



Quality assessment and fast geometry prediction in paperboard forming

Klara Liesegang¹ · Patrick Link¹ · Georg Steinert² · Lena Berthold² · Steffen Ihlenfeldt^{1,3}

Received: 7 November 2025 / Accepted: 7 February 2026
© The Author(s) 2026

Abstract

Paperboard forming is a productive process that meets the increasing demands for recyclability in packaging. However, it remains underutilized in mass production due to open challenges, for instance volatile material behavior and control challenges. Moreover, wrinkles are inevitable during paperboard forming. To produce numerous small wrinkles that do not compromise part quality and can improve its stability, the process must be set up appropriately. Achieving this efficiently requires minimal testing, which is complicated by rapid variations in material properties due to external factors. One solution to improve process design and stabilize a running process is integrating machine learning (ML) models, which represent the forming process, into a stroke-to-stroke control. Therefore, this study presents an approach consisting of Finite Element (FE) simulations to generate data, an ML model trained on them, and quality assessment for the predicted parts with an optimization strategy to determine process parameters. FE simulations were built to represent the forming process and ML models were trained on their results to establish a fast correlation between material properties, press settings, and the resulting part geometry, including wrinkles. An automated quality assessment using Fast Fourier Transform quantifies the quality of the part in terms of wrinkles. Results show that the ML model is integrated into an optimization strategy to propose enhanced process parameters if quality deviations occur. The ML model accurately represents FE simulation results and collaborates with quality assessment and optimization fast enough to ensure high part quality in a forming process with cycles of up to 60 strokes per minute.

Keywords Machine Learning · Forming · Paperboard · FE Simulation

1 Introduction

The packaging industry is making great efforts to transition towards biobased packaging solutions and reduce plastic materials. To achieve this, cost-effective and efficient processes must be developed and established that meet the special requirements of biobased materials. One example of a process that has already been successfully established is the press forming of paperboard for the production of trays and

caps. However, this process can only be used to produce parts with limited forming heights. The situation is different with deep drawing, which is a forming process where a blank is radially drawn into a forming die [1]. This can be done with forming tools in a press, active media or active energy. In this paper deep drawing always refers to the process with forming tools, that are die, punch and blank holder. The deep drawing tools can be heated to increase the ductility of the material and reduce the forming forces. Deep drawing allows for deep shapes with nearly perpendicular walls, and good stability. It can therefore extend the variety of shape and applications for 3D-forming of paperboard.

Although the deep drawing process of paper board is scientifically well understood, it remains under-utilized in mass production due to control challenges and is so far only used for low quality products. The reason for the control challenges lies in the material properties of the paperboard. On the one hand, these are dependent on the ambient conditions (especially humidity), which means that the part

✉ Klara Liesegang
klara.liesegang@iwu.fraunhofer.de

¹ Fraunhofer Institute for Machine Tools and Forming Technology IWU, 09126 Chemnitz, Germany

² Institute of Natural Materials Technology, TUD Dresden University of Technology, 01062 Dresden, Germany

³ Institute of Mechatronic Engineering, TUD Dresden University of Technology, 01062 Dresden, Germany

quality fluctuates even if the process parameters remain constant. Secondly, paper is a highly anisotropic material with poor flow properties. This means that the mobility of the individual fibers inside the fiber network of paperboard is very limited and consequently in-plane material redistribution during forming is only marginal. The geometric conditions of the drawing process result in excess material along the wall and flange region which accumulates into wrinkles. The wrinkles strongly contribute to compression in the drawing gap. Compression is necessary to consolidate the desired shape by forming new bonds in the fiber network [2]. A fine and uniform distribution of wrinkles around the circumference of a part gives good stability and geometric accuracy due to uniform compression. In addition, a finer distribution gives a more pleasing appearance than a few large wrinkles. Examples of inadequate and suitable wrinkle distributions are shown in Fig. 1.

Accordingly, the process must be set up in such a way that intact parts with a highest possible number of wrinkle are produced even if the initial moisture content of the material fluctuates. Both stabilization and optimization of the process can be achieved experimentally only with great effort and do not necessarily lead to high part quality. At the same time, a lot of waste is produced during the process design and also during the running process, which is critical in terms of sustainability. A promising approach to countering these problems is the in-line use of real-time data-driven models to predict part quality. However, several components must be combined in order to be able to use such data-driven models.

This paper answers the scientific question of how to set up a digital twin for simulating a deep drawing process of paperboard that incorporates an advanced ML model and optimization algorithm. It is used for enabling a stroke-to-stroke control for an industry-oriented demonstrator system with a stroke rate of up to 60 strokes per minute. The development and implementation of all required methods

are presented. This includes finite element (FE) simulations to create a database, the design and training of data-driven models, an algorithm to evaluate part quality (especially the wrinkle distribution), and the integration of data-driven models into an optimization strategy to recommend adapted process parameters for better part quality.

Mapping the resulting wrinkle distribution with FE simulations is a major numerical challenge. A variety of approaches have been documented in the literature, yet none have achieved a comprehensive mapping of the wrinkles that occur during the deep drawing process. As part of this study, an FE simulation is developed, that maps the wrinkle distribution accurate enough to serve as a solid foundation for the overall routine. As these FE simulations are not sufficient for in-line predictions due to their computing time of up to several hours, they are used to generate a database, on which data-driven machine learning (ML) models are trained. These models identify the relevant relationships and can predict the resulting part quality in real-time based on the process parameters.

This paper is structured as follows: first, the relevant literature and the state of the art are presented in the following section. As outlined in Section 3, the forming process setup, and boundary conditions are described, along with the developed method. It consists of three components: FE simulation, ML model, and part quality assessment with optimization. The results of each of these steps are presented in Section 4. Finally, the results are discussed in Section 5 and Section 6 summarizes the research and gives an outlook on future applications.

2 Related work

In paperboard forming the formability of the paper limits the geometry of the parts. In order to produce advanced three-dimensional deep-drawn parts, it is therefore advantageous to specifically extend the forming limits of the paperboard [3]. One such approach is the moistening of the paperboard blank before or during forming which can be beneficial for the forming process in several ways. For example, Franke et al. [4] shows that moistening the parts with steam during deep drawing improved wrinkle formation and reduced shape deviations. In [5], the moistening of blanks before deep drawing was found to reduce the necessary forming forces. There are various other process parameters that can be modified, although only a small number of these have a significant impact on part quality. Wallmeier et al. [6] claimed blank holder force, the drawing gap, tool surface and punch velocity as influencing factors for the occurrence of cracks in deep drawing. Müller et al. [7] later

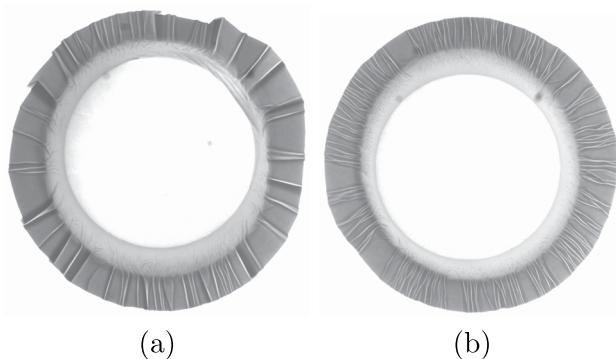


Fig. 1 Top view on parts after deep drawing (a) Uneven surface due to few, large wrinkles (b) Even surface due to fine wrinkle distribution

found blank holder forces, material moisture content and die temperature to influence part quality in deep drawing.

Paperboard forming can be modeled with FE simulations. Integral components of the modeling are the material model and the representation of the wrinkles. In [8] an explicit FE model is developed with a material model to represent anisotropy and plasticity of paperboard. This model is not capable of mapping the occurring wrinkles. The material model used in Linvill et al. [9] is extended by a hinge model. In this approach, wrinkle distribution is made visible by coloring the elements that have wrinkled and show a compressive strain above a critical value. This makes it possible to determine the number of visible wrinkles after forming, but no statements can be made about the shape and height of the wrinkles. An orthotropic plastic material model is used in Awais et al. [10]. Here, wrinkles are visible in the visualization of the simulation results, but kinks were specifically added to the paperboard before forming, which determine the later position of the wrinkles. For the FE solver LS-Dyna in this study, there are two material models that have been explicitly implemented for paperboard: *MAT_274 (*MAT_PAPER) [11] and *MAT_279 (*MAT_COHESIVE_PAPER) [13]. In [14] *MAT_274 is calibrated and shows good results for the comparison with a compression test of a paperboard part. However, the disadvantage of this model is the complex calibration, in which at least 26 parameters must be determined in 8 tests.

Depending on the computing power, FE models can take several hours to compute, which prevents them from being directly integrated into a stroke-to-stroke control. Instead they can be replaced by reduced models or real-time capable data-driven models which are trained on their results. Tercan and Meisen [15] states that for the present use case mainly supervised learning methods are useful. This includes ML methods like support vector machine and fully connected artificial neural networks as well as more complex models like convolutional neural networks (CNNs), long short term memory networks (LSTMs) as a subset of recurrent neural networks and transformer networks. In the specific field of sheet metal forming processes two of these approaches dominate. Ståhl et al. [16] uses LSTMs for modeling dependencies between the part geometry at different times of the forming process. Meyes et al. [17] presents the use of LSTMs to predict the occurrence of failure for deep drawn car body parts. Several other literature show that also CNNs achieve good results in the prediction of part quality for forming processes [18–21]. Their strength is the consideration of spatial relationships [22]. Therefore the inputs, boundary conditions and the outputs (equaling the results of the FE simulations) are represented as images. In [20], CNNs are trained on the FE simulation data from a forming process

and predict part geometry and thinning depending on variable tool geometry. Such models support the early design phase of a process. However, the models can also contribute to the stabilization and optimization of a running process. For example, Liu et al. [23] uses them to predict the part geometry after spring back depending on the process parameters for sheet metal bending processes. Link et al. [22] uses process parameters of a deep drawing process such as blank holder force, drawing depth and sheet thickness as inputs. In [24], a location-dependent friction coefficient is also taken into account by an CNN incorporated into a digital twin. The outputs in both cases are the complete part geometry and the thinning.

The quality of the parts predicted by the ML models has to be evaluated automatically, accurately and reproducibly in order to implement a stroke-to-stroke control. In this study, the computing time of the algorithm is also limited to less than one second, which results from the intended stroke rate. Since the representation of wrinkles in FE simulation is difficult in the existing literature, there is also no work on wrinkle counting algorithms specifically for simulation data. However, there are numerous approaches for counting wrinkles and assessing the quality of real parts. The approaches that work with image recognition could be transferred to simulation data with reasonable calibration effort. In ([25–28]), parts without a top rim are examined, so the wrinkles must be evaluated in the flange. This requires all-round images or measurements. In [25], an all-round image is created and evaluated using multi-stage image processing. The result is a scalar: the number of wrinkles. The entire process takes 6 to 10 minutes. In [26], the manufactured part is measured using laser triangulation. Subsequently, all local maxima of the topographical height value that exceed a material-specific threshold value, as determined through experimental means in advance, are identified. However, since several local maxima always belong to one wrinkle, a length value is determined in advance that represents the material-specific wrinkle width. Once a local maximum has been found, all others in this range are ignored. The outputs of the method are the number of wrinkles and the standard deviation of the average wrinkle distance. The major disadvantage of the method is that the evaluation of the signal alone takes up to 15 seconds. Added to this is the time-consuming measurement of the part, which takes at least 20 seconds for simple parts with a diameter of up to 110 mm. Müller et al. [27] presents a method in which images of the part are preprocessed (using Fourier filtering, among other things) and then compared with the template of an ideal wrinkle. The cross-correlation is calculated. If this exceeds a threshold value determined empirically in

advance, the area is counted as a wrinkle. The outputs are the number of wrinkles and the standard deviation of the wrinkle distance. This method requires 1.1 seconds for image preprocessing and wrinkle detection. But again the time for capturing the image needs to be added. Leng et al. [28] describes an approach with image acquisition and contour detection, but also requires 20 seconds even for parts with rim.

