INTERNATIONAL DESIGN CONFERENCE - DESIGN 2014
Dubrovnik - Croatia, May 19 - 22, 2014.

GAUSSIAN PROCESS BASED APPROACH FOR AUTOMATIC KNOWLEDGE ACQUISITION

T. Breitsprecher, P. Kestel, C. Dingfelder and S. Wartzack

Keywords: knowledge based engineering, knowledge acquisition

1. Introduction

Saving resources is a crucial objective in the 21st century and is motivating intensive work in various research fields such as renewable energy, hybrid (or fully electric) vehicles or smart electrical grid. Lightweight design is another example that is particularly important for the automotive sector and can be achieved by different approaches: Decreasing the volume of specific parts (key components) or increasing the functional density per part are only two possibilities [Mallick 2010]. However, these lightweight approaches place new demands on key components which have to operate at their limiting capability due to increased loads and/or stresses. For example, the synchronizer rings in vehicle transmission units which feature precisely arranged gear teeth are in general made of brass. But to cope with higher loads they should in future be made of steel to benefit from higher strength and improved wear resistance [Song and Jie 2008]. In many cases current production processes are only able to deliver these new high performance components by means of many different sub-process steps and thus at high costs [Merklein and Schneider 2011]. This motivates the research for new forming processes to produce high quality sheet metal components with heavily loaded functional elements. One possibility is to apply bulk forming operations to sheet metals which has led to a new class of forming processes with the overall designation sheet-bulk metal forming (SBMF) [Merklein and Allwood 2012].
In order to establish this new technology the potential of SBMF has to be revealed. Most of this potential resides in the substitution of existing manufacturing processes by SBMF and thereby decreasing the actual use of resources. However, in the sense of integrated product development (e. g. according to [Andreasen and Hein 1987] this requires an early acquisition of manufacturing-related, design-relevant knowledge and its implementation within the design process (see Figure 1).

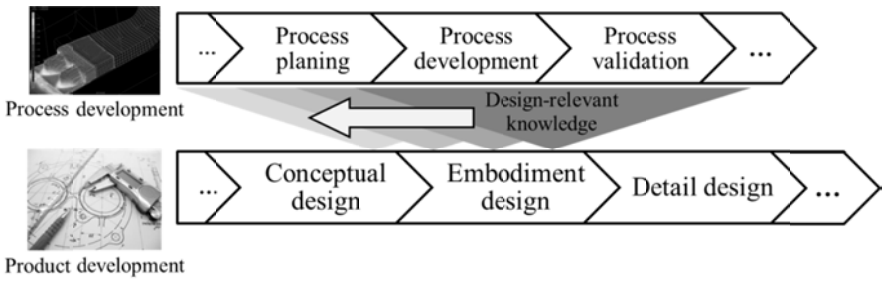

Figure 1. Shifting the acquisition of design-relevant knowledge towards earlier phases is the overall research challenge

This objective is pursued through the development of the **S**elf-**L**earning Engineering **A**ssistance **Sy**stem [Breitsprecher and Wartzack 2012] (referred to as SLASSY) that supports the design engineer

during the design process of sheet-bulk metal formed parts. The knowledge necessary for the analysis of a product regarding its manufacturability is acquired automatically and stored in SLASSY's knowledge base. In summary, the development of the self–learning engineering assistance systems addresses the well–known challenge of knowledge acquisition in the field of expert systems. The term *self-learning* refers to the implemented Knowledge Discovery in Database (KDD) process which uses data from the manufacturing process development [Röhner et al. 2011] (in this case SBMF). The data sets are created by parameter variation studies carried out with (virtual) process simulation or real experiments. Here a conflict arises: While the process engineer tries to decrease the number of simulation or experiments to decreases costs and use of resources the KDD process requires as much data as possible to extract reliable design-relevant knowledge from this data. A possible solution is to search for KDD algorithms that can increase the retrieved information per data set, thus improving the knowledge acquisition. This paper reports about the automatic acquisition of manufacturing-related and design-relevant knowledge by utilizing Gaussian processes, a method for statistical modelling. To offer the necessary background knowledge it starts with a brief description of sheet-bulk metal forming and presents the actual development stage of SLASSY with a strong emphasis on the implemented KDD-based self-learning process (Chapter 2). It continues with some basics of Gaussian processes (Chapter 3) and shows their use for the purpose of automatic knowledge acquisition with a SBMF-related case-study (chapter 4). Finally, a conclusion and an outlook are presented.

