



Critical joint identification for efficient sequencing

Downloaded from: <https://research.chalmers.se>, 2026-04-04 23:00 UTC

Citation for the original published paper (version of record):

Sadeghi Tabar, R., Wärmefjord, K., Söderberg, R. et al (2021). Critical joint identification for efficient sequencing. *Journal of Intelligent Manufacturing*, 32(3): 769-780.

<http://dx.doi.org/10.1007/s10845-020-01660-4>

N.B. When citing this work, cite the original published paper.



Critical joint identification for efficient sequencing

Roham Sadeghi Tabar¹ · Kristina Wärmefjord¹ · Rikard Söderberg¹ · Lars Lindkvist¹

Received: 29 April 2020 / Accepted: 23 August 2020
© The Author(s) 2020

Abstract

Identifying the optimal sequence of joining is an exhaustive combinatorial optimization problem. On each assembly, there is a specific number of weld points that determine the geometrical deviation of the assembly after joining. The number and sequence of such weld points play a crucial role both for sequencing and assembly planning. While there are studies on identifying the complete sequence of welding, identifying such joints are not addressed. In this paper, based on the principles of machine intelligence, black-box models of the assembly sequences are built using the support vector machines (SVM). To identify the number of the critical weld points, principle component analysis is performed on a proposed data set, evaluated using the SVM models. The approach has been applied to three assemblies of different sizes, and has successfully identified the corresponding critical weld points. It has been shown that a small fraction of the weld points of the assembly can reduce more than 60% of the variability in the assembly deviation after joining.

Keywords Critical joint · Sequence · Machine learning · PCA · Assembly · SVM

Introduction

In the manufacturing industry, nominal geometries do not exist. Every component to be assembled is subjected to geometrical variation. Identifying the sources of variation and controlling them is a common challenge. Robust design perspectives [14] and geometry assurance activities [16,28] are developed to control the effect of geometrical variation in the final product. Joining operation is one of the most critical processes causing the final geometrical variation. For sheet metal assemblies, where the sheets are compliant, the joining operation is mainly spot welding or riveting. With the recent development within digitalization and availability of data, a self-compensating assembly line for sheet metal assemblies is developed based on the digital twins [18]. In this concept, a Computer-Aided Tolerancing (CAT) tool is taking advantage of local optimizers to improve the geometrical quality of the assemblies in different disciplines. For the joining operation, the optimal sequence of the spot welding process is proposed using the digital twin for individualized assemblies [22]. For each assembly, a number of the most important welds is needed to be selected and set in a sequence

and secure the geometry in each assembly station. Therefore, first, the number of welds that are critical for the assembly needs to be determined. Secondly the sequence of such welds considering the rest of the weld points should be decided on. The question of the sequence of welding is categorized as a classical Operation Research (OR) problem. It is an NP-hard combinatorial optimization problem with an expensive evaluation function. In this paper, the above-mentioned problems are tackled with a state of the art clustering approach for weld identification and a novel machine learning perspective on the sequence optimization.

Assembly evaluation

To evaluate the geometrical deformation of the sheet metal assemblies, given the part tolerances and the material properties, CAT-tools based on the Finite Element Method (FEM) are used. With the Method of Influence Coefficients (MIC), a linear relationship between the part deviation and assembly deviation can be achieved by constructing a sensitivity matrix as a medium. This method has shown to perform faster than other direct methods for this purpose [11]. The accuracy of the method is increased by combining the MIC with a contact modeling algorithm. The contact points in this method avoid the parts penetrating in the adjacent areas during the assembly [9]. The calculation time of the MIC and contact modeling

✉ Roham Sadeghi Tabar
rohams@chalmers.se

¹ Department of Industrial and Materials Science, Chalmers University of Technology, Gothenburg, Sweden

has been reduced with a new method to calculate the sensitivity matrix of the spot welded assemblies in a sequence [12,13].

Joining sequence optimization

Finding the optimal joining parameters has been studied over time. Usage of neural networks in the early 90s has been addressed [3,6]. In more recent studies, a deformation prediction approach for continuous welding applications has been introduced using the support vector machines (SVM) and evolutionary optimization [25]. Principle component analysis (PCA) have been studied together with the optimization approaches for the optimal design of joining parameters, such as welding time, current, and electrode force [31].

Using the state of the art MIC and contact modeling, the sheet metal assemblies are evaluated for a specific joining sequence [12]. To identify which sequence of spot welding results in the optimal geometrical outcome of the assembly is a NP-hard combinatorial optimization problem. Assembly deformations after welding, calculated using MIC and contact modeling, is the objective function. The evolutionary algorithms, specifically Genetic Algorithm (GA), have been studied for this problem [5,8,17]. The performance of other evolutionary algorithms has been compared with each other with the same approach [19]. Rule-based approaches have been introduced and have shown to improve the performance of the evolutionary algorithms on the problem [21,24]. Using the neural networks, a surrogate modeling approach has been introduced, representing the sequence-deformation relationship of the assembly [22]. A new stepwise algorithm has also been introduced to identify the optimal welding sequence with the minimum number of evaluations with the method presented above [23]. In most of the studies, the aim is to present a complete sequence that results in the minimum assembly deformations after joining. It has been studied in [20] that some of the weld points are more important in the sequence than the others. A sensitivity analysis approach with regards to the weld gaps, has been shown to be effective. However, the number of critical welds in the sequence has not been discussed. Identifying such welds, accurately, help reduce the problem size for sequence optimization. For industrial applications, the weld points that are set initially on the assembly to lock the geometry, in the processes downstream, are referred to as geometry weld points or geo-spots [29]. These weld points are often identified based on physical experimentation over time, and the sequence of such weld points are critical for the geometrical outcome. The rest of the weld points on the assembly are referred to as re-spots, where the sequence of these weld points do not impact the final geometrical outcome. The challenge is to identify the number of the geometry points, the location of these points among a set of alternatives, and their sequence to minimize

the geometrical deviation of the assembly. The identification of the critical weld points helps efficient assembly process planning while improving the geometrical quality [27].

To identify the most critical weld point in a sequence, with respect to geometrical deviation, a method where joining sequences can be evaluated time-efficiently, and an approach to distinguish between the sequence elements are required. While previous studies have mainly focused on introducing an optimal complete sequence, where the problem is finding the optimal sequence for an assembly with a fixed number of joints, the problem of identifying the most critical joints, their number and their location from a set of joints has not been addressed. Therefore, in this paper, an approach to evaluate the sequences time-efficiently is introduced, and a method to identify the number and location of the critical joining elements in a sequence is proposed.

Scope of the paper

In a sheet metal assembly, a number of weld points, known as geometry weld points, play an important role in the final geometrical quality. To identify the number, the location of these weld points out of a number of candidates, and their sequence leads to reduced problem size for sequencing and efficient assembly process planning [29]. Previous studies focus on presenting an optimal sequence considering all the joints, while there is no generic approach for the identification of critical weld points. In this paper, the accurate number, location and sequence of these critical joints are identified with an input-output SVM model for sequence evaluation, and PCA to define the number of the critical joints. The first Section provided an introduction to the problem. The rest of the paper is structured as follows. In Sect. 2, the proposed method for identification and sequence determination of the critical weld points is introduced. Section 3 introduces the reference sheet metal assemblies. The method is evaluated on the sheet metal assemblies, and the results are presented in Sect. 4. Finally, the conclusions and future work plans are presented in Sect. 5.

