THE INSTITUTE OF STATISTICS

UNIVERSITY OF NORTH CAROLINA SYSTEM

ANALYSIS OF SIGNALS AND NOISES FROM UNREPLICATED FACTORIAL EXPERIMENTS VIA GRAPHICAL PROCEDURES

by

Jye-Chyi Lu and Timm Clapp

Institution of Statistics Mimeograph Series No. 1990

December, 1990

NORTH CAROLINA STATE UNIVERSITY
Raleigh, North Carolina

ANALYSIS OF SIGNALS AND NOISES FROM UNREPLICATED FACTORIAL EXPERIMENTS VIA GRAPHICAL PROCEDURES

Jye-Chyi Lu
Department of Statistics
North Carolina State University
Raleigh, NC 27695-8203

Tim Clapp
Textile Engineering, Chemistry, and Science Department
North Carolina State University
Raleigh, NC 27695-8301

SUMMARY

Statistical experimentation is the key to make the process robust to environmental noises, to produce the products at a lower cost and to enhance the productivity. Analysis of means and dispersions are rational extensions of the Shewhart control chart for comparing products from different design configurations in off-line quality improvement. In this article, we utilize these graphical procedures to analyze signals and noises from an unreplicated factorial experiment to improve the control of a textile manufacturing process. The method of maximum likelihood is employed to estimate the mean and variance functions for predicting the average and variance of yarn strengths. Finally, we exercise a constrained optimization procedure to find the best combination of controllable process variables.

KEY WORDS: Analysis of means, Analysis of dispersion effects, Variance function, Weighted regression, Unreplicated factorial experiment.

1. INTRODUCTION

In the past few years, Japan has shown the world the improving quality leads to lower costs and enhanced productivity. At present, we are in the midst of an international revolution in quality improvement. What is new in quality improvement is the idea introduced by Dr. W. E. Deming (in Deming's book *Quality, Productivity,* *and Competitive Position*, Cambridge, Mass., 1982) that a product is never good enough and should be continually improved through the use of statistical methods. Now, the ultimate goal of quality improvement is to design quality into every product and process and to follow up at every stage from design to final manufacture and sale.

There are quite a few publications on explanations and critiques of Taguchi's philosophy (Taguchi and Wu, 1979), for instances, Kacker (1985), Baker (1986), Box, *et. al.* (1988), Leon, *et. al.* (1987), etc. Chapter 14 of the book, *Statistical Methods for Quality Improvement,* written by Dr. T. P. Ryån (1989) summarizes the advantages and disadvantages of Taguchi's procedures. A special issue in *Quality and Reliability Engineering*, (April–June, 1988) is focused on the study and application of Taguchi's robust design methodology. One of Taguchi's main contributions is to utilize statistically planned experiments to find ways to reduce various sources of product variations and at the same time make the mean level on target. In today's highly competitive global market for textile products, it is essential to apply methods of robust design to textile manufacturing processes. Johnson, Clapp and Baqai (1989) utilized an experiment of factorial design to study the influence of four processing factors in high-speed weaving on fabric-breaking tenacity. Based on the estimated mean effects, they concluded that there was a significant difference in fabric tenacity across the width of the fabric (factor A: side to side). As the pick density (factor C) increased, fabric tenacity increased. Pick density interacted with air-jet pressure (factor D) and yarn type (factor B) to influence the fabric tenacity. Surprisingly, in their analysis of mean effects the interactions BC and CD are the most important factors, which is quite unusual, especially in the case that the main effects B and D are not significant. In their article the study of product variability and the strategy of finding optimal process setup have not been addressed.

Graphical analysis serves as a powerful tool for engineers to present and interpret their experimental results. Nowadays, most engineers know how to use statistical process control (SPC) chart to monitor the process parameters. Ott (1967) suggested the use of analysis of means (ANOM), to present and analyze experimental data. ANOM is a rational extension of the Shewhart control chart and is an useful supplement (or alternative) to the analysis of variance (ANOVA). Applications of ANOM to solve various industrial problems are provided in Ott (1975). Recently, the ANOM technique has been extended from the study of one-way layout (Raming, 1983) to testing interactions from multifactor experiments (Nelson, 1988) and analysis signal and noise (Ullman, 1989). Since ANOM is inherently a graphical procedure similar to control charts, it is apt to receive increasing interests from engineers and other industrial personnel for off-line quality improvement (c.f. Hogg, *et. al.*, 1985).

To reduce the process variations, if the variance is to be estimated from replications at each experimental point in a fractional factorial design or orthogonal arrays, an excessive number of runs may be needed. The expense of repeating measurements can sometimes be avoided by using unreplicated experiment to identify factors that affect dispersions and ones that affect the location effects. The objective of this article is to show examples that use simple graphical methods to analyze mean and dispersion effects from unreplicated experiments in order to direct the process to an optimal condition.

In Section 2, we introduce charts of the ANOM and ANOD (analysis of dispersion effects) to analyze the data collected from the experiment done by Johnson, *et. al.* (1989). Various strategies of searching optimal process control setup, such as mean-variance plot and optimization procedure applied on predictions from mean and variance functions, are addressed. In Section 3, since the calculated dispersion effects are functions of the averages, we compute the variances based on residuals from the mean function. In Section 4, we employ the weighted least squares method to obtain the maximum likelihood estimation of the mean and dispersion effects. The final decision on the optimal process condition is provided and compared to the choices made from the other sections. Since the true optimal point may not be located at any of the particular experimental run, we suggest a confirmation experiment based on a central composite design on the region around the pre-selected optimum. Section 5 gives the conclusion of this article.

2. ANOM AND ANOD CHARTS FOR YARN STRENGTH DATA

A complete 2^4 factorial experiment was conducted by Johnson, Clapp and Baqai (1989) to study the tensile strength properties of the yarns woven from different combinations of processing variables in high-speed air-jet weaving. Their experimental setup and the corresponding measurements, average tenacities (breaking tensile strength per unit linear density), are listed in Table 2.1. Treating the data as unreplicated responses, we propose the following intuitive approach to study the mean and dispersion effects of the process variables A, B, C, D described in Section 1 and Table 2.1.

[Please put Table 2.1 here]

In order to analyze unreplicated (fractional) factorials to identify factors that affect mean and ones that affect the variance, we combine the observations collected at the same level of a factor together (c.f. Box and Meyer, 1986). Specifically, considering the factor A we compute the average and variance of the eight data points in the high level as $\overline{X}_{AH} = (23.55 + 25.00 + 24.51 + \cdots + 24.23)/8 = 24.0863$ and $\hat{V}_{AH} = 1.2230$, respectively. Similarly, the mean and variance of the observations in the low level are calculated as $\overline{X}_{AL} = 24.7257$ and $\hat{V}_{AL} = .5673$, respectively. The estimated mean effect of factor A is the difference of averages from the high and low levels, that is $\hat{\mu}_A =$ 24.0863 – 24.7275 = -0.6412. To estimate the dispersion effect of factor A, we compute the difference of the log-variances from these two levels of factor A, that is, $\hat{\mathfrak{D}}_A = \hat{\mathfrak{D}}_{AH} - \hat{\mathfrak{D}}_{AL} = \log(1.2230) - \log(0.5673) = 0.7682$. This estimation method is quite similar to the procedure used in the study of mean effect. The choice of log-variance is due to the ease of interpretation and also the fact that its distribution can be approximated by normal, which serves as the theoretical basis for latter studies.