2. Background and related work

2.1 Sheet-bulk metal forming

The manufacturing technology sheet–bulk metal forming (SBMF) will be developed within the transregional collaborative research centre 73 (SFB/TR 73), funded by the German Research Foundation (DFG). This technology will unite the advantages of sheet and bulk metal forming processes to manufacture geometrically complex parts with variants and functional elements from thin sheet metal through forming. The objective is to manufacture these high–precision elements with close geometrical tolerances in which the geometrical details of the variants are in the range of the sheet thickness. The variants to manufacture are carriers and gearings derived from synchronizer rings and seat slide adjusters. The manufacturing of such variants out of sheet metals requires the overlapping or the sequence of 2– and 3–axis strain and stress states. To realize this, various sheet and bulk metal forming processes have to be combined [Merklein and Allwood 2012]. For the development of SBMF processes, the process combinations "deep drawing – upsetting ", "deep drawing–extrusion" and "cutting–deepdrawing" will be investigated within (SFB/TR 73). In this paper, the process combination "deep drawing–extrusion" will be used as an application example. This combination aims at the manufacturing of a part similar to synchronizer rings (Figure 2, left upper corner).

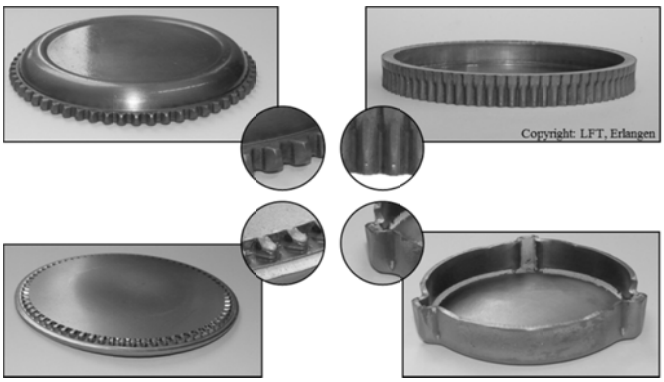

Figure 2. Exemplary sheet-bulk metal formed parts with magnified secondary design features

2.2 SLASSY at a glance

SLASSY is an engineering assistance system developed for the purpose of helping the product developer to design parts that are to be manufactured by sheet-bulk metal forming. The assistance is in

accordance to WEBER's understanding of the design process that consists of iterations between the phases *synthesizing* and *analysing* [Weber 2005]. The synthesis step is supported by offering feature elements both for the primary design features (PDF) and the secondary design features (SDF) to the design engineer. It furthermore enables him to analyse such a part regarding its manufacturability. During this analysis specific target values that allow an assessment of the manufacturability are predicted, based on the PDF and SDF geometry parameters. Such values might be the forming force necessary to form the designed part, the total equivalent plastic strain of the material after the forming process or the contact ratio between the material and the forming cavity. The prediction is based on mathematical models which represent the explicit form of the manufacturing-related and design-relevant knowledge [Röhner et al. 2011] and are acquired by means of a KDD process. Each metamodel is assigned to a specific product instance of a sheet-bulk metal formed part whereas each variant has specific target values (results) to assess its manufacturability. The architecture of SLASSY is shown in Figure 3 and explained in detail in [Breitsprecher and Wartzack 2012]

