we obtain

$$s_0 \approx \surd(2S_0)\left(1 + \frac{1}{4}S_0 + \frac{13}{96}S_0^2 + \frac{43}{384}S_0^3\right), \tag{3}$$

where $S_0 = 1 - \bar{R}$ is assumed to be small. This supports the conclusion drawn in Mardia (1972), p. 24, that for small values of S_0 the circular standard deviation reduces to a multiple of the square root of the variance. However, for data from a von Mises distribution with moderate concentration parameters (say $k \in (1, 20)$), the additional terms of the expansion will not be negligible and will influence the shape of the distribution. Simulated von Mises data show that the distribution of the circular standard deviation is quite nearly normal for a variety of sample sizes and dispersions. See Anderson (1993).

The circular standard deviation and the circular variance are two strong choices for a dispersion modelling, both satisfying the criteria established earlier. It will subsequently be convenient to utilize the connection between them, namely $s_0 \approx \sqrt{(2S_0)}$ for concentrated data. In the next section, we examine how the dispersion measures can be utilized for understanding variation.

3. Dispersion Modelling

Analogously to the study of dispersion effects for linear data as described by Nair and Pregibon (1988) and Box (1988), this section describes a method for determining the effects of factor levels on the spread of directional data.

First, we determine whether there are significant differences between the estimates of dispersion for the groups. Stephens (1982) described Bartlett's test for the homogeneity of concentration parameters for von Mises data. For each of the groups in a 2^r factorial design with m replicates, define $Q_l = mS_l$ and $q_i = m - 1$ where l ranges from 1 to 2^r for the various groups and S_l is the circular variance of group l. We also define $T = \Sigma_l Q_l$ and $t = \Sigma_l q_l$. The test statistic for testing whether there is a difference between groups is Z/C, where

$$Z = t \log T - \sum_l q_l \log Q_l - t \log t + \sum_l q_l \log q_l \tag{4}$$

and

$$C = 1 + \frac{1}{3(s-1)}\left(\sum_l \frac{1}{q_l} - \frac{1}{t}\right), \tag{5}$$

where $s = 2^r$ is the number of groups.

The test statistic Z/C is approximately χ^2 with $s - 1$ degrees of freedom under the null hypothesis of no difference between groups. Therefore, for a test of size $1 - \alpha$, we would reject that hypothesis if

$$P(\chi^2_{(s-1)} \geqslant Z/C) \leqslant \alpha. \tag{6}$$

However, it is important to note that Bartlett's test is sensitive to assumptions of normality, here data originating from a von Mises distribution. Therefore, this test should be viewed primarily as a diagnostic method for determining whether there are large differences between groups. If there is no evidence against the hypothesis that the variance estimates for all groups are constant, then the remaining analysis will probably not bear fruitful results. However, if differences between group dispersions are noted, as in the flywheel example, we proceed with further analyses.

Using the circular variance as our starting point for the choice of a dispersion measure, we obtain the resultant length for each combination of factors in the experiment and calculate the circular variance. A suitable transformation of the data is sought by using the one-parameter Box–Cox power transformation family of the form

$$(1 - \bar{R})^\lambda = X\Upsilon + \epsilon \tag{7}$$

where λ is the transformation power (if $\lambda = 0$, then the natural logarithm is used). X

is the design matrix, Υ is the vector of parameters and $\epsilon \sim \text{MVN}(0, \sigma^2 I_n)$ is the vector of error terms. Because we are using a dispersion statistic measured on a linear scale, the model has the same form as dispersion analyses for traditional linear data and the error term can be assumed to be normal, rather than from a directional distribution. This transformation to a linear scale is essential, because we typically have little intuitive feel for dispersion measures on a directional scale, and we can use the existing methods for a linear response.