The means and variances of the observations in the high and low levels of other factors and their interactions are usually reported in a table such as Table 2.2. Alternatively, we use the following ANOM and ANOD charts to display the means and variances. At the same time we study the statistical significance of effects of changing factor levels. For the analysis of mean effects, we first plot all the averages on the chart Figure 2.1(a) and draw a central line at the overall average $\overline{\overline{X}} = 24.4069$ of the 16 observations. Then, we put the upper and lower decision limits, UDL and LDL, on the chart to detect the significance of the factors. To determine the decision limits, assuming that the data are independently drawn from the normal population with mean depends on the design point and variance σ^2, we note the distribution of the statistics $(\overline{X}_H - \overline{\overline{X}})/S_{\overline{X}_H - \overline{\overline{X}}}$ is a t-distribution with the degree of freedom (df) decided from the information related to the estimate of σ^2. The decision lines are computed using the formulas (c.f. p. 408, Ryan, 1989) given as UDL = central line + $t_{\alpha/2,\nu}\,[s_M^2/(2n)]^{1/2}$ and LDL = central line $-$ $t_{\alpha/2,\nu}\,[s_M^2/(2n)]^{1/2}$, where $s_M^2/(2n)$ is the estimated variance of the statistics $\overline{X}_H - \overline{\overline{X}}$ and $\overline{X}_L - \overline{\overline{X}}$ and s_M^2 is the estimated variance of experimental error σ^2, n is the number of observations and $t_{\alpha/2,\nu}$ is the upper $\alpha/2$ percentiles of t distribution with ν df. For example, in this data set we obtain $s_M^2 = .03536$ from pooling the insignificant location effects such as effects from four third- and one fourth-order interactions, that is, $s_M^2 = [(-.1712)^2 + (.1862)^2 + (.2162)^2 + (-.2562)^2 + (-.0212)^2]/5$, and the df $\nu = 5$. Hence, the t table values are $t_{.025,5} = 2.571$ and $t_{.005,5} = 4.032$ for $\alpha = .05$ and $.01$ respectively. Then, the LDL and UDL are calculated as (24.3214, 24.4924), (24.2729, 24.5409) for the significance levels $\alpha = .05$ and $.01$, respectively. This procedure of detecting significance factors is indeed the two-sample t test and has the same power as the F test in ANOVA (c.f. p. 408 of Ryan, 1989).

[Please put Table 2.1 and Figure 2.1(a) here]

To construct the ANOD chart for analyzing the dispersion effects, we note that the overall average $\overline{\overline{X}}$ can be calculated from the average of two means $\overline{X}_H$ and $\overline{X}_L$. For instance, for factor A, the overall average $\overline{\overline{X}} = \overline{X}_A = (\overline{X}_{AH} + \overline{X}_{AL})/2$. Similarly, we can set the central line of ANOD chart at the average of two log-variances as $\hat{\overline{\mathfrak{D}}}_A = (\hat{\mathfrak{D}}_{AH} + \hat{\mathfrak{D}}_{AL})/2$. The central lines for other factors are obtained as $\hat{\overline{\mathfrak{D}}}_B = (\hat{\mathfrak{D}}_{BH} + \hat{\mathfrak{D}}_{BL})/2$, $\hat{\overline{\mathfrak{D}}}_C = (\hat{\mathfrak{D}}_{CH} + \hat{\mathfrak{D}}_{CL})/2$, etc. Then, the decision limits can be computed as $\hat{\overline{\mathfrak{D}}}_A \pm t_{\alpha/2,\nu}\,[s^2_{\mathfrak{D}}/(2n)]^{1/2}$ for any factor, say A, where $s^2_{\mathfrak{D}} = .5896$ is obtained from average of the squares of the dispersion effects of the third- and fourth-order interactions. For example, for factor A we have the central line at $\hat{\overline{\mathfrak{D}}}_A = [.2014+(-.5668)]/2 = -.1827$ and the decision limits as (–.5317, .1663), (–.7300, .3646) for the significance levels $\alpha = .05$ and .01, respectively. We thus display results of every individual central line and decisions limits from all factors on the ANOD chart in Figure 2.1(b). From Figure 2.1(b) we note that the central lines fluctuate up and down and are generally not equal to each other. Since there is no replication in each experimental run to provide information about product variability, we estimate variances from grouping observations into low and high levels according to the particular design layout for each factor. Different design layout causes variances and central lines change from factor to factor.

[Please put Figure 2.1(b) here]

In order to obtain an overall central line and decision limits for all factors in the ANOD chart, we consider the following approach. We first note that in the ANOM chart the grand average $\overline{\overline{X}}$ is equal to the average of all factor averages, that is, $\overline{\overline{X}} = (\overline{X}_A + \overline{X}_B + \cdots + \overline{X}_{ABCD})/15$. We thus take the average of all central lines computed above as the overall central line $\hat{\overline{\overline{\mathfrak{D}}}}$ for ANOD chart. Then, the decision limits are calculated as $\hat{\overline{\overline{\mathfrak{D}}}} \pm t_{\alpha/2,\nu}\,[s^2_{\mathfrak{D}}/(2n)]^{1/2}$. Note that the t-table value and the standard error here are the same as the ones in the individual decision limits. For example, in our case we obtain the overall central line as –.1770 and the decision limits as (–.5260, .1720), (–.7243, .3703) for α = .05 and .01, respectively. These new decision limits are then displayed in Figure 2.1(b) for comparison with the one constructed from individual decision lines. Since the overall central line –.1770 is rather close to the central line –.1827 for factor A alone, the decision limits constructed from these two approaches are quite close in factor A. From Figure 2.1(b) we find that the overall central line describe the general pattern of the dispersion effects for all 15 factors. Although the decision lines are different in these two approaches, the final decisions of the significant effects are more or less the same in this case. Hence, for simplicity one can consider the approach, which uses the overall decision lines in ANOD chart, to identify the important dispersion effects.

In conclusion, from the ANOM chart in Figure 2.1(a), we note that the main effects A and C as well as interactions AC, BC, BD, CD are significant at $\alpha = 0.01$. In particular, the mean effects from interactions BC and CD are larger than any other main effects and interactions, which is quite unusual in factorial experiments. For the analysis of dispersion effect, we conclude that the main effect C and the interactions AB, AD, BD, ABC, ACD are significant at $\alpha = 0.01$. Note that the dispersion effect of BD is quite large compared to other factors. Moreover, two (ABC and ACD) out of five higher order interactions are significant at $\alpha = .01$. This means that the $s^2_{\mathfrak{D}} =$ .5896, calculated from the dispersions in the higher order interactions, is overestimated and hence the decision limits in ANOD charts are wider than them should be. One might consider to pool the dispersion effects from the insignificant factors such as ABCD, BCD, ABD, CD, BC, AC to compute the variance $s^2_{\mathfrak{D}}$. For example, the average of squares of the dispersions from factors AC, BC, CD, ABD, BCD and ABCD gives the $s^2_{\mathfrak{D}}$ as .1393, which is comparably smaller than $s^2_{\mathfrak{D}} = .5896$ used in constructing of ANOD chart. Formal procedures for identifying inactive factors and estimating experimental error are addressed in Daniel (1976) and Box and Meyer (1986a).

From both the ANOM and ANOD charts we note that the significant factors of mean and dispersion effects are quite different. This makes the search of optimal process condition difficult. The conventional approach is to plot the mean effects against dispersion effects as the one given in Figure 2.2 to see the trade-off between these two effects. For example, there are five points, A, C, BC, BD and CD, outside the combined 99% confidence region of mean and log-variance effects, where the 99% confidence intervals of mean and log-variance effects are calculated from $\pm\ 4.032\, s_M/2$ and $\pm\ 4.032\, s_{\mathfrak{D}}/2$, respectively, where $s_M/2$ is an estimate of $\mathrm{sd}(\overline{X}_H - \overline{X}_L) = (2\sigma^2/8)^{1/2}$. From Figure 2.1(a) we know that BC has to be set in the low level to increase the average, that is, the combinations of factors B and C should be (+-) and (-+). From Figure 2.1(b), we note that BD has to be set at the high level to reduce the variance, *i.e.* the combinations of factors B and D should be (++) and (--). Moreover, from Figure 2.1(a) we learn that the main effects A and C should be set in the low(-) and high(+) levels, respectively, to increase the average. When C is set at the high level, we fix the BC combination at (-+) and the BD combination as (--) after B being set at low(-) level. In summary, we recommend the process parameters be set at A = Low, B = Low, C = High, D = Low, which is the same as the setup in run #5. However, this suggestion conflicts to the requirement of CD being at high level, (++) or (--), to increase the product average. We, therefore, propose the next approach of using predictions of means and variances from all design points to select the optimum.