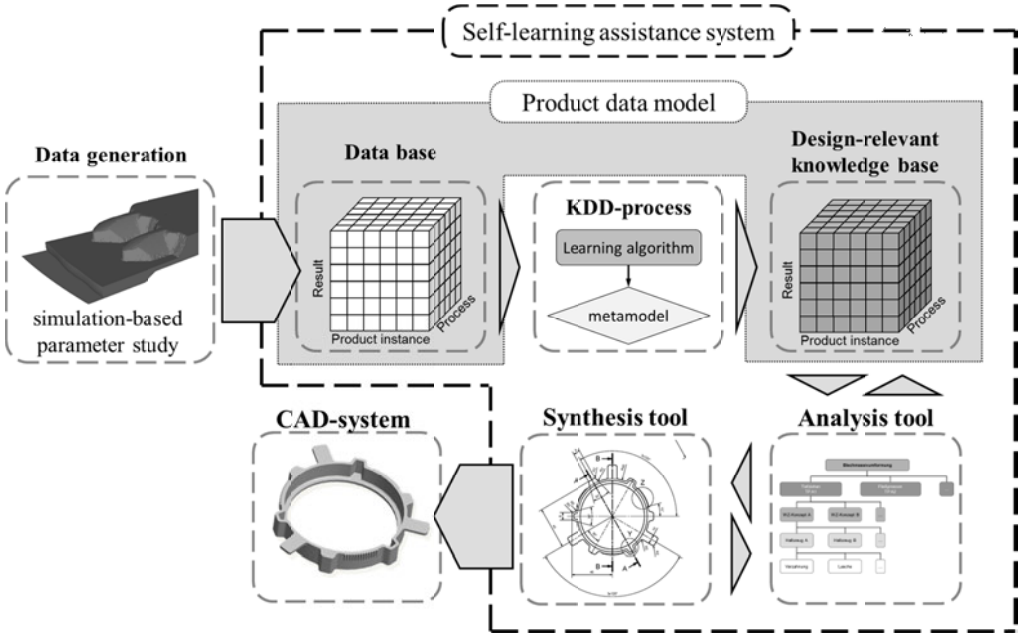

Figure 3. The overall architecture of SLASSY with product data model, KDD-process based knowledge acquisition tool, synthesis and analysis tool and interface to the CAD-system

As mentioned the data for the KDD process is derived from SBMF simulations or experiments with parameter variation studies which require simulation and/or experiment models. Therefor the predicting mathematical model can be regarded as the "model of a model", whereas the more common term is metamodel. The self-learning process itself is described in detail in [Breitsprecher and Wartzack 2012].

The result of the self-learning process is a single metamodel that has been chosen among a variety of metamodels. It is characterized by the fact that it shows the least prediction error, given a set of PDF/SDF geometry parameters. However, this prediction error depends largely on the size of the data set that was used for the self-learning process [Witten and Eibe 2011], [Röhner et al. 2011]. Common metamodels that SLASSY can acquire from manufacturing data so far are [Witten and Eibe 2011], [Evans 1996]:

- linear and polynomial regression function
- the W-M5Rules model, a decision list for regression problems that has been generated using separate-and-conquer algorithms
- the W-M5P model, that bases on the M5 model tress of QUINLAN and is equipped with a special pruning algorithm
- artificial neural networks, a special computational model that can be used to solve prediction problems

However, the potential of using different metamodelling techniques and thus making the knowledge acquisition component more flexible could only be revealed if SLASSY was able to choose the best fitting learning algorithm by itself without the intervention of the designer or a third party. A KDD-process model for that purpose was not available so it was developed within this research project and implemented in SLASSY [Breitsprecher and Wartzack 2012].

3. Gaussian processes for machine learning

A Gaussian process in general is a multivariate Gaussian distributions consisting of an infinite amount of random variables, any finite number of which are normal distributed. Therefore this stochastic process is fully specified by a mean and a covariance function. In a one-dimensional normal distribution with one input and one output quantity overall only scalars are given for the mean value and covariance. For higher dimensions the normal distribution is defined by a mean vector of and a covariance matrix whereas the Gaussian process is defined over functions: For the function f which is distributed as a Gaussian process with mean function m and covariance function k we write

$$f \sim GP(m,k) \tag{1}$$

This function shall now be demonstrated by an example. An argument x is used to index the individual random variables $f(x)$ in an infinite vector. The Gaussian process is defined by the functions

$$m(x) = \frac{1}{4}x^2 \text{ and } k(x,x') = \exp\left(-\frac{1}{2}(x-x')^2\right) \tag{2}$$

In order to illustrate the behaviour of the resulting stochastic function we generate some samples for a distinct finite number n of random variables. Therefor the index set is enumerated by the natural numbers for a given domain resulting in a regular Gaussian distribution defined by the vector of mean and covariance matrix:

$$m(x_i) = \frac{1}{4}x_i^2 \text{ and } k(x_i,x_j) = \exp\left(-\frac{1}{2}(x_i-x_j)^2\right); \quad i,j \in \mathbb{N}^+ \wedge i,j < n \tag{3}$$

Indexed by the argument x_i the normal distributed function values $f(x_i)$ are stored in a random vector

$$\mathbf{f} \sim \mathcal{N}(\mathbf{m},\mathbf{K}) \tag{4}$$

For $n = 3$ variables this vector results in three samples (see Figure 3), generated in MATLAB. Note that these "point clouds" will look different if they were drawn a second time, a third one and so one. Although one could think of three different functions, the course of each point cloud shall not be interpreted as such. Imagine it rather as lots of Gaussian distributions at different values of x_i. For $x_i = -1$ this distribution consist for example of the three values $f(x) = [0.05; 0.8; 2.4]$ (see arrows in Figure 4). For a higher number of n each distribution will consist of more samples.