In order to enable stroke-to-stroke controls to stabilize the process and achieve improved part quality, it is advantageous to integrate the ML model into an optimization strategy. If the ML model predicts insufficient part quality, the optimization algorithm determines a modified input for the ML model to produce high-quality parts. The variables that can be optimized are therefore the inputs of the ML model. As Gad [29] demonstrates, Particle Swarm Optimization (PSO), for example, is suitable for such tasks. Weichert et al. [30] shows that Genetic Algorithms and optimization based on least squares method are used in similar scenarios. Dokeroglu et al. [31] states that metaheuristics are widely used for optimization problems. For small dimension optimization problems grid searches as presented in Menke [32] show an adequate cost-benefit ratio.

3 Method

The aim of this study was to develop a method to enable a robust paperboard production process in preparation for mass production using a digital twin. This method consists of three main steps:

1. Development of a validated FE simulation to generate a database,
2. Training of ML models to predict part quality,
3. Development of a method for automated quality assessment with optimization strategy to determine improved process parameters.

Before proceeding with these steps, this section will provide a comprehensive overview of the investigated process, relevant process parameters, and boundary conditions.

3.1 Process setup

An industry-oriented demonstrator press with automated paperboard handling is used to investigate the deep drawing process of paperboard parts. This press contains several sensors that deliver measurement data for the digital twin and it includes the boundary conditions for part and tool geometry, material used, and process

parameters. The demonstrator tool consists of a fixed die and movable punch and blank holder, see Fig. 2. The demonstrator part is a round paper cup with a diameter of 80 mm and a drawing depth of 40 mm. It is made from a blank with a diameter of 180 mm. The basic quality requirement for parts is that they should not have any cracks. However, even an intact part is not necessarily considered satisfactory because it should also meet certain requirements in terms of its wrinkle distribution and shape. The presence of a high number of uniformly distributed wrinkles is considered favorable, along with a predominantly circular shape and the absence of deformations.

The material used is Trayforma 310 + 15 PE from Stora Enso [33]. Its thickness is 0.446 mm. This paperboard has a Polyethylene (PE) coating on one side and is therefore suitable for water repellent food packaging. As is the case with the majority of papers, the material is highly anisotropic. This phenomenon can be attributed to the industrial paperboard manufacturing process, in which the fibers are predominantly aligned in the machine direction of the paper web.

The demonstrator press has a moistening module prior to deep drawing, which only moisturizes the uncoated side of the paperboard with variable moisturization duration. In addition to the material moisture content, in this study, only the blank holder force and the temperature of the die are varied. Therefore, the die can be heated. In order to adjust the process parameters specifically in response to the fluctuating initial moisture content, the demonstrator press measures the moisture of each blank before forming. For each blank, the digital twin predicts the geometry of the part, that arises when the blank is formed with the process parameters that are currently set on the press. If an intact part with satisfactory wrinkle distribution is predicted, the currently set process parameters are retained. If, on the other hand, material failure or poor wrinkle distribution are predicted, the digital twin determines

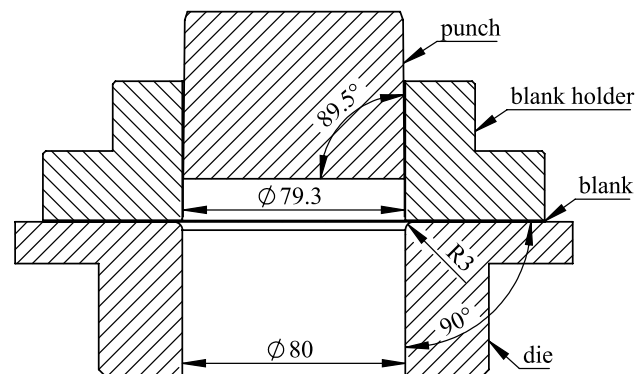


Fig. 2 Cross section of forming tool set, measures are in mm

improved process parameters. These are communicated to the control via an OPC UA interface and the settings on the demonstrator press are adjusted accordingly, even before a defective part could occur.

3.2 FE simulation

An FE model is developed to simulate the deep drawing of paperboard in order to generate a database for the digital twin. Due to the symmetries, only a quarter of the part and the tools is modeled in the following to shorten the calculation times. The thermo-mechanically coupled explicit solver LS-DYNA R12.0.0 is used for simulating the deep drawing process. The tool components are assigned rigid body behavior. The blank geometry is discretized with fully integrated shell elements with five integration points in the thickness direction. In contrast to previous publications, this study does not use a material model developed specifically for paperboard, but the material model *MAT_036 (*MAT_3-PARAMETER_BARLAT) [13], which was originally developed for sheet metals. The rationale behind this choice is that the primary objective of this research is to focus on the digital twin, which is designed to adjust the process parameters for deep drawing if required. This necessitates the execution of an FE simulation that at least qualitatively maps the relationships between the process parameters and the resulting number of wrinkles for blanks where no creases were added in advance. A review of the extant literature reveals that material models exhibit either significant shortcomings with regard to the representation of wrinkles and their geometry [8, 9] or require an inordinate amount of parameterization effort [11, 1312] for the present application, in which the FE simulation is merely a means to an end. *MAT_036 is notable for its consideration of anisotropy and its reduced reliance on parameters when compared to material models specific to paperboard. Consequently, *MAT_036 was calibrated, yielding favorable outcomes in terms of wrinkle mapping. The necessary parameters are Young's modulus E , R -values in 0° -, 45° -, 90° -direction, Poisson's ratio ν and flow curves. In this study, the experimental investigation focused on Young's modulus and flow curves as for them standardized test procedures exist for paperboard. In order to facilitate the digital twin to avert material failure, it is imperative to incorporate a failure criterion into the FE simulation. This aspect is often given a relatively minor role in the majority of publications concerning the simulations for paperboard forming. Only [34] thoroughly addresses the development of a failure criterion for paperboard and proposes a forming limit diagram for this purpose. The disadvantage of this approach is the necessity of three distinct tests (tensile tests with notched samples, Nakazima tests, and bulge tests

with a carrying layer and perforated samples) in addition to ordinary tensile tests. Awais et al. [10] mentions the investigation of failure strains but given the implementation of *MAT_036, the precision of the strains in the simulation is constrained. As a first approach, it is reasonable to employ maximum thinning (the reduction in material thickness as a percentage) as an indicator of failure, as is customary in the context of sheet metal forming. This assertion is not entirely accurate for paperboard, as paper does not undergo a reduction in thickness during the process. Instead, if the material-specific strain limits are exceeded, failure occurs. Nonetheless, the investigation yielded results indicating that maximum thinning accurately identifies material failure for the majority of the parameter space, as determined by the utilization of a validation dataset. Maximum thinning as a failure criterion is notable for its ability to be implemented without the necessity for additional calculations or experiments, thereby ensuring its suitability for demonstrating the comprehensive methodology for establishing a digital twin in paperboard forming.

In principle, all the material parameters are dependent on temperature or moisture. This must be taken into account in their experimental investigation. After evaluating the experimental data, any trends that emerged were taken into account in the modeling, while constants could be assumed for other variables for simplification. The following basic assumptions were made: The material moisture of a single blank is assumed to be constant during deep drawing instead of being modeled time-dependent. This simplification is attributable to the brief duration of the deep drawing process, which has a maximum duration of one second. A targeted reduction of the material moisture is not possible. Additional moisturizing is only possible on the uncoated side of the blank. The material moisture is assumed to be constant over the entire surface of the blank in the x and y direction. To determine the basic properties of paperboard the specimens were chosen with regard to EN ISO 186:2002 [35] to represent average quality. Before the tests, the samples were stored at standard atmosphere for conditioning following EN ISO 187:2022 [36]. The moisture content was determined following EN ISO 287:2017 [37]. Tensile tests were conducted following EN ISO 1924-2:2008 [38]. With the resulting force-strain curves, Young's modulus could be derived. Either temperature (through storage of the test specimen in an oven) or moisture content (through spraying) were varied. Normally, the mechanical properties of the paperboard would always be investigated under well-defined climate conditions due to its hygroscopic properties. The fact that precisely these climate conditions have to be varied here makes the tests more difficult and leads to large standard deviations, which can distort trends in the material properties. For example, there were only very slight

changes in the flow curves as a function of moisture, with no indication of a clear trend. It was therefore decided to include the tensile strength in the FE model as a function of temperature, but not as a function of moisture.

In addition to the material model, the friction model has a substantial influence on the precision with which the FE simulation maps the deep drawing process. In the present study Coulomb friction law is used to model the different contacts between paperboard and tools as well as the self-contact of paperboard. Therefore only the coefficients of friction between the different surfaces have to be determined. It is imperative to note that only one side of the paperboard has undergone PE-coating, necessitating the differentiation between the two sides. Consequently, the determination of the following three coefficients of friction is essential:

- μ_1 : uncoated side against polished stainless steel (die vs. blank)
- μ_2 : PE-coated side against polished stainless steel (punch vs. blank and blank holder vs. blank)
- μ_3 : PE-coated side against itself (self-contact of paperboard)

The parameter μ_1 is moisture- and temperature-dependent. As the PE-coated side of paperboard is not moisturized, μ_2 is not moisture-dependent but only temperature-dependent. μ_3 is modeled independently from moisture and temperature for simplification. To determine the coefficients of friction, frictional tests were completed following the test setup of Lenske et al. [39]. Tests were performed with varying tool temperature and different moisturizing through spraying.

Experimental tests have shown that a continuous decline of blank holder force during deep drawing is advantageous for wrinkle distribution and the prevention of cracks [40]. Only the blank holder force at the beginning of deep drawing is explicitly varied. It decreases linearly piece by piece until it is a constant 500 N from a drawing depth of 30.7 mm. The progression is illustrated in Fig. 3 and calculated using the following equation, where dd is the drawing depth and $z := -26.3 \cdot (F_{bh_0} - 500)/40 + F_{bh_0}$:

$$F_{bh}(dd) := \begin{cases} dd \cdot (F_{bh_0} - 500)/40 + F_{bh_0} & -26.3 \leq dd \leq 0 \\ (26.3 \cdot dd) \cdot (500 - z)/4.4 + 500 & -30.7 < dd < -26.3 \\ 500 & -40 \leq dd \leq -30.7, \end{cases} \quad (1)$$

3.3 ML model

The next important step of the method improved automated paperboard part production using a digital twin is the ML model \mathcal{M} for fast quality predictions. It predicts the wrinkle

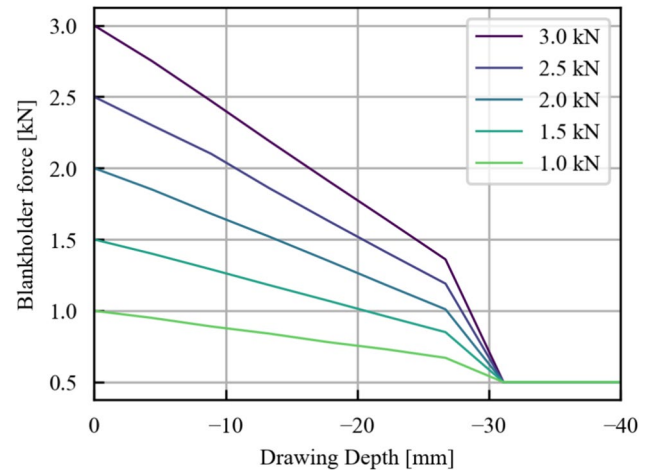


Fig. 3 Pattern for declining blank holder force with different starting blank holder forces

distribution and cup quality W as a function of the process parameters starting blank holder force F_{bh_0} , material moisture m and die temperature T_d :

$$\mathcal{M}(F_{bh_0}, m, T_d) \rightarrow W. \quad (2)$$

As presented in Section 2 literature shows that CNNs are well suited for comparable tasks. They depict spatial relationships well even with relatively few training samples. The latter is important due to the high computing time of the FE simulation. In [24], very good prediction accuracies are achieved for a deep drawing process of metal with a CNN. Therefore, the U-Net architecture with residual squeeze and excitation (RES-SE) blocks presented there is used in this study and the following hyperparameters are adopted without change: reduction ratio equals 8 for all SE blocks, kernel size is 4, stride equals 2, and padding is 1. A rectified linear unit activation function is applied after each convolution and transposed convolution layer and a hyperbolic tangent activation function is used after the last transposed convolution layer.

