Introduction

Quality improvement in industrial processes, using robust design, has recently received much attention. Robust design originates in Japan and is associated with the name of Taguchi. Process quality is evaluated by the loss function, which measures the process deviation from the target. Quality improvement can be achieved both by moving the process mean in the right direction and by minimizing the process variation. “Traditional” experimental design approach provides methods of identification of factors with important location effects. The robust design methodology goes farther – it aims to find both dispersion and location effects of factors in order that the process variation can be minimized.

The process variation is due to noise. It can be caused e.g. by uncontrolled changes of process parameters such as temperature or concentration over time or over location. Keeping such variables fixed during the normal process may be either impossible or costly. Though they cannot be controlled during the normal process, some of them can be included in the experiment as so called noise factors. The robust design methodology distinguishes two main types of factors. Control factors are variables whose values remain fixed during the process once they are chosen. Noise factors are hard to control during the normal process but they are controllable in the experiment. Different units occurring in a process are another source of noise. Operators, machines or parts are a typical example. Though these units may be the same throughout the normal process, we can consider them as levels of the noise factor, too, what means that the definition of two types of factors is not exact. The experiment is carried with both the control and noise factors. Beside changes of the control factor settings, levels of the noise factors are varied in order to imitate noise during the normal process. The aim is to find the control factor settings that establish the process output less sensitive to noise variation. It is assumed that the set of factors can be divided into factors that affect only the process mean and those that affect both the mean and variation. The robust design consists in several steps:
- Identification of important location and dispersion effects of the control factors,
- Selection of the control factor settings minimizing the process variation,
- Selection of factors with location effects that have not dispersion effect in order the process can be adjusted.

Two experimental strategies are used, i.e. crossed and single (or combined) array, where mostly fractional factorial design is used. Relative merits of each experimental approach were discussed in literature but we will concentrate on modelling strategies. These are location and dispersion modelling or response modelling. The aim of the paper is to review methods used in robust design. As an example that illustrates the problem the data reported in [6] are reanalyzed.

Example – Layer Growth Experiment [6]

One of the steps in fabricating integrated circuit devices is to grow an epitaxial layer on polished silicon wafers. The silicon wafers are mounted on a six-faceted cylinder (two wafers per facet), which is placed inside a metal bell jar where the deposition process takes place. The jar is injected with chemical vapours and heated. The process continues for a specific
time. The target layer thickness is 14.5 µm. In the experiment eight control factors are considered – susceptor-rotation method (A), code of wafers (B), deposition temperature (C), deposition time (D), arsenic flow rate (E), hydrochloric acid etch temperature (F), hydrochloric acid flow rate (G), and nozzle position (H). All the control factors have two levels in the experiment. The facets and location are taken as levels of noise factors. Data reported in [6] are truncated and only four facets are considered. Factor facet (M) has four levels, factor location (L) has two levels. The experimental plan uses fractional factorial design with sixteen control factor settings and full factorial design with eight noise factor settings. That means there are eight observations per control factor setting. Two out of sixteen control factor settings along with noise factor settings and experimental results are shown in the following table. Levels of the control factors are denoted – and +.

<table>
<thead>
<tr>
<th>Noise Factor</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A  B  C  D  E  F  G  H</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L2 (top)</td>
<td>15.318</td>
<td>15.428</td>
<td>15.266</td>
<td>15.406</td>
</tr>
</tbody>
</table>