To use Gaussian process for machine learning at first a prior has to be determined. Based on the mean and covariance functions the task of a prior is to define assumptions and properties of a stochastic function independent of training data, e.g. the course of the function in figure 1 seem to be quadratic. Subsequently the prior has to be adapted to a set of training data. After the Gaussian process has been trained with this data set, the posterior defines a mean function and the covariance matrix. This covariance function can be tuned via specific (mathematical) properties. Leaving mathematical sophistications aside the problem of learning in Gaussian processes is exactly the problem of finding suitable properties for the covariance function.

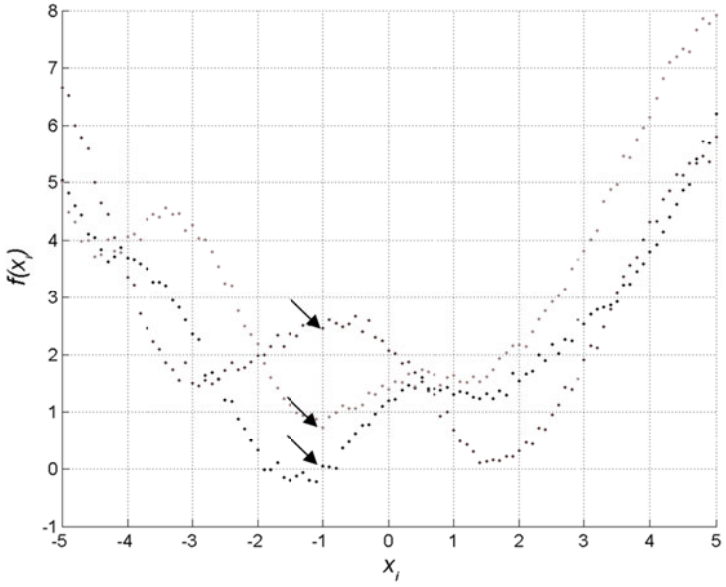

Figure 4. Three resulting random function values of the given Gaussian distribution

Now what brings in issues of machine learning? The basic challenge of machine learning is simple: We assume a given data set D of z observations from a parameter variation study (with z runs). This set consists of input parameters x_j that are changed during the variation study and output parameters y_i which represent the system's behaviour. $D = \{(x_j;\ y_j)|i = 1, \dots, z\}$ is now our training data and we want to make predictions y_* for the system's behaviour at new input parameters x_* which are not part of D.

The prediction with a posterior defined Gaussian process is, however, interpreted as follows: For the example in Figure 3 we assume, that the training data did not contain a data tupel at this point ($x_i = -1$), that is, we now want to predict the value for f(x) at ($x_i = -1$). The prediction result we will get would simply be the mean of the Gaussian distribution, which is defined through the Gaussian process at $x_i = -1$. In addition a Gaussian process based prediction always delivers a standard deviation at the point of prediction.

The utilization of Gaussian processes for this contribution covered the development of a machine learning tool that uses a Gaussian modelling toolbox and, more important, to ease the use for the design engineer our tool can automatically find optimal mean and covariance function specification for a given training data set [Rasmussen and Williams 2006].

Gaussian processes have been the basis of solutions to various engineering problems. [Yu et al. 2013] used Gaussian processes for long-term wind speed forcasting. By integrating local Gaussian process models and combining them via Bayesian inference, they are able to ensure a forecasting with high accuracy and model flexibility. Further research works try to improve the efficiency of combustion machines by evaluating combustion experiments [Chen et al. 2013].

