Int J Adv Manuf Technol (2018) 94:2255–2265
DOI 10.1007/s00170-017-0980-z

ORIGINAL ARTICLE

Principal component analysis and singular value decomposition used for a numerical sensitivity analysis of a complex drawn part

Christian Schwarz[1] · **Patrick Ackert**[1] · **Reinhard Mauermann**[1]

Received: 24 May 2017 / Accepted: 14 August 2017 / Published online: 6 September 2017

Abstract The numerical forecasting of car body construction processes is already being used in industry to provide support in the ramp-up process. However, long calculation times are stretching the finite element method (FEM) to the limit, in particular when analyzing the effect of the variation of an input variable on one or more dependent variables. Moreover, there is still a need for experienced users to separate relevant from irrelevant parameters and to determine their variation. This paper presents a method that makes it possible, based on stochastic experimental design (DOE) in combination with both principle component analysis (PCA) and singular value decomposition (SVD), to create mathematical models that separate relevant from irrelevant input variables and that represent the effect of individual variables on all part or assembly areas by means of a variance-based sensitivity analysis. The method is verified in a case study based on realistic front hood geometry. The study examines the deep-drawing process steps as well as the geometrical accuracy in a measuring device. It is shown that it is possible to represent the effects of the most important variables from these processes on the strain and geometry parameters of the car body part and to vary these, based on a model function, interactively.

✉ Christian Schwarz
christian.schwarz@iwu.fraunhofer.de

Patrick Ackert
patrick.ackert@iwu.fraunhofer.de

Reinhard Mauermann
reinhard.mauermann@iwu.fraunhofer.de

[1] Fraunhofer Institute for Machine Tools and Forming Technology IWU, Nöthnizer Strasse 44, 01187 Dresden, Germany

Keywords Finite element method (FEM) · Principal component analysis (PCA) · Singular value decomposition (SVD) · Car body process · Sensitivity analysis

1 Introduction

A middle-class car body is made of up to 500 sheet metal components which are aligned and joined in up to 120 fixtures [1] with the objective of realizing a tolerance of between ± 0.35 and 1 mm in geometrical accuracy [2]. Many iterative adjustments of the deep-drawing process parameters and tool surfaces are needed to achieve this. Nearly 30% of the [2] overall tooling costs are caused during these optimizations [3].

To reduce iteration and the associated costs, the finite element method (FEM) is being used increasingly in industry to forecast failure and springback for deep-drawing processes, even if calculation speed and accuracy are not always satisfactory at present [4]. Accordingly, the producibility of components via deep drawing is already the subject of numerical analysis, yet there are currently almost no conclusions from the automotive industry with regard to springback [5]. Aside from the long calculation times for springback steps, the main factor responsible for this is result inaccuracies caused by excessive model simplifications (for example, simplification of the forces acting on the metal sheet) [6]. Thus, it has been shown that when the same paths are used in experiment and simulation, the represented forces vary widely, but the quality of results improves by up to 80% [7]. Even though it is not possible to draw conclusions about the absolute dimensional stability of a component on this basis, it is possible to identify areas of influence of parameter variations. The quality of results achieved in the clamping process step is significantly higher due to lower model complexity. Lee [8] demonstrated that the frictional contact behavior between clamping

elements and component can be adequately represented using FEM. Based on these findings, shell-based FEM was used to develop methods that make it possible to represent dimensional stability influences in the clamp-and-join car body construction process with sufficient accuracy [9].

Siekirk [10] identified 30 process variables for the deep-drawing process and showed experimentally that blank size, blank position, material thickness, and binder force are the most significant factors for the drawn part strain. Similarly, Zhang [11] investigated the influence of process forces and shut height on the waviness of the car body shell part. Majeske and Hammet [12] presents a variance analysis-based system, which makes it possible to classify variations that occur during deep drawing and thus to initiate measures to compensate for dimensional variations.

Naceur [13] and Wei [14] used the response surface method to conduct initial parameter studies based on simple sample geometries to minimize springback behavior during deep drawing. Naceur focuses on the optimization of active tool surfaces, whereas Wei optimizes process parameters. Meinhardt [15] presents an option for interactively optimizing tool geometry parameters based on a side panel segment. The focus is on the forecast of producibility of the deep-drawn part for various parameter combinations. Input and output variables are linked via a metamodel generated on the basis of FE = finit element variant calculations. Emrich [16] investigated the robustness of the deep-drawing process with a focus on the influence of material characteristics. Both used the strain of the nodes in relation to the FLC = forming limit curve as the output.

Wärmefjord [17] used FEM-based variant simulations to forecast the effect of process and control variables in deep drawing (e.g., friction, forming speed) on the component geometry as well as the eventual effect on the assembly process. The component variation is evaluated on the basis of PCA. The PCA results are then used as source data for the functional description of the relationships between component deformation and input parameters via a metamodel. The metamodel makes it possible to generate Monte Carlo simulations to determine the 6-sigma levels, as well as to identify parameter influences by correlation analysis.

Wolff [18] and Gerbino [19] presented, independently of one another, an investigation of the effect of deep-drawing process variations on the drawn part, based on a PCA. Wolff based his investigations on the numerical variation of material characteristics, friction, sheet thickness, position, and geometry variation. Gerbino determined his data experimentally. Both were able to demonstrate that PCA is capable of describing variations that occur using a small number of orthogonal error modes. The geometric error modes are then used as a starting point for determining process interrelationships and sensitivities.

Youcef-Toumi [20] carried out the first optimization for the low-stress fixation of parts. This was followed by various methods of need-based optimization of clamping points, such as those presented by Menassa and DeVries and Franciosa [21]. The metamodel-based optimization of the dimensional stability of an assembly at discrete points by means of clamping point adjustment was first published by Schwarz [22].

Previous literature has demonstrated that the ramp-up of a car body production line can be particularly time and cost intensive, but FEM can be used to reduce this effort and to generate good results. The main problem is that, even though FEM analyses are much faster than experimental tests, they still require a large amount of time. The approach to solve this problem is to substitute single variance simulation by representing the complete parameter space with DOE-based models.

The use of numerical sensitivity analysis to analyze the relationship of inputs and outputs in car body production is widely known. For example, it is used in a user-friendly format in the AutoForm-Sigma software. The use of an SVD-based PCA in combination with a variance-based sensitivity analysis has not previously been documented.

This paper presents the mathematical formulation and verification of a method that makes use of this combination. The objective is to determine the influence of variable input parameters with respect to their effects on any number of post variables from deep-drawing simulations for the entire component area. So, it is possible to differentiate relevant and irrelevant parameters. In this case, relevant parameters are defined as having more than 2% influence on the local or global variation value. Parameters with less than 2% influence are defined as irrelevant, as they are highly likely to be some form of noise from the FEM model or the solver.