[Please put Figure 2.2 here]

First, we utilize the least squares method to build a mean function to link the average tenacity to the significant factors such as A, C, D, AC, BC, BD and CD. The estimate mean function and their standard errors are reported in Table 2.3. According to the reported t-values and multiple R-Square (= .9370), we see that the model shows a good fit. The estimate of overall standard deviation σ is computed as .3342, which implies the estimated variance s_M^2 is .1117 with the df = 8 for $t_{\alpha/2,\nu}$ in ANOM formulas. Then, predictions of average tenacity at each design point can be made from the regression function,

$$\text{prediction} = \hat{y}_i = 24.7038 - .6413\,\text{A} + .5713\,\text{C} - .3812\,\text{D} + .4813\,\text{AC} \qquad (2.1)$$
$$- 1.1213\,\text{BC} - .3938\,\text{BD} + .8913\,\text{CD}, \qquad i = 1, 2, \ldots, 16,$$

where the values of the factors are equal to 0 if they are designed to be run in the low levels and equal to 1 if they are run in the high levels. The results are given in the third column of Table 2.4.

[Please put Table 2.3 and Table 2.4 here]

For the prediction of variances, since this is an unreplicated experiment, there is no replication and, therefore, no variance for each experimental run. If there is only one (or two) significant dispersion effect(s), we can split the original 16 observations into 2 (or 4) group(s) of 8 (or 4) observations and estimate variances from them to construct variance function by using the approach in mean function. However, in this data set there are about six significant factors, which makes this approach difficult to apply. We thus propose to use the variance effects computed in Table 2.2 as variance function for prediction. Specifically, the predictions of log-variances can be made from the equation,

$$\hat{\mathfrak{D}}_i = \hat{\mathfrak{D}}_A \times \text{A} + \hat{\mathfrak{D}}_B \times \text{B} + \cdots + \hat{\mathfrak{D}}_{ABCD} \times \text{ABCD}, \qquad i = 1, 2, \ldots, 16, \qquad (2.2)$$

where $\hat{\mathfrak{D}}_A$ is the dispersion effect of the factor A, and A, B, ... are assigned as 0 and 1 according to the low and high levels of the factors. We note that this function is quite similar to the mean function (2.1), except there is no intercept term. This means that the predicted variances should be proportional to the correct variances with the constant scale $\varsigma = exp(\text{intercept})$. However, since we consider the variances in a relative sense for finding the optimal process condition, we can treat these predictions of variances as the correct variances in the decision process. The predictions of variances are listed in the fifth column of Table 2.4.

The strategy of finding the optimal experimental condition is to pick up the highest predicted average from a set of experimental runs, where the variances are limited inside a reasonable bound. This procedure is a simple version of the constrained optimization algorithm. For example, if we put the bound of variance at .25, we rule out the runs #3, #10, #13, #14. From the other runs, we can see that the run #4 has the highest predicted mean 24.9538 for A = H, B = H, C = L and D = L. This choice is different to run #5, A=L, B=L, C=H, D=L, obtained from the mean-variance plot, where the predicted average tenacity is 24.8813 and the predicted variance is .0502.

However, in the analysis of unreplicated experimental data, the variance calculations are incorrect because the variance effects are associated to the mean effects (c.f. p. 21 of Box and Meyer, 1986). We should eliminate the mean effects and consider the residuals to obtain estimates of dispersion effects as well as predicted variances. The details of this treatment is presented in next section.

3. ANALYSIS OF DISPERSION EFFECTS BASED ON RESIDUALS

The residual is the difference between the observation and the prediction. From the estimated mean function given in Table 2.3 we can make predictions and compute the residuals. The results are given in the fourth column of Table 2.4. Considering the residuals, we can apply the aforementioned variances calculation method to work out the dispersion effects. For example, from the eight residuals, .1112, .0463, -.2112, -.0637, -.5100, .3525, .1550, .1200, collected at the high level of the factor A, we find the variance is .06943 and the log-variance is -2.6675. With the log-variance calculated as -2.8440 for the low level of factor A, we thus estimate the dispersion effect of factor A as $\hat{\mathfrak{D}}^*_A = -2.6675 - (-2.8440) = .1765$. Similarly, the log-variances and dispersion effects for the other factors can be obtained (see Figure 3.1). To construct the ANOD chart, we estimate the variance as $s^2_{\mathfrak{D}} = .1905$ by using the dispersions in the third and the fourth order interactions such as -.3130, .2583, -.2236, -.1388, .8478 for ABC, ABD, ACD, BCD, ABCD, respectively. The overall average of the log-variances at low and high levels from all factors is computed as -2.9359. The overall decision limits LDL and UDL are then calculated as (-3.1243, -2.7275), (-3.2370, -2.6148) for $\alpha = .05, .01$, respectively, by using the formulas given in Section 2. We can also construct the ANOD chart using the individual central lines and the corresponding decision limits. For example, for factor A the average of the log-variances -2.6675 and -2.8440 at its low and high levels respectively, is computed as -2.7558. This gives the decision limits as $(-2.9541, -2.5574)$, $(-3.0668, -2.4447)$ for $\alpha = .05$ and .01 respectively. These two approaches of ANOD chart are both presented in Figure 3.1.

[Please put Figure 3.1 here]

From this ANOD chart we conclude that the dispersion effects AB, AC, AD, BC, BD, CD, BCD and ABCD are significant at $\alpha = .01$. This conclusion is quite different to the report given in Section 2, where B, C, AB, AD, BD, ABC and ACD are significant. Based on these estimates of dispersion effects, we predict the variances at all design points as .1150, .1297, .1304, .1007, .6702, .2602, 43.9831 (run #7), .1127, .1057, 29.3055(run #10), .1362, 2.4484 (run #12), .1253, .1647, .1806, .1307. In particular, three runs #7, #10 and #12 give very large predicted variances. To decide the optimal process setup, since the experimental run #3 has a relative small variance .1304 and highest mean 26.0763 (given in Table 2.4), we recommend the recipe A = L, B = H, C = L and D = L for the setup of the process variables. Note that this choice, run #3, is different to the selections, run #4 and run #5, suggested in Section 2. Although run #4 gives a smaller variance .1007, the average tenacity is predicted as 24.9538, which is considerably low compared to 26.0863 from run #3. The choice of run #5 based on the intuitive approach does not give good predictions in both mean and variance.

When we use the least squares method to estimate the mean effects in Table 2.3, we assume that the variances of the observations in the high and low levels of all factors are the same. However, from the ANOD chart we see that the variances from the high and low levels of AC, AD and CD are significantly different. Hence, we should estimate the mean function with the consideration of possible unequal variances in the model. Next, we use the weighted least squares method to obtain the maximum likelihood estimates of mean and variance functions.

4. MAXIMUM LIKELIHOOD ESTIMATION OF MEAN AND VARIANCE FUNCTION

Once the significant factors and the model have been identified, a more precise fitting is possible using maximum likelihood. The maximum likelihood estimates of location and dispersion effects can be obtained through the conditional approach (c.f. Hartley and Jayatillake, 1973), where the mean and variance effects are estimated iteratively. Specifically, in our problem, conditioning on variances predicted from the variance function constructed from dispersion effects given in Figure 3.1, we estimate the mean effects from weighted regression, where the weights are the reciprocal of predicted variances. The estimates of mean function is reported in Table 4.1. Based on the updated mean function, we can make predictions and work out the residuals. The weights, predictions of mean and residuals are reported in Table 4.2. From these residuals we update the variance estimates. We then go on to update the weights from the variances calculated from the new residuals, and update the estimates of mean function. We loop this procedure eight times to obtain the converged estimates of mean and variance effects. The final estimates of mean function are given in Table 4.3.