4. Use case: automatic acquisition of design-relevant manufacturing knowledge

In this use case the simulation data of the process combination "deep drawing – extrusion" will be analysed for the automatic acquisition of design-relevant knowledge via our Gaussian process machine learning tool. This data arose by a parameter study on the basis of a three-dimensional FE-model, varying merely the geometric parameters of the teeth. For reducing the computation time, just a 10° sector of the demonstrator was modeled justified by the rotational symmetry of the demonstrator (Figure 5). Theoretically regarding the symmetry condition, the modeling of a half tooth would be sufficient, but therefore the possibility concerning the verification of a faultless model would be dropped. A model can be recognized as faultless if both teeth are shaped nearly identically and the stress and strain values of both teeth are equal to each other. Furthermore, the modeling of the tool design as rigid bodies reduced the computation time. The friction conditions in the FE-model were

defined differently. Between the contact bodies blank and die a friction factor of 0.3 was defined. The definition of the friction factor between the contact bodies blank and punch represents a special characteristic because the punch was partitioned into two sectors to influence the mould filling and the punch force. The sector "punch 1" was assigned with a friction factor of 0.05 for the increase of the mould filling, whereas the sector "punch 2" was assigned with a friction factor of 0.3 for the reduction of the punch force as well as for the increase of the mould filling. The semifinished part was a circular blank with a thickness of 2 mm and with the material characteristics of DC04.

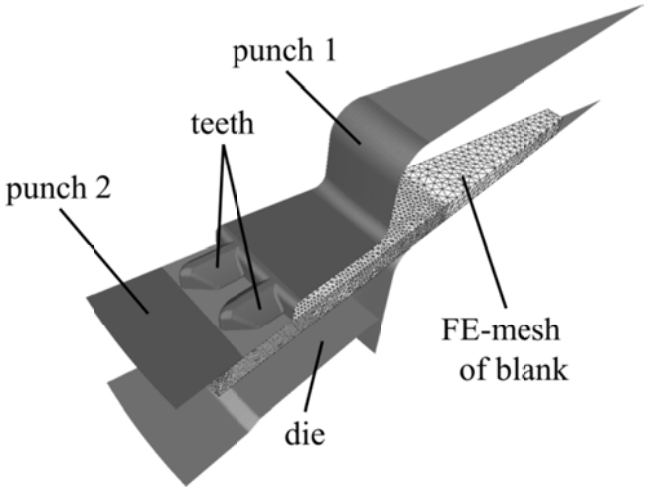

Figure 5. The FE-model of the simulated use case manufacturing process

The input parameters x_j for this parameter variation study were a selection of product characteristics that can be influenced by the design engineer (see Table 1). As an exemplary simulation output (y_j) for our use case we chose the forming force that was calculated by the simulation system (Simufact Forming®) for each simulation run. Table 1 shows an overview of the input and output parameters. The index "X" denotes that this is a parameter from a SDF, the index "T0" denotes the SDF "Teeth variant 0". There are further variant of SDF which are analysed within different studies.

Table 1. Input and output parameters of the case study

Parameter	Parameter meaning	Unit	Input / output	Mean w.r.t. the data set	Standard deviation	Range
X_T0_A0	addendum angle of tooth	°	input	55.833	3.128	[52.5 ; 60.0]
X_T0_W0	tooth width	mm	input	2.500	0.410	[2.0 ; 3.0]
X_T0_L0	tooth length	mm	input	2.750	0.251	[2.5; 3.0]
X_T0_R1	top radius	mm	input	0.600	0.246	[0.3 ; 0.9]
X_T0_R2	radius at tooth addendum	mm	input	0.367	0.170	[0.2 ; 0.6]
F_{form}	forming force	kN	output	1905	287.93	[2567 ; 1233]

We created a total data set of 162 samples with Design for Experiment methods [Fisher 1990] whereas some simulations did not converge due to typical errors (e.g. re-meshing problems, lost contact between punch and blank mesh) and that gave us a final data set of 151 samples.

To show the advantage of Gaussian processes for acquiring manufacturing-related and design-relevant knowledge from this data we set up a KDD-process in RapidMiner® to train common metamodels that have been used within the self-learning component of SLASSY. The KDD-process includes an operator which allows defining a sample ratio. At a sample ratio of $s = 1.0$ all 151 samples from the final data set are used for the KDD-process training. Decreasing this ratio, i.e. to $s = 0.8$, means that a sub-data-set with a decreased size, i.e. 121 samples, is randomly drawn from the final data set and used for the training process. The sample ratio is decreased iteratively until it reaches a minimum of $s_{min} = 0.1$. Each sub-sample is used to train both the common metamodels in RapidMiner® and a Gaussian process with our implemented tool.