2 Multivariate analysis of variance

2.1 Principal component analysis

Principal component analysis (PCA) is a multivariate statistical method for identifying trends and patterns in large amounts of data [23]. It was formulated by Pearson 1901 [24] and Hotelling 1933 [25]. The basic idea of the PCA consists in a linear transformation of the high-dimensional data into a new coordinate system with lower dimension. The variance-covariance structure of the source data can be described with a small number of "new" variables, the so-called principal components, in the result. The mutually orthogonal principal components are uncorrelated and are sorted in such a manner that the first principal components can represent the largest proportion of the total variation of the original variables [26].

Let $\boldsymbol{X}$ be a data matrix of dimension $n \times d$ and zero mean, where d is the number of variables and n is the number of

observations. The objective of the PCA is the orthogonal linear transformation of the high-dimensional $\boldsymbol{X}$ into a low-dimensional score matrix $\boldsymbol{Y}$

$$\boldsymbol{Y} = \boldsymbol{X}\,\boldsymbol{V} \tag{1}$$

where $\boldsymbol{Y}$ is an $n \times n$ matrix and $\boldsymbol{V}$ is a $d \times n$ matrix. To determine the score matrix $\boldsymbol{Y}$, first the principal axes or even principal components of the new coordinate system, which are compiled in matrix $\boldsymbol{V}$ as column vectors, must be found. The principal components are arranged by variance proportion. The first principal component in the first column of $\boldsymbol{V}$ explains the largest proportion of the total variation, the second principal component in the second column of $\boldsymbol{V}$ has the second largest variance, which is orthogonal to the first principal component, and so on. The principal components are equal to the eigenvectors of the matrix of all paired covariances of the variables of data matrix $\boldsymbol{X}$. To analyze the variance-covariance structure of data matrix $\boldsymbol{X}$, the covariance matrix $\boldsymbol{C}$ is established first

$$\boldsymbol{C} = \frac{1}{1-n}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}. \tag{2}$$

In this case, $\boldsymbol{C}$ is a symmetric matrix of dimension $d \times d$. Note that the dimension of the covariance matrix $\boldsymbol{C}$ has an exponential dependence on the number of variables of the data matrix $\boldsymbol{X}$. Therefore, for a large number of variables it can be difficult in terms of computation time to calculate the d^2 elements of $\boldsymbol{C}$ explicitly. One option for calculating the eigenvectors of covariance matrix $\boldsymbol{C}$ without forming the matrix product $\boldsymbol{X}^T\boldsymbol{X}$ from (2) is the method known as *singular value decomposition (SVD)*. By this method, the data matrix $\boldsymbol{X}$ can be represented as the product of three matrices [27]:

$$\boldsymbol{X} = \boldsymbol{U}\,\boldsymbol{\Gamma}\,\boldsymbol{V}^{\boldsymbol{T}} \tag{3}$$

In this case, $\boldsymbol{U}$ is the matrix of the left singular vectors with the dimension $n \times n$, $\boldsymbol{V}$ is the matrix of the right singular vectors with the dimension $d \times d$, and $\boldsymbol{\Gamma}$ is the $n \times d$ -diagonal matrix of the singular values. Plugging (3) into (2) yields

$$\boldsymbol{C} = \frac{1}{1-n}\,\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} = \frac{1}{1-n}\left(\boldsymbol{V}\,\boldsymbol{\Gamma}^{\boldsymbol{T}}\boldsymbol{U}^{\boldsymbol{T}}\right)\left(\boldsymbol{U}\,\boldsymbol{\Gamma}\,\boldsymbol{V}^{\boldsymbol{T}}\right). \tag{4}$$

The orthogonality of $\boldsymbol{U}$ ($\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{I}$) and the symmetry of $\boldsymbol{\Gamma}$ ($\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T$) finally result in the following spectral decomposition of covariance matrix $\boldsymbol{C}$ [28]:

$$\boldsymbol{C} = \frac{1}{1-n}\,\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X} = \left(\boldsymbol{V}\,\boldsymbol{\Gamma}^{\boldsymbol{T}}\boldsymbol{U}^{\boldsymbol{T}}\right)\left(\boldsymbol{U}\,\boldsymbol{\Gamma}\,\boldsymbol{V}^{\boldsymbol{T}}\right) = \boldsymbol{V}\,\boldsymbol{\Gamma}\,\boldsymbol{\Gamma}^{\boldsymbol{T}}\boldsymbol{V}^{\boldsymbol{T}} \tag{5}$$

$$\boldsymbol{C} = \boldsymbol{V}\,\boldsymbol{\Gamma}^2\,\boldsymbol{V}^{\boldsymbol{T}} \tag{6}$$

Formula (6) shows the spectral decomposition of the covariance matrix. Accordingly, the vector of the back singular values $\boldsymbol{V}$ corresponds exactly to the eigenvectors of the covariance matrix $\boldsymbol{C}$. Furthermore, square root $\sqrt{\sigma_i} = \lambda_i$ of the singular values of $\boldsymbol{\Gamma}$ gives the corresponding eigenvalues of $\boldsymbol{C}$. A useful requirement for applying the SVD is the symmetric property of covariance matrix $\boldsymbol{C}$ from (2), which means that $\boldsymbol{C}$ is positive semidefinite and only n eigenvalues of $\boldsymbol{C}$ are greater than or equal to zero. Since $\sqrt{\sigma_i} = \lambda_i$, the dimension of $\boldsymbol{\Gamma}$ can be reduced from $n \times d$ to $n \times n$ and the dimension of $\boldsymbol{V}$ from $d \times d$ to $d \times n$. The size of the n eigenvalues from $\boldsymbol{C}$ exactly matches the proportion of the total variance of the data matrix $\boldsymbol{X}$, which can be described via the associated principal axis direction in $\boldsymbol{V}$. Given that, based on Formula (7), the highest-percentage variance occurs with the first p principal components, a further reduction of the data can be achieved by specifically ignoring the low-variance principal components. The proportion of the ith principal component in relation to the total variance is given by:

$$Var\,(\lambda_i) = \frac{\lambda_i}{\sum_{j=1}^{n}\lambda_j}\cdot 100\% \tag{7}$$

After determining the new transformation basis, the original data must be projected to the principal components in the final step. Using Formulas (1) and (3), this data projection is defined as [29]

$$\boldsymbol{Y} = \boldsymbol{X}\,\boldsymbol{V} = \boldsymbol{U}\,\boldsymbol{\Gamma}\,\boldsymbol{V}^{\boldsymbol{T}}\,\boldsymbol{V} = \boldsymbol{U}\,\boldsymbol{\Gamma} \tag{8}$$

An important characteristic of the PCA is that the original data can be fully recovered from the projection data in the following way:

$$\boldsymbol{X} = \boldsymbol{Y}\,\boldsymbol{V}^{\boldsymbol{T}} \tag{9}$$

The eigenvectors $\boldsymbol{V}$, arranged according to their explicable variance, are called principal components or even geometric eigenmodes.