### Location and dispersion modelling

a) The sample mean \( \bar{y} \) and log sample variance \( \ln s^2 \) over noise replicates are computed. The noise replicates are not true replicates. They represent runs at the same control factor setting but with varying noise factor levels. The location and dispersion model have the form

\[
y_{avg,i} = x_i^T \beta + \varepsilon_i \quad (1)
\]

\[
y_{lns^2,i} = z_i^T \gamma + \varepsilon_i , \quad (2)
\]

where \( x_i^T \) and \( z_i^T \) denote settings of the control factors. They have some elements in common if some control factors affect both the mean and variation. Parameters \( \beta \) and \( \gamma \) denote location and dispersion effects, respectively. Errors \( \varepsilon_i \) are assumed to have the normal distribution with zero mean and constant variance. It is obvious, that if there some non-zero dispersion effects exist in (2), this assumption cannot be satisfied and heteroscedasticity is present in (1). Neglecting heteroscedasticity does not bias estimate of \( \beta \), so that results of graphical method of effect identification should not be essentially affected. On the other hand, estimates of effect variances are biased and this may be a disadvantage when the effects are tested for significance.

b) As an alternative, joint generalized linear models (GLM) for mean and dispersion are considered. The model for the mean has the identity link function, so that

\[
E(y_i) = \mu_i = x_i^T \beta . \quad (3)
\]

The error distribution is normal with dispersion parameter \( \Phi_i = \sigma_i^2 \) and variance function \( V(\mu_i) = 1 \). The mean over replicates \( y_{avg} \) serves as a response. The dispersion model has the logarithmic link function, as multiplicative heteroscedasticity is assumed, i.e.

\[
\sigma_i^2 = \exp(z_i^T \gamma) . \quad (4)
\]
Deviance components, i.e. the squares of residuals from normal model (3) can be used as a response. They have a gamma distribution with dispersion parameter \( \Phi = \sigma_i^2 \), variance function \( V(\mu_i) = 1 \) and a scale factor of 2. When \( r \) replicates exist, \( \sigma_i^2 \) can be used as a response. They have a gamma distribution, too, but the scale factor is \( 2/(r-1) \). Models can be fitted by maximum likelihood using the log-likelihood function

\[
l(\beta, \gamma) = -\frac{1}{2} \sum \ln \sigma_i^2 - \frac{1}{2} \sum (y_i - x_i^T \beta)^2 / \sigma_i^2 ,
\]

where \( \sigma_i^2 \) is given by (4). As the estimated expected information matrix is block-diagonal, estimation reduces to two separate algorithms for \( \beta \) and \( \gamma \). For given \( \sigma_i^2 \), estimate of \( \beta \) is obtained by the weighted least squares method with weights \( 1/\sigma_i^2 \), for given \( \beta \), maximum likelihood estimate of \( \gamma \) in the GLM for dispersion is calculated. That is, joint maximum likelihood estimates of \( \beta \) and \( \gamma \) can be performed by iteratively reweighted least squares, as Aitkin [1] showed.

**Response modelling**

The response \( y \) is modelled as a function of both the control and noise factors and important main effects and interactions are identified. The model has a form

\[
y_i = x_i^T \beta + z_i^T \gamma + (xz)_i^T \delta + \varepsilon_i ,
\]

where \((xz)_i^T\) denotes a part of the i-th design matrix’s row corresponding to the control-by-noise interactions, \( \delta \) denotes the interaction effects. Errors \( \varepsilon_i \) are assumed to have the normal distribution with zero mean and constant variance. Significant effects of the noise factor confirm the presumption about them being the important source of variation. Large control-by-noise interactions show the control factors that have dispersion effects. These control factors are then used to minimize the process variation. The control factors without dispersion effects can be used to adjust the process mean.

In order to select acceptable settings of dispersion factors (i.e. the control factors with significant dispersion effects), interaction plots are used and decisions are based on visual inspection. Engel and Huele [3] mention more formal method consisting in building a model for the process variance that is derived from the response model. Wu [6] calls it the transmitted variance model. The approach is rather suspicious and so its description is omitted here.

Yet another approach was applied in this paper. Absolute residuals from a preliminary fit of the response model are modelled as a function of control factors to find a model of heteroscedasticity caused by other than the noise factors

\[
y_{\mid \beta, \gamma} = x_i^T \gamma + \varepsilon_i .
\]

Model (6) is refitted by weighted least squares. Weights are estimated by \( 1/ \hat{\sigma}_i^2 \). Errors \( \varepsilon_i \) are assumed to be approximately normally distributed.