[Please put Table 4.1, Table 4.2 and Table 4.3 here]

The ANOD chart for dispersion effects based on the updated residuals is given in Figure 4.1 with the overall central line at -3.0831 and overall decision limits (-3.2847, -2.8814), (-3.4175, -2.7487) for $\alpha = .05$ and .01, respectively. The estimated variance is $s^2_{\mathfrak{D}} = .1688$ from the dispersion effects $-.5311$, .5447, .3464, $-.3813$ of the third-order interactions. We did not include the fourth-order interaction in computing $s_{\mathfrak{D}}^2$ due to its large dispersion 8.3659. The $t_{\alpha/2,\nu}$ values with the degree of freedom $\nu =$ 4 are reported as 2.776 and 4.604 for $\alpha = .05$ and .01 respectively. The individual central lines and decision limits are computed similarly and displayed in Figure 4.1 for comparison with the overall decision lines. We note that the dispersion effects from the main effects A, B, C, D are all above the overall UDL's and within the decision limits derived from the individual central lines. In general, we conclude that the interactions AB, AC, AD, BC, BD, CD and ABCD are significant and other factors are not significant at $\alpha = .01$. Comparing the ANOD chart (Figure 4.1) constructed using weighted least squares to the chart (Figure 3.1) constructed from the unweighted regression, we conclude that the significance of the important dispersion factors is clearer in the case of using weighted regression, which is expected because this approach of including the unequal variances in the model fits the data better.

[Please put Figure 4.1 here]

Based on the prediction of average yarn strengths given in Table 4.4, we compute the overall average as 24.4174 and the estimated variance as $s_M^2 = .01540$ from the third- and fourth- order mean effects. We compute the decision limits as (24.3610, 24.4738) and (24.3290, 24.5059) for $\alpha = .05$ and .01 respectively. Then, we display the ANOM chart in Figure 4.2 for comparison with the chart given in Figure 2.1(a) based on the original 16 observations. In general, the identified significant factors are the same, but the difference of significant and insignificant effects is much clearer in Figure 4.2. This implies that the maximum likelihood method gives more precise estimates, and the model with unequal variances should be utilized to analyze this data set.

[Please put Figure 4.2 here]

To find the optimum design point, since the decision procedure based on plotting location against dispersion effects is inconclusive, we need the predictions of average and standard deviation of yarn strengths from the mean and variance functions similar to (2.1) and (2.2), respectively. The predictions from the mean function is reported in the second column of Table 4.4. The standard deviations computed from the residuals are listed in the fifth column of Table 4.4. We recommend the experimental condition run #3, A = L, B = H, C = L and D = L due to its high predicted average 25.9854 and low standard variation .0701. From the data (see Table 2.1) itself, of course, one will pick up the highest average tenacity at run #3 and suggest it as the best design plan. However, without the study of dispersion effects, we have no idea if this particular run will guarantee a small process variation or not. For example, run #4 and #13 both give high average tenacities, but their standard deviations, .3370 and .3531, respectively, are quite large compared to the standard deviation .0701 in run #3. Hence, selecting the optimal design point from the sample averages alone does not necessarily give us the best process control, especially in the case that the consistency of process is of much concern.

When the direction of process improvement is determined, we should conduct a small follow-up experiment to confirm the finding of optimal process setup. This confirmation experiment is particularly important in the case that the true optimum is not at any particular design point and, in the case that some of the parameters, especially the continuous variable, are expected to have non-linear (especially non-monotone) effects on the response. In this experiment there are two class variables, factor A (side to side) and factor C (yarn type) and two continuous variables, factors C (pick density) and D (air pressure). We recommend the follow-up experiment be conducted around the pre-selected optimum, that is, experiment #3, A = L (Nozzle), B = H (Ring spun), C = L (35 ppi) and D = L (30 psi). We suggest to fix the factor A at Nozzle and B at Ring spun and consider the following central composite design for the continuous variable in order to utilize the response surface method for finding the true optimum. Specifically, we treat the optimal levels 35 ppi and 30 psi of factors C and D, respectively as the central points and run a complete 2^2 experiment at the combinations (30 ppi, 25 psi), (30 ppi, 35 psi), (40 ppi, 25 psi) and (40 ppi, 35 psi) as the star points. Then, the response surface with two input variables C and D can be formed and the path of steepest ascent can be established to reach the true optimal point for pick density and air pressure. Chapter 15 of the book written by Box, Hunter and Hunter (1978) on response surface methodology gives a good reference and detail of computational procedure.

5. CONCLUSION

Based on the economic consideration, it is very often that some highly fractionated factorial experiments without replications are recommend to screen the inactive location and dispersion factors. In this article, we propose a simple graphical procedure extended from SPC charts to make the analysis and interpretation of the unreplicated experiment easier and more intuitive. With the increasing interest in reducing the product variability, we recommend the user to utilize the weighted least square method to obtain the maximum likelihood estimates of mean and variance functions and analyze the dispersion effects based on residuals. The decision of optimal process setup should be made from the predictions of average and variance at all combinations of design configurations. With the algorithm of constrained (on variances) optimization (on mean levels) and the response surface constructed from the follow-up experiment, the true optimum can be found systematically. The presentation of this article serves as a guide for improving product or process quality.

REFERENCES

Baker, T. R. (1986), 'Quality engineering by design: Taguchi's philosophy,' *Quality Progress*, 19(12), 32-42.

Box, G. E. P., Bisgaard, S. and Fung, C. (1988), 'An explanation and critique of Taguchi's contributions to quality engineering,' *Quality and Reliability Engineering International*, 123-131.

Box, G. E. P., Hunter, W. G. and Hunter, J. S. (1978), *Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building.* Wiley, New York.

Box, G. E. P., and Meyer, D. (1986a), 'An analysis of unreplicated fractional factorials,' *Technometrics*, 28, 11-18.

Box, G. E. P., and Meyer, D. (1986), 'Dispersion effects from fractional designs,' *Technometrics*, 28, 19-27.

Daniel, C. (1976), *Applications of Statistics to Industrial Experimentation*, Wiely, New York.

Hartley, H. O. and Jayatillake, K. S. E. (1973), 'Estimation for linear models with unequal variances,' *Journal of the American Statistical Association*, 68, 189-192.

Hogg, R. V., *et. al.* (1985), 'Statistical education for engineers: An initial task force report,' *The American Statistician*, 39, 168-175.

Johnson, R., Clapp, T. and Baqai, N. (1989), 'Understanding the effect of confounding in design of experiments: A case study in high-speed weaving,' *Quality Engineering*, 1(4), 501-508.

Kackar, R. N. (1985), 'Off-line quality control, parameter design, and the Taguchi method (with discussion),' *Journal of Quality Technology*, 17(4), 176-188.

Leon, R. V., Shoemaker, A. C. and Kacker, R. N. (1987), 'Performance measures independent of adjustment (with discussion),' *Technometrics*, 29(3), 253-265.

Nelson, P. R. (1988), 'Testing for interactions using the analysis of means,' *Technometrics*, 30, 53-61.

Ott, E. R. (1967), 'Analysis of means – A graphical procedure,' *Industrial Quality Control*, 24, 101-109.

Ott, E. R. (1975), *Process Quality Control.* McGraw-Hill, New York.

Raming, P. F. (1983), 'Applications of the analysis of means,' *Journal of Quality Technology*, 15, 19-25.

Ryan, T. P. (1989), *Statistical Methods for Quality Improvement.* Wiley, New York.

Ullman, N. R. (1989), 'The analysis of means (ANOM) for signal and noise,' *Journal of Quality Technology*, 21, 111-127.