The approach of Box and Cox (1964) to data transformations strives to balance three separate goals: simplicity of structure, variance homogeneity and normality. Some simplifying assumptions about the model may be required to have some degrees of freedom available for estimating an error term. For example, only the main effects and two-factor interation terms may be considered for the initial choice of transformation. Once an optimal transformation has been identified, the full model can be examined to determine the relative importance of the factors and their interactions. Because the circular variances can be reasonably approximated by a χ^2-distribution in many cases, the optimal choice of the power transformation parameter will frequently lie in $\lambda \in (-1, 1)$ (Hawkins and Wixley (1986) showed that the optimal transformation for χ^2-distributions with at least 10 degrees of freedom is $\lambda = \frac{1}{3}$.) As a result, $\lambda = 0$ and $\lambda = \frac{1}{2}$ may frequently be contained in the 95% confidence interval.

When $\lambda = 0$, the model has multiplicative effects and multiplicative errors on the circular variance, since $\log(1 - \bar{R}) = X\Upsilon + \epsilon$ can be re-expressed in the form (here for the two-way case)

$$\begin{aligned} 1 - \bar{R}_{ijk} &= \exp \sigma_0 \exp A_i \exp B_j \exp(AB_{ij}) \exp \epsilon_{ijk} \\ &= \sigma_0^* A_i^* B_j^* AB_{ij}^* \epsilon_{ijk}^*, \end{aligned} \tag{8}$$

where σ_0^* is the base-line measure of variability of the data and A_i^*, B_j^* and AB_{ij}^* are the main and interaction effects of the factors. To interpret the factor effects, if $A_i < 0$, and hence $A_i^* < 1$, then level i of factor A decreases the circular variance. For highly concentrated data, the circular variance is approximately proportional to the usual variance for the data projected onto a straight line. In this situation, using the logarithmic transformation coincides with the method commonly used to model dispersion for linear data.

Because the range of S_0 is restricted to the range $1 - \bar{R} \in [0, 1]$, we have an additional concern for this modelling, not present for traditional linear data where the variance has no finite upper bound. For directional data, $\log(1 - \bar{R}) \in (-\infty, 0]$ which means that the linear combination of factor effects must also be restricted to lie in this range. For a general linear model, there is no convenient way to adjust the range of the $X\Upsilon$ to accommodate this restriction, since extrapolation into some regions of the design space could lead to an expected value of $\log(1 - \bar{R})$ which might be positive and hence lie outside the interpretable range of values. McCullagh and Nelder (1989) commented that a transformation is less desirable if there is a possibility of obtaining a value for $X\Upsilon$ outside defined boundaries. However, the boundary of the range corresponds to the uninteresting extreme of the data being uniformly distributed around the circle. This corresponds to the least desirable case for variation reduction strategies.

For several industrial examples considered by the authors, the common range for the resultant vector length is (0.45, 0.95), which corresponds roughly to the von Mises concentrations parameter $k \in (1, 20)$ and circular variances of (0.05, 0.55). This yields a log S_0 range of $(-3.0, -0.6)$ which is a reasonable distance from the problem area near 0. In these cases with a noticeable gap between the edge of the projected region and 0 relative to the expected spread of the data, obtaining estimates of $\log S_0$ outside the acceptable range would probably not be a major problem for interpolating between the high and low levels of the factors. In addition, if we are doing a dispersion analysis with the goal of variance reduction, then the problem area lies at the opposite extreme to our desired target region of minimal variance. As $1 - \bar{R}$ decreases, $\log(1 - \bar{R}) \to -\infty$, which is in a stable area away from the boundary. In this region we can expect the variance estimates for different factor combinations to be well defined.

If $\lambda = \frac{1}{2}$, then instead of modelling the square root of the circular variance we may choose to exploit the relationship between S_0 and the circular standard deviation. As demonstrated in the previous section, for concentrated data, the two are nearly proportional. The advantage of using the standard deviation is ease of interpretation. This model yields an additive model with an additive error on a known dispersion quantity,

$$s_0 = X\Upsilon + \epsilon, \tag{9}$$

rather than considering the square root of the circular variance, which has no inherent meaning, as the basis for the linear model.

In different industrial applications, one of the models described above may concur more closely with the physical understanding of the process, and hence be preferable.

After a suitable transformation has been selected, a half-normal plot of the factor effects can provide insights into the relative influence of different factors on dispersion and to suggest a suitable combination of factor levels to attain minimum variation.

