Predictive Analysis from numerical and experimental data in press hardening

Lars Penter*1,2, Patrick Link2, Anke Stoll2, André Albert2, Steffen Ihlenfeldt1,2

1 Institute of Mechatronic Engineering, TU Dresden, Dresden, 01062, Germany
2 Fraunhofer Institute for Machine Tools and Forming Technology, Chemnitz, 09126 Germany

* lars.penter@tu-dresden.de

Abstract. Machine learning, big data and deep learning are today’s catchphrases for how to improve reliability and productivity of your manufacturing equipment. Production companies implement a large number of sensors to record every activity within their production lines and learn as much as possible about their running processes in order to predict shifting product properties and to prevent stoppage due to failure. The successful application of machine learning algorithms to predict machine and process behavior depends on a reliable and balanced database. Since the foremost goal of every manufacturing business is to make sound parts and to avoid defects, there is a large amount of data available for smoothly running processes but only very little for failure production. One approach to solve this imbalance would be to link the production line data with simulation data. Simulation models allow for computing failure parts with no additional costs and therefore enable the exploration of the entire parameter space. We conducted press-hardening experiments with a variation of process parameters for a structural car body part on the press hardening line at Fraunhofer IWU. As an evaluation criterion, we measured the hardness of the final part at critical spots. In order to expand the experimental data, we applied FE simulations to the entire press hardening process chain. The paper explains limitations of the model and elaborates on its parameterization. As a final task, we applied a basic machine-learning algorithm to both experimental and numerical data as well as to their combination in order to evaluate the data space expansion through simulations. The results obtained through machine learning indicate significant differences in the prediction of part quality for solely experimental data and its combination with simulation data. This is especially true for press hardening because of non-linear system behavior and a large amount of uncertain and hard-to-identify parameters. We found that the most challenging parts of uniting measured and simulated data is not only to create simulations with appropriate accuracy, which allow for a meaningful extrapolation of the parameter space, but also to compare simulation and production data based on the same criterion and to have stable simulation models for the entire parameter range.

1 Introduction

When reading thought current popular science papers, one could conclude that Machine Learning will solve all problems on the shop floor. Machine learning algorithms require balanced data sets for training. Since the foremost goal of every manufacturing business is to make sound parts and to avoid defects, there is data available for robust running processes but only very little for failure production.
Compared to the sample size available to big tech companies such as Google or Booking.com, the data sets often found the production environment are small. This is also true for research institutions analysing processes with high per-part costs.

One approach to dealing with small data is the application of simple models (see overview in Figure 1). Good starting points for small data sets are linear models such as linear/logistic regression or simple Bayesian models such as Naïve Bayes (few parameters and a direct way to adjust your prior.). Another approach is to use feature selection based on domain knowledge. Furthermore, regularization is a useful technique that can help improving the accuracy of regression models for small data sets [1]. Small data sets require models that are low in complexity in order to avoid overfitting the model to the data. A model will have a low accuracy if it is over fitted. This happens because the model is trying too hard to capture the noise in the training dataset. The so called bias-variance trade-off is the conflict in trying to minimize bias and variance at the same time in order to obtain a good model [1]. In reality, models with a lower bias in parameter estimation have a higher variance of the parameter estimates across samples, and vice versa. This problem prevents machine learning algorithms from generalizing beyond their training set. Other issues are the identification of outliers and noise in small data sets [2, 3]. Data cleaning is helpful but reduces the already small data set even further. Alternatively, models which are robust to outliers can be used, such as quantile regression [4].

Imbalanced data leads to new challenges that have to be solved. Models, such as one-class support vector machines (SVM), clustering methods or Gaussian Anomaly detection [5, 6], can be used which consider the minor class as outlier class. Another approach to handle imbalanced data is to simply bring it back into proportion by either oversampling or undersampling [7] or by generating artificial data [8]. Assembling techniques that combine multiple different models have also shown some success in dealing with unbalanced data.

Our approach to solve the imbalance in press-hardening-process data is to pool production line data and artificial simulation data. Simulation models allow for computing failure parts with no additional costs and therefore enable the exploration of the entire parameter space. The simulation model can create failure patterns which could later be detected in the production data and or be used for virtual commissioning of tools and machines.

2 Methods

We conducted press-hardening experiments with a variation of process parameters for a structural car body part on the press hardening line at Fraunhofer IWU. As an evaluation criterion, we measured the hardness of the final part at critical spots. In order to expand the experimental data, we applied FE simulations to the entire press hardening process chain. As a final task, we applied a machine-learning method (supervised learning with linear regression) to both experimental and numerical data as well as to their combination in order to evaluate the data space expansion through simulations.
2.1 Experimental equipment

The press-hardening production line at Fraunhofer IWU (Figure 2) is used for direct press hardening [9]. A linear feeder loads and unloads a chamber furnace with the sheet blank. The blank is heated up above 900°C. An industrial robot lifts the red-hot sheet into the cooled forming tool. The servo-hydraulic deep drawing press enables motion and force profiles adjusted to the needs of press hardening. After forming and quenching, the robot picks the formed part from the tool and places it on the deposition table for further cooling.