To evaluate the prediction quality of each iteration we use the procedure of k-fold cross-validation. The k-fold cross-validation splits the dataset randomly into k approximately equal disjoint subsets. The learning procedure is executed r-times in which per procedure one subset is used for testing and the remainder is used for training. Finally, the resulting k error estimates are summarized and averaged to the total error estimate [Tan et al. 2006]. According to [Vercellis 2009] the recommended value for k is 10. For a more reliable error estimate, the 10-fold cross-validation is repeated ten times which leads to a ten-times ten-fold cross validation [Witten and Eibe 2011], [Küstner et al. 2013]. As error estimate we propose the root mean squared error (RMSE) because it is a good quality measure to compare prediction errors of different metamodels for a particular output [Evans 1996].
Figure 6 shows the root mean squared error of our common metamodels and the Gaussian process trained with the implemented tool.

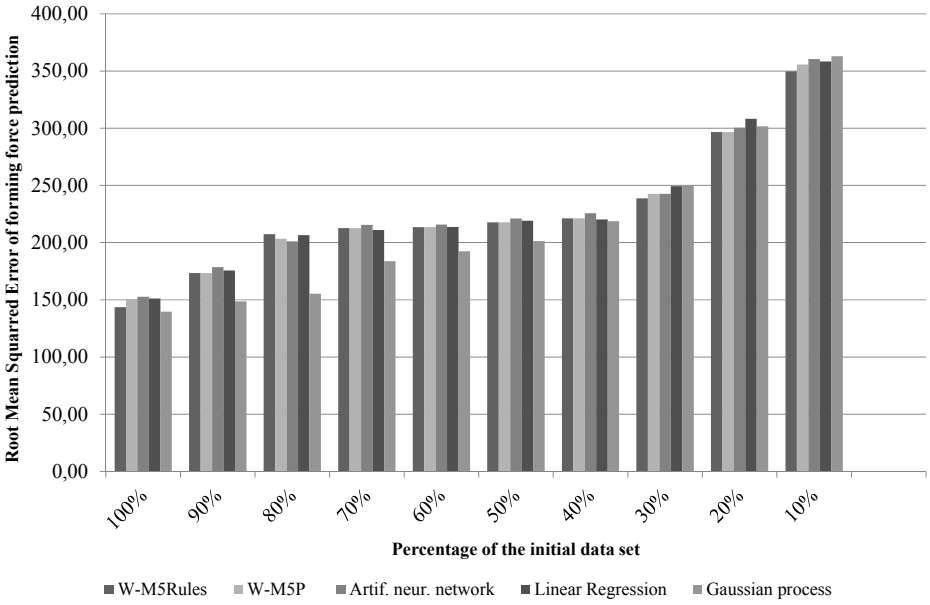

Figure 6. The RMSE of different metamodels for decreasing data sets shows that Gaussian processes intend to offer a smaller prediction error

It can be seen that the prediction quality of every metamodel decreases while the sample ratio decreases and Gaussian processes are basically no exception from this tendency. However, eye-catching is that Gaussian processes tend to offer a smaller prediction error while our common metamodels keep stable in the area of $0.8 \leq s \leq 0.4$. Within this area our Gaussian process tool tries to fit the model to the decreased data set by varying the prior and thus enabling a dynamic adaption. Only for 30% of the initial data set size (and less) the Gaussian process joins up to the prediction error of the other metamodels and performs quite bad, too. Especially the behaviour in the medium range motivates us to include Gaussian processes in order to make SLASSY and its knowledge acquisition component more flexible.

5. Summary and outlook

Within the research for new metamodelling techniques the so called Gaussian processes showed potential to acquire reliable prediction models even from small data sets. With some effort they were utilized and a computer-aided tool was developed to automatically train a Gaussian process. We used our tool for different simulation data whereas a data set from the SBMF-process "deep drawing – extrusion" was taken for a short case study to compare some common metamodels and a trained Gaussian process. We could show, that the Gaussian process was able to predict with a higher accuracy for decreased data sets until the sample ratio reached a critical value. At this point all metamodels showed a high prediction error.