4. Incorporating Noise Factor Effects

We now address a further enhancement for addressing the robustness of the process to changing noise factor levels. Factors can often be broken into two categories. Control factors are those which are relatively easy to adjust (or control). Noise factors, in contrast, are expensive or impractical to control in production but can be controlled or measured during experimentation. Desensitizing the process to noise variation is the objective of robust design.

Two common choices of general designs are available. The first involves genuine replicates for the observations at each factor combination, whereas the second involves sampling across a variety of noise factor levels to determine what levels of control factors are robust to changes in the noise factors. If the data are of the first form, then our dispersion measures within each cell give an indication of the short-term variation in the system but may not give an accurate assessment of variation over the total range of production conditions. The second approach strives to simulate a range of possible operating conditions by changing the levels of some factors which are known to be variable but are typically difficult to control in production. This approach gives a more realistic assessment of long-term variability in the process and allows the experimenter to gain information about what combinations of the control factors might reduce this variability. Both approaches are easily incorporated within the dispersion modelling framework previously described by careful selection of the repeated observations for each combination of the control factors.

For linear data, it is well established that in many situations the use of a robust design can give an overall reduction in response variability without having to control the levels of noise factors. Two major approaches are taken to study possible exploitable relationships between control and noise factors (Shoemaker *et al.*, 1991):

(a) loss modelling involves studying a measure of the dispersion as a function of control and noise effects to determine an optimal setting for control factors levels;
(b) examining the response directly can provide insights into specific relationships between control and noise factors.

For directional data, no location model for exploring the factor effects of a factorial design currently exists (see Anderson (1993) for an explanation of the difficulties associated with such a model). However, the two methods for examining the dispersion effects can be extended to the circular data situation, even in the absence of a working model.

We now clarify how the control and noise factors are incorporated directly. If we have control factors C and noise factors N, we can choose an orthogonal array with factor effects of interest for the control factors (called the control array) and a similar array for the noise factors. Subsequently, the noise array is run for each row of the control array, to give the product array. Ideally, it would be desirable to have replicates at each combination of factors. Once the standardized length of the resultant vector and hence the circular variance have been calculated for each combination, the analysis of control, noise and control-by-noise factor effects could be studied directly. However, this replication may not be feasible if the cost of multiple runs is prohibitive.

If only one observation is available from each combination of the product array, then the standardized length of the resultant vector for all observations at a given control factor setting is calculated and an overall estimate of the circular variance is obtained across all noise level settings. In this way, a measure is taken of how the variation of the process behaves under a wide variety of noise conditions. Once the circular variances have been obtained for each of the combinations of the control array, we model the dispersions by using the procedure outlined earlier. For this model we attempt to minimize the spread of the data across the range of the noise array.

The second approach parallels the examination of control-by-noise interactions for linear data, but because of the absence of a model a quantitative assessment of control-by-noise interactions using the response model approach is not possible. However, a qualitative comparison of different relationships between control and noise factors can give insight into these relationships and suggest dispersion reduction strategies. The key to these qualitative methods is an interaction plot for this type of response (see Fig. 4, later, and Anderson (1994)). As Shoemaker *et al.* (1991) noted, the absolute magnitude of the control-by-noise interaction is not of primary interest, but rather the existence of a control factor combination which gives responses that are robust to changes in the noise factor levels.

Recall that, for the linear case, interest lies not in the particular slopes of the interaction lines but rather with the overall range of responses across all noise combinations for a given set of control factor levels. A qualitative graphical analysis of the range of data over the set of noise levels considered allows us to assess robustness. Because of the lack of an underlying model for the directional response from a multiway design, the loss model approach is superior. It gives a quantitative assessment of which control factor combinations are best for minimizing the spread of the data over the range of noise factor levels. However, analysing dispersion through a single statistic, like the circular variance or the circular standard deviation, can sometimes disguise interesting attributes in the data. The graphical presentation of control-by-noise interactions can complement the more formal quantitative methods obtained by modelling a function of the dispersion directly.

5. Automotive Example

In this section we consider a complete analysis of the flywheel data and illustrate the previously described techniques. As noted in Section 1, because interest lies in reducing the variation in the process, the focus of the analysis is strictly on the dispersion attribute of the data. For further details about how location effects could be measured for data of this form, see Anderson and Wu (1995).