The input of the model is a 256×256 pixel image with three channels, see Fig. 4. All three inputs represent process parameters on the demonstrator press. Channel 1 contains the starting blank holder force F_{bh_0} and thus also the geometry of the blank holder. Channel 2 contains the expected moisture m of the blank. This is the sum of the moisture measured specifically for the current blank m_m and the amount of additional added moisturizer \hat{m} currently set on the demonstrator press: $m := m_m + \hat{m}$. This representation of the material moisture also incorporates the geometry of the blank into the ML model. Channel 3 contains the temperature of the die T_d and its geometry. All three inputs are constant across the blank

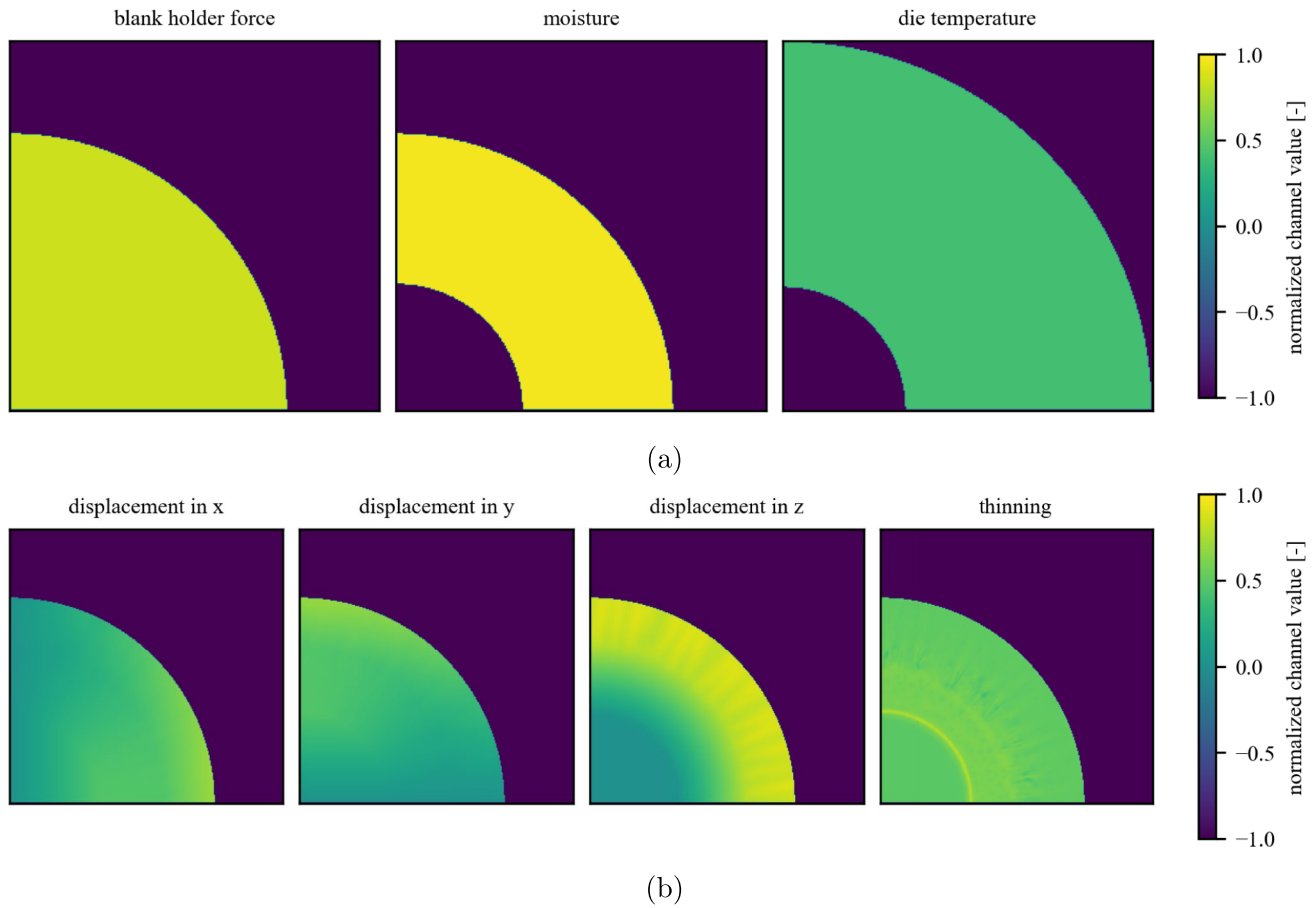


Fig. 4 Visualization of (a) The three input channels incorporating tools and blank geometry and (b) The four output channels

or the tool. Pixels that do not belong to the respective blank or tool are assigned the value -1 . As the blank and tool geometry are already set for the demonstrator press, the position of pixels with value -1 is uniform for all samples.

The output of the ML model is a 256×256 image with four channels. Channels 1 to 3 contain the displacement of each pixel in the x , y and z direction. Channel 4 contains the thinning $s := (t_0 - t)/t_0 \cdot 100\%$ where t_0 denotes the initial blank thickness and t the thickness after deep drawing. All inputs and outputs are linearly interpolated so that they are provided on an equidistant grid with 256×256 pixels and scaled to $[0, 1]$.

The loss function has a significant influence on the training success of ML models and should therefore be adapted to the specific inputs and outputs. For the evaluation the simple loss function mean absolute error (MAE) and modified versions of it were used. The modifications aimed on penalizing deviations in the z -coordinate more as

this might support the mapping of wrinkles. Therefore, there were defined MAE_2 , where the MAE in the z -coordinate is multiplied by 1.2,

$$MAE_2 := MAE_x + MAE_y + MAE_s + 1.2 \cdot MAE_z \quad (3)$$

and MAE_3 , where an exponential function is used to give more weight to deviations in the z -coordinate,

$$MAE_3 := MAE_x + MAE_y + MAE_s + MAE_z + 0.05 \cdot \exp(MAE_z). \quad (4)$$

Hereby, MAE_s describes the error in thinning. Additionally, two more complex losses MixL1 and MixL2 from Link et al. [24] were used. MixL1 describes an equal weighting of an L1 loss and the structural similarity index measure where as MixL2 is an equal weighting of MAE, and structural similarity index measure.

3.4 Quality assessment and optimization

In this section, first a method is presented, that automatically evaluates the outputs of the ML model \mathcal{M} with regard to the quality criteria declared below. It's quality assessment forms the basis for determining whether sufficient parts will be produced or whether optimization of the process parameters is necessary. The optimization strategy is presented in the second part of this section.

The aim of the quality assessment is to evaluate the number of wrinkles and the uniformity of their distribution with one scalar. The components of the quality assessment that process a signal are developed in such a way that they can also be used for the quality assessment of real parts after an additional calibration. With this predictions from the digital twin can be compared with parts produced on the demonstrator press to give a feedback on the prediction power of the digital twin.

3.4.1 Signal acquisition:

The round paper cup used in this paper is made with a rim so that it can be sealed. The wrinkle distribution is analyzed as centrally as possible on the rim. For this purpose, the radius $r = 53$ is set and the z coordinate value of all k pixels located on this circle are used. The corresponding pixels are colored purple in the Fig. 5. Note that the values after rescaling to the original value range are used. The z-coordinates form the discrete signal to be examined, hereafter referred to as $h(x), x = 0, 1, \dots, k - 1$. This signal exhibits numerous minute fluctuations and as Müller et al. [26] describes, several local maxima belong to one wrinkle, so that not all local maxima can simply be counted, see Fig. 6. However, unlike in Müller et al. [26], a standard wrinkle width should not be assumed, as this is too restrictive due to the anisotropy of the material. Instead, the signal is initially rectified as explained below.

3.4.2 Peak filtering:

At first all values before the first local maxima and after the last one are deleted. Then the sets \hat{E} of all u local maximizer $\hat{E} := \{\hat{x}_p | \hat{x}_p \text{ local maximizer of } h(x), p = 0, 1, \dots, u - 1\}$ and \bar{E} of all w local minimizer $\bar{E} := \{\bar{x}_q | \bar{x}_q \text{ local minimizer of } h(x), q = 0, 1, \dots, w - 1\}$ are determined. The function value $h(\hat{x}_p)$ of every local maximizer with $p \in [4, u - 4]$ is compared to the average of the function values of the three subsequent minima to the left $A_{l_p} := (h(\bar{x}_{q-1}) + h(\bar{x}_{q-2}) + h(\bar{x}_{q-3}))/3$ and the three subsequent minima to the right $A_{r_p} := (h(\bar{x}_{q+1}) + h(\bar{x}_{q+2}) + h(\bar{x}_{q+3}))/3$ with

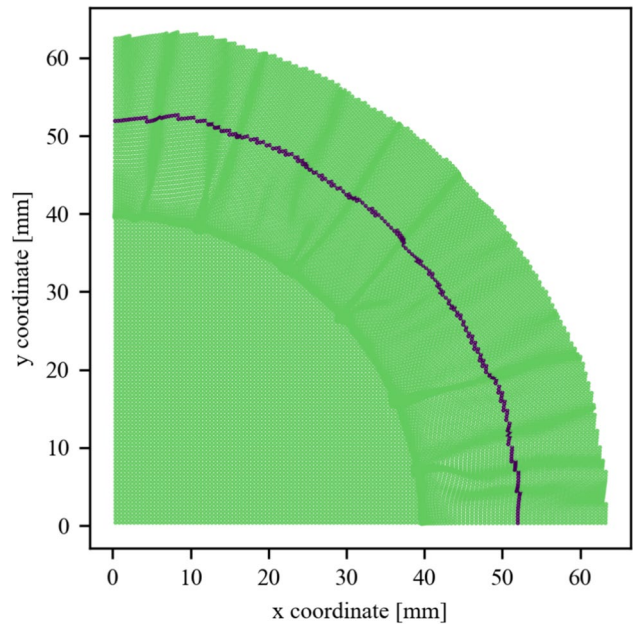


Fig. 5 Pixels for the quality assessment marked in purple

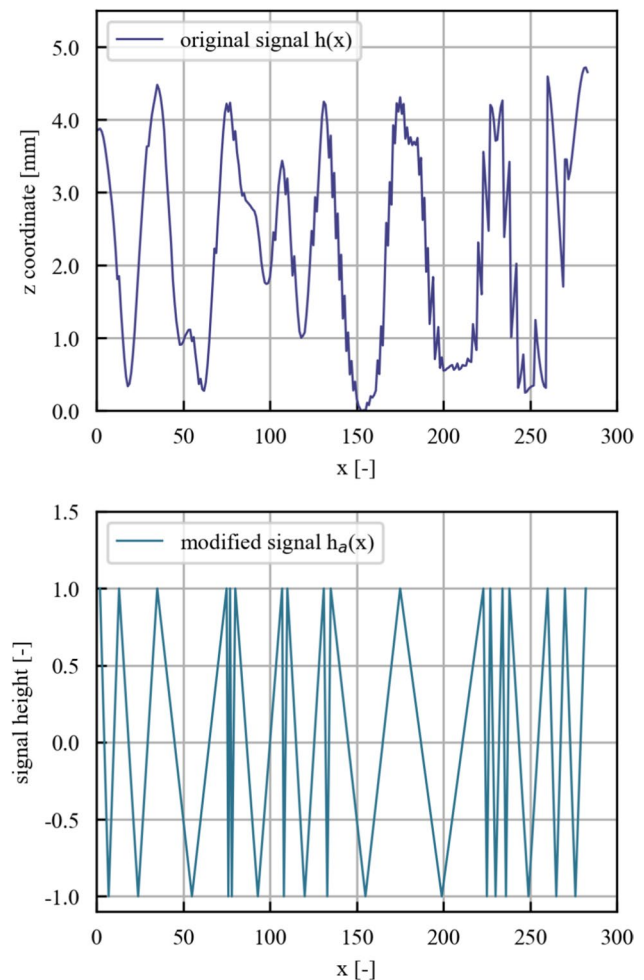


Fig. 6 Top: original signal $h(x)$ and Bottom: modified signal $h_a(x)$

$q = p$. Only those local maximizer $\hat{x}_p \in \hat{E}$ whose function values exceeds the averages A_{l_p} and A_{r_p} by a specified amount D_1 are retained. They form the set

$$\tilde{E} := \{\tilde{x}_r | \exists x_p \in \hat{E} : \tilde{x}_r = x_p \wedge (h(\hat{x}_p) - A_{l_p} > D_1) \wedge (h(\hat{x}_p) - A_{r_p} > D_1), r = 0, 1, \dots, \tilde{u} - 1 < u\}.$$