Proposed identification method

In this section, the proposed method, including the sequence prediction method and the critical weld point identification, is described. The prediction method is built on the surrogate modeling approach studied in [22]. In this study, to predict the geometrical outcome of the spot welded assemblies, a radial basis function (RBF) network is designed and trained with a specified sample. The approach is for the purpose of the complete sequence optimization, considering all the weld points in the assembly. In this paper, accurate evaluation of the sequences needed for critical weld point identifica-

Table 1 List of variables and their description

Variable	Description
w_i	Weld point i
n	Number of weld points
Π	Sorted matrix of permutations of n
π	A row vector of Π
Π'	Input-response matrix of Π
π'	A row vector of Π'
D	Differences of π' in Π'
μ_D	Mean of D
σ_D	Standard deviation of D
x_i	Position i in sequence [1 to n]
Π_s^i	Sample matrix of interval i
e	Root mean square of error
u	Defined error limit
L_i	Length of the interval i
$\hat{\Pi}_s^i$	The updated matrix Π_s^i
$f(x)$	Deviation estimation function to a sequence x
w_l	Support vector weights
ϕ_l	Support vector transfer function
b	Support vector bias
$G(x_1, x_2)$	Support vector kernel function
α_l	Lagrangian multipliers
q	PCA initial sample matrix
Q	PCA complete matrix
m	Number of rows in Q
Σ_Q	Covariance matrix of Q
U	Matrix of eigenvectors of Σ_Q
k	PCA data dimension
Z	PCA approximation in k dimension
J	Percentage of the retained variance by PCA
P_c^w	Number of principle component weld points
y	Real observed assembly deviation of a sequence
\hat{y}	Assembly deviation estimate of the sequence

tion is essential. Therefore, the above approach is further expanded with SVM networks tuned to predict the outcome of the sequences accurately, and PCA for critical weld point identification. The list of the variables that are introduced throughout the paper and their description are presented in Table 1.

Sequence data processing

Evaluating the initial sequence elements, specifically the initial triplet, has shown to describe the overview of the changes between the sequence with respect to the geometrical deviation of the assembly after welding [22]. Consider the assembly with seven weld points, $[w_1, w_2, w_3, w_4, w_5,$

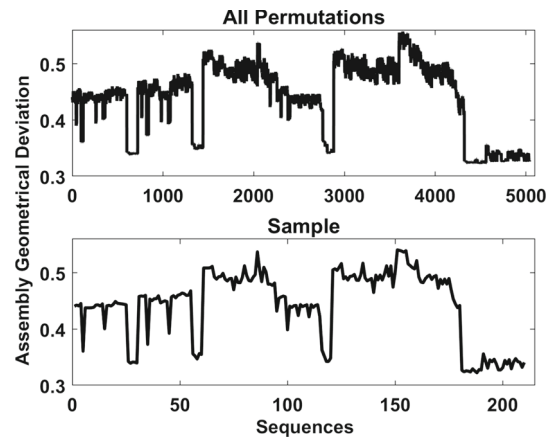


Fig. 1 Comparison between all the permutations and sample of first triplets

$w_6, w_7]$, then all the combinations of the first three elements, w_1, w_2, w_3 , with any combination from the other elements, provides a clear understanding of the major differences between the sequences. Figure 1 shows all the geometrical deviation in an assembly with seven weld points and the selected sample, as stated above. To represent the geometrical deviation of the assembly the root mean square of the deviations of all the mesh nodes in a specific direction, here normal direction, y , are calculated using the state of the art MIC and contact modeling approach presented by [12]. Any other preferred measure of deviations can be used in the approach.

This information can be used to divide the sequences to different intervals, based on the differences between the sequence intervals. Within each interval, the refinements of the observations can be performed to capture the variability among all the sequence elements.

To identify such intervals of permutations, the initial elements of the sequence is evaluated, and a surrogate model is built with the previously introduced approach in [22]. Now finding the outliers in the data with such information can be achieved comparing the differences between the surrogate model responses in each interval. Defining $\Pi_{n! \times n}$ the sorted matrix of all the permutations of n , with a unique sequence of welds in each row, $\Pi'_{n! \times n+1}$ is the input-output matrix, where the output π' , assembly response to each sequence retrieved from the surrogate model is in the last column of the matrix. The vector of the difference of the simulated responses is gathered in;

$$D = \sum_{i=1}^{n!} \pi'_{i+1} - \pi'_i. \tag{1}$$

Consider the assembly with seven weld points in Fig. 1, the vector D reveals the interval indices between the sequences.

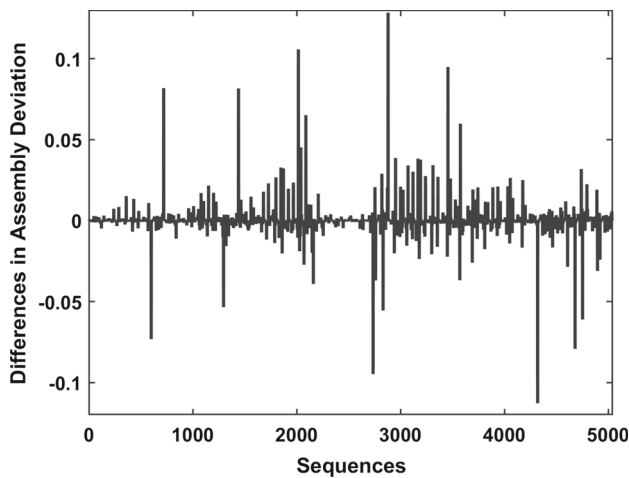


Fig. 2 Differences in the assembly deviation

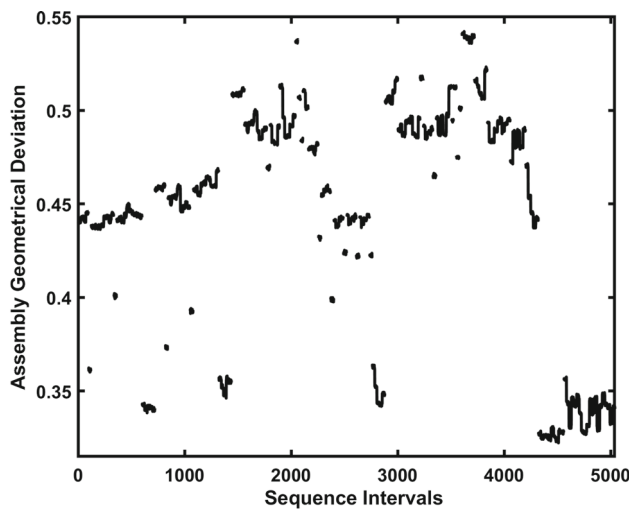


Fig. 3 The calculated intervals in the permutations and responses

Figure 2 is visualizing the D vector for such an assembly. The drastic changes in D represent the interval indices of the divisions in Π .

The limit for the changes in the D can be set to three standard deviations from mean. Let μ_D , and σ_D be the mean and standard deviation of vector D . Elements of D , are represented by d . If $d \geq \mu_D \pm 3\sigma_D$, then the index of d in vector D is the data division point in Π . The outcome of this process is the divided data where in each interval the response of the assembly to the sequences is within the specified limit. Figure 3 shows the generated intervals on the data set presented in Fig. 2. On this data set, 49 sequence intervals are generated where, for each interval, the assembly response lie within the threshold defined.

