

GENERALIZED LINEAR MODELS FOR SIMULATED HYDROFORMING PROCESSES

U. Gather, S. Kuhnt

Chair of Mathematical Statistics and Industrial Applications, Department of Statistics,
University of Dortmund, 44221 Dortmund, Germany

Summary

In high-pressure sheet metal hydroforming (HBU) simulations provide a useful tool for off-line quality improvement. We analyse such finite element simulations by using extended generalized linear models. Especially, we determine settings of the design parameters which minimize the estimated variance of the quality measures in these models. At the same time the parameter setting will fulfil the estimated mean model equation at the desired target. The optimal process is carried out simultaneously for the two quality characteristics blank thinning and geometrical accuracy. This is done using a multivariate loss function approach.

1 Introduction

Off-line planning methods for hydroforming processes can nowadays rely on modern finite element simulations. Thereby they save expensive test series. Nevertheless it remains unfeasible to simulate the hydroforming process for all possible parameter settings. In order to find an optimal setting with respect to quality measures for the formed workpiece we can use simulations according to a suitable design of experiments. The stochastic nature of the real process should be taken into account here by including so-called noise parameters. These parameters can be set in the finite element simulation but are uncontrollable in the actual hydroforming process. The task of finding values of the controllable parameters for which the mean of the quality measures hits the desired target with minimal variance defines a typical robust parameter design problem [6]. The classical solution given by Taguchi [13] is based on experimental designs which are a product of a design for the controllable parameters and a design for the uncontrollable or noise parameters. The variance is estimated by means of repeating experiments with fixed design points for the controllable part. As an alternative we may employ generalized linear models for modelling the mean and the variance. This approach can be based on experimental designs without repeated settings of the controllable parameters. Hence, less experiments or simulations are needed. Generalized linear models can also be used to predict the mean and variance for all possible parameter settings, where the prediction is optimal within the experimental region. The model also provides estimates of the expected loss function. This loss function must be minimized in the Taguchi approach.

In this paper, we review the idea of modelling the mean and variance of a quality measure simultaneously. We will apply these methods to the hydroforming process. As the hydroforming process cannot be sensibly measured by only one quality characteristic, we

also investigate the use of the estimated models in a multivariate robust parameter design strategy.

2 Multi-Response Parameter Design using Loss Functions

We consider the situation of an industrial process, the outcome of which is given in terms of quality characteristics Y_1, \dots, Y_ℓ . We further assume a vector of finite target values $\tau = (\tau_1, \dots, \tau_\ell)'$ to be given. The random vector $Y = (Y_1, \dots, Y_\ell)'$ is assumed to depend on a vector $v = (v_1, \dots, v_k)'$ of design factors (covariates) by a function f , that is we have $Y = f(v, \epsilon)$, where ϵ describes the noise. The conditional distribution of Y given v has expectation $E(Y|v) = \mu(v)$ and covariance matrix $\text{Cov}(Y|v) = \Sigma(v)$. Hence, the response mean and the covariance matrix depend on the design factors.

The single response case is treated by the classical Taguchi approach [13]. There, the overall quality of the product is modelled in terms of a loss resulting from the deviation of a quality characteristic Y from its target τ . Taguchi takes this loss to be $\text{loss}(y) = c(y - \tau)^2$, the quadratic loss, where c is some constant.

As a straightforward extension to multiple quality characteristics the quadratic form

$$\text{loss}(y) = (y - \tau)'C(y - \tau)$$

has been proposed in [10]. Here C denotes a $\ell \times \ell$ symmetric cost matrix. For the expected loss, the so-called risk function, one can write

$$\begin{aligned} R(v) &= E(\text{loss}(Y|v)) = E(((Y|v) - \tau)'C((Y|v) - \tau)) \\ &= \text{trace}(C'\Sigma(v)) + (\mu(v) - \tau)'C(\mu(v) - \tau) \end{aligned} \quad (1)$$

Hence, minimizing the risk function will, roughly speaking, bring the mean on target and simultaneously minimize the variances. On the basis of designed experiments, different strategies have been proposed to find combinations of the design factor (parameter) values which minimize the risk function (see e.g. [10, pp. 8-13]). For the hydroforming process analysed in Section 4, we will follow the approach of [3] to fit generalized linear models for the mean and variance of each response. Thus, we can minimize the expected loss resulting from these models for a sequence of cost matrices.