2.2 Variance-based sensitivity analysis

According to Saltelli [30], the objective of a global sensitivity analysis is to be able to quantify the influence of varying input factors on the variance of the associated output independently of the model. In this case, as opposed to a local sensitivity analysis, the influence of the input parameters on the variation of the source variables is not determined for an individual point but rather across the entire range of variation of the input factors [30]. For a robust estimate of global sensitivity, it is particularly advisable to use variance-based sensitivity analyses that can determine the global sensitivity independently of the linearity or additivity of a model [31]. In summary, variance-based methods have the following advantages [30]:

- Variance-based methods are independent of the model
- Determination of influence is possible across the entire range of input factors

- It is possible to identify interactions among the input factors
- Determination of non-influent input factors

Variance-based sensitivity methods have a long history of development, from the analysis of variance using multiple Fourier series [32, 33], through Sobol's [34] variance decomposition formulas, to the quantification of variance using total sensitivity indices discussed in [35]. According to Sobol [34], the output variance $V(Y)$ of a model output $Y=f(X_1, X_2, \cdots, X_k)$ with k -uncertain input factors is decomposed as:

$$V(Y)=\sum_i V_i+\sum_i\sum_{j>i} V_{ij}+\cdots+V_{12\ldots k} \tag{10}$$

Where

$$V_i=V_{X_i}[E_{\boldsymbol{X}_{\sim i}}(Y|X_i)]$$

$$V_{ij}=V_{X_{ij}}\left[E_{\boldsymbol{X}_{\sim ij}}\left(Y\,\middle|\,X_i,X_j\right)\right]-V_i-V_j$$

and so on. The terms V_i and V_{ij} in function (10) can be interpreted as follows [36]:

- V_i is the reduction of variance to be expected if the ith input factor X_i of the model could be fixed. The mean $E_{\boldsymbol{X}_{\sim i}}\left(Y\,\middle|\,X_i\right)$ is taken over all possible values of $\boldsymbol{X}_{\sim i}$ while X_i keep fixed. The outer variance V_{X_i} is taken over all possible values of X_i. In this context $\boldsymbol{X}_{\sim i}$ denotes the matrix of all factors but X_i.
- V_{ij} is the reduction of variance to be expected if both the i^{th} input factor X_i and the jth input factor X_j of the model could be fixed simultaneously. The mean $E_{\boldsymbol{X}_{\sim ij}}\left(Y\,\middle|\,X_i,X_j\right)$ is taken over all possible Values of $\boldsymbol{X}_{\sim ij}$ while X_i and X_j keep fixed. The outer variance $V_{X_{ij}}$ is taken over all possible values of X_i and X_j. In this context, $\boldsymbol{X}_{\sim ij}$ denotes the matrix of all factors but X_i and X_j.

Dividing both sides of Eq. (10) by the total variance $V(Y)$ leads to variance-based sensitivity indices [34]:

$$\sum_i S_i+\sum_i\sum_{j>i} S_{ij}+\cdots+S_{12\ldots k}=1 \tag{11}$$

where S_i is the well-known first-order sensitivity index. The first-order sensitivity index measures the main effect contribution of each input factor to the output variance. S_{ij} is the second-order index, which represents the interaction effect of the factor pair (X_i, X_j) on the output variance $V(Y)$. An analogous procedure is done for estimating higher-order indices. The development in (11) provides an estimation of the sensitivity of all the input factors and their interactions, whereby for k -input factors a total number of 2^k-1 indices have to be evaluated. However, when the number of input factors is large, the estimation of all first-order and interaction effects from (11) can be very intensive in terms of computing time. To reduce the computational afford the total sensitivity index S_{T_i} described in [35, 37] has proven to be an effective measurement tool for the determination of global sensitivities:

$$S_{T_i}=\frac{E_{\boldsymbol{X}_{\sim i}}\left[V_{X_i}\left(Y\,\middle|\,\boldsymbol{X}_{\sim i}\right)\right]}{V(Y)}=1-\frac{V_{\boldsymbol{X}_{\sim i}}\left[E_{\boldsymbol{X}_i}\left(Y\,\middle|\,\boldsymbol{X}_{\sim i}\right)\right]}{V(Y)} \tag{12}$$

Here, $E_{\boldsymbol{X}_{\sim i}}\left[V_{X_i}\left(Y\,\middle|\,\boldsymbol{X}_{\sim i}\right)\right]$ is the variance to be expected if all input parameters except X_i could be fixed. $V_{\boldsymbol{X}_{\sim i}}\left[E_{X_i}\left(Y\,\middle|\,\boldsymbol{X}_{\sim i}\right)\right]$ is the reduction of variance to be expected if all input parameters except X_i could be fixed [36]. Like Formula (11), the total sensitivity index S_{T_i} provides the total contribution of the input factor X_i to the output variance $V(Y)$, including its interactions with all other input parameters [38]. In contrast to 2^k-1 indices from (11), only the evaluation of k indices is necessary. A condition of S_{T_i} is:

$$\sum_{i=1}^{k} S_{Ti}\geq 1 \tag{13}$$

The sum of all terms of the total sensitivity index S_{T_i} from condition (13) is either equal to 1 when the model is purely additive or greater than 1 when there are interaction effects among the input factors. Furthermore, a total sensitivity index $S_{T_i}>0$ indicates that the input variable X_i has a quantitative influence on the output variance. In contrast, a total sensitivity index $S_{T_i}=0$ is a necessary condition for X_i having no portion of the output variance across its entire range of uncertainty, which makes it possible to contemplate a reduction of the non-influent input factors from the model [31]. The calculation of the total sensitivity index based on a Monte Carlo integration procedure first proposed by Homma and Saltelli [35] and extensively discussed in [31, 36]:

$$\begin{aligned} S_{T_i}&=1-\frac{V_{\boldsymbol{X}_{\sim i}}\left[E_{\boldsymbol{X}_i}\left(Y\,\middle|\,\boldsymbol{X}_{\sim i}\right)\right]}{V(Y)}=1-\frac{y_B\cdot y_{C_i}-f_0^2}{y_A\cdot y_A-f_0^2}\\ &=1-\frac{\frac{1}{N_S}\sum_{j=1}^{N_S} y_B^{(j)}y_{C_i}^{(j)}-f_0^2}{\frac{1}{N_S}\sum_{j=1}^{N_S}\left(y_A^{(j)}\right)^2-f_0^2} \end{aligned} \tag{14}$$

Where

$$f_0=\frac{1}{N_S}\sum_{j=1}^{N_S} y_A^{(j)} \tag{15}$$

and $y_A=f(\boldsymbol{A}), y_B=f(\boldsymbol{B}), y_{C_i}=f(\boldsymbol{C}_i)$ are three vectors of model outputs generated from the associated matrices $\boldsymbol{A}, \boldsymbol{B}$, and

$\boldsymbol{C}_i$. $\boldsymbol{A}$ and $\boldsymbol{B}$ are two matrices with random samples of the k input factors formed by a low-discrepancy sequence such as Latin hypercube sampling. $\boldsymbol{C}_i$ is a matrix that contains all columns from $\boldsymbol{B}$ except the column of the ith input factor, which is from $\boldsymbol{A}$. N_S is the number of random samples.