Additionally the local maximizer $\hat{x}_p \in E \setminus \tilde{E}$ which are greater than the subsequent maximizer $\tilde{x}_p > \tilde{x}_{p-1} \wedge \tilde{x}_p > \tilde{x}_{p+1}$ and have a difference to the subsequent minimizer smaller than D_2 are added to \tilde{E} . This captures the wrinkles which are represented by a larger plateau instead of high peaks.

3.4.3 Signal Normalization:

The function values for all $\tilde{x}_r \in \tilde{E}$ are set to 1. The adjusted signal $h'_a(x)$ for $x \in [0, k - 1]$ is thus defined by:

$$h'_a(x) := \begin{cases} 1 & \exists \tilde{x}_r \in \tilde{E} : x = \tilde{x}_r \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The values of D_1 and D_2 depend on the specific data and are determined in Section 4.3. The current signal has only local maxima but no local minima and is therefore unsuitable for the following Fast Fourier Transform. To change that the function values are set to -1 for the points in the middle between the maximizer and the final version of the adjusted signal is defined by:

$$h_a(x) := \begin{cases} -1 & \exists \tilde{x}_r, \tilde{x}_{r+1} \in \tilde{E} : x = I_r \\ h'_a(x) & \text{otherwise,} \end{cases} \quad (6)$$

where $I_r := \lfloor (\tilde{x}_r + \tilde{x}_{r+1})/2 \rfloor$. The final version of the adjusted signal is exemplary illustrated in Fig. 6.

3.4.4 Subdivision of signal into sections:

The signals shown, generated from the ML predictions, show areas of different intensities, meaning that the wrinkles are unevenly distributed. This effect is even more pronounced for the signals from real parts, especially for those with few wrinkles. The reason for this is the strong anisotropy of the paper. To be able to assess the distribution of the wrinkles, the signal obtained is now divided into any number n of sections $S_j := [\tilde{x}_{r_{j-1}}, \tilde{x}_{r_j}]$, $j = 1, \dots, n$ depending on the peak spacing, see top left of Fig. 7. Fewer areas indicate a more even distribution of wrinkles.

For the subdivision of $h_a(x)$ into sections, there is an iteration over the local maximizer $\tilde{x}_r \in \tilde{E}$, $r = 1, \dots, \tilde{u} - 1$. It applies $\tilde{x}_{r_0} = \tilde{x}_0$ and $S_0 = [\tilde{x}_0]$. The maximizer \tilde{x}_r is added to the current section $S_j = [\tilde{x}_{r_{j-1}}, \tilde{x}_{r-1}]$ as long as the absolute

value of the difference between the distance $\tilde{x}_r - \tilde{x}_{r-1}$ and the average distance of the current section $B_j := (\tilde{x}_{r-1} - \tilde{x}_{r_{j-1}})/(r_{j-1} - (r - 1))$ stays below a predefined limit L^* . If $|(\tilde{x}_r - \tilde{x}_{r-1}) - B_j| > L^*$ then \tilde{x}_r becomes the first entry of a new section S_{j+1} . Nevertheless there is one exception: if the following distances are very close to the current average $|(\tilde{x}_{r-2} - \tilde{x}_r)/2 - B_j| < 0.6 \cdot L^*$ no new section is started. For higher average distances higher deviations should be allowed so L^* is defined as follows:

$$L^* := \begin{cases} 0.5 \cdot L_1 & B_j \leq L_1 \\ 0.5 \cdot L_2 & L_1 < B_j \leq L_2 \\ L_3 & L_2 < B_j \end{cases} \quad (7)$$

with $0 < L_1 < L_2 < L_3$. The values of L_1, L_2 and L_3 depend on the specific data and are determined in Section 4.3.

3.4.5 Fast Fourier Transform:

The following steps aim on evaluating the number of wrinkles. Since it is not always clear whether a peak is a wrinkle or not, the characterization of the parts based purely on the absolute number of wrinkles is not very robust. This can be remedied by evaluating the Discrete Fourier Transform (DFT) of the modified signal $h_a(x)$ individually for each section $S_j, j \in 1, \dots, n$. This value fluctuates less if the number of peaks differs slightly. The DFT is calculated for each section, which contains at least two maxima, using the Fast Fourier Transform (FFT). The magnitude of the DFT $b_j(f_j) := |FFT(h_a(x))|$, $x \in S_j$ on the frequency spectrum f_j is considered, see top right and bottom of Fig. 7. For all samples and sections the highest magnitude never appeared for frequencies $f_j > 0.2$. That's why in the following only $f_j \in [0, 0.2], j \in 1, \dots, n$ is considered.

Every section S_j is characterized by it's dominant frequency f_j^* . In general this is the frequency with the highest amplitude: $f_{j_1} := f_j \in [0, 0.2]$ with $b_j(f_j)$ maximum. However, since this reduces the entire section to one scalar, the characterization would greatly distort the final result as soon as a second frequency f_{j_2} exists whose amplitude $b_{j_2}(f_{j_2})$ is only insignificantly smaller than $b_{j_1}(f_{j_1})$. That's why the dominant frequency f_j^* is from now on defined as

$$f_j^* = \begin{cases} 0.5 \cdot (f_{j_1} + f_{j_2}) & \exists b_{j_2}(f_{j_2}) > 0.94 \cdot b_{j_1}(f_{j_1}) \\ f_{j_1} & \text{otherwise.} \end{cases} \quad (8)$$

The specific limit of 94 % was chosen in a way that in the underlying data set for all samples there was at maximum one frequency with a magnitude above $94\% \cdot b_{j_1}(f_{j_1})$. The section wise FFTs b_j with the dominant frequencies f_j^* are exemplary illustrated in Fig. 7.

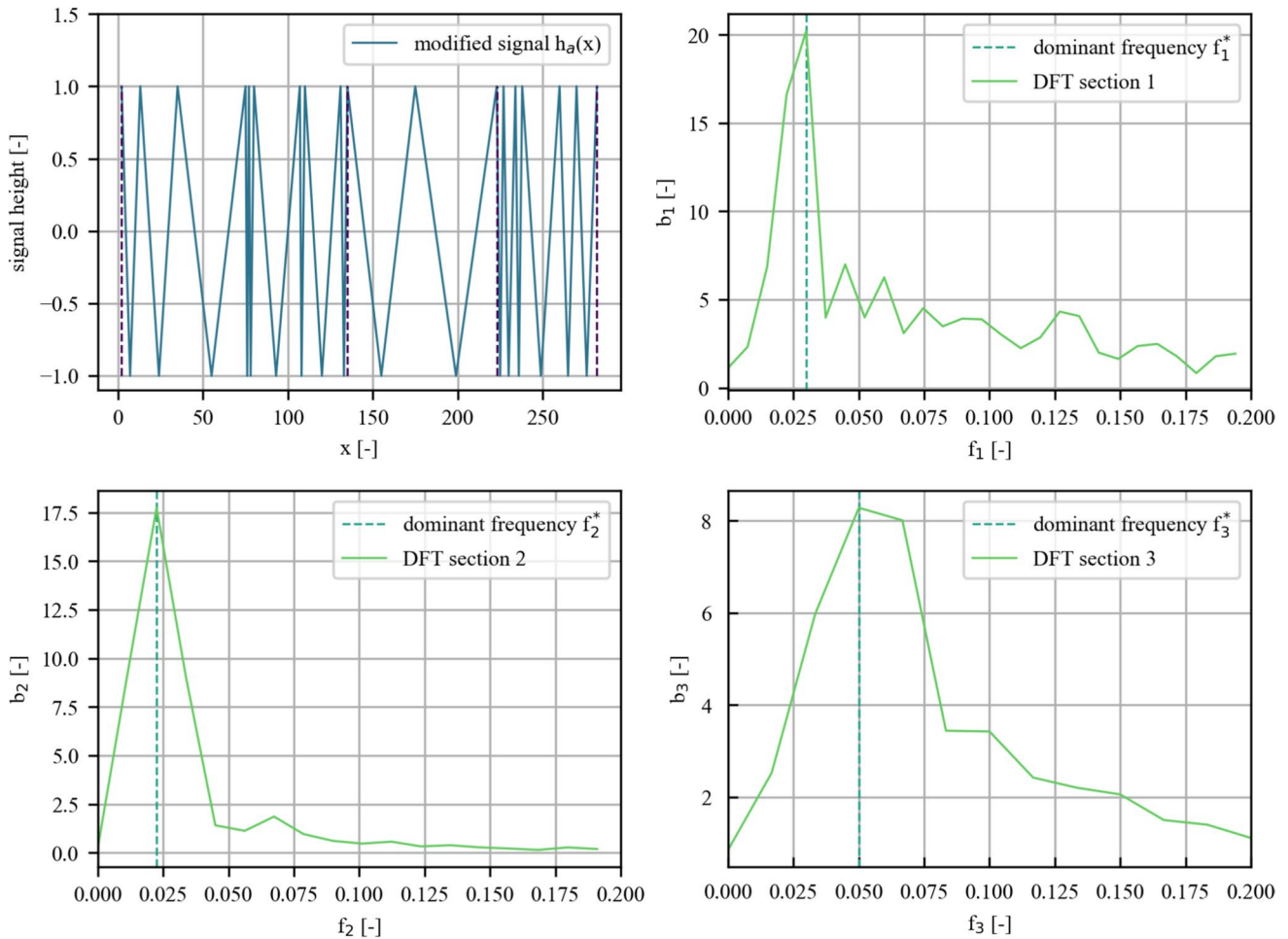


Fig. 7 Top left: modified signal $h_a(x)$ divided into sections and Top right, bottom: their DFTs b_j with the dominant frequencies f_j^*

3.4.6 Criterion

The criterion c finally results in

$$c := \sum_{j=1}^n s_j \cdot f_j^* \cdot l_j / \tilde{x}_{\tilde{u}-1} \cdot 1000, \tag{9}$$

$$s_j = \begin{cases} +1, & f_j^* \geq 0.013 \\ -1, & f_j^* < 0.013 \end{cases}$$

where $l_j := \tilde{x}_{r_j} - \tilde{x}_{r_{j-1}}$ is the length of a section and $\tilde{x}_{\tilde{u}-1}$ gives the total length of the signal. It is therefore summed over the sections and for each section the dominant frequency is weighted with the length of the section. Particularly low frequencies indicate almost wrinkle-free sections. As this is undesirable, frequencies which are smaller than 0.013 are interpreted as penalty terms and given a negative sign. The specific limit of 0.013 was determined by examining the mean dominant frequencies for sections with almost no wrinkles in

samples from the underlying dataset. The final value is multiplied by 1000 to simplify the comparison with the real word wrinkle numbers. Due to the penalty terms, negative values for c are possible. The larger c is, the better the cup is.

