Virtual End-of-Line Test—Prediction of the Acoustic Behavior of Gearboxes Based on Topographic Deviations Using Neural Networks

Marius Willecke, Dr. Jens Brimmers, and Prof. Christian Brecher

Tooth contact analysis is an integral part of the gear design process. With the help of these simulation tools, it is possible to calculate the excitation caused by a tooth contact (Ref. 1). Usually, the load-free transmission error or the total transmission error under load is used for this purpose. However, the calculation with the tooth contact analysis ZAKO3D allows only a quasi-static consideration of the excitation. To better evaluate the behavior in the overall system, it is therefore necessary to perform a dynamics simulation. However, the main disadvantage of such dynamics simulations is the much longer computing time compared to quasi-static tooth contact analyses due to the high computational effort.

In the context of measures such as the increase of resource efficiency as well as due to increasing demands of the end customers, the pressure on gearbox manufacturers to produce fewer rejects and at the same time to meet higher demands on the quality of the gearboxes is growing. To ensure that an assembled gearbox meets the acoustic requirements, tests are therefore carried out on the fully assembled units in end-of-line (EoL) tests (Ref. 2). If an anomaly is detected, the unit must be disassembled and overhauled as far as necessary or otherwise scrapped. It would be a

significant advantage if assemblies, that are relevant for the acoustic behavior, such as the tooth meshes, could already be examined for their excitation behavior before assembling them.

Theoretically, this would be possible by simulating the acoustic behavior of the gearbox in a dynamics model, as it is currently used in the gear design process, using the real topographies of the gears to be installed. However, the main problem here is the calculation time of these dynamics models, which prevents a calculation parallel to the manufacturing process time.

To solve this challenge, a way must be found to predict the excitation behavior of the system much faster. One approach that is to be investigated for this purpose is the usage of metamodels, which allow a considerably faster calculation of the target variables. It is relevant here that a description of the tooth flanks is used, that is as exact as possible to be able to determine a realistic assessment of the excitation behavior that arises from the two wheels in the tooth mesh.

Willecke et al. (Ref. 3) have already shown that it is possible to use a substitute model for the quasistatic description of the problem. The approach developed there offers the potential to be used for dynamic calculations as well.

State of the Art

The work in this publication is mainly based on the preliminary work of Willecke et al. (Ref. 3). There is also an explanation of the structure and the application of deep neural networks (DNN) for the use of topographic deviations as input values. In the state of the art of this publication, generative adversarial networks (GAN) are explained.

Extending Datasets Through Generative Adversarial Networks

GANs are a subfield of Machine Learning and belong to the area of unsupervised learning. The basic idea is the use of two competitive DNNs. The two networks are a generator network and a discriminator network, see Figure 1. The generator network takes random input values and generates data from them. The discriminator network takes this generated "artificial" data as well as a set of "real" data. Based on patterns in the "real" data, the network tries to decide whether the "artificial" data is fake or real data. If the discriminator can recognize the generator's "artificial" data as such, the generator is informed about this and subsequently tries to generate more realistic data. If the discriminator is not able to recognize the "artificial" data of the generator as such, the discriminator is informed

about it and improves itself afterward. This process continues until equilibrium occurs and the discriminator classifies $p = 50$ percent of the generated data as "artificial" and $p = 50$ percent of the generated data as "real" (Ref. 4).

Current challenges concerning GAN lie in achieving this converged state. Depending on the dataset, training method, and hyperparameters (see Figure 1), this is not always the case (Ref. 5).

GANs are often used in the generation or post-processing of images. Here, for example, based on a set of "real" images, new "artificial" images can be generated, or new "artificial" pixels can be created in an existing image, thus resharpening the image (Ref. 6).

Objective and Approach

The objective is to determine the characteristic values of the differential acceleration of a deviation-affected gear pair orders of magnitude faster than is possible with a conventional elastic multibody simulation model (EMBS) under load. Therefore, a DNN is developed that approximates the EMKS and achieves this time advantage. The method to be developed for calculating the characteristic values of the transmission error (TE) with this model is shown in Figure 2. It is assumed that each of the two gears is provided with a deviation. In the first step, these two deviation surfaces (AWF) of gear 1 and gear 2 are combined to a sum deviation surface ($\sum$ AWF). Thus, an equivalent tooth contact is created in which only one gear has deviations, and the other gear has an ideal involute. In the second step, the parameter pi is derived from the $\sum$ AWF. These parameters should be able to describe the $\sum$ AWF as accurately as possible with as small of a number of control points as possible. In the third step, the parameter pi is given as input to the model, which then returns the TE characteristics.

Corresponding to the individual steps of the method for calculating the TE characteristic values for two gears with deviations, the procedure for achieving this goal can also be divided into several sub-steps. In the first substep, the $\sum$ AWF must be set up. For this purpose, the already developed procedure according to Willecke et al. based on the approach of Brimmers can be used (Refs. 3, 7).

In a third substep, the actual DNN for calculation is developed. For this purpose, a training data set with $n = 3{,}000$ data points is created. First, a pool of variants is generated. These are not combined fully factorially but distributed with the help of a Latin hypercube experimental design so that fewer variants are needed to cover the complete variation space. An EMBS calculation under load is performed with each variant of this variant pool. The acoustic parameters are then calculated for the differential accelerations calculated in this way. For each calculated variant of the pool, not only the parameters pi as input of the model are known, but also the acoustic parameters as output of the model. With the help of the data set created in this way, the DNN can be trained. A dataset created in analog with $n = 480$ data points, which are explicitly not part of the training dataset, is used to validate the created network.

Subsequently, it will be investigated how well the trained DNN is suited to predict data outside the trained range and how sensitive the network reacts to a reduction in the size of the training dataset. The GAN represents one way in which the dataset could be extended and refined for this purpose. The extent to which GANs can be used to selectively increase the size of the dataset for training is being investigated. The goal is to generate additional realistic training data.