3 Methodology

The basis for the method is data reduction by means of SVD-based PCA technique. The use of PCA allows the element-rich FE meshes of the deep-drawing simulation to be replaced by a few principal components. Based on the identified principal components, a variance-based sensitivity analysis is then performed in the reduced characteristic space. The back-transformation of the reduced data finally enables the element-specific visualization of the sensitivities on the entire component surface. Figure 1 shows the flow of the method.

3.1 Stamping simulation

The first step is to define the input parameters (e.g., blank-holder force, draw beads, initial blank position, etc.) that have to be varied during the various deep-drawing simulations. A sample matrix $\boldsymbol{S}_{Input}$ with n simulation variants is created on the basis of the selected i input parameters.

$$\boldsymbol{S}_{Input} = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1i} \\ s_{21} & s_{22} & \cdots & s_{2i} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n1} & s_{n2} & \cdots & s_{ni} \end{bmatrix} \tag{16}$$

The random numbers for the sample matrix can be generated, for instance, by Latin hypercube sampling. Next, the n entries of the sample matrix are calculated. After every simulation run, the distances of the simulation meshes in normal direction as well as process-relevant post variables such as thinning, principal strains, and principal stresses are evaluated and compiled in a field-data matrix. The field-data matrix $\boldsymbol{X}_u$ is defined as follows for the uth post variable

$$\boldsymbol{X}_u = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1d} \\ x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \tag{17}$$

where n is the number of simulation variants and d is the number of mesh elements.

3.2 Dimension reduction by PCA

Due to the high dimensionality of field-data matrix $\boldsymbol{X}_u$, principal axis transformation is performed by SVD-based PCA. The projection of the field-data matrix $\boldsymbol{X}_u$ onto the related principal components $\boldsymbol{V}$ is given by Formula (8):

$$\boldsymbol{Y}_{Score} = \boldsymbol{X}_u \, \boldsymbol{V} = \boldsymbol{U} \, \boldsymbol{\Gamma} \tag{18}$$

The result is a low-dimensional projection-data matrix $\boldsymbol{Y}_{Score}$ of dimension $n \times n$. Since the first m principal components according to Formula (7) explain most of the variance, it is advantageous in terms of calculation time to limit the number of principal components. A constraint for the reduction of the principal components is that at least 99% of the source variance of $\boldsymbol{X}_u$ should be maintained. Therefore, $\boldsymbol{Y}_{Score}$ can be written as

$$\boldsymbol{Y}_{Score} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{bmatrix} \tag{19}$$

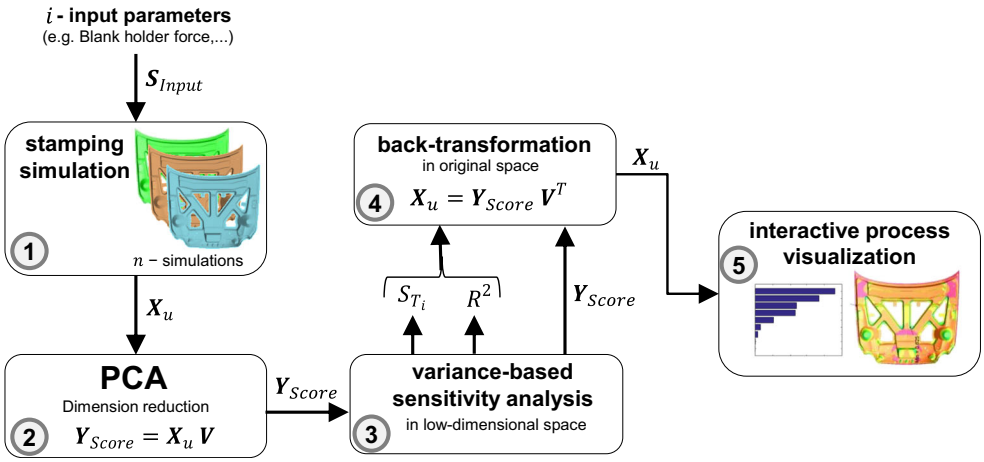

Fig. 1 Method description

where m is the number of principal components and n is the number of samples.

3.3 Variance-based sensitivity analysis based on PCA data

Next, the sensitivity analysis is performed on the basis of the reduced PCA data set $\boldsymbol{Y}_{Score}$. To determine how a change of the input parameters affects the dependent variables under consideration, a regression analysis is used to establish a functional connection between the sample matrix $\boldsymbol{S}_{Input}$ and the projection-data matrix $\boldsymbol{Y}_{Score}$. The regression approach gives the regression formula for the projection vector of the jth principal component in matrix notation:

$$\hat{\boldsymbol{Y}}_{Score_j} = \hat{\boldsymbol{S}}_{Input} \cdot \beta \tag{20}$$

where

$$\hat{\boldsymbol{Y}}_{Score_j} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \hat{\boldsymbol{S}}_{Input} = \begin{bmatrix} 1 & s_{11} & s_{12} & \cdots & s_{1i} \\ 1 & s_{21} & s_{22} & \cdots & s_{2i} \\ 1 & \vdots & \vdots & \ddots & \vdots \\ 1 & s_{n1} & s_{n2} & \cdots & s_{ni} \end{bmatrix},$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_i \end{bmatrix}.$$

$\hat{\boldsymbol{S}}_{Input}$ pt?>is the regression basis formed from the sample matrix $\hat{\boldsymbol{S}}_{Input}$, β is the vector with the estimated regression coefficients, and $\hat{Y}_{Score_j}$ is the jth approximated projection vector in the jth row of Y_{Score}. The regressors β of the model are determined by least-squares minimization [39]. It must be noted that not everyregression basis can equally explain the interactions between the projection vector $\hat{Y}_{Score_j}$ and the sample matrix $\hat{\boldsymbol{S}}_{Input}$. For this reason, depending on the complexity, it may be necessary to use linear polynomial approaches as shown in Formula (20) as well as non-linear polynomial approaches for the approximation. For example, the regression basis $\hat{\boldsymbol{S}}_{Input}$ of a second-order polynomial regression model is defined as follows:

$$\hat{\boldsymbol{S}}_{Input} = \begin{bmatrix} 1 & s_{11} & s_{12} & \cdots & s_{11}^2 & s_{12}^2 & \cdots & s_{11}s_{12} & \cdots \\ 1 & s_{21} & s_{22} & \cdots & s_{11}^2 & s_{12}^2 & \cdots & s_{21}s_{22} & \cdots \\ 1 & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & s_{n1} & s_{n2} & \cdots & s_{n1}^2 & s_{n2}^2 & \cdots & s_{n1}s_{n2} & \cdots \end{bmatrix} \tag{21}$$

The model selected for the paper is the one that can best explain the relationship between projection vector $\hat{Y}_{Score_j}$ and the input parameters, meaning it has the best forecasting accuracy. To estimate the prognostic quality of each model the multiple coefficient of determination (COD) R^2 is used [39]:

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{22}$$

Here, $\bar{y}$ denotes the mean, y_i is the ith score of the jth projection vector $\hat{Y}_{Score_j}$ and y_i is the forecast of the regression model for the ith score of the jth projection vector $\hat{Y}_{Score_j}$. If the underlying regression model can explain a majority of the variance of the model output, the result is an R^2 value close to 1. However, if the R^2 values are small, the output variation cannot be explained via the examined regression model and the R^2 value tends toward 0 [30]. In addition, as opposed to the COD, cross-validation methods can be used to further check the model quality of the regression models. It has to be taken into account that for each projection vector $\hat{Y}_{Score_j}$ a total number of m regression models must be build.