To increase the accuracy of the variability representation in each arbitrary interval, the sequences are analyzed to identify the common elements. In each interval, the sequences are represented by the matrix Π^i . The size of the matrix is the

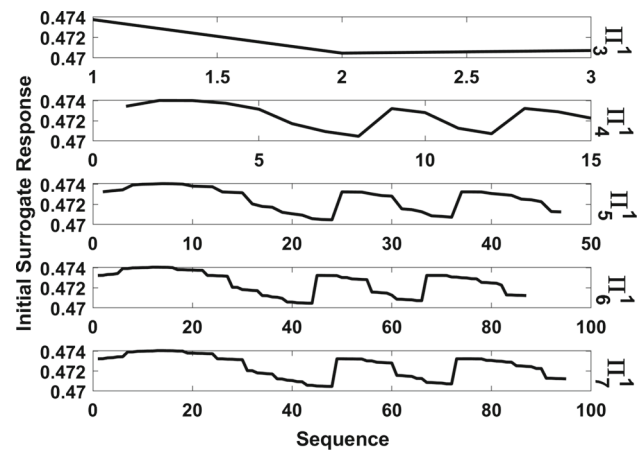


Fig. 4 The selected samples in an interval

same as the length of the interval, and number of the welds in the assembly. If each row of Π^i is represented by the vector $\pi^i = [x_1, x_2, \dots, x_n]$, then, there are n positions available for each weld to appear in. The common elements in each position in the interval are identified. This task can be performed by a simple indexing over the positions x_1, \dots, x_n and rows of Π^i . With this approach, for an assembly with n welds, n sets of samples are generated. Defining these sets of samples with different number of rows as $\Pi_s^i, s = \{1, \dots, n\}$, the n^{th} sample set is identical to Π^i . With a brute force approach, a suitable Π_s^i is selected. One approach to achieve this is to evaluate how many elements of the sequence in each interval can describe the variability between the sequences, within the specified error from a reference. The initial surrogate response is considered as the reference response for each sequence. To clarify this point further, consider the first interval of the sequences Π^1 , the response of this interval, from the initial surrogate is visualized in Fig. 3. This interval is made of 96 sequence vectors of length $n = 7$. Finding the common elements in each position x_1 to x_7 between the 96 sequences, give us seven sets of samples, Π_1^1 to Π_7^1 . The responses of each of these samples are shown in Fig. 4. In this interval the first two sample sets did not have any members. The rest of the samples are presented. It can be seen that a model fitted to Π_4^1 , could describe the variability in this interval. To identify which set of sample is suitable for each interval, a model is fitted to each sample at each interval. The model that results in the root mean square error, e , below the specified limit, u is chosen to be evaluated using the CAT-tool.

$$e = \sqrt{\frac{1}{L} \sum_{i=1}^L (\Pi^i - \hat{\Pi}_s^i)^2} \leq u, \quad \forall s \in \{1, \dots, n\}. \quad (2)$$

The difference in the response of the initial surrogate to the sequences in the interval Π^i , and the response of the model

built by sample Π_s^i , shown by $\hat{\Pi}_s^i$ is calculated to build e , where L is the length of each interval. The error limit, u , can be set based on the computation time window. After evaluating the sample of the interval using the CAT-tool, the interval response set $\hat{\Pi}^i$ is updated with the chosen $\hat{\Pi}_s^i$ observations.

Depending on the size of the problem, the necessary intervals can be selected. For larger problems, the intervals with the minimum assembly deviation, and the intervals necessary for PCA analysis are updated.

Support vector machines

With the data processed, non-linear models need to be built for the intervals. There are several approaches available for this purpose, while it has been shown that usage of the radial basis functions can provide an accurate estimation for modeling the different deviation among the sequences [22]. SVM is a method that incorporates the RBF kernels to transform the sequence input to the required assembly deviation form. It has also been shown that the performance of such models can be improved by hyper-parameter optimization [10] which can be performed in parallel for time-efficiency. Having a set of observations in each arbitrary interval of the sequences and the initial surrogate model, an SVM network is built for the interval, to represent the variability between the sequences. SVM has been introduced by [26] for data classification [7], and has been extensively used for non-linear modeling [4,30]. Considering the welding sequence as the multiple input, and, the assembly response as one output, then SVM can be used for nonlinear function estimation [26].

$$f(x) = \sum_{l=1}^L w_l \phi_l(x) + b \tag{3}$$

The function $f(x)$ is the estimation of the assembly response to weld sequence input, x . Where w_l is the support vector weights and ϕ_l is the non-linear transfer function of the input x , and b is the bias. With a Langrange dual formulation, the kernel function $G(x_1, x_2)$, expresses the dot product of the transfer functions ϕ_1 and ϕ_2 . The Lagrangian function with the multipliers α_l and α_l^* for each weld sequence assembly response x_l is expressed as:

$$L(\alpha) = \frac{1}{2} \sum_{r=1}^L \sum_{u=1}^L (\alpha_r - \alpha_r^*)(\alpha_u - \alpha_u^*)G(x_r, x_u) + \dots + \epsilon \sum_{r=1}^L (\alpha_r - \alpha_r^*) - \sum_{r=1}^L f(x_r)(\alpha_r^* - \alpha_r) \tag{4}$$

Through regularization, the weight factors are determined. Ultimately, the estimation function can be expressed as :

$$\hat{y} = f(x) = \sum_{l=1}^L (\alpha_l - \alpha_l^*)G(x_l, x) + b, \tag{5}$$

where the values of $(\alpha_l - \alpha_l^*)$ can be determined through quadratic programming. In this study, the Gaussian RBF kernels are used in the SVM model, as the previous studies specified the applicability of the RBF on such problems [22]. The RBF kernel function is expressed as:

$$G(x_l, x) = \exp\left(-\frac{\|x_l - x\|^2}{2\sigma^2}\right), \tag{6}$$

where σ is the spread of the kernel function.

With the above formulation, for each interval i , $\hat{y}_i = f_i(x)$ is trained using the generated sample, to estimate the response to the sequences Π^i . As a result, a non-linear SVM model represents the response of the sequences in the interval. Now to evaluate a given sequence, x_l , initially the interval corresponding to the sequence is identified, and the $f_i(x)$ is used to estimate the response.

For problems of a larger size, the SVM models can be built for the intervals necessary for PCA analysis and intervals resulting in the minimum mean assembly deviation.