Because of the strong evidence obtained by applying Bartlett's test (here $Z/C \approx 44$ with significance level 0.0001; see Section 1), we proceed with the formal analysis to determine the specific influence of individual factors. First, the standardized resultant length $\bar{R}$ (Table 2) for each of the factor combinations is obtained, and the circular variance calculated.

The full model to be fitted to the data follows equation (7) where X is the design matrix with a column for the overall mean plus 15 orthogonal columns, one for each

TABLE 2
Flywheel dispersion summaries by group

Group	$\bar{R}$	S_0	*log* S_0	s_0	*Predicted* S_0
1	0.812	0.188	−1.672	0.645	0.261
2	0.203	0.797	−0.228	1.783	0.858
3	0.510	0.490	−0.713	1.161	0.551
4	0.054	0.946	−0.055	2.418	0.438
5	0.815	0.185	−1.687	0.640	0.177
6	0.434	0.566	−0.568	1.293	0.581
7	0.936	0.064	−2.752	0.363	0.123
8	0.915	0.085	−2.460	0.423	0.098
9	0.604	0.396	−0.925	1.005	0.261
10	0.154	0.846	−0.168	1.933	0.858
11	0.234	0.766	−0.267	1.703	0.551
12	0.836	0.164	−1.806	0.599	0.438
13	0.845	0.155	−1.863	0.581	0.177
14	0.349	0.651	−0.429	1.452	0.581
15	0.809	0.191	−1.653	0.652	0.123
16	0.861	0.139	−1.972	0.547	0.098

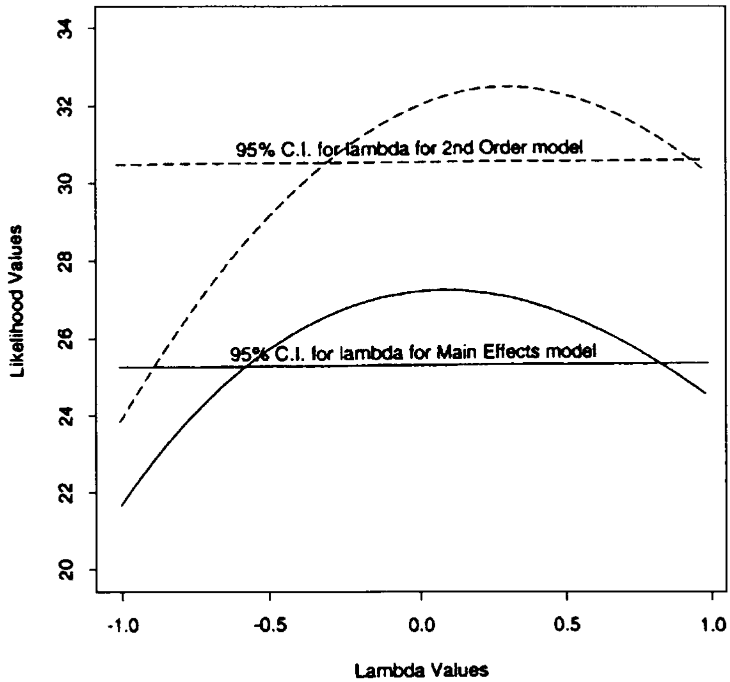

Fig. 1. Box–Cox maximum likelihood estimate of λ

main and interaction effects. To implement the Box–Cox procedure, we must make a few simplifying assumptions about the model. If we assume the full model with all possible effects, no degrees of freedom are available for an error estimate and the model for any value of λ will fit the data perfectly. If we eliminate only the four-way interaction, that gives only 1 degree of freedom for error, and we have most probably overfitted the data with too complicated a model. Therefore, to carry out the method we consider two possible design matrices:

(a) only the four main effects and
(b) the four main effects and their six two-way interaction terms.