After the regression analysis, the variance-based sensitivity analysis has to be performed. The projection vector $\hat{Y}_{Score_j}$ has approximately the same variance structure as the field matrix $\boldsymbol{X}_u$, so the total contribution of the ith input parameter to the output variance of the jth approximated projection

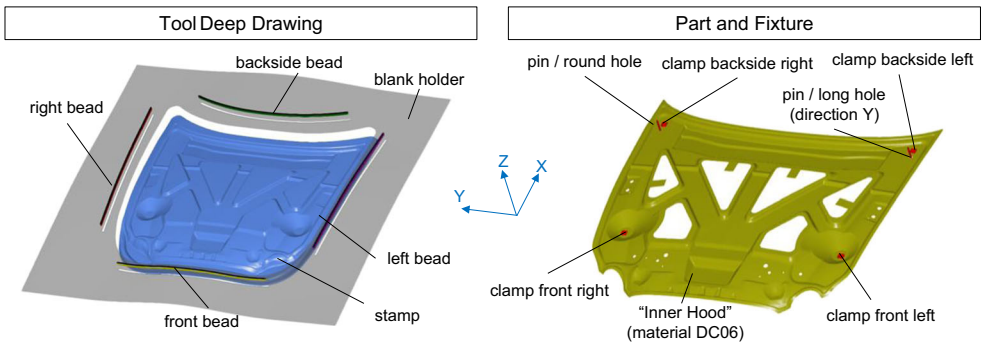

Fig. 2 Experimental setup

Table 1 Parameter variation

Parameter	Variation range	Parameter	Variation range
Thickness	0.95 to1.05 mm	Friction (blank to all) surfaces)	0.1 to 0.2 (Coulomb)
Blank holder/die distance	2 to 3.5 mm (*Z* direction)	Clamp level, front left	− 0.5 to 0.5 mm (*Z* direction)
Right bead level	− 0.5 to 2.5 mm (*Z* direction)	Clamp level, rear left	− 0.5 to 0.5 mm (*Z* direction)
Left bead level	− 0.5 to 2.5 mm (*Z* direction)	Clamp level, front right	− 0.5 to 0.5 mm (*Z* direction)
Front bead level	− 0.5 to 2.5 mm (*Z* direction)	Clamp level, rear right	− 0.5 to 0.5 mm (*Z* direction)
Rear bead level	− 0.5 to 2.5 mm (*Z* direction)	Blank insertion X position	− 5 to 5 mm (*X* direction)
Blank insertion Y position	− 5 to 5 mm (*Y* direction)		

vector via the total sensitivity index S_{T_i} from Formula (12) can be directly estimated in the low-dimensional characteristic space using the regression model from Formula (20):

$$S_{T_{ij}} = 1 - \frac{V\left[E\left(\hat{Y}_{Score_j} \mid \boldsymbol{S}_{Input\sim i}\right)\right]}{V\left(\hat{Y}_{Score_j}\right)} \tag{23}$$

Formula (23) also makes it possible to determine the relevant and irrelevant process parameters. If the value for $S_{T_i} \approx 0$, the *i*th input parameter of the sample matrix $\boldsymbol{S}_{Input}$ has no influence on the total output variation and can be removed from the regression model.

3.4 Back-transformation in original space

To visualize the sensitivity results of the *i*th input parameter on the component geometry in an element-specific manner, a back-transformation of the data from all *m* principal components into the original space is necessary using Formula (9) and (23):

$$S_{T_i}^{Field} = \sum_{j=1}^{m} \left(S_{T_{ij}} \cdot \frac{V\left(Y_{Score_j} V_j^T\right)}{V(\mathbf{X}_u)} \right) \tag{24}$$

$S_{T_i}^{Field}$ is a $(d \times 1)$ vector of elementwise sensitivity results of the *i*th input parameter. $S_{T_{ij}}$ is the total sensitivity index of the *j*th principal component from (23). $V\left(Y_{Score_j} V_j^T\right)$ is the proportion of variance that is explained by the j-principal component and $V(\boldsymbol{X}_u)$ is total variance of the field matrix $\boldsymbol{X}_u$.

3.5 Interactive process visualization

The mathematical description of the projection matrix via a regression model (see Formula (20)) also offers the opportunity to connect the process parameters in their range of variation interactively by means of graphical animation tools such as sliders (see Fig. 7, left) with visualization of results. The back-transformation of the model-based projection vector $\hat{Y}_{Score_j}$ is given by Formulas (9) and (20):

$$\hat{\boldsymbol{X}}_{u_j} = \hat{Y}_{Score_j} \boldsymbol{V}^{\boldsymbol{T}} = \left(\hat{\boldsymbol{S}}_{Input} \cdot \beta\right) \cdot \boldsymbol{V}^{\boldsymbol{T}} \tag{25}$$

Formula (25) creates a direct functional relationship between input and output variables. If the slider is used to set a previously unknown parameter combination, the model-based field-data matrix $\hat{\boldsymbol{X}}_{u_j}$ can be interpolated and updated in the graph of results in milliseconds. The elementwise evaluation of deep-drawing sensitivities can help in attaining a profound process understanding and in deriving faster measures to