Development of a Dynamic Model Considering the Topological Deviations in the Simulation of the Complete System

In this section, a model is set up with which the acoustic excitation of an e-drive transmission can be investigated for various topographical deviations. *Simpack* 2023 from Dassault Systems is used as the software
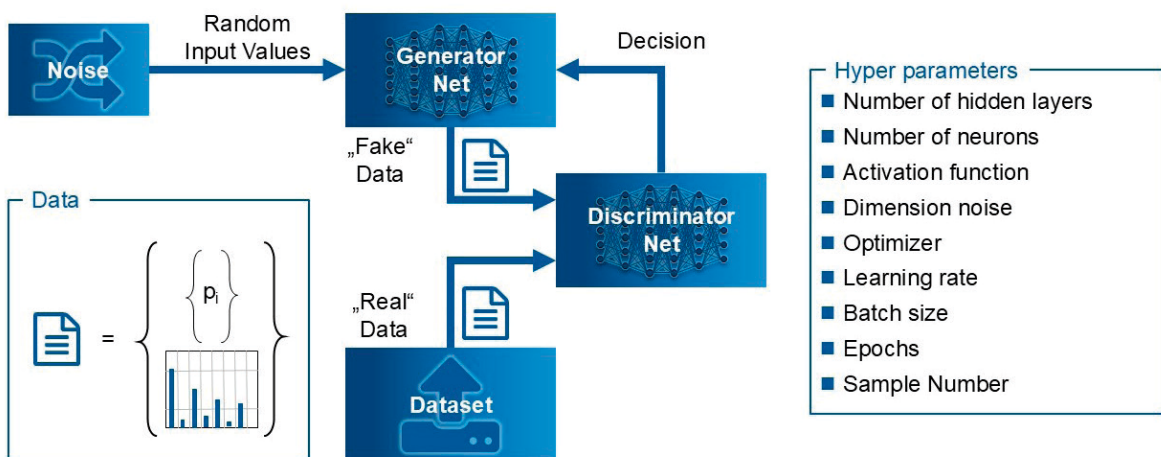

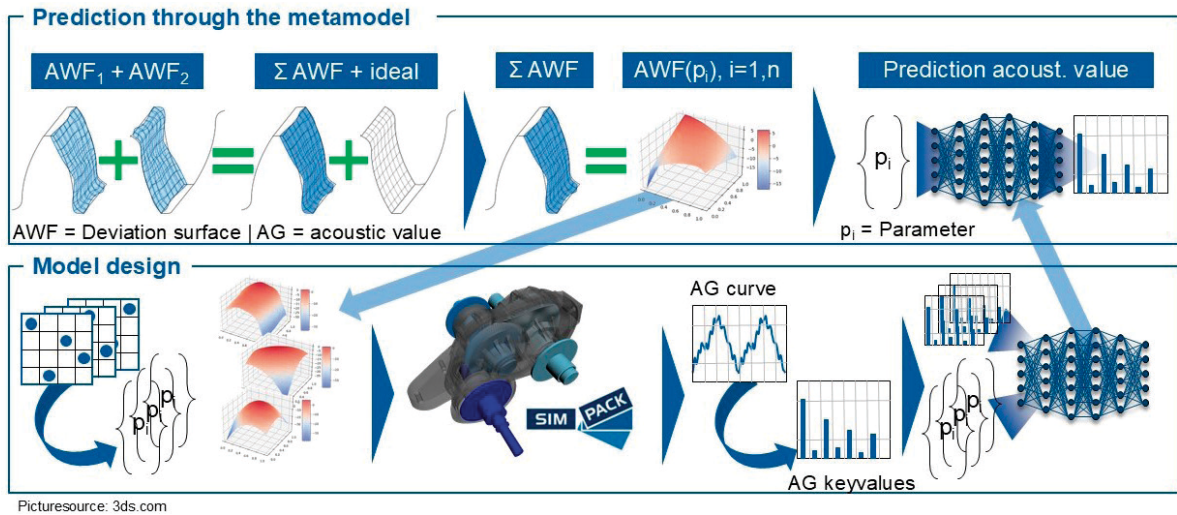*Figure 1—Structure of a generative adversarial network.*

Figure 2—Procedure for the construction of the deep neural network.

for the simulation environment. The exact use case is the high-speed demonstrator transmission of the WZL Gear Research Circle (Ref. 8). In the elaboration of this report, however, not both stages of the transmission are considered, but only the fast-rotating first stage. Experimental data on the test bench can be obtained for one stage as well as for two stages.

Modeling the Tooth Contact
The tooth meshing in the simulation model is modeled by the *Force Element 225: Gear Pair* (Ref. 9). The calculation method used is the DIN 3990 method B / Steiner approach. The number of slices in the width direction is limited to n_{Slices} = 5 for performance reasons (Ref. 9). The microgeometry can be specified in *Simpack* on the *Gear Pair* as a two-dimensional function of the amount of deviation over the tooth width and gear diameter. The deviation specifies the amount at the corresponding location that is removed from the ideal geometry. Negative values are not permitted. Therefore, no material can be added.

Modeling of Further Gearbox Components
In addition to the gearing, the other peripheral components of the transmission must also be modeled. These include all components that are in the power flow. The shafts, bearings,

and housing as well as the fixation of the housing are considered here. The modeling of the corresponding components is described below. The structure is based on the structure of a dynamics model for a two-stage e-drive transmission described by Willecke et al. (Ref. 8).

Shaft and Housing
Both the shaft and the housing are prepared as modally reduced bodies. To create the modally reduced bodies, the geometries are exported as STEP files from the CAD model and imported into the CAE program *Abaqus* 2021, where they are meshed (Ref. 10). For this purpose, the shafts are manually partitioned to be able to use hexahedron-shaped elements in as large a volume as possible. These elements allow finer meshing with the same number of elements compared to tetrahedron-shaped elements. Thus, better results can be achieved with the same computing power. Beam models are not suitable for that use case hence they can't model the asymmetric modal behavior caused by keyways. However, due to the complexity of the housing, it will be meshed with tetrahedron-shaped elements. To ensure the quality of the meshing, a FE convergence study is performed, see Figure 3, left side.

If the edge length of the elements is reduced from l_k = 5 mm to l_k = 3 mm, the frequency of the

first eigenmode changes by only Δf = 0.38 percent. It is thus assumed that with an edge length of l_k = 3 mm the modal behavior can be calculated with sufficient numerical stability. A further reduction of the edge length leads beyond the capacity of a node (RAM = 196 GB) of the used high-performance computer due to the resulting too-high number of elements, which cannot be solved any longer with the installed RAM. The modally reduced bodies are subsequently imported into *Simpack* and modeled there as linear flexible bodies. The eigenmodes are taken over from the calculation in *Abaqus*. Since all the components are still in the manufacturing stage, validation of the components modes was not yet possible. The model can later be updated with measuring data. This does not influence the following steps.