3 Generalized Linear Models with Varying Dispersion

Based on an experimental design with n trials we assume that the corresponding responses Y_i , $i=1, \dots, n$, for a single quality characteristic are random variables with means $E(Y_i) = \mu_i$ and variances $\text{Var}(Y_i) = s_i^2$. An appropriate function $x = (x_1, \dots, x_p)'$ of

the original covariates $v = (v_1, \dots, v_k)'$ is used in the modelling process. The values of x for the individual experiments are denoted by $x_i = (x_{i1}, \dots, x_{ip})'$, $i = 1, \dots, n$.

Consider now classical linear regression models

$$Y_i = \sum_{j=1}^p x_{ij} \beta_j + \varepsilon_i, \quad i = 1, \dots, n,$$

where $\beta \in \mathbb{R}^p$ denotes an unknown parameter vector. The responses Y_i , $i = 1, \dots, n$, are supposed to be normally distributed with constant variances ($\sigma_i = \sigma$, $i = 1, \dots, n$). The systematic effects are supposed to be linear ($x_{i1} \beta_1 + \dots + x_{ip} \beta_p$). The errors ε_i , $i = 1, \dots, n$, are independent and normally distributed with zero mean, such that

$$E(Y_i) = \sum_{j=1}^p x_{ij} \beta_j \quad i = 1, \dots, n.$$

Sometimes a transformation of the response helps to fulfil these assumptions, at least approximately. However, often this cannot be achieved by a single transformation.

In this case generalized linear models [4,8] can be applied, which extend the assumptions in two ways. Firstly, for each Y_i , $i = 1, \dots, n$, the conditional distribution of Y_i given x_i is assumed to belong to an exponential family with mean μ_i and, in some cases, on a common scale parameter ϕ not depending on i . These distributional families includes for example families of gamma, inverse Gaussian, Poisson and binomial distributions.

Secondly, for each i , $i = 1, \dots, n$, the expectation μ_i is related to the linear predictor $\sum_{j=1}^p x_{ij} \beta_j$ by a so-called link function g , which is a known one-to-one, sufficiently smooth function,

$$g(E(Y_i)) = \sum_{j=1}^p x_{ij} \beta_j.$$

Further, the variance of the response can be written as

$$\text{Var}(Y_i) = \phi V(\mu_i), \quad i = 1, \dots, n,$$

where V is a known variance function. The usual linear regression model is part of this model class; then the responses are normally distributed, g is the identity function, $\phi = \sigma^2$ and $V(\mu) = 1$.

The aim in quality-improvement experiments is the selection of covariate values, which bring the mean on target while keeping the variability of the product at a minimum. We must allow the variances to be influenced by covariates as well. For this purpose the model can be extended to a joint model of mean and dispersion, a so-called double generalized

linear model [2,7,11,12]. In this sense we now assume that the mean model given above is extended to

$$g(E(Y_i)) = \sum_{j=1}^p x_{ij} \beta_j \text{ with } \text{Var}(Y_i) = \phi_i V(\mu_i), \quad i = 1, \dots, n. \quad (2)$$