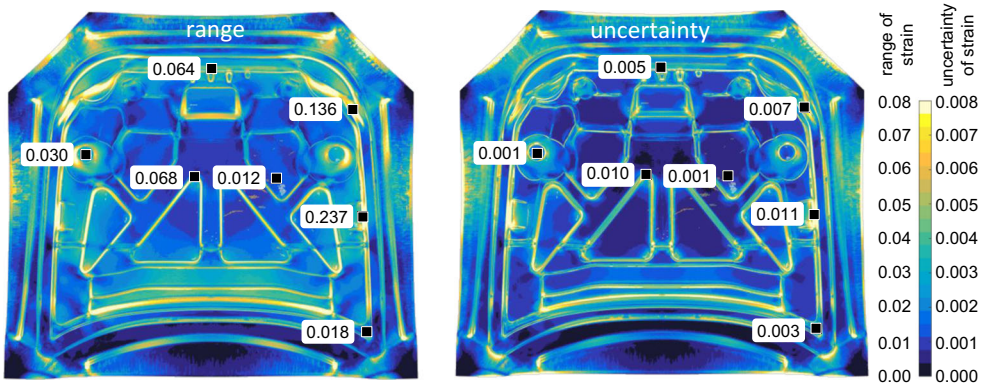

Fig. 3 Strain distribution

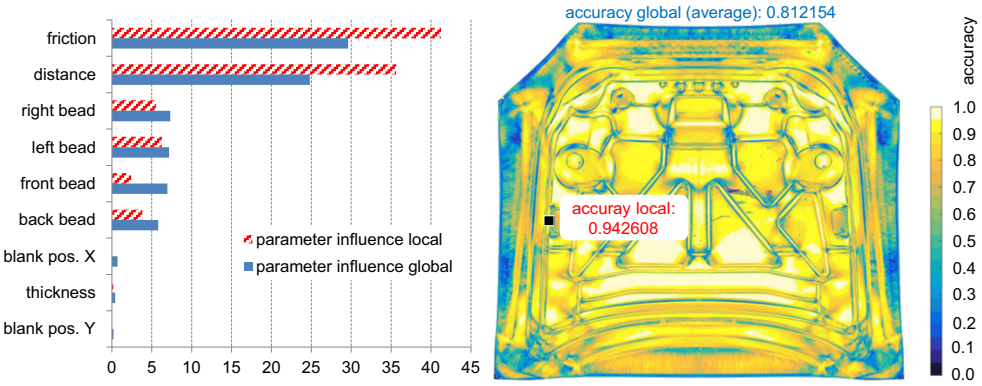

Fig. 4 Accuracy and parameter influence overview for strain distribution

compensate for dimensional variations in the deep-drawing process.

4 Case study

4.1 Experimental setup

The presented methods were verified based on the deep-drawing and clamping process steps for the inner part of a front hood. The basis was the geometry from the ESI PamStamp 2G Tutorial "Standard Forming Hood Inner." Four beads were added in the deep-drawing tool. In addition, during the cutting process, a long hole and a round hole were introduced, which enable definition in space in a measuring device with four variable clamping points and two pins (see Fig. 2).

The model was calculated in the PamStamp 2G 2015.1 solver with explicit, non-linear contact algorithms and adaptive mesh refinement. The component size of 1600 × 1500 mm, represented by up to 1.1 million elements makes it possible to compare the results with real analysis in car body ramp-up. The simulation attribute file was built parametrically, so that an external ASCII-based modification of the input parameters by external scripts was possible. The evaluation of the observed output parameters strain and accuracy was realized by an element-based contour plot.

In conformity with other experiments from the technological state of the art, 12 process input parameters as well as the sheet thickness were evaluated with respect to the influence on strain and geometrical accuracy of the clamped part (see Table 1). Variations in material characteristics were intentionally omitted, since according to Berry [40], they are less relevant to deep-drawing results than are other parameters. To ensure a high forecasting quality for the PCA, the DOE of the 13 process parameters was based on Latin hypercube sampling (LHS) with 100 simulation samples (see Chapter 3.1).

4.2 Sensitivity analysis of component strain

Figure 3 shows the distribution of numerical post variables, using plastic strain as an example, on the uncut deep-drawn part in the bottom dead center. Similar evaluations using other post variables (thinning, stress, etc.) are also possible. Here, the range is the absolute value of the difference between the largest and smallest component strain value across the 100 simulations. The uncertainty describes the expected average error of the mathematical model. This is calculated as the square root of the variance proportion that cannot explained by the COD (see Formula (22)).

It is clear that some of the given parameters (see Table 1) and their variation have a significant effect on component strain and some do not. By using cross-validation methods,

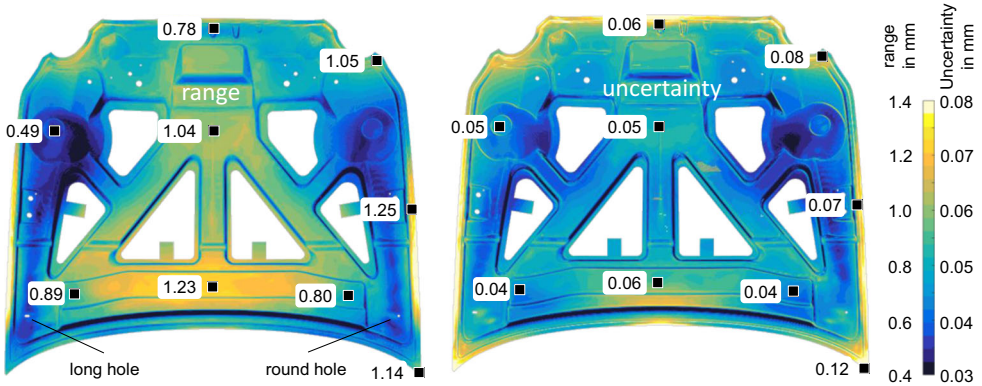

Fig. 5 Springback and tolerance analysis

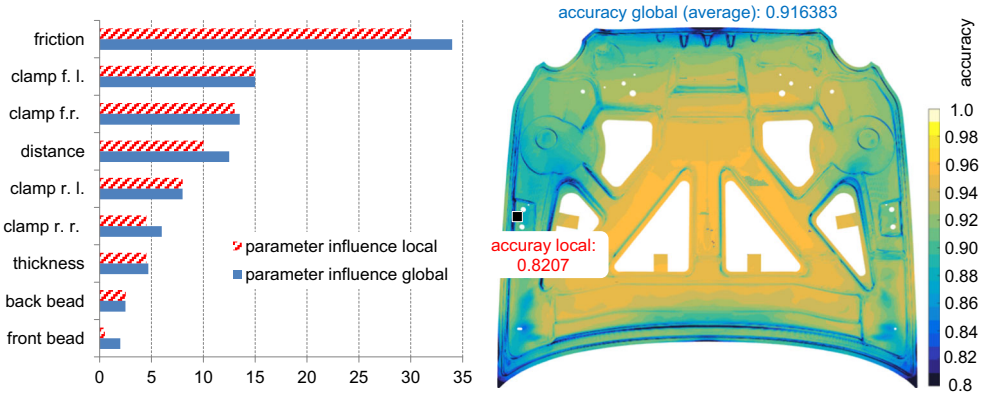

Fig. 6 Accuracy and parameter influence and accuracy for springback analysis

it was possible to isolate irrelevant from relevant parameters and to find the best-fitting model which has uncertainties of, at maximum, approximately 10% of the range achieved.

This can also be seen in Fig. 4, right. The prognostic quality (see Formula (22)) within the area of the beads is between 85 and 100%. The waviness in the outer plate area that occurs arbitrarily depending on the parameter settings, as well as the different plate indentation, means that it is not always possible to assign the post variables to a specific node of the reference, with the result that there is reduced model accuracy in these areas in particular.

Figure 4, left, shows and compares the principal influencing variables for global strain distribution and for a local point. As previously stated, it is clear that friction and the distance from the blank holder to the die have the greatest influence on the strain, both globally and locally. However, these parameters have a significantly higher effect on the selected local point than on average across the entire component. Furthermore, it is found that although the selected point is on the left side of the component, the left bead has a scarcely higher influence on strain than the right bead. This is also reflected overall in the relatively low influence of the beads on the average strain across the component as well as on the local selected point. It can be seen that friction blurs the influence of the other parameters due to the high range selected and the resulting high influence on the component strain. It can be concluded that the method presented is able to reliably represent, in high quality, the post variable distribution on a deep-drawn component; however, the selection of input parameters and their range has a substantial influence on the significance.