Table 2.1. A Complete 2^4 Design and Yarn Strength Data

Run	A(1)	B(2)	C(3)	D(4)	AVERAGE TENACITY (g/den)
1	−	−	−	−	24.50
2	+	−	−	−	23.55
3	−	+	−	−	25.98
4	+	+	−	−	25.00
5	−	−	+	−	24.63
6	+	−	+	−	24.51
7	−	+	+	−	24.68
8	+	+	+	−	23.93
9	−	−	−	+	23.73
10	+	−	−	+	22.05
11	−	+	−	+	24.52
12	+	+	−	+	23.64
13	−	−	+	+	25.68
14	+	−	+	+	25.78
15	−	+	+	+	24.10
16	+	+	+	+	24.23

(1): A = Side to side; Low (−) = Nozzle, High (+) = Opposite.

(2): B = Yarn Type; Low (−) = Air spun, High (+) = Ring spun.

(3): C = Pick density (ppi); Low (−) = 35, High (+) = 50.

(4): D = Air Pressure (psi); Low (−) = 30, High (+) = 45.

Table 2.2. Location and Dispersion Effects

Main Effects

	A	*B*	*C*	*D*
mean H	24.0863	24.5100	24.6925	24.2163
mean L	24.7275	24.3038	24.1213	24.5975
diff m ($\hat{\mu}$)	−0.6412	0.2062	0.5712	−0.3812
Sd H	1.1059	0.7339	0.6909	1.1925
Sd L	0.7532	1.2093	1.1668	0.7212
log var H	0.2014	−0.6189	−0.7395	0.3522
log var L	−0.5668	0.3802	0.3086	−0.6536
diff v ($\hat{\mathfrak{D}}$)	0.7682	−0.9991	−1.0481	1.0058

Two-Factor Interactions

	AB	*AC*	*AD*	*BC*	*BD*	*CD*
mean H	24.4175	24.6475	24.4363	23.8463	24.2100	24.8525
mean L	24.3963	24.1663	24.3775	24.9675	24.6038	23.9613
diff m ($\hat{\mu}$)	0.0212	0.4812	0.0588	−1.1212	−0.3938	0.8912
sd H	0.6903	0.8147	1.2341	0.8167	0.4174	0.8954
sd L	1.2444	1.1087	0.7074	0.8000	1.3276	0.8773
log var H	−0.7414	−0.4099	0.4206	−0.4050	−1.7474	−0.2209
log var L	0.4374	0.2065	−0.6922	−0.4463	0.5668	−0.2618
diff v ($\hat{\mathfrak{D}}$)	−1.1788	−0.6164	1.1128	0.0413	−2.3142	0.0409

Three- and Four-Factor Interactions

	ABC	*ABD*	*ACD*	*BCD*	*ABCD*
mean H	24.3213	24.5000	24.5150	24.2788	24.3963
mean L	24.4925	24.3138	24.2988	24.5350	24.4175
diff m ($\hat{\mu}$)	−0.1712	0.1862	0.2162	−0.2562	−0.0212
sd H	1.2333	0.9199	0.7076	1.1272	1.0464
sd L	0.6983	1.0767	1.2239	0.8469	0.9645
log var H	0.4194	−0.1670	−0.6918	0.2394	0.0907
log var L	−0.7183	0.1479	0.4041	−0.3323	−0.0724
diff v ($\hat{\mathfrak{D}}$)	1.1377	−0.3149	−1.0959	0.5717	0.1631

Table 2.3. Initial Estimates of Mean Function

Parameter	Coefficient	Standard Error	t-value
Intercept	24.7038	.2363	104.55
A	−.6413	.1671	−3.84
C	.5713	.1671	3.42
D	−.3812	.1671	−2.28
AC	.4813	.1671	2.88
BC	−1.1213	.1671	−6.71
BD	−.3938	.1671	−2.36
CD	.8913	.1671	5.33

$s = .3342$, Multiple $R^2 = .9370$, $F_{7,8} = 16.9944$

Table 2.4. Predictions from Mean and Variance Functions

Run	Observation	Prediction of Mean	Residual	Variance
1	24.50	24.5613	−0.0612	0.0535
2	23.55	23.4388	0.1112	0.1475
3	25.98	26.0763	−0.0963	2.1297
4	25.00	24.9538	0.0463	0.1485
5	24.63	24.8813	−0.2513	0.0502
6	24.51	24.7213	−0.2112	0.0514
7	24.68	24.1538	0.5263	0.0985
8	23.93	23.9938	−0.0637	0.1259
9	23.73	23.6825	0.0475	0.1713
10	22.05	22.5600	−0.5100	101.8802
11	24.52	24.4100	0.1100	0.0552
12	23.64	23.2875	0.3525	0.1228
13	25.68	25.7850	−0.1050	0.6898
14	25.78	25.6250	0.1550	0.9903
15	24.10	24.2700	−0.1700	0.0562
16	24.23	24.1100	0.1200	0.0549

Note: The variances are predicted from the variance function constructed in Section 2.

Table 4.1. Weighted Regression on Mean Responses

Parameter	Coefficient	Standard Error	t-value
Intercept	24.5442	.9520	462.90
A	−.4416	.0358	−12.33
C	.3868	.0359	10.77
D	−.1933	.0362	−5.34
AC	.5250	.0364	14.53
BC	−1.1165	.0332	−33.68
BD	−.3906	.0328	−11.92
CD	.9237	.0363	25.44

$s = .1475$, Multiple $R^2 = .9069$, $F_{7,8} = 371.8004$

Table 4.2. Weights and Predictions from Mean Function

Run	Weight	Observation	Prediction	Residual of Mean
1	8.6939	24.50	24.4858	0.0142
2	7.7115	23.55	23.5192	0.0308
3	7.6707	25.98	25.9929	−0.0129
4	9.9272	25.00	25.0263	−0.0263
5	1.4920	24.63	24.5404	0.0896
6	3.8434	24.51	24.6239	−0.1139
7	0.0227	24.68	23.8145	0.8655
8	8.8741	23.93	23.8980	0.0320
9	9.4567	23.73	23.7595	−0.0295
10	0.0341	22.05	22.7929	−0.7429
11	7.3405	24.52	24.4853	0.0347
12	0.4084	23.64	23.5188	0.1212
13	7.9782	25.68	25.6615	0.0185
14	6.0732	25.78	25.7449	0.0351
15	5.5371	24.10	24.1544	−0.0544
16	7.6499	24.23	24.2378	−0.0078

Table 4.3. Final Estimates of Mean Function

Parameter	Coefficient	Standard Error	t-value
Intercept	24.5739	.0216	1135.68
A	−.3871	.0060	−64.03
C	.3855	.0060	63.77
D	−.2095	.0060	−34.65
AC	.5401	.0210	25.68
BC	−1.1799	.0061	−193.07
BD	−.3349	.0210	−15.94
CD	.8698	.0060	143.89

$s = .1241$, Multiple $R^2 = .9999$, $F_{7,8} = 12454.18$

Table 4.4. Final Predictions from Mean and Variance Functions

Run	Observation	Mean Prediction	Residual	Weight	Standard Deviation
1	24.50	24.4705	0.0295	9.4328	0.3256
2	23.55	23.5433	0.0067	199.4230	0.0708
3	25.98	25.9854	−0.0054	203.3128	0.0701
4	25.00	25.0582	−0.0582	8.8044	0.3370
5	24.63	24.6260	0.0040	210.4214	0.0689
6	24.51	24.7789	−0.2689	0.0787	3.5640
7	24.68	23.7810	0.8990	0.0000	3223.355
8	23.93	23.9339	−0.0039	210.6589	0.0690
9	23.73	23.7261	0.0039	210.6138	0.0689
10	22.05	22.7989	−0.7489	0.0000	534.9210
11	24.52	24.5711	−0.0511	0.1603	2.4978
12	23.64	23.6439	−0.0039	210.4337	0.0689
13	25.68	25.6214	0.0586	8.0191	0.3531
14	25.78	25.7743	0.0057	201.0705	0.0705
15	24.10	24.1065	−0.0065	203.1570	0.0702
16	24.23	24.2595	−0.0295	9.8812	0.3181

Note: The standard deviation is predicted from the final variance function.