3.4.7 Optimization:

The ML model \mathcal{M} from Section 3.3 is integrated into an optimization strategy. This determines optimized inputs for the ML model (meaning in this case adapted process parameters) to ensure a stable process at the demonstrator press. The combination of two criteria is decisive in determining whether the system settings need to be changed for the current blank or if it can be formed with the current settings. If the maximum predicted thinning $s_{max}(F_{b_0}, m, T_d) := \max(\mathcal{M}_s(F_{b_0}, m, T_d))$ exceeds a critical value s_c that indicates material failure. Secondly

if the quality evaluation from Section 3.4 calculates a low value $c < c_c$, an insufficient wrinkle distribution is indicated. If at least one out of these two criteria indicate poor quality, the optimization is carried out. It determines new values for the three inputs of the ML model: starting blank holder force F_{b_0} , moisture of the blank m and temperature of the die T_d , so that the thinning is reduced below the critical value and the quality criterion is increased. Meanwhile, it is ensured that all three inputs remain within the given process limits. For reasons of efficiency and sustainability, it makes sense to use as little humidification as necessary. These conditions result in the following loss function

$$\mathcal{L} = 2 \cdot (s_{max}(F_{b_0}, m, T_d) - s_c) + (c_c - c) + m, \quad (10)$$

which is minimized by the optimization to obtain improved input parameters

$$F_{b_0}^*, m^*, T_d^* := \min_{F_{b_0}, m, T_d} \mathcal{L}. \quad (11)$$

The thinning is weighted with factor 2 to ensure that intact parts are produced.

The range within which the three input parameters are varied is relatively constrained, a consequence of the narrow process limits. This observation is particularly evident in the case of die temperature, given the relatively short time span between optimization and part production, which is approximately four seconds. During this time, it is impractical for the real die to achieve larger temperature differences than 5°C. Under the given conditions, the computationally intensive PSO is therefore not necessary. Also metaheuristics which are designed for high performance on large problem sizes are not suitable. Instead, a grid search [32] is proposed that systematically examines equidistant points in the space formed by F_{b_0} , m and T_d . Taking into account all boundary conditions, the grid search varies the parameters within the following intervals: $F_{b_0} \in [1000, 3500]$, $m \in [m_m, 22.0]$ and $T_d \in [\max(\hat{T}_d - 5, 23), \min(\hat{T}_d + 5, 120)]$, where the variable \hat{T}_d denotes the current temperature of the die at the demonstrator press. The input with the lowest loss is output by the optimization. The employment of an increased number of sampling points per interval has the potential to yield more optimal solutions. However, it should be noted that this approach also results in a substantial increase in the duration of the optimization process. Consequently, the maximum time available for execution of the optimization process is a critical factor in determining the proximity of the solution to the global maximum.

4 Results

4.1 FE simulation

The experiments described in Section 3.2 were carried out to calibrate the FE simulation. The value for Young's modulus calculated from the flow curves for different temperatures and moisture contents varies between 3500 and 5600 N/mm². To stabilize the FE simulation this parameter was slightly adapted to $E = 3100$ N/mm². R values were not examined experimentally but varied for further calibration in terms of wrinkle distribution and thinning. The chosen values are: $R_{0^\circ C} = 1.1$, $R_{45^\circ C} = 1.05$, $R_{90^\circ C} = 1.2$. For the flow curves only the one for 23°C is taken from the experimental data and converted from force - strain curve into true stress - true strain curve. The flow curve has been approximated, using the combined hardening law Hockett-Sherby $\sigma_y = \sigma_\infty - (\sigma_\infty - \sigma_0) \exp(-m\varepsilon^p)$ [41]. Then the curve was scaled for higher temperatures respecting the trend of increasing true stress but being monotonic increasing and without intersection points. The resulting true stress - true strain curves are illustrated in Fig. 8.

The coefficients of friction for PE coated side against polished stainless steel, denoted by μ_2 , are adopted from the friction tests (see Fig. 9). The coefficient of friction for the self-contact of paperboard, denoted by μ_3 , exhibited a range of 0.24 to 0.39 during the experimental phase. In the final model, μ_3 , is set to 0.40, a modification aimed at enhancing the occurrence of wrinkles. For the coefficients of friction μ_1 uncoated side against polished stainless only the values for 7.6 % moisture content have been used without change. The experimental values for 11.7 % moisture content were used for the highest moisture content of 16.1% instead. For moisture contents in between there was performed linear interpolation between these

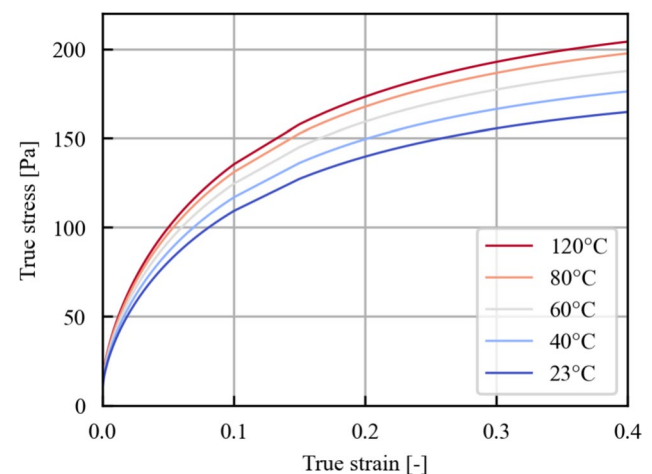


Fig. 8 Extrapolated temperature dependent true strain - true stress curves

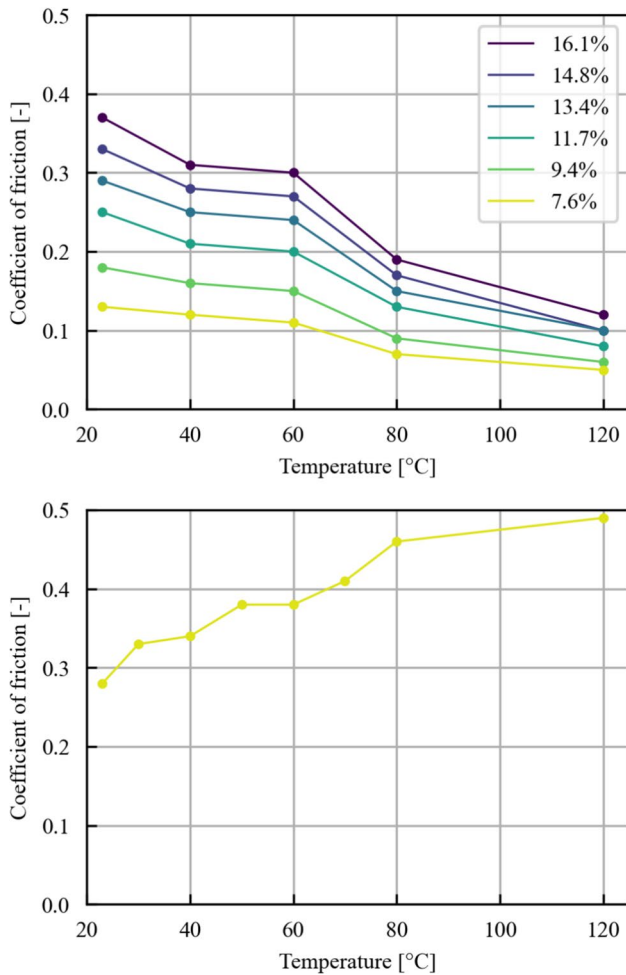


Fig. 9 Coefficients of friction. Left: moisture and temperature dependent μ_1 , and Right: temperature dependent μ_2

two curves, see Fig. 9. This was necessary because the FE model contains several simplifications (for example by using Coloumb friction law) what made it unstable with the original coefficients of friction.

Deep drawing trials on a press were performed to produce parts with varying starting blank holder force, material moisture, and die temperature. These trials were used to validate the FE simulation. The three process parameters were varied using a full-factorial design of experiments. However, some parameter combinations were skipped because the parts would have cracked, based on the results of previous parameter combinations. Additionally, the database was slightly reduced because the FE simulation was too unstable for simulating the full drawing depth for some samples. Therefore, the final experimental database \mathcal{E} consists of 39 samples. The specific parameter values for all samples are listed in Appendix I. Three parts were produced for each sample. If at

least two out of the three parts are intact, the sample state is labeled as intact; otherwise, it is labeled as broken.

Two criteria were considered for the validation on \mathcal{E} : the failure and the number of wrinkles for intact parts. For the visualization of the validation on \mathcal{E} in terms of failure, intact real parts are assigned the thinning 0 %, while broken parts are marked with 100 % thinning. For the simulation data the calculated maximum thinning is plotted in Fig. 10.

In general, it can be seen that the simulation predicts significantly higher thinning for most of the broken samples from \mathcal{E} than for the intact ones. Increased thinning for intact cups is only calculated by the simulation for parameter combinations where the die temperature is 80°C. Therefore, the FE simulation fulfills the requirement of providing a sufficient indication with regard to failure and simulated cups with a thinning $s \leq s_c := 40\%$ are considered intact.

The threshold value s_c was determined through a meticulous examination of the confusion matrices for five thresholds on the data set \mathcal{E} , as illustrated in Fig. 11. As the values 40 %, 50 %, and 60 % yielded equivalent outcomes, the lowest value was selected to mitigate the risk of samples being erroneously classified as intact.

Examples for the resulting wrinkle distributions are displayed in Fig. 12. To validate the FE simulation in terms of the number of wrinkles, the wrinkles were counted manually for the intact parts from the experiments and the average was calculated for each intact sample from \mathcal{E} . For these samples also the wrinkles on the simulated quarters were counted manually and multiplied by 4. The result is shown in Fig. 13.