As a second generalized linear model a dispersion model is formulated for a suitable statistic d_i as a measure of dispersion for each $i=1, \dots, n$. Here

$$\phi_i = E(d_i), \quad h(\phi_i) = \sum_{j=1}^m z_{ij} \gamma_j, \quad \text{Var}(d_i) = t V_D(\phi_i), \quad i = 1, \dots, n, \quad (3)$$

where $z = (z_1, \dots, z_m)'$ is a further function of the original covariates v , h a dispersion link function, $\gamma \in \mathbb{R}^m$ an unknown parameter vector, $V_D(\phi)$ the dispersion variance function, and $t \in \mathbb{R}$ a known scalar. In [8] we find two possible choices for the dispersion statistic d_i , namely the generalized Pearson contribution

$$r_P^2(i) = \frac{(Y_i - \mu_i)^2}{V(\mu_i)}$$

and the contribution to the deviance of unit i , respectively,

$$r_D^2(i) = 2 \int_{\mu_i}^{Y_i} \frac{Y_i - t}{V(t)} dt \quad i = 1, \dots, n.$$

To fit such a model, we must choose a suitable dispersion variance and suitable link functions. In the case of normally distributed Y the statistics $r_D^2(i)$ and $r_P^2(i)$ have a $\phi_i \chi_1^2$ distribution, such that a gamma model with $V_D(\phi) = 2\phi^2$ is appropriate. The most natural link functions are the identity, giving additive variance components, and the log function, giving multiplicative effects of the covariates.

The mean model (2) and the dispersion model (3) are related in the sense that the mean model requires $1/\phi_i$ to be used as a prior weight, while the dispersion model requires μ_i to be known in order to form the dispersion response variable d_i . The estimation of the unknown parameter vectors β and γ is usually achieved by the obvious algorithm, which alternates between fitting the model for the means for given weights $1/\hat{\phi}_i$ and fitting the model for the dispersion using the response variable $d_i(Y_i, \hat{\mu}_i)$. The individual estimates are based on extensions of the maximum likelihood principle. Hence they aim at maximising the joint probability function of the given sample [1,2,8,12].

4 Application to a hydroforming process

In the development stage of a sheet metal forming process, often both process parameters and the geometry of the workpiece itself can be varied within a given range [5]. To study the effect of geometrical features on the quality of the hydroforming result a die with a square pocket is considered, see Figure 1.

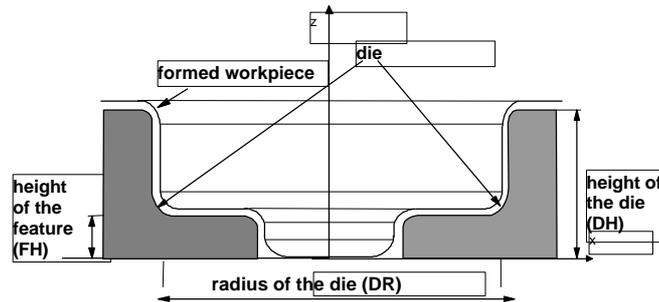


Figure 1: Die with a square pocket as workpiece

For every simulated high-pressure sheet metal hydroforming process the resulting virtual workpiece is described by two quality characteristics, whose values are to be optimised simultaneously. They describe the form and geometric accuracy (GA) as well as the maximal thinning of the blank sheet (MaT) caused by the forming operation. The target values are 0 and 0.15 respectively. The die diameter and height as well as the pocket dimensions and position can be varied, such that screening experiments with 10 covariates have been carried out. Five covariates have been identified as having large effects on the quality characteristic:

- PW: Pressure of the working media (600 Bar - 1000 Bar),
- FH: Height of the feature (1.67 mm - 7.33 mm),
- DH: Height of the die (20 mm - 40 mm),
- DR: Radius of the die (140 mm - 180 mm),
- TB: Original blank sheet thickness (0.9 mm - 1.1 mm).

Simulations following a so-called central-composite design have been carried out, where $-\sqrt{2}$ and $\sqrt{2}$ are the coded values corresponding to the settings given in blankets. The resulting data set is given in Tab.1.