Bearings
The modeling of the bearings is done in *Simpack* with *Force Element 49: Bearinx Roller Bearing*. The force element is a user force extension of the company Schaeffler for *Simpack*. The bearings themselves are modeled in the Force Element based on maps. The maps were created by Schaeffler. The maps take the positions and velocities of the housing bore and the shaft as input and calculate the resulting reaction forces and torques.
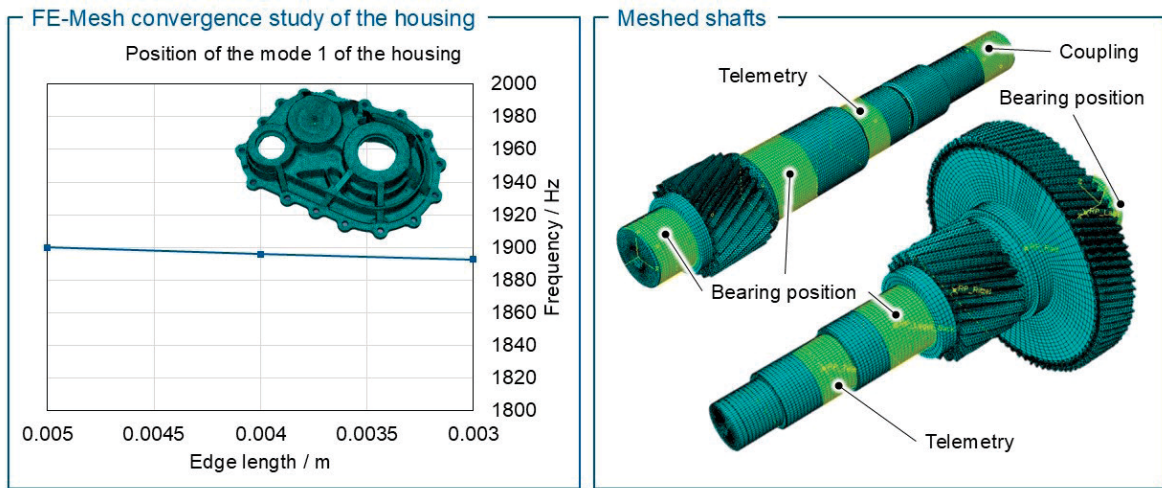
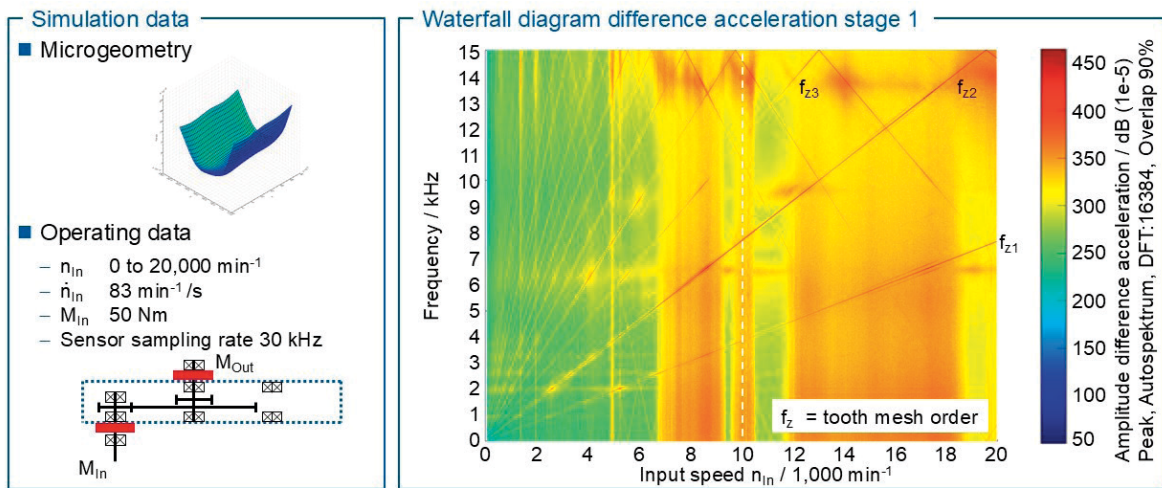Figure 3—Mesh convergence study and networked components.

Figure 4—Calculated waterfall diagram of the gearbox up to n_{in} = 20,000 min^{-1}.

Assembling the Model

The individual components created are connected in *Simpack* by different *Force Elements* and thus assembled to form the overall model. The application of the torque is realized in two components. The first component applies tension with constant torque, which is balanced by the ratio of the transmission and thus does not lead to any acceleration of the drive train. The second component is imposed by a PI controller on the drive shaft. This controller adjusts the drive train to the specified speed curve. To drive the model efficiently, the control parameters of the controller must be set sensibly. To do this, the P component is first set to be minimally subcritical (K_p = 0.025)

and then the I component is also set (K_I = 0.05).

The overall model created in this way can now be operated at any desired operating point. To obtain an overview of the behavior of the gearbox in the operating range, the gearbox is clamped at a torque of M_{in} = 50 Nm and then a pinion speed ramp-up from n_{in} = 0 min^{-1} to n_{in} = 20,000 min^{-1} is simulated. The calculated waterfall diagram of the speed ramp-up is shown on the right side of Figure 4. The placement of the two rotational acceleration measurement systems is marked in red on the left side of Figure 4. In the calculated waterfall diagram three significant areas can be seen where the tooth meshing frequencies go into resonance with natural frequencies of the overall

system, resulting in increased response of the system. The speed n_{in} = 10,000 min^{-1} is selected as the operating point for further investigations, which is in the middle of the speed range of the gearbox. It is easy to see that this rotational speed lies in a range of increased resonance behavior of the gearbox and thus increased response of the system is to be expected at this operating point.