4.3 Sensitivity analysis of springback

Furthermore, the method presented was investigated with respect to its prognostic quality for springback processes and dimensional stability analysis. To this end, the virtually deep-drawn components were also cut, inserted into a measuring device, clamped, and loaded with gravity. As shown in Fig. 5 a high level of forecasting ability was also achieved in this case. In turn, the absolute uncertainty of the model was less than 10% of the range.

It was even possible to represent the influences of the asymmetrical determination by mounting on a round hole on the right side and on a long hole that permits relative movement on the left side. This is significant in that, despite a different range in the area of the holes, the uncertainty remains identical.

An influence analysis and an accuracy analysis were also performed for the springback process step (see Fig. 6, right). It can be seen that the global prognostic quality, at 92%, has improved, while the local prognostic quality at the selected point has worsened to 82%.

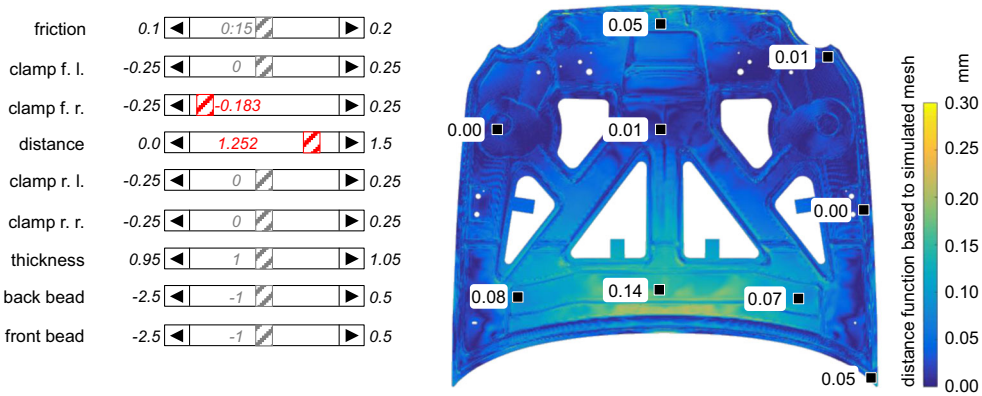

Fig. 7 Potential of using sliders for interactive influence analysis

It is also clear that the adjustment of the front clamp has reduced the importance of the blank holder-to-die distance parameter; thus, its influence on the dimensional stability of the component is greater. The influence of the sheet thickness on the springback behavior of a component is also clearly higher than it is on the component's percentage strain (see Fig. 6, left).

4.4 Interactive influence analysis

Using the calculation method presented as Formula (25), it is possible to manipulate the input variables manually and to evaluate the influence, individually or in combination with each other, on the dependent variables. The function-based representation of the interactions allows an interpolation between recorded values and thus enables a continuous selection of values. In Fig. 7, this was carried out using the example of the dimensional stability of the component. The influence of the presented parameter variation on the original condition was evaluated. The schematic representation of sliders is intended to serve as an example for a user-friendly implementation of this method.

A back-transformation of the modes can be used at any time to regenerate the original data, or the post variable-related simulation mesh; thus, it is possible to export a manually varied mesh and to use it as input data for further simulations (e.g., for assembly process simulation).

This was used to generate the comparison shown in Fig. 7. Where a set of input values was defined arbitrarily (see Fig. 7, left) and following the response (geometrical set of nodes, which represents the body part) was generated by running a simulation and by using the described mathematical approach on the previous generated model. Fig. 7, right, shows the distance between the simulated mesh and the function-based mesh for the given parameter combination in Fig. 7, left. It can be seen that both meshes are in good agreement.

5 Discussion and conclusion

A mathematical approach of using the SVD-based PCA in combination with a variance-based sensitivity analysis and its capability for analysis of car body processes was documented. The case study showed that the mathematical and numerical methods presented can be used to represent interactions in car body construction, based on mode-based models.

The low uncertainty levels show that it was possible, with the PCA method, to generate a model which represents the interaction of the nine continuous input parameters with reference to the output parameters strain and geometry based on 50 simulations. The applicability of the prognosis model for real problem-solving in the ramp-up of a car body line, of course, depends on how well the respective simulation model is able to depict reality. As explained in the state of the art, this is certainly possible for strain, stress, and thinning, but only a qualitative—not quantitative—statement is possible for springback. Another positive aspect of using the PCA is that it was possible to perform the calculation of the eigenmodes (100 meshes with 1.2 million elements each) on a standard Quad-Core PC with 8 GB RAM. The total computation time was approximately 2 min.

In the next steps, the method should be verified based on further car body components. The extension of the numerical process chain to clamping and joining or hemming is also planned.

As shown in Chapter 4.4, with sufficient model quality, it is possible to generate new design variants mathematically instead of numerically, as has previously been done. So, after a teaching phase of 100 simulations that can be run in parallel, it is possible to obtain information regarding effects on car body process adjustments in a matter of seconds instead of hours of simulation time or days of experiments. This would allow car body construction measures to be checked just in time before implementation based on metamodels, thus saving on cost-intensive experiments.

Furthermore, there is the potential to generate realistic variations of drawn parts for the simulation of assembly processes or later process steps without simulating the drawing process.

As presented in the introduction chapter, FEM simulation is capable of representing influences of variation in car body processes, but it requires too much time to be useful for urgent problems. Using the method described, it is possible to transfer FEM-generated knowledge into a mathematical model that is able to generate answers just in time. This will increase the use of FEM during the planning or ramp-up processes of car body production lines, because it can now be used to solve a problem instead of just to analyze it.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Hu M, Lin Z, Lai X et al (2001) Simulation and analysis of assembly processes considering compliant, non-ideal parts and tooling variations. Int J Mach Tools Manuf 41(15):2233–2243. https://doi.org/10.1016/S0890-6955(01)00044-X
2. Long Y, Hu SJ (1998) A unified model for variation simulation of sheet metal assemblies. In: ElMaraghy HA (ed) Geometric Design Tolerancing: Theories, Standards and Applications. Springer US, Boston, pp 208–219
3. Schulte S (1999) Umformsimulation als fester Bestandteil des Produktentstehungsprozesses bei Audi. In: Doege E (ed) 16.