Nr.	PW	FH	TB	DH	DR	GA	MaT
1	-1	-1	-1	-1	1	29.148	0.04
2	-1	-1	-1	1	-1	33.111	0.09
3	-1	-1	1	-1	-1	30.846	0.05
4	-1	-1	1	1	1	38.834	0.07
5	-1	1	-1	-1	-1	33.533	0.23
6	-1	1	-1	1	1	36.107	0.27
7	-1	1	1	-1	1	39.585	0.17
8	-1	1	1	1	-1	48.591	0.26
9	1	-1	-1	-1	-1	27.001	0.05
10	1	-1	-1	1	1	27.700	0.08
11	1	-1	1	-1	1	28.907	0.05
12	1	-1	1	1	-1	31.642	0.10
13	1	1	-1	-1	1	28.316	0.23
14	1	1	-1	1	-1	31.389	0.31
15	1	1	1	-1	-1	31.110	0.25
16	1	1	1	1	1	32.967	0.28
17	$-\sqrt{2}$	0	0	0	0	33.809	0.13
18	$\sqrt{2}$	0	0	0	0	26.229	0.14
19	0	$-\sqrt{2}$	0	0	0	25.297	0.07
20	0	$\sqrt{2}$	0	0	0	32.972	0.37
21	0	0	$-\sqrt{2}$	0	0	21.889	0.13
22	0	0	$\sqrt{2}$	0	0	32.126	0.15
23	0	0	0	$-\sqrt{2}$	0	45.878	0.07
24	0	0	0	$\sqrt{2}$	0	34.128	0.19
25	0	0	0	0	$-\sqrt{2}$	30.310	0.14
26	0	0	0	0	$\sqrt{2}$	24.602	0.10
27	0	0	0	0	0	29.467	0.13

Tab 1: Data resulting from a central-composite design

We look for a model that explains the data satisfactorily and fits with the mechanical background, but is also parsimonious at the same time. This design allows for the estimation of main and interaction effects, as well as quadratic effects.

For each of the two quality characteristics we separately consider double generalized linear models as described in Section 3. The thickness of the original blank sheet affords a noise factor which cannot be controlled in the actual hydroforming process. Therefore this factor is not included actively in the modelling process. We use the mean model with link function equal to the identical and with normal distribution. Further, we use the gamma distribution and the log-link for the dispersion model. Unknown model parameters are estimated by the restricted maximum likelihood approach described in [12]. Starting with the model including all possible effects we successively remove the least significant covariates found by t-tests. Concerning the maximal thinning we find the following reduced mean and dispersion model equations

$$\begin{aligned}\hat{E}(\text{MaT}) &= 0.135 + 0.003\text{PW} + 0.101\text{FH} + 0.025\text{DH} - 0.011\text{DR} + 0.039\text{FH}^2 - 0.006\text{DR}^2 \\ &\quad + 0.0032\text{PW} \cdot \text{DH} + 0.0054\text{FH} \cdot \text{DH} - 0.004\text{FH} \cdot \text{DR} - 0.0053\text{DH} \cdot \text{DR} \\ \hat{\text{V}}\text{ar}(\text{MaT}) &= \exp(-10.481 - 1.26\text{PW} + 2.362\text{FH} - 0.592\text{DH} - 0.672\text{DR} - 2.451\text{PW}^2 \\ &\quad + 1.553\text{DH}^2 - 1.813\text{DH} \cdot \text{DR}).\end{aligned}$$

The diagnostic plot in Figure 2 shows that the selected model is adequate. For both, the complete model as well as the selected model the observed values are almost equal to the predicted values from the mean model.