Principle component analysis

Accurate number of the sequence elements contributing to the majority of the variation among the sequences needs to be determined. This is in line with the definition of the PCA used to exploratory data analysis, to summarize the main characteristics of the data in lower dimensions [2]. Being able to estimate the response of the assembly to the sequences using the SVM models, PCA is used to determine the number of the features describing the variability in the sequence of the weld points. The PCA method intends to map the explanatory data in higher dimensional space to lower dimensions. The approach is based on eigenvalue decomposition of a data covariance matrix [1]. To evaluate the number of features describing the behavior between the sequences, a meaningful data set is needed. As shown in Fig. 1 the triplets of the sequences can describe the major effects between the sequences. Generating a data set that considers all the possible triplets in all the positions, describes the variability among all the possible sequences. For an assembly with n number of welds, there are n possible positions in the sequence. Consider $x_l = [x_1, x_2, x_3, \dots, x_n]$, each element can have the value 1 to n . Now if all the combination of the first three triplets is $q = [x_{1:n}, x_{2:n}, x_{3:n}]$, this matrix can be cycled through x_1 to x_n . $q_1 = [q, \dots, x_n]$,

$q_2 = [x_1, x_2, x_3, q, \dots, x_n]$, $q_{n/3} = [\dots, x_{n-3}, q]$. To clarify this point further, consider an assembly with six weld points, the matrix q is expressed as follows:

squared projection error form n dimensions to k , on the total variation in the data.

$$q = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 1 & 2 & 5 \\ \vdots & \vdots & \vdots \\ 2 & 1 & 3 \\ 2 & 1 & 4 \\ 2 & 1 & 5 \\ \vdots & \vdots & \vdots \\ 6 & 5 & 2 \\ 6 & 5 & 3 \\ 6 & 5 & 4 \end{bmatrix} \quad q_1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 2 & 4 & 3 & 5 & 6 \\ 1 & 2 & 5 & 3 & 4 & 6 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2 & 1 & 3 & 4 & 5 & 6 \\ 2 & 1 & 4 & 3 & 5 & 6 \\ 2 & 1 & 5 & 3 & 4 & 6 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 6 & 5 & 2 & 1 & 3 & 4 \\ 6 & 5 & 3 & 1 & 2 & 4 \\ 6 & 5 & 4 & 1 & 2 & 3 \end{bmatrix} \quad q_2 = \begin{bmatrix} 4 & 5 & 6 & 1 & 2 & 3 \\ 3 & 5 & 6 & 1 & 2 & 4 \\ 3 & 4 & 6 & 1 & 2 & 5 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 4 & 5 & 6 & 2 & 1 & 3 \\ 3 & 5 & 6 & 2 & 1 & 4 \\ 3 & 4 & 6 & 2 & 1 & 5 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 3 & 4 & 6 & 5 & 2 \\ 1 & 2 & 4 & 6 & 5 & 3 \\ 1 & 2 & 3 & 6 & 5 & 4 \end{bmatrix}$$

The matrix $[q_1; q_2; \dots; q_{n/3}]_{m \times n}$ is evaluated using the SVM models created, and the response feature is added as a column to the matrix. The creation of the cyclic matrices for non-divisible dimensions can be performed by considering all the combinations of the last three elements, $[n - 2, n - 1, n]$.

Defining the data matrix as Q , the PCA is performed on such a matrix.

After normalization of the Q matrix, and ensuring zero mean for each feature x_1 to x_{n+1} , the covariance matrix, Σ_Q , of the matrix Q , with m input vectors of $n + 1$ feature, where x_1 to x_n are the sequence of welds, and x_{n+1} is \hat{y} , is established:

$$Q = \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 & \dots & x_{n+1}^1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_1^m & x_2^m & x_3^m & \dots & x_{n+1}^m \end{bmatrix},$$

$$\Sigma_Q = \frac{1}{m} \sum_{i=1}^{n+1} (x_i)(x_i)^T \tag{7}$$

The eigenvectors of the covariance matrix Σ_Q represents the singular value decomposition of the data set. Defining $U_{n \times n}$ as the matrix of the eigenvectors for each feature, the approximation of the data, Z from $x \in \mathbb{R}^n$ to $x \in \mathbb{R}^k$ is:

$$Z = U_{(n+1) \times k}^T X_{k \times (n+1)} \tag{8}$$

To identify which k number of features represents the variability in the data, a limit J can be set for the ratio of average

$$\frac{\frac{1}{m} \sum_{i=1}^m \|x_i - z_i\|^2}{\frac{1}{m} \sum_{i=1}^m \|x_i\|^2} \leq J, \quad x_i \in X, z_i \in Z \tag{9}$$

The minimum k that satisfies (9), to retain the variability in the data corresponds to the number of the critical weld points in the assembly.

Reference assemblies

The proposed method is evaluated on three automotive sheet metal assemblies. The assemblies have been modeled in the CAT-tool RD&T [15] according to the assembly process, for compliant variation simulation. The positioning system for the individual components and assembly inspection is defined. Scanned data of the components are used as part deviation input. Contact modeling is performed in all the assemblies. The welds are introduced to the models. The material properties of the assemblies are also included. All the assemblies are steel with different sheet thicknesses. The elastic modulus for all the parts has been 210 GPa. The details of the assemblies are provided below.

Assembly I

This assembly is constructed with two parts. Seven spot welds join the parts together. The sheet thicknesses for Part 1 is 1.6 mm (mm) and Part 2, 1.2 mm. Contact modeling is performed using 154 contact points. The positioning system of the assembly is shown in Fig. 5 with the arrows. The spot weld locations and their numbering are also visualized.

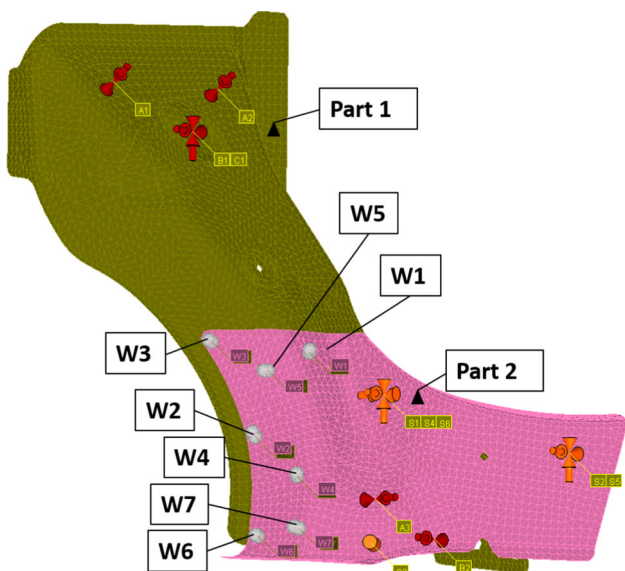


Fig. 5 Assembly I

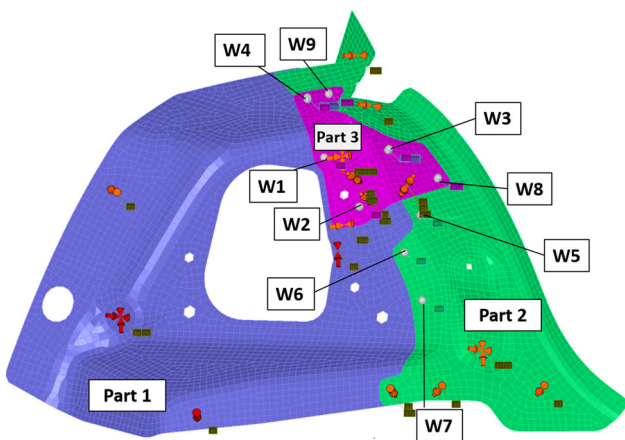


Fig. 6 Assembly II

Assembly II

In this assembly, three parts are joined together with nine weld points. The contact modeling is performed using 194 contact points to avoid the penetration of the parts. The sheet thicknesses for all the parts are 0.8 mm. The positioning system and the welds points are visualized in Fig. 6.