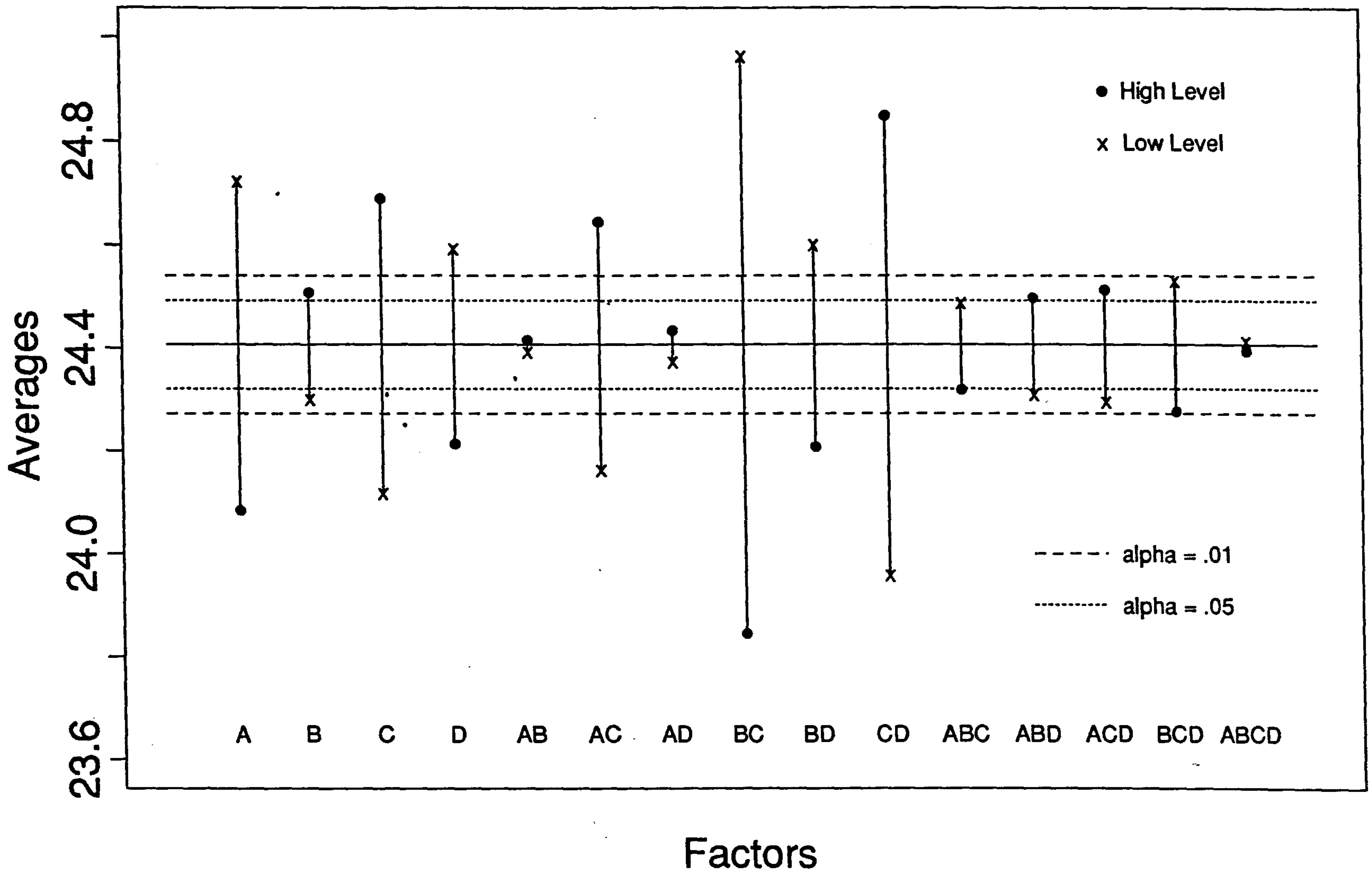

Figure 2.1(a), Analysis of Mean Effects

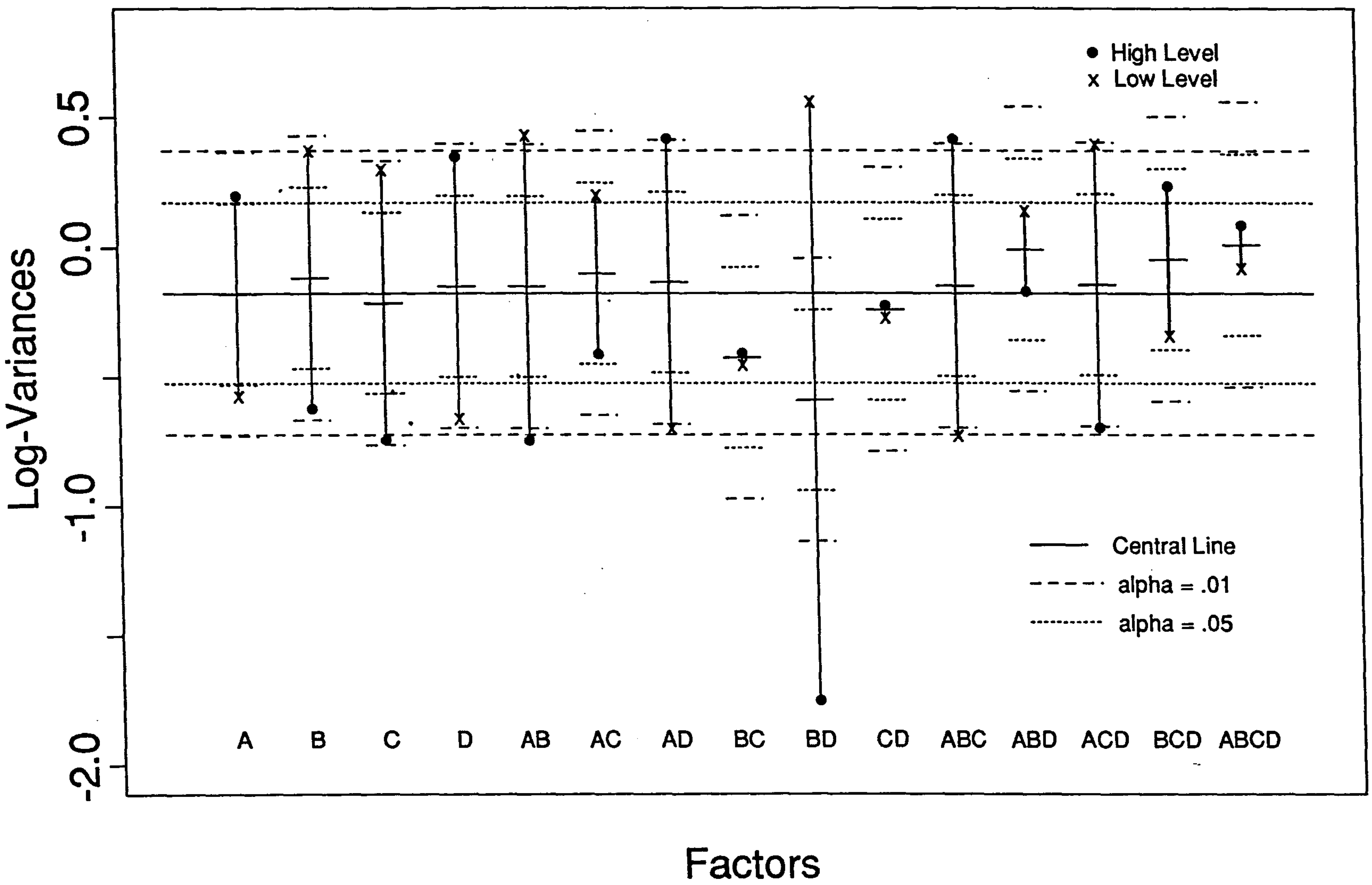

Figure 2.1(b), Analysis of Dispersion Effects

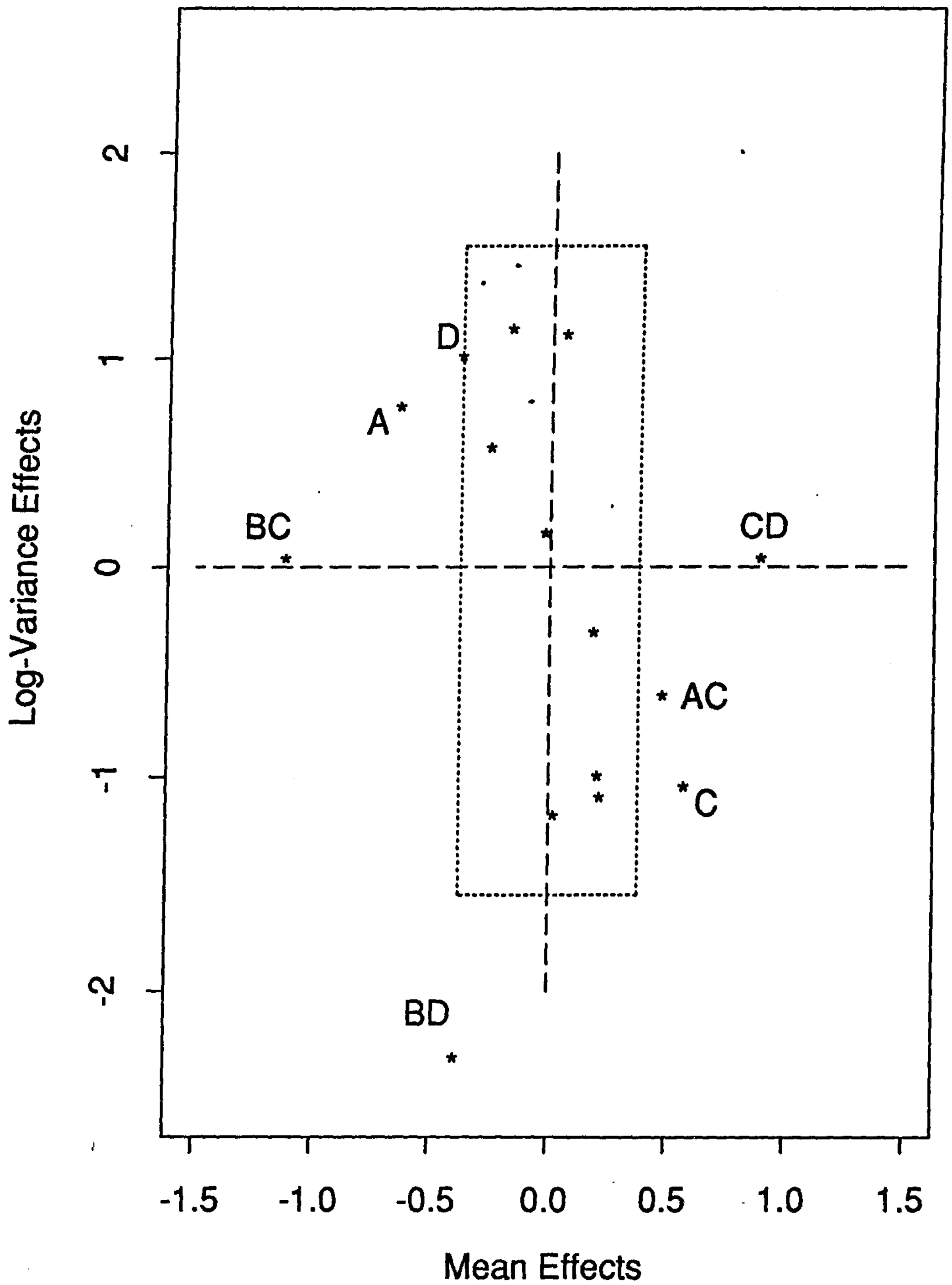