It is true that the simulations have significantly fewer wrinkles than the real cups. Nevertheless, the trend of when

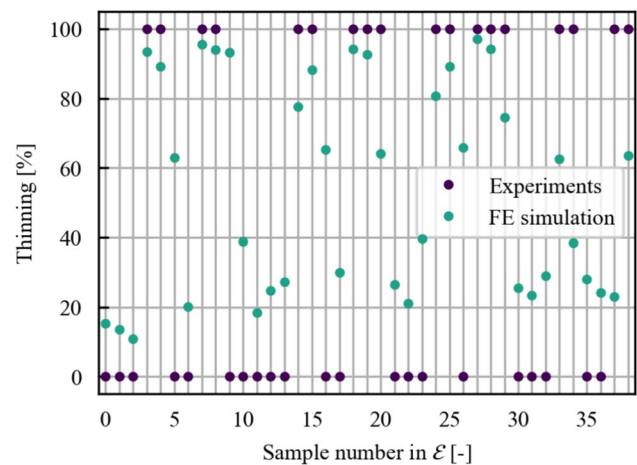


Fig. 10 Maximum thinning in FE simulation compared to experimental part state for \mathcal{E}

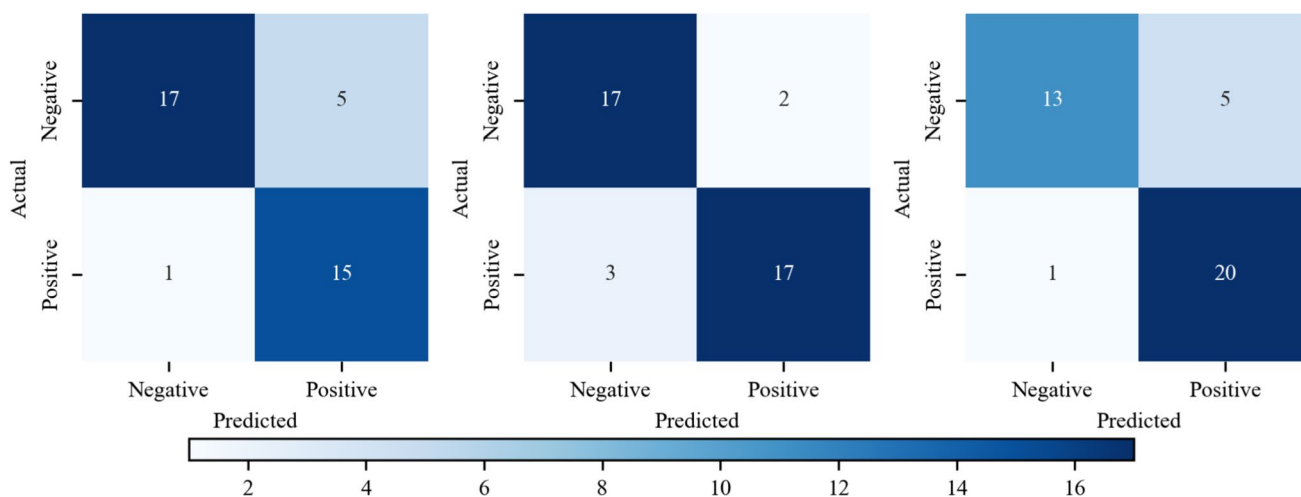


Fig. 11 Confusion matrices for different thresholds for s_c on the data set \mathcal{E} . Left: 30 %, Middle: 40 % / 50 % / 60 %, and Right: 70 %

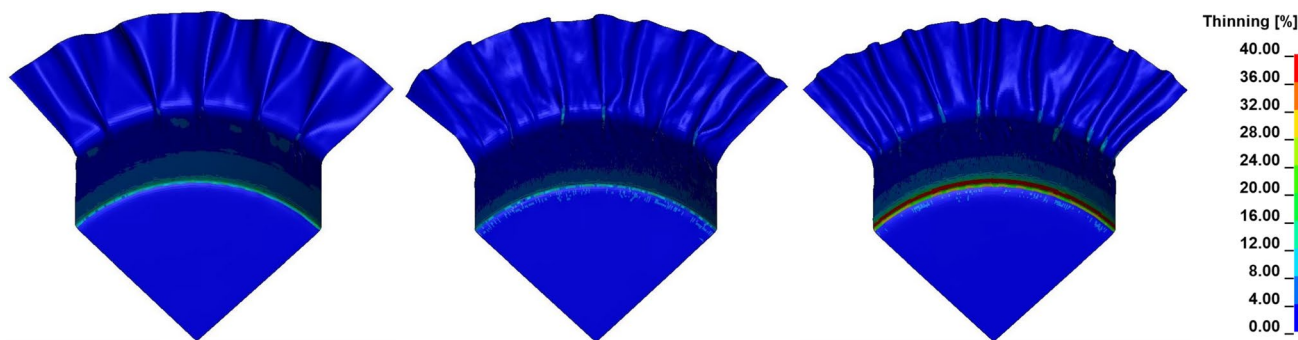


Fig. 12 Three samples indicating the appearance range in the deep drawing simulation with regard to the wrinkle distribution colored by its thinning

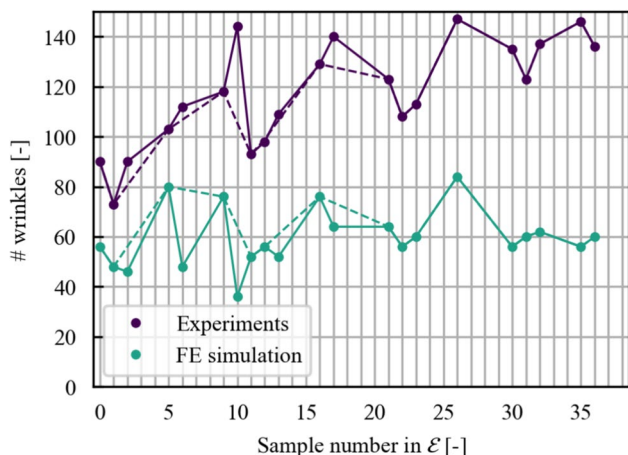


Fig. 13 Number of wrinkles in FE simulation compared to the experimental results for intact samples of \mathcal{E}

more and when fewer wrinkles occur is mostly consistent. In particular, if the number of wrinkles for cups at 120°C die temperature is neglected, the trend matches almost exactly. The outliers for precisely this die temperature can

be explained by the fact that the number of wrinkles is significantly influenced by the coefficient of friction and no coefficients of friction were determined experimentally for 120°C. Instead, these values had to be estimated.

4.2 ML model

In this section the dataset used to train the ML models is described. Then the identification of suitable hyperparameters and the outcomes derived from a particular model are presented.

Latin hypercube sampling method was employed to distribute 3660 samples over the space of three variables with boundaries $F_{bh_0} \in [1000, 3500]$, $m \in [7.6, 22.0]$, $T_d \in [23, 120]$. The calibrated FE model was used to simulate deep drawing. As the mapping of the wrinkles in combination with the thinning is in general numerical challenging and this effect amplifies with parameter combinations that lead to broken parts, for some samples the FE model was unable to simulate the full drawing depth of 40 mm. To ensure that these samples do not affect the predictions of

the ML model, only samples with full drawing depth are considered resulting in the final data set consisting of 2971 samples.

As described in Section 3.3, the influence of several loss functions was evaluated. The loss functions were varied together with batch size and learning rate in a small grid search. Tested batch sizes were {64, 128, 256} and {0.00075, 0.001, 0.0025, 0.005, 0.01} for the learning rates. For all parameter combinations models were trained for 200 epochs and a training/validation/test split of 2200/350/421 samples. The investigation revealed that the batch size exhibited minimal influence, while high learning rates were found to be associated with elevated error rates.

In an iterative process, the channel-specific best performing models were further trained until the overall best performing model was selected. The final selected model was trained with the MAE₃ loss, a batch size of 128 and a learning rate 0.001 for 2000 epochs. The development of the training and validation loss of this model \mathcal{M} is shown in Fig. 14.

For samples with no wrinkles, \mathcal{M} predicts their absence correctly. For samples with up to medium number of wrinkles \mathcal{M} predicts nearly the exact number of wrinkles and their original shape. Only for some of the samples with a high number of very small wrinkles the predictions of \mathcal{M} show higher deviations from the ground truth. In terms of thinning \mathcal{M} shows good predictions for all three scenarios. Examples are illustrated in Fig. 15.

4.3 Quality assessment with optimization strategy

The variable parameters of the quality assessment method are primarily the tolerances D_1, D_2 for when a peak is included in the adjusted signal and the limits for the distances between the peaks L^*, L_1, L_2, L_3 , which decide when a new section

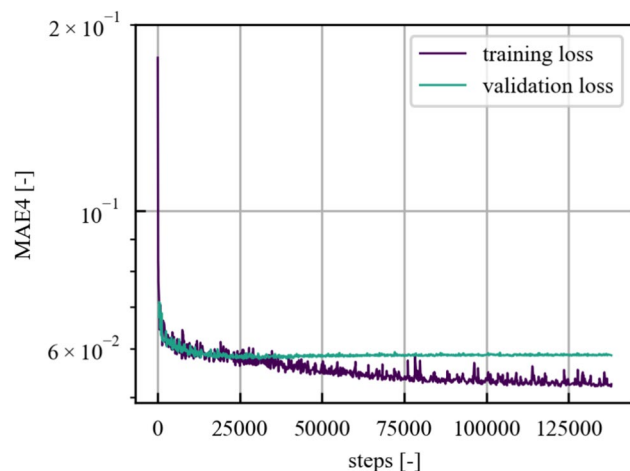


Fig. 14 Training and validation loss of the best performing ML model over 2000 epochs

is started in the adjusted signal. These parameters were varied in a small grid search. It compares the trends in the results of the quality assessment method to the trends occurring for the wrinkle numbers counted on the FE simulation results for intact samples of \mathcal{E} . The results of the best version after grid search are shown in Fig. 16. The calibration was performed with the results of the quality assessment method on the outputs of the FE simulation as the ML model outputs still have some deviations with regard to the wrinkle distribution.

The critical value c_c is now defined with the final calibration. Parts with $c \geq c_c$ are considered acceptable in terms of wrinkle distribution and shape. To determine c_c , the first step was to evaluate which real parts for intact samples from \mathcal{E} show satisfactory wrinkle distributions. From these samples, only those were kept for which the simulation also gave good results, and the quality criterion c was determined for the simulation results. The minimum of these values is defined to be the critical value c_c , resulting in $c_c := 25.0$. The quality assessment requires on average 9 milliseconds per cup with the method presented. It can therefore be used in a stroke-to-stroke control for the demonstrator press with a stroke rate of up to 60 strokes per minute.

With the calibrated quality assessment and the critical thinning from Section 4.1, the best performing ML model \mathcal{M} from Section 4.2 is integrated into the optimization strategy described in Section 3.4. As described the results of the optimization depend strongly on the number of grid points for each interval in which the process parameters are varied. For several samples the optimization was performed with different number of grid points to examine the quality of the solutions with the lowest loss and the required time. This showed that the interval for F should be discretized into at least six points each to ensure the existence of a feasible solution. Conversely, only three grid points were necessary for the intervals for m and T_d what conserves time. In this way 54 parameter combinations are evaluated for every initial solution and the optimization always reaches a feasible solution. The required time is 860 milliseconds on average. In scenarios where less stringent time constraints are in place, particularly in cases involving a slower process cycle during forming, it is recommended to augment the number of support points. This approach is expected to yield enhanced part quality. An exemplary initial solution with predicted material failure and the prediction for the process parameters suggested by the optimization are illustrated in Fig. 17. The initial process parameters are $F_{bh_0} = 1513$ N, $m = 14.9\%$, $T_d = 108^\circ\text{C}$. To avoid material failure, the optimization suggests the reduction of the moisturization duration, resulting in a lower material moisture content, and the increase of the blank holder force. The resulting process parameters after optimization are $F_{bh_0} = 2600$ N, $m = 10.9\%$, $T_d = 108^\circ\text{C}$.