Assembly III

This assembly is a larger model, with three parts. Ten weld points join these three parts together. The thickness for Part 1 is 2.3 mm, Part 2, 1.6 mm, and Part 3, 1.5 mm. The model is prepared with 813 contact points. The positioning system and the weld points are shown in Fig. 7.

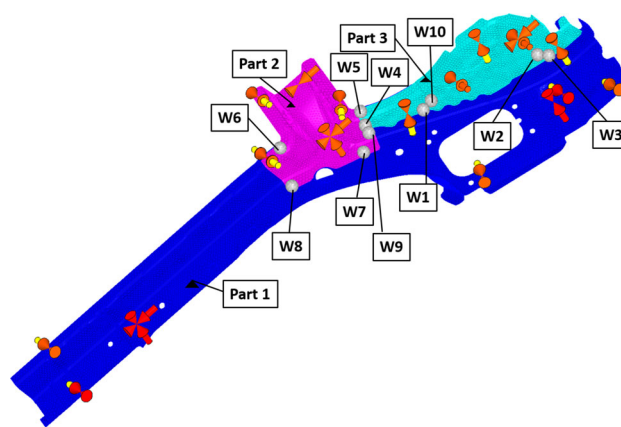


Fig. 7 Assembly III

Method evaluation

The proposed method is applied to the three assemblies, and the critical weld points for each assembly are identified. The results are presented in this section. The summary of the proposed method for critical weld point identification is as follows:

- Build an initial surrogate model, evaluate the sequences, and divide them into intervals, Sect. 2.1.
- Generate a sample for the necessary intervals, and build SVM models for those intervals, Sect. 2.2.
- Generate the sequence sample set, Q , for PCA. Estimate assembly deviation of the sample, using the SVM models, and apply PCA to identify the critical weld points 2.3.

SVM

Applying the method on the assemblies results in models of sequence intervals for each assembly. The proposed SVM approach is set up in MATLAB, and parallel Bayesian optimization is performed for each interval model for optimal SVM parameter settings. With the generated SVM models, assembly deviations for each sequence is estimated. The estimates of the assembly deviation for Assembly I is presented in Fig. 8. For this Assembly, since the problem is smaller compared to the other assemblies, all the intervals are modeled for a better visualization. The generated SVM captures higher variability due to the sample updates in each interval, see Sect. 2.2.

These models are now used to estimate the required sequences for the PCA analysis.

For Assembly II, the sequencing problem is much larger, with 362,880 alternatives. The initial model for all the sequences relies on the mean numbers of the achieved observation in each interval. In total, 421 intervals are identified. Since covering all the intervals is time-consuming for prob-

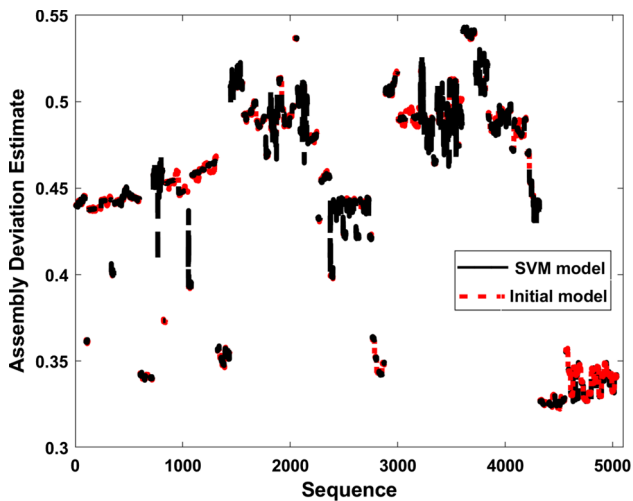


Fig. 8 Assembly I: the SVM models for each interval compared to the initial model

lems with larger sizes, the intervals with minimum mean assembly response from the initial model, and intervals necessary for PCA analysis are built. Figure 9a shows the mean assembly deviation estimation for each interval with the initial model. The interval with the minimum mean assembly deviation is shown as well. Figure 9b shows the SVM model estimates for this interval and compares it to the initial model response. As can be seen, the initial model is dependent on the mean value achieved from the initial sample evaluations, while the interval specific SVM model captures the variability between the sequences more clearly. The SVM models helps to capture the variability between the sequences in the interval with the minimum mean assembly deviation.

Assembly III, includes 10 weld points with more than three million possible permutations. The initial model divides the data to 8761 intervals. Among these intervals, 690 intervals have the lowest mean assembly deviation. By building the SVM models for these candidates, the interval resulting in the minimum mean assembly deviation is identified. Figure 10a shows the mean assembly deviation for each interval. The interval resulting in the minimum mean is updated with the SVM model. Figure 10b shows the updated model compared to the initial model estimate. The initial model underfits the responses in the interval, while the updated SVM model is capturing the variability between the sequences in the interval.

With these information the PCA analysis is performed and the critical weld points are identified.

PCA

The sequences required for PCA analysis for each assembly is generated and estimated using the SVM models, Sect. 2.3. The PCA analysis is set up in MATLAB. The cumulative sum

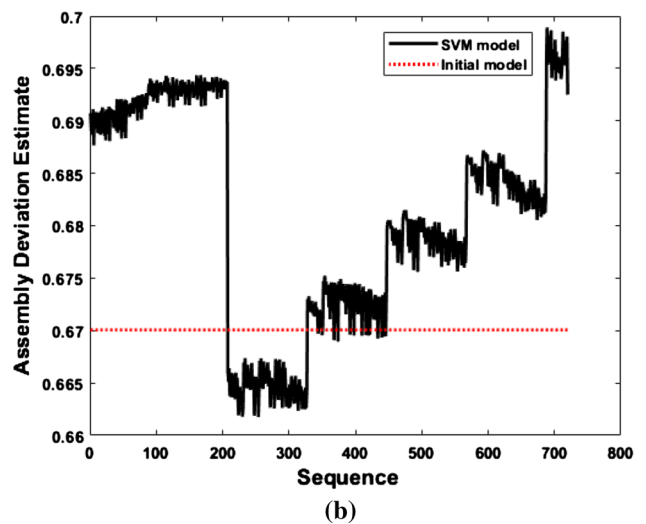
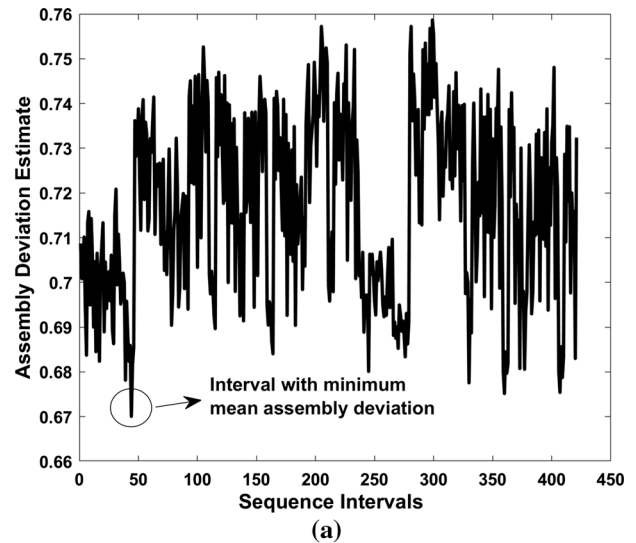


Fig. 9 Assembly II **a** the initial surrogate model mean assembly deviation for each interval **b** the SVM model for the interval with minimum mean assembly deviation compared to the initial model

of the explained variances by each component for Assembly I is shown in Fig. 11a.