An interesting property of Gaussian processes motivates further research that can decrease the effort for the manufacturing parameter variation studies. Beside the predicted value e.g. for the forming force, a Gaussian process outputs an uncertainty of the prediction at the specific point where the prediction was made, that is, we have local prediction error. This separates a Gaussian process from other metamodels which always offer a global prediction error. The derived research question for our further work is now: Is it possible to find a position where the local prediction error of the Gaussian process model has a maximum and to derive a set up for the input values at this point? This would lead us to a method for adaptive sampling where the simulation set up is not fixed from the beginning but is dynamically adapted for each simulation or experiment run. Loosely speaking the information retrieval per simulation is improved and reliable metamodels are acquired with less simulation.

Acknowledgement

This work was supported by the DFG within the scope of the Transregional Collaborative Research Centre on sheet-bulk metal forming (SFB/TR 73).

References

Andreasen, M. and Hein, L., "Integrated Product Development", Bedford IFS Publications, 1987.

Breitsprecher, T., Wartzack, S., "Architecture and realization of a self–learning engineering assistance system for the use within sheet–bulk metal forming", Proceedings of the 9th Norddesign, Kyvsgård Hansen, P.; Rasmussen, J.; Jørgensen, K. A.; Tollestrup, C. (Ed.), Aarlborg University, Denmark, 2012, pp. 79-86.

Chen, J., Chan, L., Cheng, Y., "Gaussian process regression based optimal design of combustion systems using flame images", Applied Energy, Vol. 111, pp. 153-160, 2013.

Evans, J. D., "Straightforward Statistics for the Behavioral Sciences", Duxbury PressBelmont, 1996.

Fisher, R. A., "Statistical methods, experimental design, and scientific inference", Oxford University, 1990.

Küstner, C., Breitsprecher, T., Wartzack, S., "Die Auswirkung der Reihenfolge von Mess- und Simulationsdaten auf das Ergebnis der Kreuzvalidierung in KDD Prozessen", Design for X - Beiträge zum 24. DfX-Symposium 2013,19.-20. September 2013, Krause, D., Paetzold, K., Wartzack, S. (Ed.), 2013, pp. 175-186.

Mallick, P., "Materials, design and manufacturing for lightweight vehicle", Boca Raton CRC Press, 2010.

Merklein, M., Allwood, J., "Bulk forming of sheet metal", CIRP Annals – Manufacturing Technology, Vol. 61, No. 2, 2012, pp. 725–745.

Merklein, M., Schneider, T., "Functional Integration of Sheet Metal Parts by Deep Drawing", Proceedings of 10th International Conference on Technology in Plasticity, Aachen (Germany), 25.09.2011, Düsseldorf, Stahleisen, 2011, pp. 1036-1041.

Rasmussen, C. E., Williams, C. K., "Gaussian Processes for machine learning", MIT Press Cambridge,2006.

Röhner. S.; Breitsprecher, T.; Wartzack, S., "Acquisition of design–relevant knowledge within the development of sheet–bulk metal forming", Proc. of ICED11,Technical University of Denmark, 2011, pp 365-374.

Song, X., Jie, Z., "A Research on Steel Synchronizer Gear Ring's Precision Forging", Materials Science Forum, Vol. 575-578, 2008, pp. 226–230.

Tan, P., Steinbach, M., Kumar, V., "Introduction to Data Mining", Pearson Education, Boston, 2006.

Vercellis, C., "Business Intelligence: Data Mining and Optimization for Decision Making", Chichester, 2009.

Weber, C., "CPM/PDD – An Extended Theoretical Approach to Modelling Products and Product Development Processes", Proceedings of the 2nd German-Israeli Symposium on Advances in Methods and Systems for Development of Products and Processes, TU Berlin / Fraunhofer IPK, 07.-08.07.2005, pp. 159-179.

Witten, I. H., Eibe, F., "Data Mining – Practical Machine Learning Tools and Techniques", Morgan Kaufmann San Francisco, 2011.

Yu, J., Chen, K., Mori, J., Rashid, M. M., "A Gaussian mixture copula model based localized Gaussian process regression approach for long-term wind speed prediction", Energy, vol. 61, 2013, pp. 673-686.

Dipl.-Ing. Thilo Breitsprecher
Friedrich-Alexander University Erlangen, Chair of Engineering Design
Martenstraße 9, 91058 Erlangen, Germany
Email: breitsprecher@mfk.uni-erlangen.de