For further investigations, the simulations are now carried out at the operating point. Several phases are provided for moving the model cleanly from standstill to the operating point and performing measurements there. These are shown in Figure 5. In phase I from t_{model} = 0 s to t_{model} = 0.1 s, the load with which the
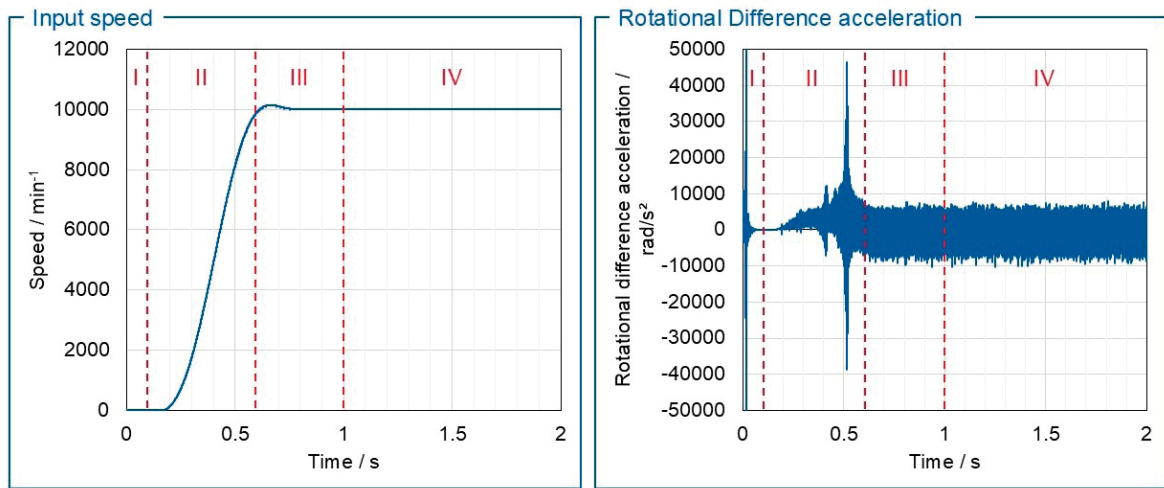
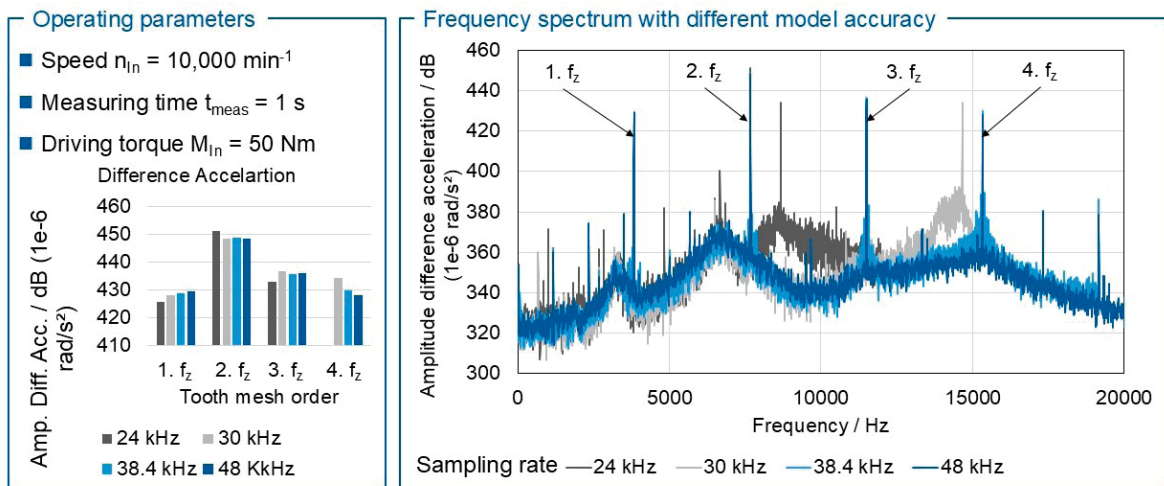**Figure 5—Time histories with sectors at the selected operating point.**

Figure 6—Influence of the considered modes on the system behavior.

system is braced is applied to the stationary system at a constant load change rate of $\dot{M}_{in}$ = 500 Nm/s. The load increases from M_{in} = 0 Nm to M_{in} = 50 Nm. The tension torque is constant in the further phases.

In phase II from t_{model} = 0.1 s to t_{model} = 0.6 s, the target drive speed n_{in},target, which is the setpoint of the PI controller, is increased with a constant target speed change rate of $\dot{n}$ = 20,000 min⁻¹/s. This results in a torque imprint by the controller that accelerates the model from rest from n_{in} = 0 min⁻¹ to n_{in} = 10,000 min⁻¹. In phase III from t_{model} = 0.6 s to t_{model} = 1.0 s, no external variables are changed, and the model is given time to settle. In the course of the speed on the left side of Figure 5, the overshooting of the speed, caused by

the discontinuity, in the acceleration specification can be seen well at the beginning of phase III. The model's behavior is stabilized in phase IV from t_{model} = 1.0 s to t_{model} = 2.0 s. The time course of this phase is used for the measurement, which is later used to calculate the acoustic characteristic values.

The eigenmodes of the flexible bodies were calculated up to a maximum eigenfrequency of $f_{eig,max}$ = 20 kHz to cover the entire human hearing range. However, since only the amplitudes of the tooth meshing orders are of interest for further evaluation, it is investigated whether modes can be removed to save calculation time without significantly changing the result. Therefore, at the intended operating point for the EoL, see Figure 6 top

left, different sampling rates as well as maximum considered natural frequencies were investigated.

The right side of Figure 6 shows the frequency spectra at the selected operating point for four different sampling rates. The maximum considered natural frequency is in each case 0.42 times the sampling frequency to comply with the sampling theorem according to Shannon (Ref. 11). The amplitudes of the marked tooth meshing orders are shown separately in the bar chart at the bottom left of Figure 6 for better comparability. It can be seen well that at a sampling frequency of down to f_{sample} = 30 kHz there is no significant change in the amplitudes. No significant differences can be seen here in the entire frequency
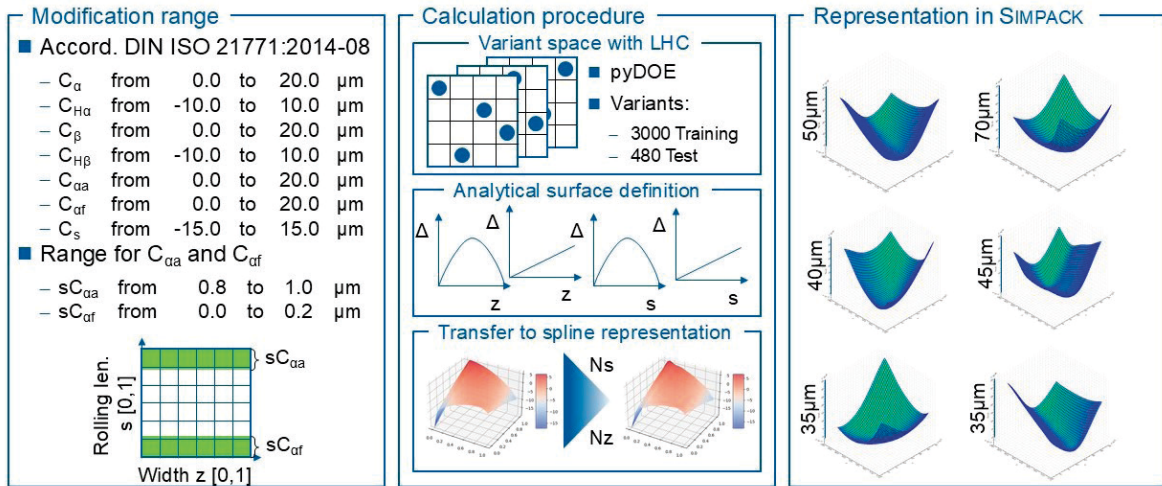
Figure 7—Modification range and application of the modifications in Simpack.