Figure 2.2, Mean Effects against Log-Variance Effects

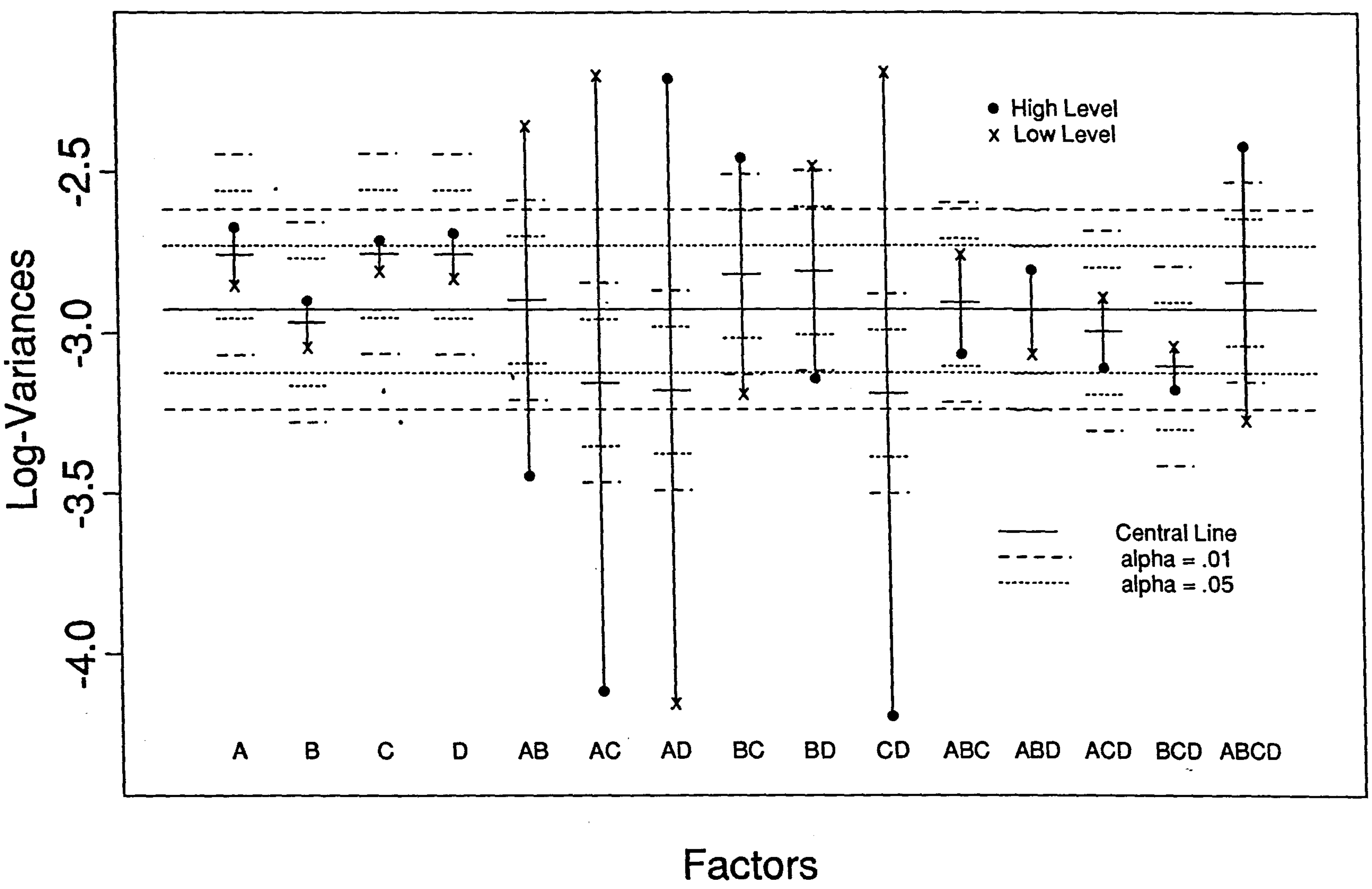

Figure 3.1, Analysis of Dispersion Effects from Residuals

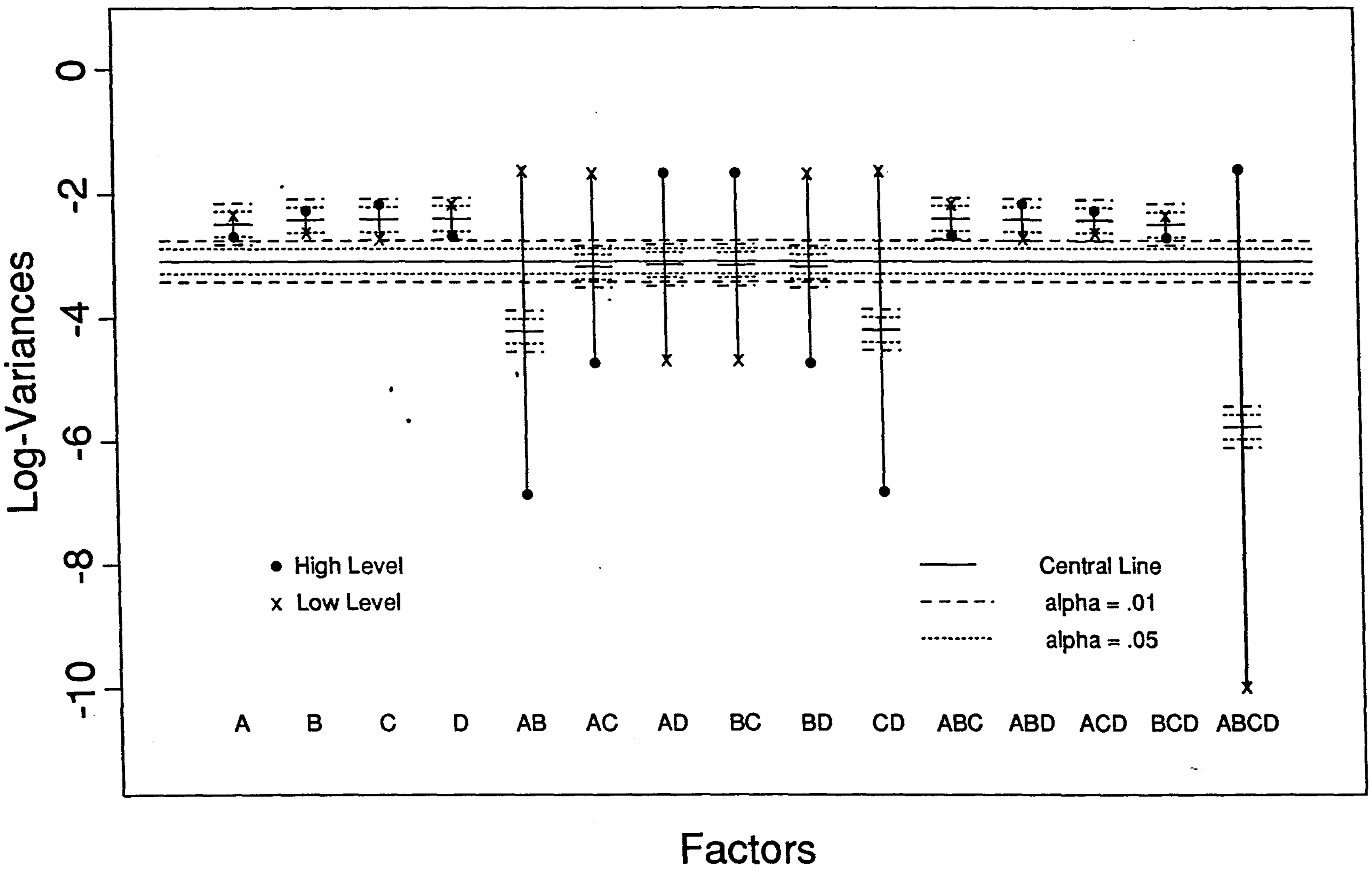