Fig. 1 shows the plot of two Box–Cox transformation analyses for the flywheel data assuming models (a) and (b). For model (a), the maximum value for the likelihood occurs for $\lambda = 0.09$ with the 95% confidence intervals covering the range $(-0.56, 0.83)$. For model (b), the maximum value occurs at $\lambda = 0.32$ and the 95% confidence interval includes $(-0.28, 0.95)$. The difference between the two curves for any given λ gives the improvement in the likelihood function by extending the model to include the two-way interactions. Both the circular standard deviation and the log-transform models lie well within the confidence intervals for acceptable values for λ and have sensible interpretations. The fourth and fifth columns of Table 2 contain the log-circular-variance and circular standard deviation values. The current levels of production correspond to group 1, with a resultant length for the group of size 0.812. This is the seventh smallest dispersion of the 16 groups, so there is promise of

TABLE 3
Dispersion analysis

Source	$\log S_0$		s_0	
	Sum of squares	*Rank*	*Sum of squares*	*Rank*
A	0.069	(12)	0.004	(14)
B	3.563	(1)	1.909	(1)
C	1.071	(4)	0.337	(8)
D	0.924	(5)	0.706	(3)
AB	0.263	(10)	0.114	(10)
AC	0.015	(15)	0.057	(12)
AD	0.632	(7)	0.380	(5)
BC	1.123	(3)	0.524	(4)
BD	0.091	(11)	0.001	(15)
CD	2.004	(2)	1.029	(2)
ABC	0.872	(6)	0.355	(6)
ABD	0.418	(8)	0.343	(7)
ACD	0.371	(9)	0.299	(9)
BCD	0.016	(14)	0.011	(13)
ABCD	0.021	(13)	0.080	(11)

possible improvement by selecting an optimal set of factor combinations.

With the set of transformations selected, we revert to the full model including all main effects and interactions. When the logarithm of the circular variance ($\log S_0$) and the standard deviation s_0 are modelled separately against the 15 effects we obtain the analysis-of-variance results summarized in Table 3. The half-normal plots for the factor effects (Figs 2 and 3) for the respective analyses show considerable overlap of dominant factors. Engineers involved with the flywheel process felt that the multiplicative model for the variance matched their intuition of the mechanism explaining variation more closely than did the circular standard deviation model. Therefore, for the remainder of the analysis, $\log S_0$ was used.

Lines have been drawn for each plot showing one approach to identifying important factors. Both of these lines appear to indicate that only factor B is important. However, the choice of line and which factors are judged to be significant is quite subjective, and hence here a final model has been selected to balance both adequate fit and parsimony. For both transformations, factor B is the most influential effect, with the two-way interaction CD also appearing to contribute significantly. Assuming that a hierarchical model is suitable, we would include the main effects B, C and D, with the two-way interaction CD. Since these effects represent four of the five largest effects, we also consider including the third-largest factor effect, interaction BC.

This yields a final model of the form

$$\log(1 - \bar{R}) = -1.34 - 0.94B_i - 0.52C_j + 0.48D_k + 0.78(CD)_{jk} + 0.56(BC)_{ij} + \epsilon_{ijkl} \tag{10}$$

where the indicator variables have value 0 at the low level and 1 at the high level of each factor. The final column of Table 2 gives the predicted values of the circular variance under this model. The optimal choice is high–high–high for BCD, which