spectrum either. Only at a sampling frequency of f_{sample} = 24 kHz significant differences become visible. In the following, a sampling frequency of f_{sample} = 30 kHz and a maximum natural frequency of $f_{eig,max}$ = 12.5 kHz are used. The calculation time can thus be reduced from t_{sim} = 21.5 min to t_{sim} = 6 min by p = 72 percent.

Development of the Surrogate Model to Predict the Acoustic Characteristic Values

This chapter explains how the neural network is built. For this purpose, it is first described how the necessary data sets, which are later used for the generation and validation of the network, are generated. Furthermore, it is explained how the hyperparameters for the design of the network are selected. Subsequently, the validation of the prediction of the neural network is performed. For this purpose, the ability of the neural network to make predictions about previously unknown data is evaluated.

Variant Calculation with the Dynamic Model to Generate the Training and Test Dataset

The creation of the neural network, which is used as a surrogate model for the prediction of the acoustic parameters, requires a training data set. In addition, a test dataset is needed later for the validation of the prediction quality. Both data sets consist of different variants of topographic deviations. It should be noted that the variants of the test dataset are not part of the training dataset. The topographies are generated by free spline surfaces with N_z x N_s = 3 x 8 grid points out of analytical surfaces (Ref. 3). The procedure, see Figure 7 middle frame, is described by Willecke et al. (Ref. 3). To obtain freeform topographies as realistic as possible, they are generated from a superposition of standardized modifications (Ref. 3). According to Willecke et al. (Ref. 3), the range of these modifications is shown in Figure 7 on the left side.

The analytical surface is then converted into a free-form surface. This free-form surface is the topography that is subsequently used for the tooth mesh. To be able to imprint the topography on the *Gear Pair* in *Simpack*, it must be converted beforehand, see the section "Modelling the Tooth Contact." The right side of Figure 7 shows examples of some topographies and their processing in *Simpack*.

The data is transferred to *Simpack* as an "Input Function afs File" (Ref. 9). In this file the topography is stored with a resolution of 20 x 20 interpolation points. The intermediate points of the topography are mapped with a bivariate spline with two dimensions in both directions. A total of n_{var} = 3000 variants for the training data set and n_{var} = 480 variants for the test data set are generated with a Latin hypercube.

The boundaries are designed to be wider than a possible occurring manufacturing deviation. The variants are subsequently simulated in *Simpack*. A variant of the training data set computes on average t_{sim} = 965.16 s and a variant of the test data set computes on average t_{sim} = 943.67 s. The generation of the training data set thus takes t_{sim} = 2,895,480 s, which corresponds to d = 33.5 days of computing time, and the generation of the test data set requires t_{sim} = 463,277 s, which corresponds to d = 5.24 days of computing time. However, the calculations can be performed in parallel, so that the real time required can be reduced by the factor of the parallel calculations if the corresponding IT infrastructure is available.

The scatter of the simulation time of the individual variants in the data sets is shown on the left side of Figure 8. All calculations are performed as single-core calculations on a workstation (CPU: i7-8700 / 64GB Ram).

However, n_{par} = 10 variants are calculated simultaneously, since the time the user has to wait for the results can be reduced efficiently this way. After the variant calculation is completed, the differential acceleration between t_{model} = 1 s and t_{model} = 2 s in the time domain is calculated for each variant. This time signal forms the basis for the calculation of the acoustic parameters. The characteristic values are used, whose applicability Willecke et al. could already
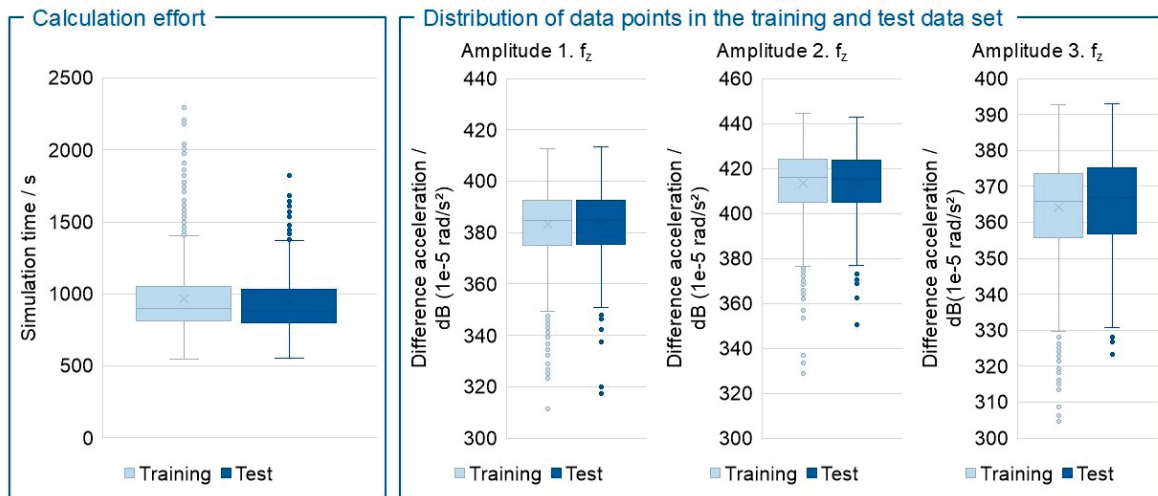
Figure 8—Distribution of acoustic parameters in the training and test data set.
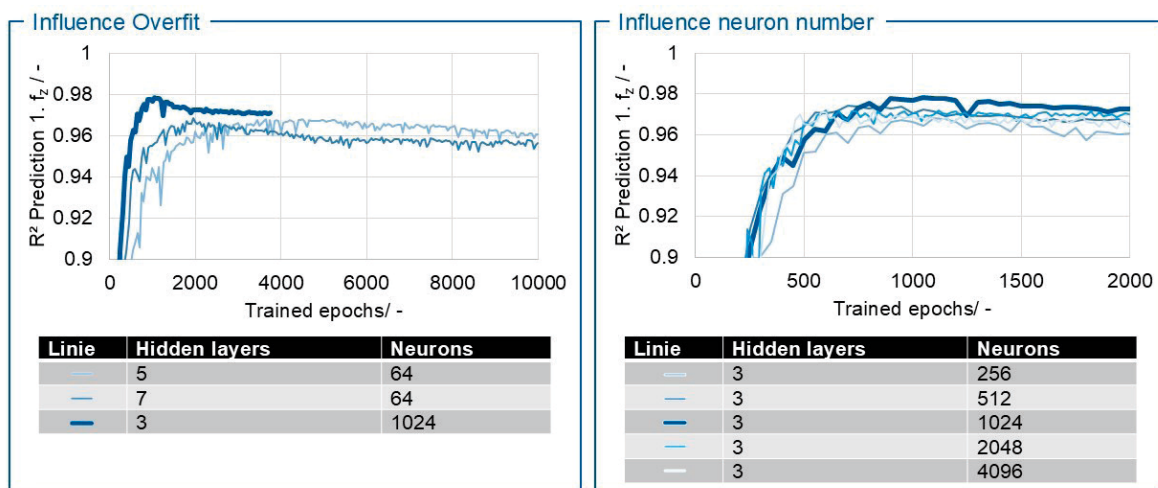