Figure 2: Press hardening line at Fraunhofer IWU, Chemnitz.

The 3MA test was applied to measure the Vickers hardness of the final part at six critical locations (see Figure 3). Point 1 lies in a plane plateau, points 2 and 6 in radius areas, point 3 in the area of the greatest drawing depth, point 4 in the blankholder area and point 5 in the transition from the plateau to the component area in a radius.

Figure 3: Location of hardness and thickness measurements on final part.

The experiments were designed with a set tool temperature of 20 °C, a press force of 2,000 kN and a set furnace temperature of 950 °C. 2 mm shims create a 0.5 mm fixed gap between blank and tool in the blankholder area.
2.2 Simulation models

The analysis of press hardening requires a multi-step simulation model which comprises the heating of the blank, the transfer from oven to press hardening tool, the forming and process and subsequent spring back as well as free cooling to measuring temperature (Figure 4). As of now, the simulation model is capable of computing the sheet thickness and the hardness after forming, quenching and passive cooling to room temperature.

We used the LS-DYNA solver to compute the process chain. In order to investigate model complexity and its influence on the prediction accuracy, we defined two FE simulation models with a) simple and b) intermediate complexity. We used both simulation models to compute the final part hardness at the exact same positions as on the experimental parts.

The blank was modeled with type 16 shell elements with 5 integration points through thickness. The material (Table 1), interface (Table 2) and parameters of the simulation models were obtained from literature. We used the LS-DYNA material model 244 which is based on [10]. Chemical composition of the blank, the activation energy for phases and latent heat were obtained from [11]. The tool surface is rigid and a constant surface temperature was assumed for both models.

Table 1: Material parameters for blank model

| parameter                  | unit                 | values           | source         |
|----------------------------|----------------------|------------------|----------------|
| Sheet thickness            | mm                   | 1.5              | measured       |
| Young’s modulus            | MPa                  | 100.000          | [12]           |
| True stress curves         | MPa                  | f(true strain, phase, T) | Dynamore     |
| Emission coefficient       | -                    | 0.8              | [12]           |
| Heat capacity              | mJ/(tonK)            | f(T)             |                |
| Thermal conductivity       | mW/(mm K)            | f(T)             |                |

Figure 4: Simulation models for computing the final part properties

1 Solver version: mpp d R11.0.0, revision 129956
Table 2: Contact interface parameters

| parameter                          | unit     | values   | source                           |
|------------------------------------|----------|----------|----------------------------------|
| Heat transfer coefficient (upper)  | W/mm²K   | 0.00945  | [11]                             |
| Heat transfer coefficient (lower)  | W/mm²K   | 0.00945  | [11]                             |
| Friction coefficient               | -        | 0.4      | NUMISSHEET-benchmark 2008        |
| Heat transfer coefficient (gap)    | W/mm²K   | 0.0546   | assumption: air                  |

2.3 Design of experiments and machine learning model

For the experiments, we choose transfer time and quenching time as controllable variables. The press hardening line additionally allows for actively adjusting furnace time (which leads to different blank temperatures) and for changing the forming pressure. The machine and individual part data was acquired for 12 specimens with 4 different configurations (Figure 5). To obtain a broader idea of the design space, the boundaries for input variables were expanded for the computer design of experiments (Figure 5). The DOE is a space filling design which maximizes the minimal distance between any two input points. The design was created with the LS-OPT software.

Figure 5: DOE for experiment and simulations

Since the experimental data set is small, supervised learning with input and output data and a simple model were applied to the experimental data, the simulation data and a combination of both data sets. The relationship between the two independent and one dependent variable was modelled with a multiple linear regression fitted with the least square approach. Experimental data set A, B and C plus all simulation results were used as training data to predict the hardness for experimental parameter set D (34.5 seconds of handling time and 6 seconds of quenching time).

3 Results

The following section presents experimental as well as numerical results and the predictions made via linear regression based on two data sets.

3.1 Experimental results

The slide motion reflects the variations of handling time and quenching time for parameter sets A to D (see Figure 6 left). The furnace temperature falls as soon as the furnace door opens. Also the temperature during blank austenitization falls 20 K short of the set temperature. The handling time includes opening, the linear feed and the robot transfer (see Figure 6 right). The tool temperature rose during each forming and quenching step from ca. 20 °C to 42 °C (see Figure 6 right). To guarantee the same initial conditions
for every experiments we made sure to cool the tools back down to room temperature before starting a new run.

The measured Vickers hardness varies within a range of 336 to 502 HV. The hardness value is the average for three repetitions. The hardness decreases with rising transfer time (see Figure 7 for results at measuring point 1 and 2). This trend could be observed for all measuring locations. The quenching time had no systematic influence on the final part hardness. At measuring location 1, a considerable increase of the hardness with rising quenching time was observed for 29.5 sec of handling time. However, measuring location 2 showed no increase due to longer quenching times at all.