In Assembly I, the first two sequence elements describe approximately 60% of the variability between the sequences. Considering the range of the differences between the sequences which is 0.22 mm, optimizing the first two elements of the sequence reduces 0.13 mm of the assembly geometrical deviation. The percentage of the geometrical deviation that needs to be improved depends on the requirements of the assembly. For these types of assemblies, the requirements are within one-tenth of a millimeter. Therefore, the J value, Sect. 2.3, can be set to secure the needed requirements. For this assembly, the sequence elements three to seven, reduces only 0.088 mm of the total variability range.

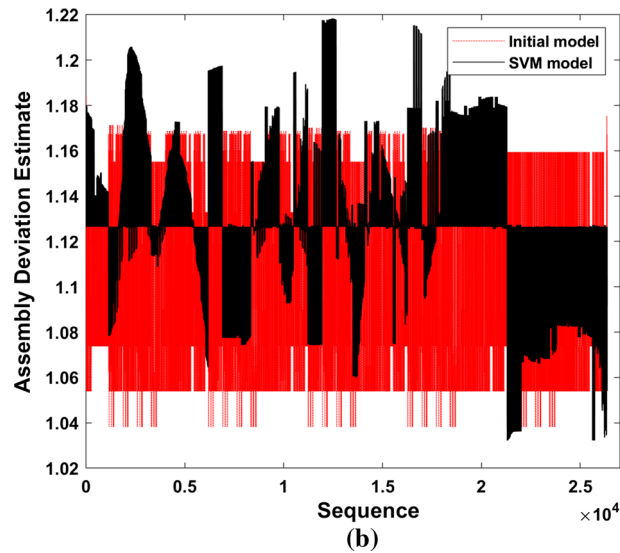
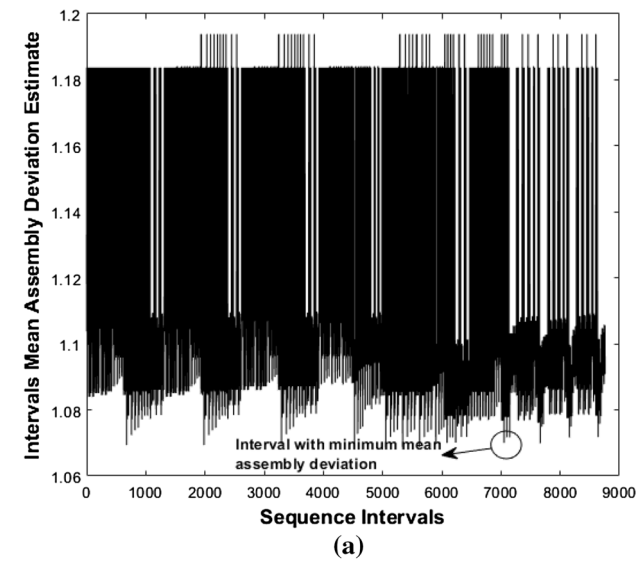


Fig. 10 Assembly III **a** the initial surrogate model mean assembly deviation for each interval **b** the SVM model for the interval with minimum mean assembly deviation compared to the initial model

With this analysis, the number of the critical weld points, referred to as principle weld component, P_c^w , in this assembly, is set to two.

To identify the two welds, the interval with minimum mean assembly deviation is analyzed. In this assembly, interval 48 results in the minimum average assembly deviation. Looking at the sequences in this interval, there are 240 sequences, all are starting with the pairs (w_1, w_6) , and (w_1, w_7) . The optimum sequence for this assembly is, $[w_1, w_6, w_5, w_2, w_3, w_4, w_7]$. The most critical weld point, therefore is chosen to be (w_1, w_6) . The estimate of the above sequence is $\hat{y} = 0.3236$ mm with the exact same actual value, $y = 0.3236$. The summary of the results is presented in Table 2.

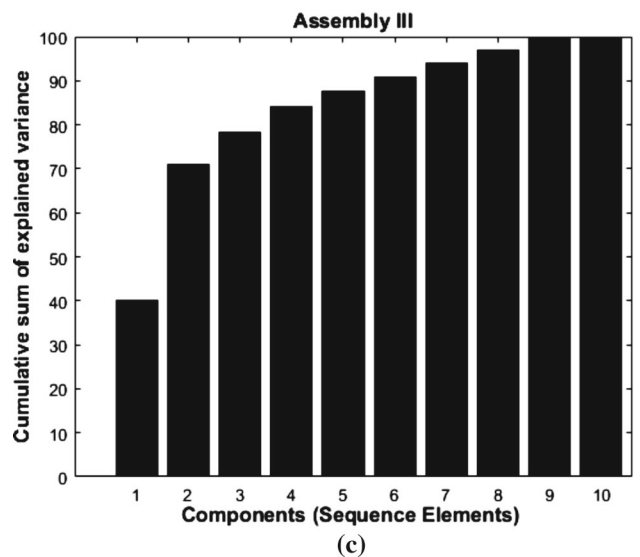
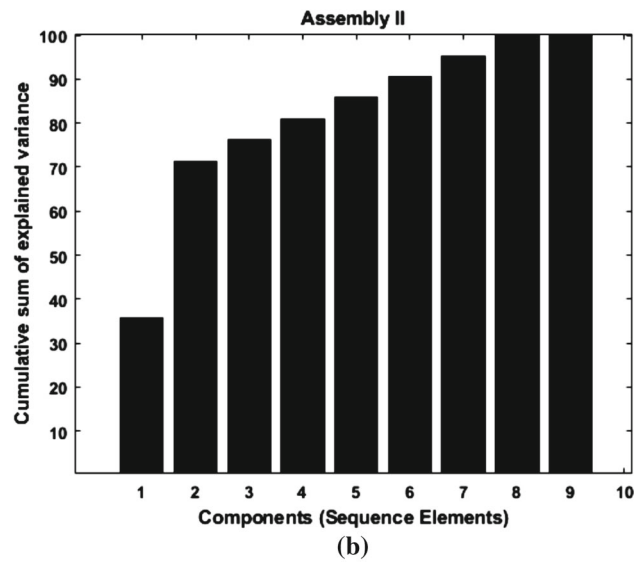
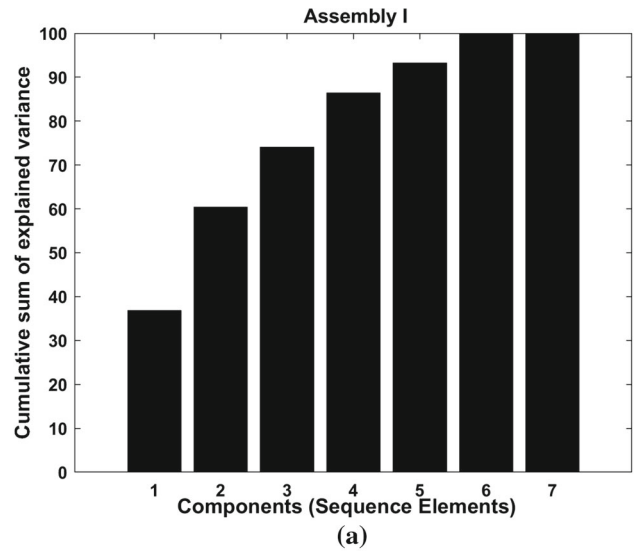