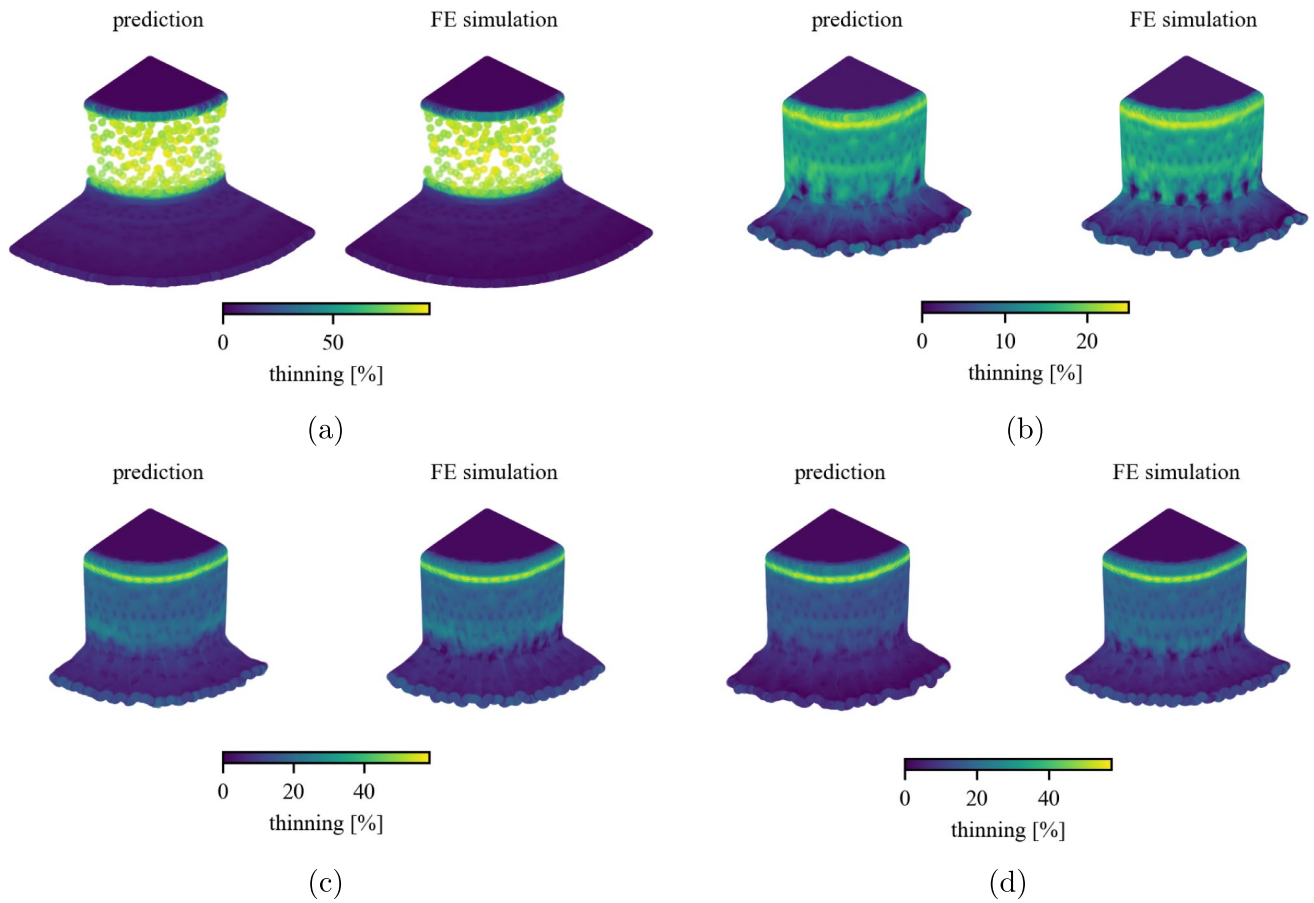


Fig. 15 ML model predictions on the test set vs. ground truth colored by it's thinning for (a) a sample without wrinkles, (b) a sample with a medium number of wrinkles, (c) a sample with a high number of

wrinkles with good prediction quality, and (d) a sample with a high number of wrinkles with poor prediction quality

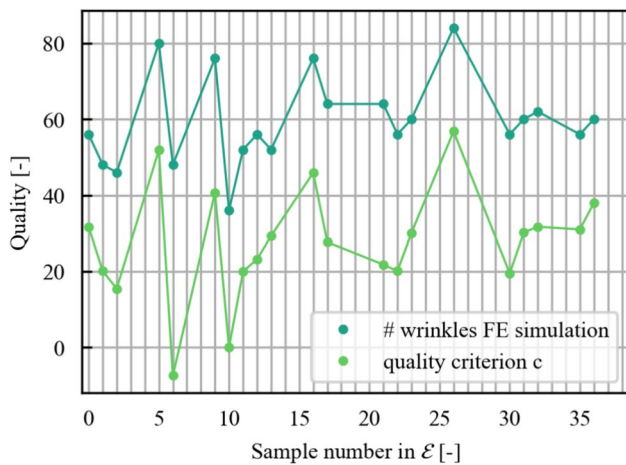


Fig. 16 Criteria c calculated by the quality assessment compared to number of wrinkles in FE simulation for intact samples of \mathcal{E}

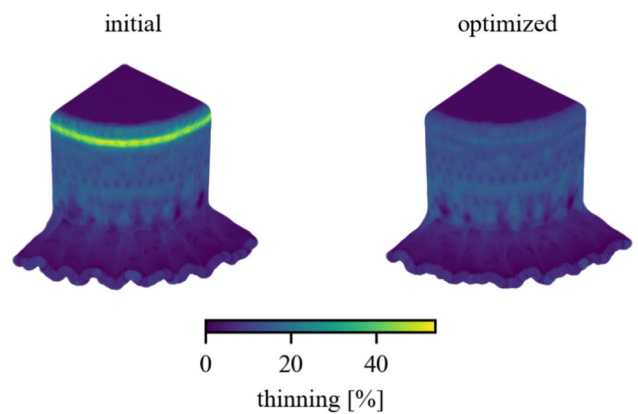


Fig. 17 Left: Predicted part quality for the initial parameter set indicating material failure, and Right: prediction for the improved process parameters suggested by the optimization strategy, both colored by it's thinning

5 Discussion

The findings demonstrate the feasibility of simulating wrinkles during the deep drawing process of paperboard, and furthermore, the capacity of CNNs to predict the geometry and quantity of these wrinkles. These outcomes provide a foundation to enhance the reliability and efficiency of paperboard manufacturing, preparing for large-scale production. The FE model utilized a simpler material model, compared to paperboard specific material models, developed for sheet metals and basic Coulomb friction law. Maximum thinning was employed as an artificial criterion for predicting material failure. The authors of this study acknowledge that the assumptions and simplifications made do not fully capture the complex material behavior of paperboard. This is one limitation of the present study. However, given the difficulty in mapping wrinkles, even with material models specially developed for paperboard, which require a significantly higher calibration effort than the one employed in this study, this research proposes a solution. It is acknowledged that larger deviations can be accepted in favor of the reduced calibration effort if the fundamental trends regarding wrinkle number and material failure are captured by the simulation. This is evidenced in the present study across the majority of the validation space. As significant deviations for the wrinkle number only occur for 120°C, these can be linked to specific coefficients of friction and flow curves. These material parameters should be critically re-examined for improvements to the simulation. Same applies for the thinning, where huge deviations mainly occur for 80°C. Especially when those outliers are neglected, the quantitative deviations do not impair the subsequent training of the ML models and the integration into a stroke-to-stroke control.

For the ML model, the RES-SE-U-Net architecture was adopted without alteration from Link et al. [24], as it was demonstrated there that the ML model \mathcal{M} attains high prediction accuracies for deep-drawn metal parts. In that study, wrinkles occurred in a very small number of parts, but despite the limited number of such samples, the CNN was able to map the rough shape of the wrinkles. Consequently, varying the architecture of the ML model was not part of this study. Instead, it investigated the potential to map wrinkles using a dataset comprising predominantly wrinkled samples, with a specific modification to the loss function. The investigation revealed that under these conditions, CNNs can predict approximate wrinkle distributions and their shape, which is relevant and important for deep drawing of paperboard. The model exhibits deficiencies in accurately determining the absolute position of the wrinkles. However, it is not the absolute position that is evaluated in the quality assessment method, but the number of wrinkles, their spacing and shape. As these values are

well predicted even for the majority of samples with errors in the absolute position, these deviations are negligible. Finally, challenges are associated with the mapping of the wrinkling behaviour for some of the samples that contain a substantial number of small wrinkles. To address this issue, a comprehensive parameter study for the weight of the z-component in the current MAE₃ loss function could be helpful. Furthermore, the efficacy of more sophisticated weights that impose substantial penalties for minor deviations in the z-coordinate could be assessed. Alternatively, a comprehensive investigation of the basic architecture with regard to the hyperparameters and modified CNN architectures would also be possible.

The quality assessment method has been shown to offer a reliable evaluation of quality, with a processing time of 9 milliseconds that renders it highly suitable for utilization in mass production contexts. The fact that there are fewer wrinkles in the simulation than in reality has no affect on the robustness of the quality assessment method. The reason is that the wrinkle numbers from the FE simulation are used as ground truth for the calibration of the quality assessment method. The deviations between FE simulation and reality only need to be considered when determining the critical value of the quality criterion. This means that the desired real number of wrinkles cannot be converted directly into the quality criterion. Instead, the critical value for the quality criterion needs to be determined as described in Section 4.3. Due to the observed deviations in wrinkle distribution between the FE simulation and the predictions of \mathcal{M} the quality assessment is currently calibrated on the FE outputs rather than the ML outputs, what is another limitation of the present study. Nevertheless, given that the calibration process was automated using a grid search, it can be repeated on the ML outputs at any time with minimal effort. To do this, the ML model would be retrained after modification to further reduce its deviations from the FE simulation. The grid search which varies the parameters of the quality assessment method is then repeated using the predictions of the modified ML model.

The optimization algorithm is capable of suggesting admissible solutions for all initial solutions. If deemed necessary, a more complex optimization algorithm could be employed to approach the global maximum more closely. However, given the limited number of inputs in the optimization, the development of more complex algorithms is not a primary focus in future studies.

6 Conclusion

The process of deep drawing of paperboard has been identified as a potentially viable method for the production of paperboard packaging that can exhibit a wide range of shapes and compete with plastic packaging. Nevertheless,

when endeavouring to expand the deep drawing of paperboard to mass production, control issues frequently emerge. For instance, it remains challenging to ensure the integrity of the parts and maintain high quality, defined as numerous fine, uniformly distributed wrinkles, on a continuous basis. In order to achieve this objective, it is imperative that precise process parameter settings are implemented. The quality of the resulting part can be predicted through the utilisation of FE simulations. However, due to the computational time required, these simulations are not feasible for in-line predictions. Consequently, a methodology was formulated in this study for predicting the quality of paperboard parts during the manufacturing process and suggesting optimised process parameters if necessary by employing a digital twin.

The integrity of the parts and the distribution of wrinkles are predominantly dictated by the blank holder force, the targeted application of moisture to the blanks, and the temperature of the die. The developed FE model predicts part quality depending on these three process parameters, using a material model for metal and the Coulomb friction. The FE model was calibrated through tensile and friction tests and validated through deep drawing tests on a press. Subsequently, it was employed to generate a database for the training of an ML model, specifically a CNN with a RES-SE-U-Net architecture. A methodology was formulated for the objective assessment of part quality. Finally, the ML model was incorporated into an optimization strategy via a grid search. The primary findings of this study are as follows:

1. Despite the use of a material model for metal and a simple friction law, the FE model is capable of qualitatively mapping wrinkle and material failure trends across a substantial portion of the parameter space.
2. The ML model demonstrates the capacity to predict the approximate number of wrinkles formed and their rough geometry as a function of the process parameters.
3. The evaluation of the wrinkle distribution in the predictions is robust and has an average duration of 9 milliseconds. The integration of this approach into a grid search optimization strategy suggests process parameters that can prevent the production of substandard parts within 860 milliseconds. The presented method is therefore suitable for supporting a robust production of deep drawn paperboard parts with effective stroke rates of up to 60 strokes per minute.

Future studies could increase the accuracy of the FE model by conducting further experiments on the material parameters and using an alternative material model. Additionally, the refinement of modified loss functions and modifications in the ML model architecture hold potential in enhancing the prediction accuracy of the ML model.