3.2 Simulation results
A sensitivity analysis to understand the impact of uncertain parameters on the simulated hardness suggested the heat transfer coefficient as most influential variable. We obtained an optimum HV prediction for parameter combination A at measuring location 1 (compare Figure 7 and Figure 8) using a heat transfer coefficient of 2 W/(mm²K) with a difference of less than 8 HV. The values of the Vickers hardness obtained through simulation show a similar trend as the experiment regarding the influence of handling and quenching time. The handling time has a measurable impact whereas the quenching time does not significantly affect the hardness. Due to the five-step simulation and its complexity, we
experience instabilities in the lower left and the upper right corners of the DOE (Figure 8). As of now, the model shows no sensitivity regarding a hardness variation at different measuring locations. Therefore, the deviation for parameter combination A at measuring location 2 exceeds 30 HV.

Figure 8: Simulation results for hardness after quenching

3.3 Predictions based on regression models for machine learning
Due to the small amount of experimental data, a linear regression was applied (Figure 9 left). It predicts a hardness value of 361 HV for the test point (D) based on experimental data set A, B and C. The simulation data contains results for a much wider design space and the hardness predictions for the test point based on the simulation data is significantly higher (447 HV). Pooling the experimental and simulation data leads to more information about the variable space and would allow for higher order regressions. To assure comparability, the linear model was selected. The hardness expected for the test point is 434 HV. The prediction based exclusively on experimental data is closest to the test point value.

Figure 9: Linear regressions based on different data sets

4 Discussion
The procedure to enhance the machine learning algorithm with artificial data is practical but its quality is largely dependent on the accuracy of the simulation model. For press hardening this entails a highly complicated simulation set-up with a large amount of parameters. In our case, the model contains 82 material, interface and numerical parameters. Therefore, an efficient identification procedure is a key factor. The results of the process simulation model show that the model has no sensitivity regarding the computation of the hardness distribution within the part. There are two main reasons: 1) the model does
not contain all relevant dependencies such as a pressure-dependent heat transfer coefficients and 2) the actual surface of the tool might differ from CAD data. As for machine learning, the predictions based on experimental data indicate a good estimation quality for the hardness at the test point. However, estimations based on the same data set for parameter configurations outside of the experimental data space are too extreme (575 HV for 25 sec of handling time and 2 sec of quenching time) and are not trustworthy. Expanding the data range with the simulation results seems to give a much more realistic value for points outside the original data space even with a linear regression only.

5 Conclusion

We consider the investigation as a starting point to apply more complex machine learning algorithms to a much larger amount of production data. We aim to apply simulation models to both expanding and balancing real world data. Combining both seems to be possible but leads to new challenges regarding the quality of the process simulation model and its parameters. Parameters for numerical control significantly impact the stability of the simulation and need to be adjusted with regard to the physical parameters. This is important for DOE calculations for a wide range of input parameters. In order to test the prediction of the machine learning algorithm, we will create larger experimental and numerical data sets. The simulation quality will be further improved by adding more detailed submodels for the contact area which also includes a scan of the actual tool surface. A standard procedure for parameter identification based on [13] will be implemented.

References

[1] Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning*. 2013.
[2] Robert J. Seely, Louis Munyakazi, John Haury, Heather Simmerman, W. Heath Rushing, and Thomas F. Curr. Demonstrating the consistency of small data sets. *BioPharm International*, 2003.
[3] James N. Miller. Tutorial review - outliers in experimental data and their treatment. *Analyst*, 118:455–461, 1993.
[4] Roger Koenker and Kevin F. Hallock. Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156, December 2001.
[5] Kun-Lun Li, Hou-Kuan Huang, Sheng-Feng Tian, and Wei Xu. Improving one-class svm for anomaly detection. In *Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE Cat. No.03EX693)*, volume 5, pages 3077–3081 Vol.5, Nov 2003.
[6] Varun Chandola, Arindam Banerjee, and Vipin Kumar. Anomaly detection: A survey. *ACM Computing Surveys*, 2009.
[7] F. Provost. Machine learning from imbalanced data sets 101. In *AAAI 2000 workshop on imbalanced data sets*, 2000.
[8] Jie Sun, Jie Lang, Hamido Fujita, and Hui Li. Imbalanced enterprise credit evaluation with dte-sbd: Decision tree ensemble based on smote and bagging with differentiated sampling rates. *Information Sciences*, 425:76 – 91, 2018.
[9] D. Landgrebe, J. Schönherr, N. Pierschel, S. Polster, A. Mosel, and F. Schieck. New approaches for improved efficiency and flexibility in process chains of press hardening. *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)*, pages 2A–2015, 2015.
[10] Paul Akerstrom. *Modelling and Simulation of Hot Stamping*. PhD thesis, LuleåUniversity of Technology, 2006.
[11] Bernd Hochholdinger. *Simulation des Presshärteprozesses und Vorhersage der mechanischen Bauteileigenschaften nach dem Härten*. PhD thesis, ETH Zürich, 2012.
[12] Arthur B. Shapiro. Using ls-dyna for hot stamping. In *7th European LS-DYNA Conference*, 2009.
[13] Steffen Schroeder, Bernd Kauschinger, Arvid Hellmich, Steffen Ihlenfeldt, and Damien Phetsinorath. Identification of relevant parameters for the metrological adjustment of thermal machine models. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, Jan 2019.