Figure 9—Effect of overfitting and influence of neuron number.

prove in the quasistatic tooth contact analysis (Refs. 3, 12).

In the following, the amplitudes of the first three tooth meshing orders are used for the elaboration. The distribution of the amplitudes of the individual tooth meshing orders over the variants is shown on the right side of Figure 8. It is easy to see that the amplitudes of most variants move in a narrower band and there are only a few outliers in the low amplitude range. The position of the mean values and quartiles for the training and test data set is similarly distributed, so that it can be assumed that the test data set is also representative of the area of the network for which the training data set provides input values.

Optimal Choice of Hyper-Parameters for the Neural Network

The neural network is characterized by various parameters. These parameters, which are basically the settings of the network, are called hyperparameters. To find a starting point for the selection of the hyperparameters, first the same was chosen, which led to good results in a network for the prediction of the characteristic values of a quasistatic tooth contact analysis (Ref. 3). A square-error function was used here since it is suitable for regression problems (Ref. 13). As learning rate $l = 0.0001$ is used. The network thus learns more slowly but achieves a higher quality (Ref. 3). The softplus function is used as the activation function in the hidden layers since it has proven to be particularly useful for the problem (Ref. 3). The number of hidden layers as well as the number of neurons per layer will be adjusted for this model. The training is done with a group size of $s = 32$ (Ref. 3).

A problem in training neural networks can be the overfit of the network to the data of the training data set. To investigate the occurrence of this effect, the network is trained with the training data set for a few epochs and then tested to determine the quality with which the test data set can be predicted. The quality of the prediction is defined by the coefficient of determination (R^2 value). On the left side of Figure 9, the course of the R^2 value for the prediction of
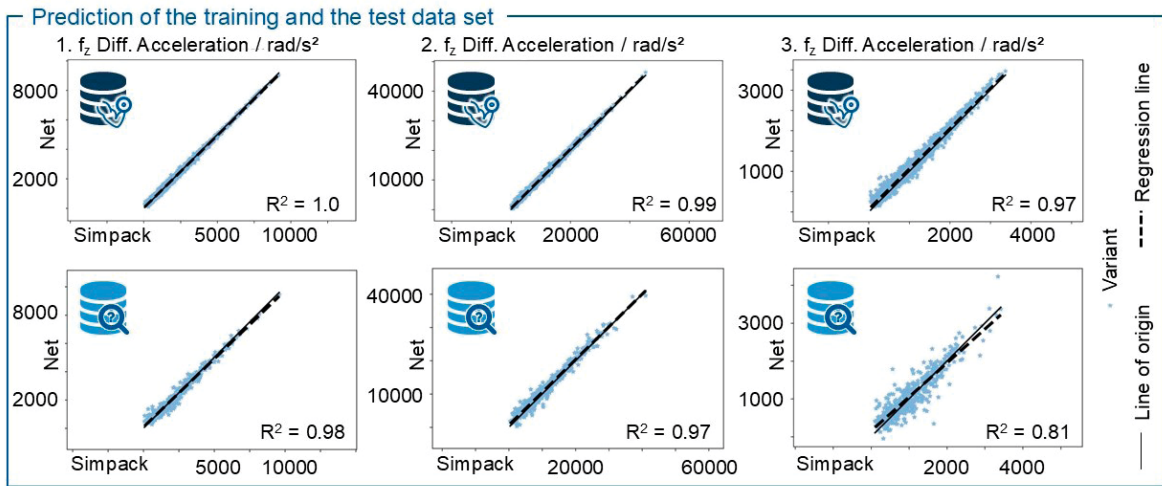
Figure 10—Prediction of the training and the test data set.

the first tooth meshing order of the differential acceleration is shown for three network topologies.

It is easy to see that the R^2 value first increases rapidly and then slowly decreases again after it has reached a maximum. The position of the maximum depends on the number of neurons and the number of hidden layers. The investigations of Willecke et al. on quasistatics as well as the evaluation of the data set of the dynamics of the model investigated here show that three hidden layers lead to the best prediction quality in principle (Ref. 3). To investigate how the number of neurons in the hidden layers affects the quality of prediction, it is varied for training dataset. The results of the variation are shown on the right side of Figure 9. It can be seen well that a neuron count of $n_z = 1,024$ leads to a better R^2 value than the other network topologies. Therefore, in the following, a network topology consisting of three hidden layers with $n_z = 1,024$ neurons is used.