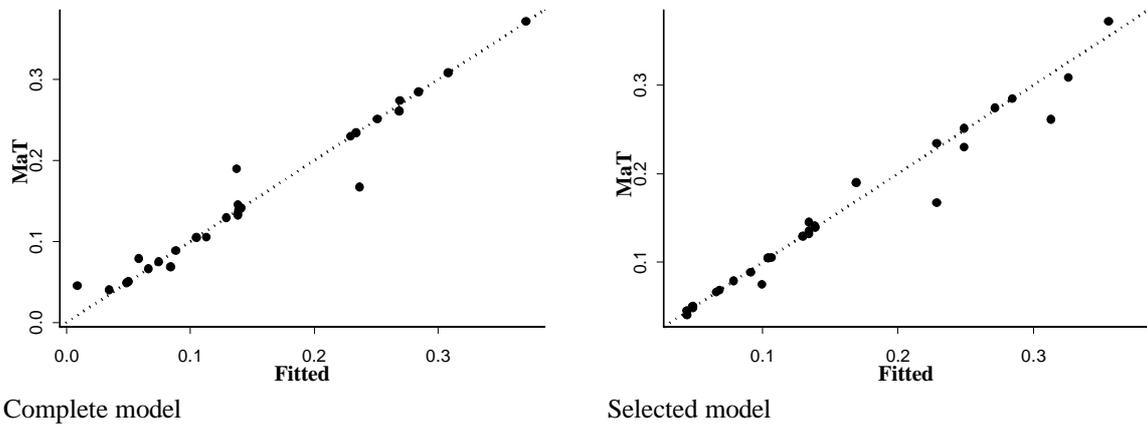


Figure 2: Diagnostic Plots

Following the above strategy we again derive a model for the mean and variance of the geometric accuracy:

$$\begin{aligned}\hat{E}(\text{GA}) &= 30.126 - 2.756\text{PW} + 0.964\text{FH} + 0.303\text{DH} + 0.532\text{DR} + 2.628\text{DH}^2 \\ &\quad + 1.992\text{PW} \cdot \text{DH} - 0.819\text{PW} \cdot \text{DR} - 1.091\text{FH} \cdot \text{DH} + 1.065\text{DH} \cdot \text{DR}\end{aligned}$$

$$\hat{\text{Var}}(\text{GA}) = \exp(2.805 - 2.859 \text{PW} - 0.611 \text{FH} - 0.744 \text{DH} + 1.163 \text{DR} - 3.901 \text{PW}^2 + 3.285 \text{PW} \cdot \text{DR} - 2.503 \text{FH} \cdot \text{DR}).$$

The estimated prediction models for the mean and variance of the quality characteristics are plugged into the expected risk function (1). The considered risk function theoretically becomes equal to zero if the process is always on target, this is the means of all quality characteristics are equal to the target values and the variances are zero. The pre-specified weighting C of the quality characteristics describes the relevance of reaching an optimal result for each characteristic. For different possible weightings values of the controllable design factors can be determined which minimize the estimated risk function. Accomplishing this for a sequence of weight matrices C as described in [3], we can summarize the result in the joint optimisation plot given in Figure 3.