Fig. 11 The cumulative sum of the explained variances of the components for **a** Assembly I **b** Assembly II **c** Assembly III

Table 2 Summary of the assembly evaluation results

Assembly	I	II	III
J	60.35	71.29	78.22
P_c^w	2	3	3
Critical welds	(w_1, w_6)	(w_9, w_1, w_4)	(w_3, w_6, w_5)
Sequence	{1, 6, 5, 2, 3, 4, 7}	{9, 1, 4, 8, 6, 2, 3, 7, 5}	{3, 6, 5, 10, 9, 8, 4, 7, 2, 1}
\hat{y}	0.3236	0.6617	1.0322
y	0.3236	0.6658	1.0294

With the same analysis in Assembly II, elements 4 to 9 of the sequence cover 0.02 mm of the assembly. Considering that there are three parts included in the assembly, by three weld points 71.29% of the assembly deviation is covered. The cumulative sum of the explained variances by each sequence element, for this assembly, is shown in Fig. 11b. The interval with the minimum mean assembly deviation is composed of 720 sequences with (w_9, w_1, w_4) , and (w_9, w_1, w_5) , as the initial three elements. The minimum complete sequence in this interval is composed of (w_9, w_1, w_4) . Therefore, this triplet is chosen as the critical weld points for this assembly.

Assembly III is composed of three parts. Considering this aspect, it can be observed that the initial three weld points in the sequence retain 78.22% of the variability between the sequences, Fig. 11c. According to the interval with the minimum assembly deviation, there are totally 26,368 sequences in the interval, where the sequences starts with the pair (w_3, w_7) or (w_3, w_6) . The complete sequence results in the minimum assembly deviation estimate of 1.0322 mm. The observed value of this sequence is 1.0294 mm. The summary of the results are presented in Table 2.

Accuracy comparison

The proposed method intends to evaluate the outcome of each sequence accurately for PCA analysis and critical joint identification. Previous studies have focused on complete sequence identification using the meta-heuristic algorithms. The advantages of the non-linear modeling over the meta-heuristic algorithms, such as GA for complete sequence optimization, are explained and discussed in previous research [22]. Other integer programming methods for joining sequence optimization, require recursive cost matrices (assembly deviation with FEM) for each element of the sequence. Based on this aspect, a stepwise algorithm is introduced for complete sequence optimization [23]. Since the focus on this paper is on sequence evaluation and identification of the critical weld points, the previous critical weld point identification method, presented in [20] is compared to the proposed approach. Rules and guidelines are used to determine the location of a predetermined number of critical weld points in the mentioned study. These rules are; the dis-

tance of the weld point to the positioning system, weld gap before welding, and weld gap sensitivity. The sequence of the selected critical weld points is determined by the meta-heuristic algorithm GA. The results have shown that weld gap sensitivity is more accurate compared to the rest of the rules. This approach is compared to the presented critical joint identification and sequence determination approach. Table 3 presents the comparison between the proposed non-linear modeling approach with SVM and critical weld point identification using PCA, and the previously introduced weld gap sensitivity analysis for critical weld points identification. One advantage of the proposed approach is identifying the number of critical weld points in a sequence. In contrast, the number of critical weld points in the weld gap sensitivity approach is predefined. In all the assemblies the proposed method outperforms the weld gap sensitivity approach with respect to proposing an accurate number of joints, location and their sequence which results in minimum assembly geometrical deviation.

Conclusion

Joining sequence determination is a time consuming combinatorial problem to solve. While some of the elements of the sequence have a dominant role in the geometrical outcome of the assembly, there has not been a prescriptive study identifying the number, location among a set of alternatives, and the sequence of such weld points, which can be related to the geometry weld points in industrial applications. In this paper, based on the principles of machine intelligence in operation research, with a time-efficient approach, SVM models are built for estimating the outcome of the assembly with specific sequences in combination with the CAT-tools. To identify the number of the critical weld points, a PCA analysis approach is proposed to evaluate the variability of each element of the sequence. The SVM model is then used to determine the sequence of such critical weld points, acting as initial weld points set in the sequence. Three reference sheet metal assemblies with different sizes are evaluated using the proposed approach. The critical weld points of the assembly is successfully identified, showing to reduce more than

Table 3 Accuracy comparison

Approach		I	Assembly II	III
SVM model and PCA	Number identified	Yes	Yes	Yes
	Number of Critical welds	2	3	3
	Welds	(w_1, w_6)	(w_9, w_1, w_4)	(w_3, w_6, w_5)
	y	0.3236	0.6658	1.0294
	Weld gap sensitivity	No	No	No
Weld gap sensitivity	Number identified	No	No	No
	Number of Critical welds	Predefined	Predefined	Predefined
	Welds	4	5	5
	Welds	(w_4, w_2, w_3, w_5)	$(w_1, w_5, w_7, w_6, w_4)$	$(w_6, w_{10}, w_7, w_1, w_5)$
	Sequence	{4, 2, 3, 5, 1, 6, 7}	{1, 5, 7, 6, 4, 2, 3, 8, 9}	{1, 7, 6, 10, 5, 8, 9, 3, 2, 4}
y	0.4323	0.7207	1.0928	

60% of the variation between the sequences. The sequence of the critical weld points are identified using the applied SVM model, and a suggestion for the complete sequence of joining is also proposed.

The future research includes expanding the approach with permutation learning for environments with independent part input models, or batch analysis, where for each assembly, there are multiple set of sequence intervals. Within a batch of incoming parts, the part deviations are expected to be correlated to each other. Among the batches, these part deviation may not be correlated, which means iterations of the proposed approach are required. Through these iterations, permutation learning can be deployed to analyze the trend of the changes in the joining sequences among the batches. The proposed approach can potentially be applied to other sequenced operations in the assembly process such as clamping.

Acknowledgements The work was carried out in collaboration with Wingquist Laboratory and the Area of Advance Production at Chalmers within the project Smart Assembly 4.0, financed by The Swedish Foundation for Strategic Research. The support is gratefully acknowledged.