Figure 4.1, Analysis of Dispersion Effects
from Maximum Likelihood Estimation of Model Parameters

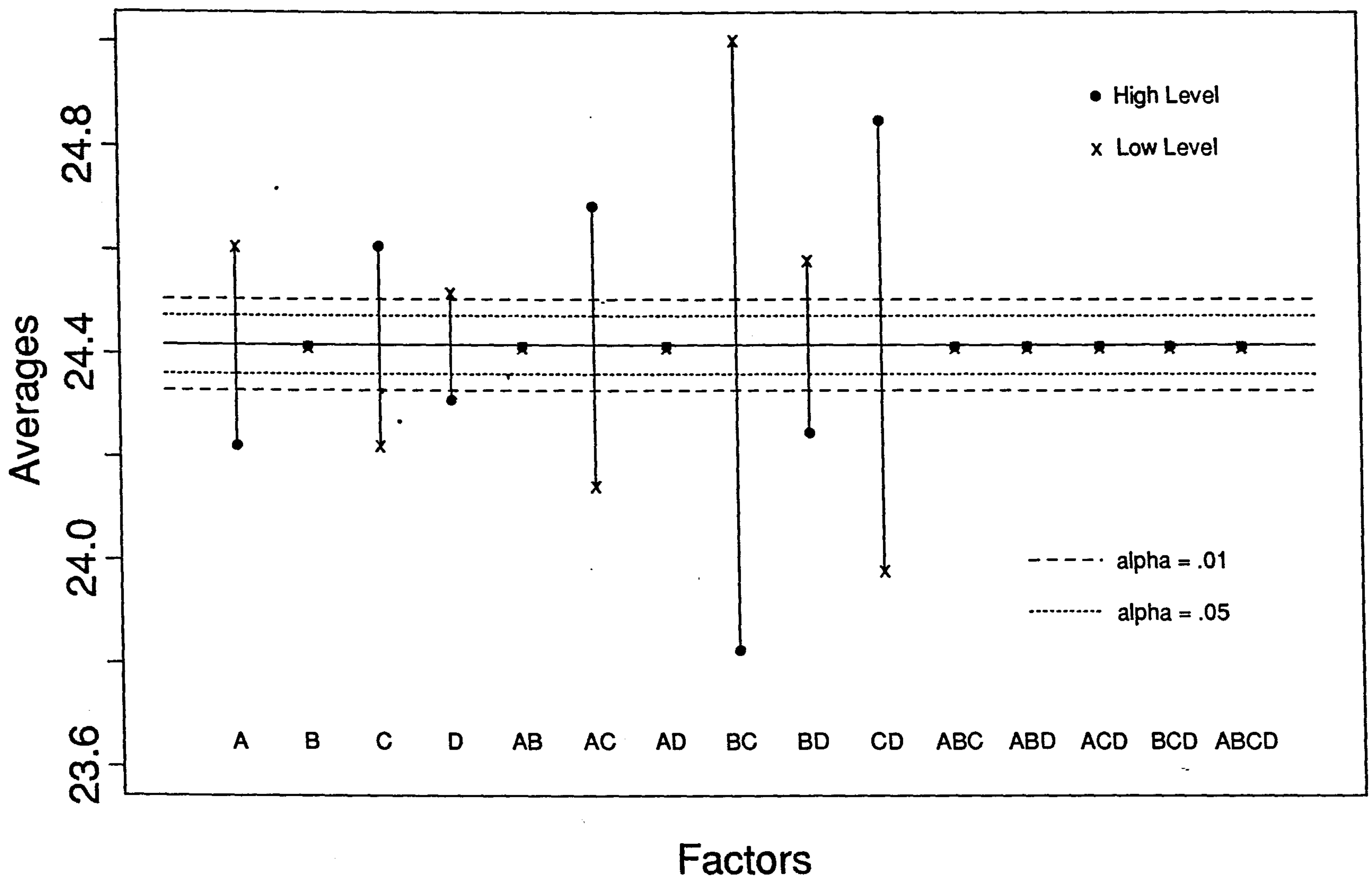

Figure 4.2, Analysis of Mean Effects from Predictions