**Model fitting**

Fitting a model based on the fractional factorial experiment, one must be aware of effects’ aliasing. In an experiment with \( n \) runs only \( n-1 \) effects are estimable. The design matrix has \( n-1 \) columns at maximum. No aliases (i.e. identical columns) can occur in it. Usually \( 2^{N-p} \)
experiments are used, i.e. the experiments with \( N \) two-level factors, where only \( 2^{N-p} \) factor settings are selected. The relation defining alias structure must be used to determine which main effects and interactions are aliased. If all possible main effects and interactions are considered and only one run per factor setting (considering both the control and noise factors) has been carried, the model is saturated and no degrees of freedom are left for the residual component estimate and, accordingly, for significance tests. The problem is solved by different ways. In case of two-level factors half normal or normal plots are used most frequently, see e.g. [6]. This informal graphical method involves visual judgement. Lenth’s method of testing effect significance is a relatively new formal test. A robust estimator of the standard deviation of the effect employs the median of effects in combination with trimming. Special tables of critical values are needed [6]. When some of the factors have more than two levels, the two methods are inapplicable. Stepwise procedure for nested models, when sums of squares are compared, can be used. Effects are added or subtracted based on a significant or insignificant \( F \) statistic or possibly a \( \chi^2 \) statistics in case of GLM.

### Application to layer growth experiment

**Location and dispersion models**

a) Models (1) and (2) are given by

\[
\hat{y}_{\text{avg},j} = 14.352 + 0.4019x_{Di} + 0.0867x_{Hi} \\
\hat{y}_{\text{ln}s^2,j} = -1.9531 + 0.6173x_{Di} - 0.9791x_{Hi}.
\]

The control factors have two levels coded -1 and +1. The subscript \( i \) denotes the \( i \)-th control factor setting \((i = 1, 2, \ldots, 16)\).

<table>
<thead>
<tr>
<th>location model</th>
<th>dispersion model</th>
</tr>
</thead>
<tbody>
<tr>
<td>effect</td>
<td>estimate</td>
</tr>
<tr>
<td>D</td>
<td>0.4019</td>
</tr>
<tr>
<td>H</td>
<td>0.0867</td>
</tr>
</tbody>
</table>

It follows, that the control factor A should be set at its – level and H at its + level so that variation is reduced. Factor D can be used to move mean towards the target according to these results.

b) Model (4) of the response \( s^2 \) is given by

\[
\hat{y}_{s^2,j} = -1.7326 + 0.4928x_{Di} + 0.4519x_{Di} - 0.8610x_{Hi}.
\]

<table>
<thead>
<tr>
<th>dispersion model</th>
</tr>
</thead>
<tbody>
<tr>
<td>effect</td>
</tr>
<tr>
<td>A</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>H</td>
</tr>
</tbody>
</table>

In comparison with the previous method, factor D in addition to A and H has been identified to have a dispersion effect.

**Response model**

Model (6) is given by
\[ \hat{y}_i = 14.3520 + 0.4019x_{Di} + 0.0867x_{Hi} + 0.3299x_{Li} - 0.0125x_{M_{12}} + 0.0461x_{M_{13}} + 0.0442x_{M_{14}} + 0.0316x_{M_{i1}}x_{Ci} + 0.0814x_{M_{i2}}x_{Ci} + 0.0016x_{M_{i3}}x_{Ci} - 0.2388x_{Hi}x_{Li} \]

All the control factors and noise factor L have two levels coded -1 and +1 and each of them (including interaction HL) is represented by one column in the design matrix. Noise factor M has four levels and three columns in the design matrix correspond to it. Helmert orthogonal contrasts were chosen in the analysis. Considering the fact that the model serves for identification of important effects and not for parameter estimation, the choice of contrasts is non-essential and other details are omitted. The subscript \( i \) denotes the \( i \)-th experimental run \((i = 1, 2, \ldots, 128)\). P-values at all selected effects were smaller than 0.0001. As control-by-noise interactions CM and HL proved to be highly significant, their plots were constructed (Fig. 1).