Umformtechnischen Kolloquium, Hannover 25. und 26. HFF, Hannover, pp 263–277
4. Li E (2013) Reduction of springback by intelligent sampling-based LSSVR metamodel-based optimization. Int J Mater Form 6(1): 103–114. https://doi.org/10.1007/s12289-011-1076-1
5. Roll K, Lemke T, Wiegand K (2004) Simulationsgestützte Kompensation der Rückfederung. Procedia of 3. LS-DYNA Anwenderforum: A-I-1 - A-I-13
6. Großmann K, Ulbricht V, Hartmann A et al. (2007) Vergleichende Bewertung der Simulation von Umformprozessen mit elastischen Randbedingungen: Ergebnisse eines Vorhabens der Industriellen Gemeinschaftsforschung (IGF) …; Abschlussbericht 2007. EFB-Forschungsbericht, vol 264. EFB, Hannover
7. Ackert P, Schwarz C, Müller J et al (2015) New approach for measurement & simulation in sheet metal forming. GOM Conference, Braunschweig
8. Lee JD, Haynes LS (1987) Finite-element analysis of flexible fixturing system. J Eng Ind 109(2):134–139. https://doi.org/10.1115/1.3187103
9. Landgrebe D, Ackert P, Grützner R et al (2015) Improving the ramp-up process of a body-construction line by numerical supported design of clamping devices and FEM based tolerance prognosis. Key Eng Mater 651-653:932–937. https://doi.org/10.4028/www.scientific.net/KEM.651-653.932
10. Siekirk JF (1986) Process variable effects on sheet metal quality. J Appl Metalwork 4(3):262–269. https://doi.org/10.1007/BF02833934
11. Zhang ZT, Pitt G, King A et al. (1998) Stamping part surface quality improvement through process optimization. SAE technical paper series 980073. doi: https://doi.org/10.4271/980073
12. Majeske KD, Hammet PC (2003) Identifying sources of variation in sheet metal stamping. Int J Flex Manuf Syst 15(1):5–18. https://doi.org/10.1023/A:1023993806025
13. Naceur H, Guo YQ, Ben-Elechi S (2006) Response surface methodology for design of sheet forming parameters to control springback effects. Comput Struct 84(26–27):1651–1663. https://doi.org/10.1016/j.compstruc.2006.04.005
14. Wei L, Yuying Y, Zhongwen X et al (2009) Springback control of sheet metal forming based on the response-surface method and multi-objective genetic algorithm. Mater Sci Eng A 499(1–2): 325–328. https://doi.org/10.1016/j.msea.2007.11.121
15. Meinhardt J, Grossenbacher K, Ganser M et al. (2007) Neue Wege zum wirtschaftlichen Leichtbau: Innovative Lösungen zur Blechumformung und mechanischen Fügetechnik; Tagungsband des EFB-Kolloquiums Blechverarbeitung 2007 am 6. und 7. März 2007 in Fellbach. Tagungsband / Europäische Forschungsgesellschaft für Blechverarbeitung e.V, T 27. Europäische Forschungsges. für Blechverarbeitung e.V. (EFB), Hannover
16. Emrich A, Liewald M, Ruf G (2012) Increasing press shop productivity and quality by the means of robustness analysis in the die engineering phase. Mini-Symposium: Robust Design in Forming Technology, IJmuiden
17. Wärmefjord K, Söderberg R, Ottosson P et al. (2013) Prediction of geometrical variation of forged and stamped parts for assembly variation simulation. In: Hora P (ed) Proceedings of iddrg Conference 2013: Conference 2013, Zurich, Switzerland, June 2–5, 2013. Institute of Virtual Manufacturing ETH Zurich, Zürich, pp 99–104
18. Wolff S (2015) Robustness analysis of metal forming simulation – state of the art in practice. In: Landgrebe D, Drossel W, Putz M (eds) 5th international conference on accuracy in forming technology: ICAFT 2015, 22nd SFU 2015: proceedings. Verlag Wissenschaftliche Scripten, Auerbach, pp 319–334
19. Gerbino S, Kriechenbauer S, Franciosa F et al (2015) Understanding whole shape variability of stamped sheet metal parts for root cause analysis. In: Landgrebe D, Drossel W, Putz M (eds) 5th international conference on accuracy in forming technology: ICAFT 2015, 22nd SFU 2015: proceedings. Auerbach, Verlag Wissenschaftliche Scripten, pp 423–438
20. Youcef-Toumi K, Liu WS, Asada H (1988) Computer-aided analysis of reconfigurable fixtures and sheet metal parts for robotic drilling. Robot Comput Integr Manuf 4(3–4):387–394. https://doi.org/10.1016/0736-5845(88)90009-9
21. Franciosa P, Gerbino S, Ceglarek D (2016) Fixture capability optimisation for early-stage design of assembly system with compliant parts using nested polynomial chaos expansion. Procedia CIRP 41: 87–92. https://doi.org/10.1016/j.procir.2015.12.101
22. Schwarz C, Ackert P, Rössinger M et al. (2015) Mathematical optimization of clamping processes in car-body production. 12th optimization and stochastic days, Weimar https://www.dynardo.de/fileadmin/Material_Dynardo/bibliothek/WOST12/22_WOST2015_Session6_Schwarz_Paper.pdf
23. Johnson RA, Wichern DW (2002) Applied multivariate statistical analysis, 5th edn. Prentice Hall, Upper Saddle River, NJ
24. Pearson K (1901) On lines and planes of closest fit to systems of points in space. Philosophical Magazin 2:559–572
25. Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Educ Psychol 24(6):417–441. https://doi.org/10.1037/h0071325
26. Jolliffe IT (2004) Principal component analysis, 2. ed., [Nachdr.]. Springer series in statistics. Springer, New York
27. Golub G, Kahan W (1965) Calculating the singular values and pseudo-inverse of a matrix. J Soc Ind Appl Math Ser B Numer Anal 2(2):205–224. https://doi.org/10.1137/0702016
28. Reris R, Brooks JP (2015) Principal component analysis and optimization: a tutorial. Proceedings of 14th INFORMS computing society conference 2015: 212–225. doi: https://doi.org/10.1287/ics.2015.0016
29. Madsen RE, Hansen LK, Winther O (2004) Singular value decomposition and principal component analysis. Neural Netw 1:1–5
30. Saltelli A (2004) Sensitivity analysis in practice: a guide to assessing scientific models. John Wiley & Sons Ltd., Hoboken
31. Saltelli A, Ratto M, Andres T et al (2008) Global sensitivity analysis: the primer. John Wiley & Sons Ltd., Chichester, Hoboken
32. Cukier IR, Levine HB, Shuler EK (1978) Nonlinear sensitivity analysis of multiparameter model systems. J Comput Phys 26(1): 1–42
33. Cukier IR, Fortuin MC, Shuler EK et al (1973) Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. J Chem Phys 59:3873–3878
34. Sobol IM (1993) Sensitivity estimates for nonlinear mathematical models. Math Model Comput Exp 1:407–414
35. Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. Reliab Eng Syst Saf 52(1):1–17. https://doi.org/10.1016/0951-8320(96)00002-6
36. Saltelli A, Annoni P, Campolongo F et al (2010) Variance based sensitivity analysis of model output - design and estimator for the total sensitivity index. Comput Phys Commun 181:259–270
37. Saltelli A (2002) Making best use of model evaluations to compute sensitivity indices. Comput Phys Commun 145(2):280–297. https://doi.org/10.1016/S0010-4655(02)00280-1
38. Saltelli A, Annoni P (2010) How to avoid a perfunctory sensitivity analysis. Environ Model Softw 25(12):1508–1517. https://doi.org/10.1016/j.envsoft.2010.04.012
39. Montgomery DC, Runger GC (2003) Applied statistics and probability for engineers, 3rd edn. Wiley, New York, Chichester
40. Berry B (1996) Inside the stamping press, where the costs are. New Steel 12(6):22–35