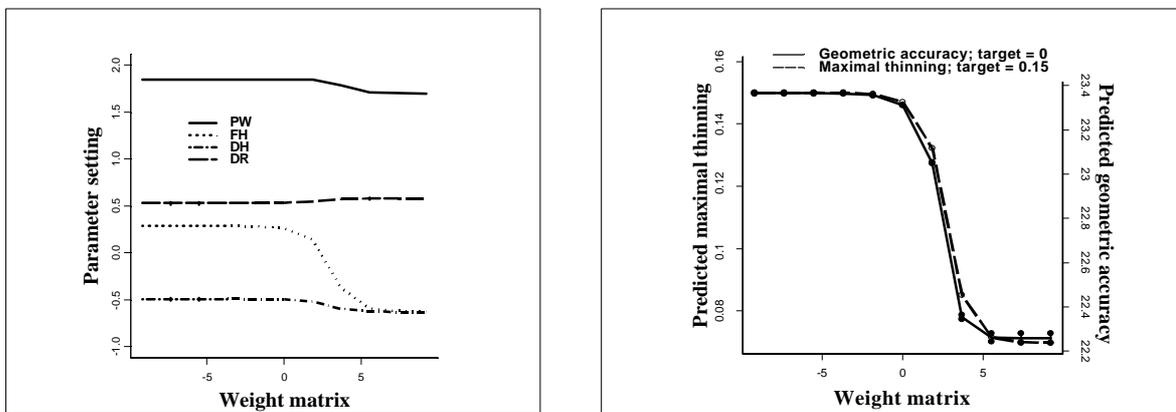


Figure 3: Joint Optimisation Plot

On the left hand side of **Figure 3**, the derived design parameter settings are plotted against the used weights. More emphasis is put on reaching the target for maximal thinning to the left and on the geometric accuracy to the right. The plot on the right hand side shows the corresponding predicted means and variances of the quality characteristic. The effects of the parameter factors on the quality characteristics are contradictory. A better geometrical accuracy evokes a thinning of the blank sheet, which is then further away from the target. But what is more important: lower values of the feature height lead to better geometric accuracy but stretch the blank sheet below the desired target value. Therefore the optimum settings have to be based on a compromise between geometrical accuracy and thinning of the blank sheet.

5 Outlook

We have presented the suitability of double generalized linear models in parameter design strategies for multiple responses. The suggested methods can be widely applied to many engineering and manufacturing processes, not only to hydroforming processes. Another advantage of the generalized linear model is its applicability in cases of continuous

responses as well as discrete responses. Future research should include the development of joint generalized linear models for the mean vector and covariance matrix of multivariate responses.

Acknowledgements

Financial support for the research project “Process and Tool Optimising Methods by Statistical Design of Experiments and FEA for Sheet Metal Forming Operations” by the central public funding organisation for academic research in Germany (DFG) is gratefully acknowledged. We also thank G. Smyth (<http://www.statsci.org/s/index.html>) for providing the S-plus function `dglm()`.

References

- [1] Engel, J. and Huele, A.F.: A Generalized Linear Modeling Approach to Robust Design. *Technometrics* 38, S. 365-373, 1996.
- [2] Engel, J. and Huele, A.F.: Taguchi Parameter Design by Second-Order Response Surfaces. *Quality and Reliability Engineering International* 12(2), S. 95-100, 1996.
- [3] Erdbrügge, M. and Kuhnt, S.: Strategies for Multi-Response Parameter Design Using Loss Functions and Joint Optimisation Plots. Technical Report, SFB 475, University of Dortmund, Germany, 2002.
- [4] Fahrmeir, L.; Tutz, G.: *Multivariate Statistical Modelling Based on Generalized Linear Models*. 2nd Edition, Springer, New York, 2001.
- [5] Gather, U.; Homberg, W.; Kleiner, M.; Klimmek, Ch. and Kuhnt, S.: Parameter Design for Sheet Metal Hydroforming Processes. In: *Proceedings of the 7th International Conference on the Theory of Plasticity (ICTP)*, Yokohama, Japan, October 17 – November 4, 2002.
- [6] Grize, Y.L.: Review: A Review of Robust Process Design Approaches. *Journal of Chemometrics*, 9, 239-262, 1995.
- [7] Lee, Y. and Nelder, J.A.: Generalized Linear Models for the Analysis of Quality-Improvement Experiments. *Canadian Journal of Statistics* 26(1), S. 95-105, 1998.
- [8] McCullagh, P. and Nelder, J.A.: *Generalized Linear Models*, 2nd Edition, Chapman & Hall, London, 1989.
- [9] Nelder, J.A. and Lee, Y.: Generalized Linear Models for the Analysis of Taguchi-Type Experiments. *Applied Stochastic Models and Data Analysis* 7, S. 107-120, 1991.
- [10] Pignatiello, J.J.: Strategies for Robust Multiresponse Quality Engineering. *IIE Transactions*, 25, S. 5-15, 1993.
- [11] Smyth, G.K.: Generalized Linear Models with Varying Dispersion. *Journal of the Royal Statistical Society B* 51(1), S. 47-60, 1989.
- [12] Smyth, G.K. and Verbyla, A.P.: Adjusted Likelihood Methods for Modelling Dispersion in Generalized Linear Models. *Environmetrics* 10, S. 695-709, 1999.
- [13] Taguchi, G. and Phadke, M.S.: *Quality Engineering Through Design Optimization*. Proceedings of GLOBECOM 84 Meeting, Piscataway. NJ: IEEE Communication Society, 1106-1113, 1984.