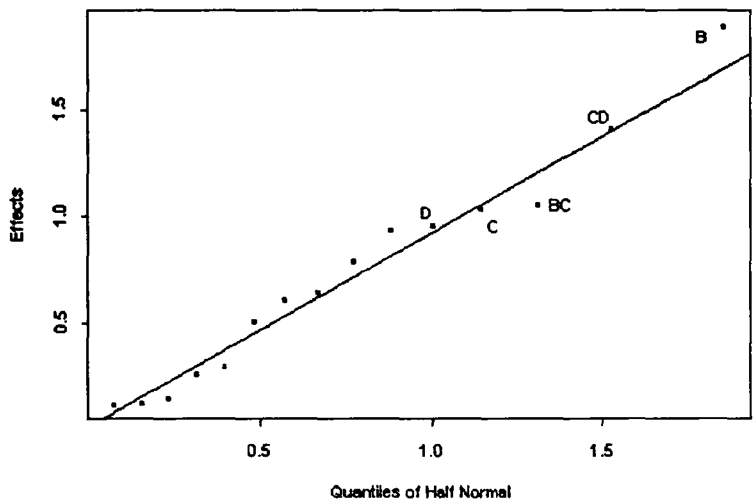

Fig. 2. Half-normal plot of effects from the $\log S_0$ model

yields a predicted value of -2.325. If these levels of the factors are selected, we would expect the circular variance to be 0.098. Alternatively, if the present production levels are used the resultant length is 0.812. By changing from the current production levels to the new set of factor combinations, we would be able to reduce variation from 0.188 to 0.098, a 48% reduction. By examining the two combinations of factors with the HHH for BCD we find that their average circular variance is 0.109, still a 42% reduction from current levels. Therefore, a substantial saving can be realized as a result of this study.

We return now to the questions raised in Section 1. By using the dispersion modelling approach outlined here, the relative influence of the factors can be

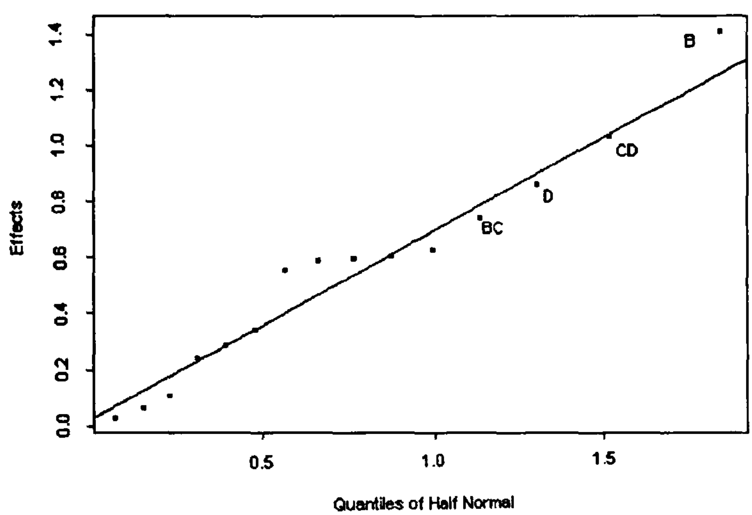

Fig. 3. Half-normal plot of effects from the s_0-model

assessed. Factor B is the most influential in affecting the dispersion of the response. Both C and D are also influential, whereas A is relatively unimportant.

Given that factor C is expensive to control, we can alternatively treat it as a noise factor and plot some of the control-by-noise interactions and see whether it may be possible to exploit one of them. We have already identified an interaction effect for the change in dispersion for combinations of factors CD and now examine the CD control-by-noise interaction plot (shown in Fig. 4), which examines the difference in location for the combinations of factor effects. In this plot the resultant directions of the four factor combinations are shown, along with an asterisk indicating the mean direction of the entire data set. It suggests a level of control factor that will reduce the range of values obtained for the angles, and hence reduce the overall variability of the process by limiting the range of responses to be expected in production where the noise factor will not be controlled by the experimenter. Recall that the goal of examining this plot is to identify whether one level of the control factor (here D) gives a smaller range of values across the noise factor (here C). Clearly, the range of the low level of factor D gives a much smaller range of responses and hence would be preferable if it were too expensive to control the level of C in production. Recall that factor C identified the amount of thickness variation that would be tolerated in the flywheel, with the low level reflecting a lower grade. One possible interpretation for the fact that the smaller counter-weight (low D) might be more robust to variation in factor C may be tied to the fact that we do not have access to where that difference lies. If there was a systematic location for the large and small thicknesses on the part, then the positioning of the counter-weight may magnify or decrease the effect of the thickness difference. The smaller counter-weight will have a less dramatic effect on the location of the imbalance in the presence of a thickness difference, and hence a more robust part is produced. Clearly this is an area where further understanding of the process could yield a greater improvement to the model.