Funding Open access funding provided by Chalmers University of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abdi, H., & Williams, L. J. (2010). Principal component analysis. *WIREs Computational Statistics*, 2(4), 433–459. <https://doi.org/10.1002/wics.101>.
- Bai, Y., Sun, Z., Zeng, B., Long, J., Li, L., de Oliveira, J. V., et al. (2019). A comparison of dimension reduction techniques for support vector machine modeling of multi-parameter manufacturing quality prediction. *Journal of Intelligent Manufacturing*, 30(5), 2245–2256.
- Fukuda, S., & Yoshikawa, K. (1990). Determination of welding sequence: A neural net approach. *Engineering Analysis with Boundary Elements*, 7(2), 78–82. [https://doi.org/10.1016/0955-7997\(90\)90024-4](https://doi.org/10.1016/0955-7997(90)90024-4).
- García, V., Sánchez, J. S., Rodríguez-Picón, L. A., Méndez-González, L. C., & de Jesús, Ochoa-Domínguez H. (2019). Using regression models for predicting the product quality in a tubing extrusion process. *Journal of Intelligent Manufacturing*, 30(6), 2535–2544.
- Huang, M.-W., Hsieh, C. C., & Arora, J. S. (1997). A genetic algorithm for sequencing type problems in engineering design. *International Journal for Numerical Methods in Engineering*, 40(17), 3105–3115.
- Karsai, G., Andersen, K., Cook, G. E., & Barnett, R. J. (1992). Neural network methods for the modeling and control of welding processes. *Journal of Intelligent Manufacturing*, 3(4), 229–235.
- Kuo, C.-F. J., Tung, C.-P., & Weng, W.-H. (2019). Applying the support vector machine with optimal parameter design into an automatic inspection system for classifying micro-defects on surfaces of light-emitting diode chips. *Journal of Intelligent Manufacturing*, 30(2), 727–741.
- Liao, Y. G. (2005). Optimal design of weld pattern in sheet metal assembly based on a genetic algorithm. *The International Journal of Advanced Manufacturing Technology*, 26(5), 512–516. <https://doi.org/10.1007/s00170-003-2003-5>.
- Lindau, B., Lorin, S., Lindkvist, L., & Söderberg, R. (2016). Efficient contact modeling in nonrigid variation simulation. *Journal of Computing and Information Science in Engineering*, 16(1), 11002–11007. <https://doi.org/10.1115/1.4032077>.
- Liu, G., Gao, X., You, D., & Zhang, N. (2019). Prediction of high power laser welding status based on pca and svm classification of multiple sensors. *Journal of Intelligent Manufacturing*, 30(2), 821–832.

11. Liu, S. C., & Hu, S. J. (1997). Variation simulation for deformable sheet metal assemblies using finite element methods. *Journal of Manufacturing Science and Engineering*, 119(3), 368–374. <https://doi.org/10.1115/1.2831115>.
12. Lorin, S., Lindau, B., Lindkvist, L., & Söderberg, R. (2018a). Efficient compliant variation simulation of spot-welded assemblies. *Journal of Computing and Information Science in Engineering*. <https://doi.org/10.1115/1.4041706.011007>.
13. Lorin, S., Lindau, B., Sadeghi Tabar, R., Lindkvist L., Wärmeffjord, K., & Söderberg, R. (2018b). Efficient variation simulation of spot-welded assemblies. In *ASME International Mechanical Engineering Congress and Exposition, vol 2: Advanced Manufacturing* (p. V002T02A110). Pittsburgh, Pennsylvania, USA, November 9–15. <https://doi.org/10.1115/IMECE2018-87454>
14. Phadke, M. S. (1995). *Quality engineering using robust design* (1st ed.). New York: Prentice Hall PTR.
15. RD&T Technology AB. (2017). RD&T Software Manual.
16. Söderberg, R., Lindkvist, L., Wärmeffjord, K., & Carlson, J. S. (2016). Virtual geometry assurance process and toolbox. *Procedia CIRP, 14th CIRP CAT 2016—CIRP Conference on Computer Aided Tolerancing* (vol. 43, pp. 3–12) Gothenburg, Sweden, 18–20 May. <https://doi.org/10.1016/j.procir.2016.02.043>
17. Segeborn, J., Carlson, J. S., Wärmeffjord, K., & Söderberg, R. (2011). Evaluating genetic algorithms on welding sequence optimization with respect to dimensional variation and cycle time. In *International Design Engineering Technical Conferences and Computers and Information in Engineering, vol 5: 37th Design Automation Conference, Parts A and B* (pp. 697–704). Washington, DC, USA. August 28–31. <https://doi.org/10.1115/DETC2011-48393>
18. Söderberg, R., Wärmeffjord, K., Carlson, J. S., & Lindkvist, L. (2017). Toward a digital twin for real-time geometry assurance in individualized production. *CIRP Annals*, 66(1), 137–140. <https://doi.org/10.1016/j.cirp.2017.04.038>.
19. Tabar, R. S., Wärmeffjord, K., & Söderberg, R. (2018) Evaluating evolutionary algorithms on spot welding sequence optimization with respect to geometrical variation. *Procedia CIRP, The 15th CIRP Conference on Computer Aided Tolerancing* (vol. 75, pp. 421–426) Milan, Italy, 11–13 June. <https://doi.org/10.1016/j.procir.2018.04.061>
20. Tabar, R. S., Wärmeffjord, K., & Söderberg, R. (2019a). A method for identification and sequence optimisation of geometry spot welds in a digital twin context. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 233(16), 5610–5621. <https://doi.org/10.1177/0954406219854466>.
21. Tabar, R. S., Wärmeffjord, K., Söderberg, R., & Lindkvist, L. (2019b). A novel rule-based method for individualized spot welding sequence optimization with respect to geometrical quality. *Journal of Manufacturing Science and Engineering*. <https://doi.org/10.1115/1.4044254.111013>.
22. Tabar, R. S., Wärmeffjord, K., & Söderberg, R. (2020a). A new surrogate model-based method for individualized spot welding sequence optimization with respect to geometrical quality. *The International Journal of Advanced Manufacturing Technology*, 106(5), 2333–2346.
23. Tabar, R. S., Wärmeffjord, K., & Söderberg, R. (2020b). Rapid sequence optimization of spot welds for improved geometrical quality using a novel step-wise algorithm. *Engineering Optimization*. <https://doi.org/10.1080/0305215X.2020.1757090>.
24. Tabar, R. S., Wärmeffjord, K., Söderberg, R., & Lindkvist, L. (2020c). Efficient spot welding sequence optimization in a geometry assurance digital twin. *Journal of Mechanical Design*. <https://doi.org/10.1115/1.4046436.102001>.
25. Tian, L., & Luo, Y. (2020). A study on the prediction of inherent deformation in fillet-welded joint using support vector machine and genetic optimization algorithm. *Journal of Intelligent Manufacturing*, 31(3), 575–596.
26. Vapnik, V. N. (2000). *The nature of statistical learning theory. Statistics for engineering and information science*. Berlin: Springer.
27. Wärmeffjord, K., Söderberg, R., & Lindkvist, L. (2010). Strategies for optimization of spot welding sequence with respect to geometrical variation in sheet metal assemblies. In *ASME International Mechanical Engineering Congress and Exposition, vol 3: Design and Manufacturing, Parts A and B* (pp. 569–577). Vancouver, British Columbia, Canada. November 12–18. <https://doi.org/10.1115/IMECE2010-38471>
28. Wells, L. J., Megahed, F. M., Camelio, J. A., & Woodall, W. H. (2012). A framework for variation visualization and understanding in complex manufacturing systems. *Journal of Intelligent Manufacturing*, 23(5), 2025–2036.
29. Wärmeffjord, K., Söderberg, R., & Lindkvist, L. (2010). Variation simulation of spot welding sequence for sheet metal assemblies. In *Proceedings of NordDesign 2010 International Conference on Methods and Tools for Product and Production Development* (vol. 2, pp. 519–528)
30. Zhang, F., & Zhou, T. (2019). Process parameter optimization for laser-magnetic welding based on a sample-sorted support vector regression. *Journal of Intelligent Manufacturing*, 30(5), 2217–2230.
31. Zhao, D., Wang, Y., Sheng, S., & Lin, Z. (2014). Multi-objective optimal design of small scale resistance spot welding process with principal component analysis and response surface methodology. *Journal of Intelligent Manufacturing*, 25(6), 1335–1348.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.