Validating the Model with a Test Dataset

The DNN is built up and trained with the hyperparameters defined in the section "Optimal Choice of Hyper-Parameters for the Neural Network." The network is used in the training state, which has just no overfit and lies in the maximum of the prediction quality, see Figure 9 on the left. To check the prediction quality, the data points of the test data set are predicted with the trained network. Subsequently, the predicted acoustic characteristic values are compared with the calculated acoustic characteristic values. This is shown in Figure 10 for the first, second and third tooth meshing order of the differential rotational acceleration.

The first row of the diagrams shows the prediction of the training data. It is easy to see that all variants lie very well on the line of origin in the diagram, i.e., the predicted values correspond almost without deviations from the originally calculated values. This also results in coefficients of determination $R^2 \approx 1$. The second row of the diagrams shows the prediction of the variants of the training data set. Again, the points of the individual variants for the first two fz are very close to the line of origin and the R^2 values indicate a good mapping of the data set. Only the prediction of the amplitudes of the third tooth meshing order shows larger deviations and thus also only fulfills a value of $R^2 = 0.81$. The prediction of the entire test data set thereby takes $t_{sim} = 0.057$ s in comparison to $t_{sim} = 463\ 277$ s in the EMBS model, see the section "Optimal Choice of Hyperparameters for the Neural Network." Thus, the prediction by the network is faster by a factor of $\Delta DNN = 8\ 127\ 666$.

Investigating the Influence of the Number of Data Points in the Training Data Set

The number of data points in the training dataset used to create the neural network has an impact on the prediction quality achieved by the network. To investigate this influence, the number of data points used from the training data set to create the DNN is varied. For this investigation, a random sample of data points is taken from the entire training data set and the DNN is only trained with these points. To exclude the possibility of an influence due to the specific chosen random points, the process is repeated $n = 15$ times for each data set size and the mean value and standard deviation of the achieved prediction quality are calculated. For the network, the same topology and the same hyper-parameters are used, which were worked out in the section "Optimal Choice of Hyper-Parameters for the Neural Network." The results are shown in Figure 11.

It is easy to see that with an increasing number of training data points, the quality of the prediction increases rapidly at the beginning and increases much more slowly with an increasing number of data points. However, the effort to generate the data points increases linearly, since for each data point in the training dataset an equally effortful simulation must be performed.
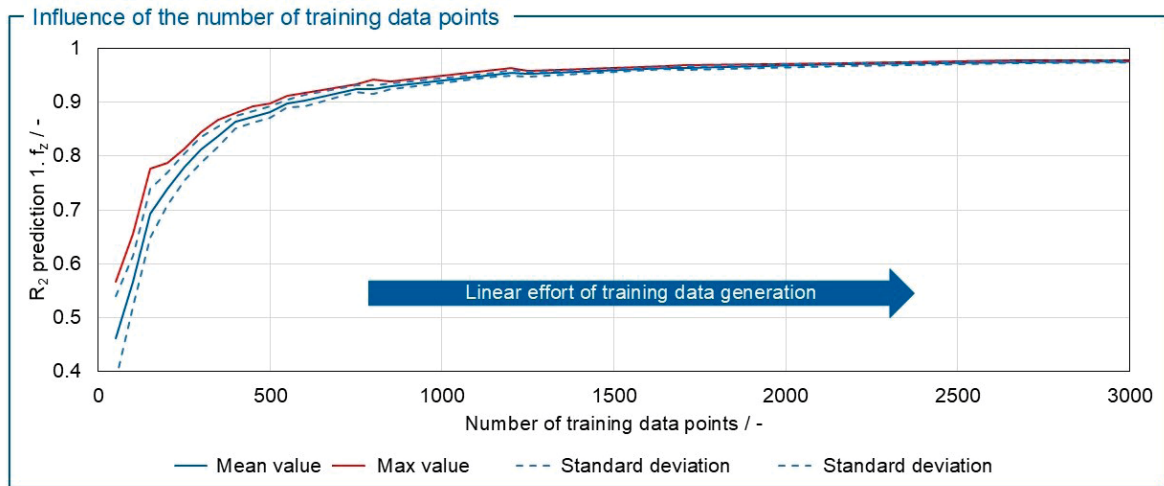
Figure 11—Influence of the number of training data points on the prediction quality.
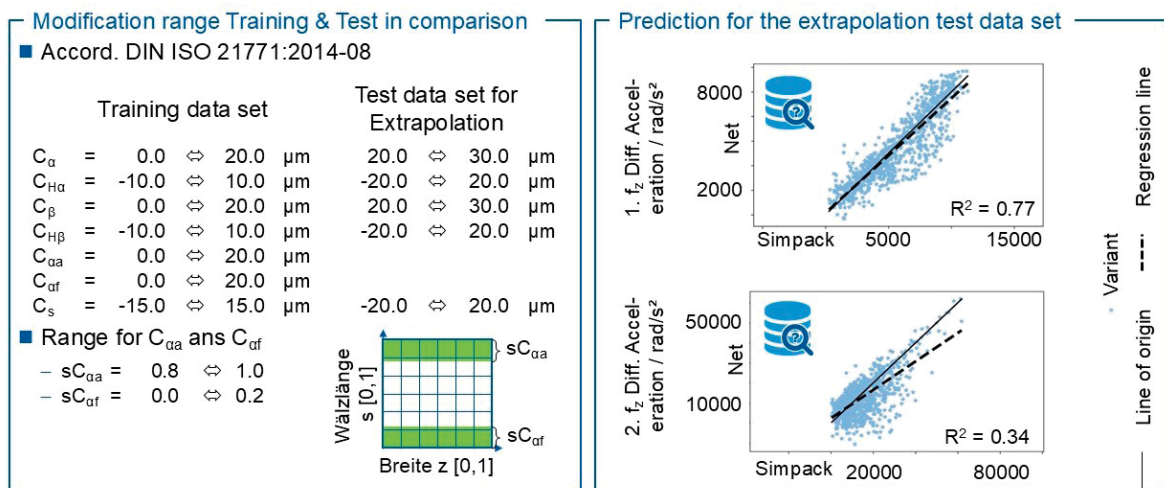

Figure 12—Range of variation and prediction quality for the extrapolation test data set.

Investigation of the Extrapolation Capability of the Neural Network

The behavior of the neural network has so far only been tested and evaluated in the range for which it has been explicitly trained. To investigate how the network behaves outside this range, a further test data set with $n = 1,000$ variants is created. The deviations of this data set are shown in Figure 12 on the left.