![Fig. 1. Plots of control-by-noise interaction CM and HL](image)

Figure suggests setting control factor C at its – level and H at its + level because the response plot is flatter at them and response variation should thus be reduced. Interaction CM is less important than interaction HL (as it follows from ANOVA table.)

Normal linear model (7) is given by

\[ \hat{y}_{i|A} = 0.2301 + 0.0823x_{Hi}. \]

P-value at effect of A was smaller than 0.0001. It follows that with the aid of control factor A the amount of random variation could be reduced. As A affects random variation, it could not be discovered in the preliminary response model.

The response model was refitted by weighted least squares

\[ \hat{y}_i = 14.3761 + 0.4176x_{Di} + 0.0835x_{Hi} + 0.3352x_{Li} + 0.0030x_{M_{12}} + 0.0334x_{M_{13}} + 0.0774x_{M_{14}} + 0.0220x_{M_{i1}}x_{Ci} + 0.0516x_{M_{i2}}x_{Ci} - 0.0017x_{M_{i3}}x_{Ci} - 0.2057x_{Hi}x_{Li} + 0.0712x_{Di}x_{Li} \]

In this updated model p-value of interaction effect DL is 0.0004, while in the unweighted alternative with the same effects included it was 0.0768. Based on the updated response model four dispersion effects have been identified, namely the effects of control factors A, C, D, and H. Factors C, D, and H affect variation due to changes of the noise factors, factor A affects the amount of random variation.

Summary and discussion:

Results of various methods differ. The differences are obviously caused by the fact, that the response variation at a control factor setting arises partly from varying noise factor levels (explained variation), partly is due to the experimental error, i.e. by changes of variables not included in the experiment and therefore out of control (random variation). The other reason
is neglecting heteroscedasticity in some models. As factor D has both location and dispersion effect, its use for robust design is questionable (oppose to the original analysis in [6]).

Absolute residuals from a preliminary fit of the response model were used in this paper. Most procedures estimating structural parameters of variance functions take transformations of absolute residuals or sample variances or standard deviations from replicates. Davidian and Carroll [2] reviewed methods of variance function estimation and explored their efficiency and robustness. They pointed out, based on asymptotic theory, that using logarithms of absolute residuals is less efficient than using absolute residuals themselves and it is less efficient than using squared residuals. When up to ten replicates are analyzed, absolute residuals and squared residuals are more efficient relative to standard deviations and variances, respectively. The logarithm method based on standard deviations, though, is more efficient relative to log absolute residuals for four and more replicates, but is always worse than analysis based on squared residuals. These features may not be true for finite samples. Our simulation study conducted as a part of the IGA project indicates, that absolute residuals perform quite well. On the other hand, tests used in GLM with gamma distribution seem to be rather progressive, i.e. the null hypothesis is rejected more frequently than it should be, whereas in log sample variance model the number of rejections correspond to the specific significance level. Some disadvantage of the logarithm method consists in its sensitivity towards inliers, i.e. very small sample variances. It is recommended to delete these inliers but the problem of insufficient degrees of freedom arises when the number of control factor settings is small.

Comparison of the analyses clearly points to a danger of misidentification of important dispersion effects and trickiness of some instructions that are recommended in literature. On the other hand, industrial practise demands simple methods which can be applied with the aid of available software. The approach suggested in this paper, i.e. response modelling with weighted least squares estimation of factor effects, seems to be both relatively simple and quite powerful. Moreover, it enables to separate variation caused by the noise factor having been identified from another noise, thus contributing to the better knowledge of the process.

References:

Eva Jarošová
University of Economics Prague
Department of Statistics and Probability
W. Churchill Sq. 4
130 67 Prague 3
Czech Republic
jarosova@vse.cz