Appendix

Table 1 Appendix I: experimental data base \mathcal{E} for validation of FE simulation

number	starting blank holder force F_{bh_0} [N]	moisture m [%]	die temperature T_d [°C]	state
0	1000	7.6	23	intact
1	1000	7.6	50	intact
2	1000	7.6	120	intact
3	1000	13.4	23	cracked
4	1000	13.4	50	cracked
5	1000	13.4	80	intact
6	1000	13.4	120	intact
7	1000	19.5	23	cracked
8	1000	19.5	50	cracked
9	1000	19.5	80	intact
10	1000	19.5	120	intact
11	1500	7.6	50	intact
12	1500	7.6	80	intact
13	1500	7.6	120	intact
14	1500	13.4	23	cracked
15	1500	13.4	50	cracked
16	1500	13.4	80	intact
17	1500	13.4	120	intact
18	1500	19.5	50	cracked
19	1500	19.5	80	cracked
20	1500	19	120	cracked
21	2000	7.6	23	intact
22	2000	7.6	50	intact
23	2000	7.6	80	intact
24	2000	13.4	23	cracked
25	2000	13.4	50	cracked
26	2000	13.4	80	intact
27	2000	19.5	23	cracked
28	2000	19.5	50	cracked
29	2000	19.5	80	cracked
30	2500	7.6	23	intact
31	2500	7.6	50	intact
32	2500	7.6	80	intact
33	2500	13.4	80	cracked
34	2500	13.4	120	cracked
35	3000	7.6	23	intact
36	3000	7.6	50	intact
37	3000	7.6	80	cracked
38	3000	13.4	80	cracked

Author Contributions Klara Liesegang was responsible for the conceptualization and methodology, as well as the further development of the digital twin (including FE simulation and ML models) to enable its application to paperboard deep drawing. She was also responsible for drafting the quality assessment. Patrick Link provided the preliminary design of the FE simulation and the ML architecture. Georg Steinert and Lena Berthold were responsible for the planning and execution of all experiments and the provision of process-specific knowledge. Steffen Ihlenfeldt was responsible for the acquisition of funding, the

supervision of the project, and the provision of the necessary resources. The initial manuscript draft was authored by Klara Liesegang, and all authors provided feedback on previous iterations of the document. All authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. This research was funded by the Federal Ministry for Economic Affairs and Climate Action as part of the program Industrial Collective Research (“Industrielle Gemeinschaftsforschung”, IGF) - 011F22573N.

Declarations

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

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

- DIN 8584–3:2003–09 (2003) Manufacturing processes forming under combination of tensile and compressive conditions - part 3: Deep drawing; classification, subdivision, terms and definitions. Norm
- Hauptmann M, Majschak J (2011) New quality level of packaging components from paperboard through technology improvement in 3d forming. *Packag Technol Sci* 24(7):419–432. <https://doi.org/10.1002/pts.941>
- Vishtal A, Retulainen E (2012) Deep-drawing of paper and paperboard: The role of material properties. *BioResources* 7(3):4424–4450. <https://doi.org/10.15376/biores.7.3.4424-4450>
- Franke W, Stein P, Dörsam S, Grochex P (2018) Formability of paperboard during deep-drawing with local steam application. In: AIP Conference Proceedings, p 100008. <https://doi.org/10.1063/1.5034948>
- Niini A, Berthold L, Müller T, Tanninen P, Majschak J-P, Varis J, Leminen V (2022) Effect of blank moisture content on forming behaviour and mechanical properties of paperboard tray packages. *J Appl Packaging Res* 14:4
- Wallmeier M, Hauptmann M, Majschak J-P (2016) The occurrence of rupture in deep-drawing of paperboard. *BioResources* 11(2). <https://doi.org/10.15376/biores.11.2.4688-4704>
- Müller T, Lenske A, Hauptmann M, Majschak J-P (2017) Analysis of dominant process parameters in deep-drawing of paperboard. *BioResources* 12(2). <https://doi.org/10.15376/biores.12.2.3530-3545>
- Wallmeier M, Linvill E, Hauptmann M, Majschak J-P, Östlund S (2015) Explicit fem analysis of the deep drawing of paperboard. *Mech Mater* 89:202–215. <https://doi.org/10.1016/j.mechmat.2015.06.014>
- Linvill E, Wallmeier M, Östlund S (2017) A constitutive model for paperboard including wrinkle prediction and post-wrinkle behavior applied to deep drawing. *Int J Solids Struct* 117:143–158. <https://doi.org/10.1016/j.ijsolstr.2017.03.029>
- Awais M, Sorvari J, Tanninen P, Leppänen T (2017) Finite element analysis of the press forming process. *Int J Mech Sci* 131–132:767–775. <https://doi.org/10.1016/j.ijmecsci.2017.07.053>
- Xia Q, Boyce M, Parks D (2002) A constitutive model for the anisotropic elastic-plastic deformation of paper and paperboard. *Int J Solids Struct* 39(15):4053–4071. [https://doi.org/10.1016/S0020-7683\(02\)00238-X](https://doi.org/10.1016/S0020-7683(02)00238-X)
- Nygårds M, Just M, Tryding J (2009) Experimental and numerical studies of creasing of paperboard. *Int J Solids Struct* 46(11–12):2493–2505. <https://doi.org/10.1016/j.ijsolstr.2009.02.014>
- Livermore Software Technology (2020) LS-DYNA® Keyword User’s Manual - Volume II Material Models - LS-DYNA R12
- Tokura S, Takekoshi K (2020) Simulation of compression behavior of paper product using *mat_paper. 16th International LS-DYNA users conference
- Tercan H, Meisen T (2022) Machine learning and deep learning based predictive quality in manufacturing: a systematic review. *J Intell Manuf* 33(7):1879–1905. <https://doi.org/10.1007/s10845-022-01963-8>
- Ståhl N, Mathiason G, Falkman G, Karlsson A (2019) Using recurrent neural networks with attention for detecting problematic slab shapes in steel rolling. *Appl Math Model* 70:365–377. <https://doi.org/10.1016/j.apm.2019.01.027>
- Meyes R, Donauer J, Schmeing A, Meisen T (2019) A recurrent neural network architecture for failure prediction in deep drawing sensory time series data. *Procedia Manuf* 34:789–797. <https://doi.org/10.1016/j.promfg.2019.06.205>
- Zimmerling C, Trippe D, Fengler B, Kärger L (2019) An approach for rapid prediction of textile draping results for variable composite component geometries using deep neural networks. In: Proceedings of the 22ND International Esaform Conference On Material Forming: ESAFORM 2019. AIP Publishing. <https://doi.org/10.1063/1.5112512>
- Attar H, Zhou H, Foster A, Li N (2021) Rapid feasibility assessment of components to be formed through hot stamping: A deep learning approach. *J Manuf Process* 68:1650–1671. <https://doi.org/10.1016/j.jmapro.2021.06.011>
- Attar H, Zhou H, Li N (2021) Deformation and thinning field prediction for hfq® formed panel components using convolutional neural networks. *IOP Conf Ser Mater Sci Eng* 1157(1):012079. <https://doi.org/10.1088/1757-899X/1157/1/012079>
- Zhou H, Xu Q, Nie Z, Li N (2021) A study on using image-based machine learning methods to develop surrogate models of stamp forming simulations. *J Manuf Sci Eng* 144(2). <https://doi.org/10.1115/1.4051604>
- Link P, Bodenstab J, Penter L, Ihlenfeldt S (2022) Metamodeling of a deep drawing process using conditional generative adversarial networks. *IOP Conf Ser Mater Sci Eng* 1238(1):012064. <https://doi.org/10.1088/1757-899X/1238/1/012064>
- Liu S, Xia Y, Shi Z, Yu H, Li Z, Lin J (2021) Deep learning in sheet metal bending with a novel theory-guided deep neural network. *IEEE/CAA J Autom Sin* 8(3):565–581. <https://doi.org/10.1109/JAS.2021.1003871>
- Link P, Penter L, Rückert U, Klingel L, Verl A, Ihlenfeldt S (2025) Real-time quality prediction and local adjustment of friction with digital twin in sheet metal forming. *Robot Comput Integr Manuf* 91:102848. <https://doi.org/10.1016/j.rcim.2024.102848>
- Wallmeier M, Hauptmann M, Majschak J-P (2015) New methods for quality analysis of deep-drawn packaging components from paperboard. *Packag Technol Sci* 28(2):91–100. <https://doi.org/10.1002/pts.2091>

26. Müller T, Lenske A, Hauptmann M, Majschak J-P (2017) Method for fast quality evaluation of deep-drawn paperboard packaging components. *Packag Technol Sci* 30(11):703–710. <https://doi.org/10.1002/pts.2315>
27. Müller T, Meyer M, Lenske A, Hauptmann M, Majschak J-P (2018) Optical inline quality assessment of deep-drawn paperboard containers. *J Mater Process Technol* 262:615–621. <https://doi.org/10.1016/j.jmatprotec.2018.08.003>
28. Leng Y, Sanjon C, Tan Q, Groche P, Hauptmann M, Majschak J-P (2024) Study of parameters influencing wrinkles in the deep drawing of fiber-based materials using automatic image detection. *J Manuf Mater Process* 8(6):237. <https://doi.org/10.3390/jmmp8060237>
29. Gad AG (2022) Particle swarm optimization algorithm and its applications: A systematic review. *Arch Comput Methods Eng* 29(5):2531–2561. <https://doi.org/10.1007/s11831-021-09694-4>
30. Weichert D, Link P, Stoll A, Rüping S, Ihlenfeldt S, Wrobel S (2019) A review of machine learning for the optimization of production processes. *Int J Adv Manuf Technol* 104(5–8):1889–1902. <https://doi.org/10.1007/s00170-019-03988-5>
31. Dokeroglu T, Sevinc E, Kucukyilmaz T, Cosar A (2019) A survey on new generation metaheuristic algorithms. *Comput Ind Eng* 137:106040. <https://doi.org/10.1016/j.cie.2019.106040>
32. Menke W (2012) Nonlinear inverse problems. In: *Geophysical data analysis: Discrete inverse theory*, pp 163–188. <https://doi.org/10.1016/b978-0-12-397160-9.00009-6>
33. Stora Enso (2025) Trayforma™. <https://www.storaenso.com/-/media/documents/download-center/documents/product-specifications/paperboard-materials> Accessed 07 Feb 2025
34. Groche P, Huttel D (2015) Paperboard forming - specifics compared to sheet metal forming. *BioResources* 11(1). <https://doi.org/10.15376/biores.11.1.1855-1867>
35. DIN EN ISO 186:2002–08: Iso 186 (2002). Paper and board - sampling to determine average quality. Norm
36. DIN EN ISO 187:2022–10: Iso 187 (2022). Paper, board and pulps - standard atmosphere for conditioning and testing and procedure for monitoring the atmosphere and conditioning of samples. Norm
37. DIN EN ISO 287:2017–11: Iso 287 (2017). Paper and board - determination of moisture content of a lot - oven-drying method. Norm
38. DIN EN ISO 1924–2:2008: Iso 1924–2 (2008). Paper and board - determination of tensile properties - part 2: Constant rate of elongation method (20 mm/min). Norm
39. Lenske A, Müller T, Hauptmann M, Majschak J-P (2018) New method to evaluate the frictional behavior within the forming gap during the deep drawing process of paperboard. *BioResources* 13(3):5580–5597. <https://doi.org/10.15376/biores.13.3.5580-5597>
40. Hauptmann M, Weyhe J, Majschak J-P (2016) Optimisation of deep drawn paperboard structures by adaptation of the blank holder force trajectory. *J Mater Process Technol* 232:142–152. <https://doi.org/10.1016/j.jmatprotec.2016.02.007>
41. Hockett J, Sherby O (1975) Large strain deformation of polycrystalline metals at low homologous temperatures. *J Mech Phys Solids* 23(2):87–98. [https://doi.org/10.1016/0022-5096\(75\)90018-6](https://doi.org/10.1016/0022-5096(75)90018-6)

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