This illustrates why both loss modelling and the control-by-noise interaction plots are usefully applied to the same data, since one may provide insights that are not revealed by the other. In this case, different results are obtained if factor C will be controlled in production or not. Hence, whereas the overall optimal combination of factors for interaction BCD is high–high–high, if factor C is not controlled in production, then D at the low level is a superior choice. As can be seen from Table 2,

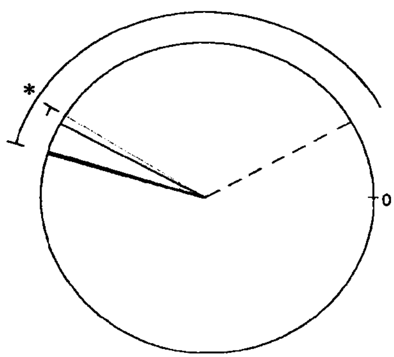

Fig. 4. Circular plot of interaction CD: ——, D low, C low; ········, D low, C high; – – – –, D high, C low; ▬▬ , D high, C high

all four circular variance estimates for interaction BD at the high–low combination (groups 5, 7, 13 and 15) are consistently small.

The estimates of variability obtained here are based on knowledge that noise factor E has been allowed to vary across its usual range of values within each group. Hence this assessment of the dispersion will probably be more indicative of the true variability than if E had not been incorporated and only pure replicates under an unmonitored range of noise factor levels were used.

This example demonstrates the methods described and also gives some practical illustrations of the insights that may be gained about the process through this type of analysis.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada, General Motors of Canada, the Manufacturing Research Corporation of Ontario and National Science Foundation grant 94-04300. The authors would like to thank the referees for their suggestions which greatly improved the exposition of the paper, as well as R. J. MacKay, University of Waterloo, and P. E. Peters, General Motors, for help in obtaining the data.

References

Anderson, C. M. (1993) Location and dispersion analyses for factorial experiments with directional data. *Thesis*. University of Waterloo, Waterloo.

——(1994) Graphical methods for circular and cylindrical data. *Technical Report TR-94-05*. Department of Statistical and Actuarial Sciences, University of Western Ontario, London.

Anderson, C. M. and Wu, C. F. J. (1994) Dispersion measures and analysis for factorial directional data with replicates. *Technical Report RR-94-08*. Institute for Improvement in Quality and Productivity, Department of Statistics and Actuarial Science, University of Waterloo, Waterloo.

——(1995) Measuring location effects from factorial experiments with a directional response. Submitted to *Appl. Statist.*

Box, G. E. P. (1988) Signal-to-noise ratios, performance criteria, and transformations (with discussion). *Technometrics*, **30**, 1–40.

Box, G. E. P. and Cox, D. R. (1964) An analysis of transformations (with discussion). *J. R. Statist. Soc.* B, **26**, 211–252.

Fisher, N. I. (1993) *Statistical Analysis of Circular Data*. Cambridge: Cambridge University Press.

Fisher, N. I. and Lee, A. J. (1992) Regression models for an angular response. *Biometrics*, **48**, 665–677.

Hawkins, D. M. and Wixley, R. A. J. (1986) A note on the transformation of chi-squared variables to normality. *Am. Statistn*, **15**, 197–224.

Mardia, K. V. (1972) *Statistics of Directional Data*. London: Academic Press.

McCullagh, P. and Nelder, J. A. (1989) *Generalized Linear Models*, 2nd edn. London: Chapman and Hall.

Nair, V. N. and Pregibon, D. (1988) Analyzing dispersion effects from replicated factorial experiments. *Technometrics*, **30**, 247–257.

Rivest, L.-P. (1982) Some statistical methods for bivariate circular data. *J. R. Statist. Soc.* B, **44**, 81–90.

Shoemaker, A. C., Tsui, K.-L. and Wu, C. F. J. (1991) Economical experimentation methods for robust design. *Technometrics*, **33**, 415–427.

Stephens, M. A. (1982) Use of the von Mises distribution to analyse continuous proportions. *Biometrika*, **69**, 197–203.

Taguchi, G. (1986) *Introduction to Quality Engineering: Designing Quality into Products and Processes*. Tokyo: Asian Productivity Organization.

Watson, G. S. and Williams, E. J. (1956) On the construction of significance tests on the circle and the sphere. *Biometrika*, **43**, 344–352.