They extend beyond the boundaries of the training data set so that the mesh is forced to extrapolate. The right side of Figure 12 shows the prediction of the amplitude of the first and second tooth meshing frequencies of the differential rotational acceleration of the extrapolation test data set. For the first tooth

meshing order, a principal grouping of the values around the origin line can still be seen. However, the R^2 value of $R^2 = 0.77$ indicates that the prediction of the extrapolated data is of low quality. The prediction of the second tooth meshing order shows only a R^2 value of $R^2 = 0.34$. Here it is clearly visible that no meaningful prediction can be made with the developed approach. The DNN can therefore only be used for extrapolating predictions to a very limited extent.

Extending the Training Data Set with the Help of a Generative Adversarial Network

The section "Investigating the Influence of the Number of Data Points

in the Training Data Set" shows that a training data set with a too-small number of training data points leads to a DNN with a worse prediction quality. Furthermore, it is shown in the section "Investigation of the Extrapolation Capability of the Neural Network" that the DNN also has a lower prediction quality for predictions that go beyond the boundaries of the trained domain. The first challenge, that a too-low prediction quality is achieved with too few training data points, can be met trivially by using more training points. The second challenge can be met by increasing the size of the training domain itself. However, this decreases the density of data points in the trained domain, so the number of training data points must also
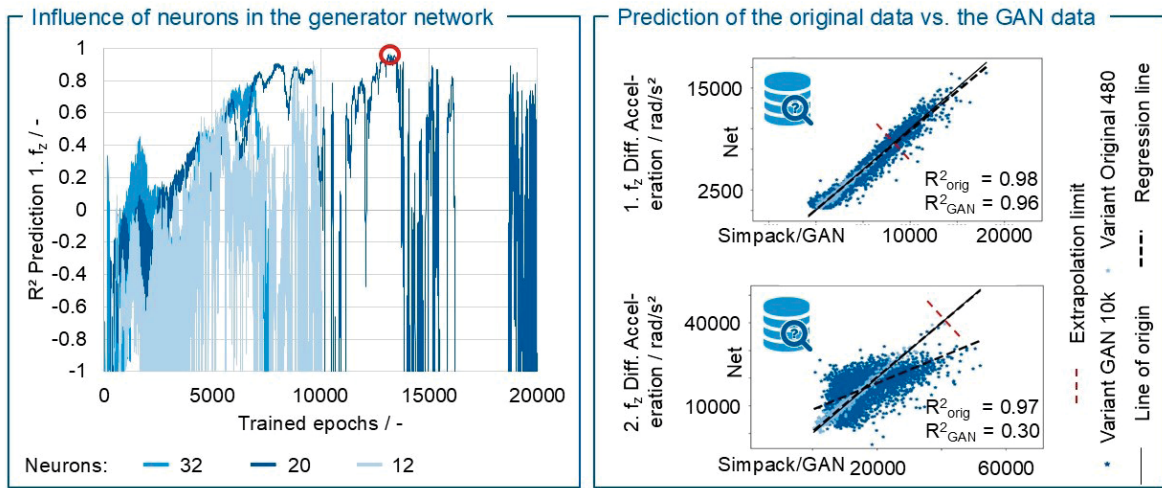
Figure 13—Influence of generator network neurons on data generation with the GAN.

be increased. A major disadvantage of generating a larger training dataset is that, classically, more simulations must be performed with the EMBS model. This is very computationally expensive.

An alternative to generating more data when an initial data set is already available can be a GAN, see the "State of the Art" section. These networks are typically used to create new data that fits the pattern of existing data. The resulting augmented dataset can then in turn be used as a training dataset for the DNN. Classical GANs are usually designed to process images (Ref. 6). The data here has the size of the number of pixels in the height direction times the number of pixels in the width direction times the number of color values per pixel.

However, the data in the use case to be investigated here is a vector with 18 grid points of the topography as input and a vector with the acoustic characteristic values of the corresponding variant as output values. GANs try to recognize patterns in existing data and generate new "artificial" data based on these patterns. Therefore, the input variables of the grid points and the output variables of the characteristic values must first be converted into a suitable format. For this purpose, the input and output vectors are stacked on top of each other to a new vector, see Figure 1, so that a data set is created that no longer consists of pairs

of input and output quantities, but only of a single vector per data point. The GAN can now attempt to generate new "artificial" vectors based on the existing vectors as "real" data.

As a test for the applicability of a GAN the data set with $n = 480$ points is used, see section "Variant Calculation with the Dynamic Model to Generate the Training and Test Dataset." If the DNN, which is used for prediction, is trained with only $n = 480$ points, the quality of the prediction is not sufficient, see section "Investigating the Influence of the Number of Data Points in the Training Data Set." The goal is therefore to enlarge the data set consisting of $n = 480$ points with a GAN in such a way that a new DNN can be trained with it, which then provides improved prediction accuracy. The correct choice of hyper-parameters has a significant influence on the results of the GAN (Ref. 5). These hyper-parameters are therefore investigated in the following.

To verify whether the GAN provides meaningful results, the generated data must be examined. When "artificial" images are generated, they can be viewed by a human and it is obvious to the human whether they look realistic or not. This is not the case with the generation of "artificial" vectors. Therefore, the "artificial" vectors are again decomposed into input and output variables and then predicted using the DNN from the section "Validating the Model

with a Test Dataset." Based on the validation in the section "Validating the Model with a Test Dataset," it is assumed that the DNN correctly represents the model behavior. Thus, if the "artificial" points generated by the GAN can be correctly predicted by the DNN, it can be assumed in principle that the GAN works. The quality of this prediction is again evaluated with the R^2 value. On the left side of Figure 13 the influence of different numbers of neurons in the generator network is shown.

It can be seen well that in the first $n = 10\,000$ epochs the R^2 value of the "artificial" data increases. However, discontinuities occur afterward and there seems to be no convergence of the GAN. In principle, however, a neuron number of $n_{Neur,Gen} = 20$ seems to provide the most promising results. For the data set from this GAN, which reaches the highest R^2 value (red circle on the left side of Figure 13), a comparison between the original data set and the "artificial" GAN data is shown on the right side of Figure 13. It is particularly noticeable that only the points of the first tooth meshing frequency are apparently correct. Furthermore, it is noticeable that the value range of the GAN data extends beyond the value range of the original data. If one now considers the fact, that the DNN, which is used as the basis for calculating the R^2 value, has a lower quality for extrapolating statements, the assessment of the quality of the
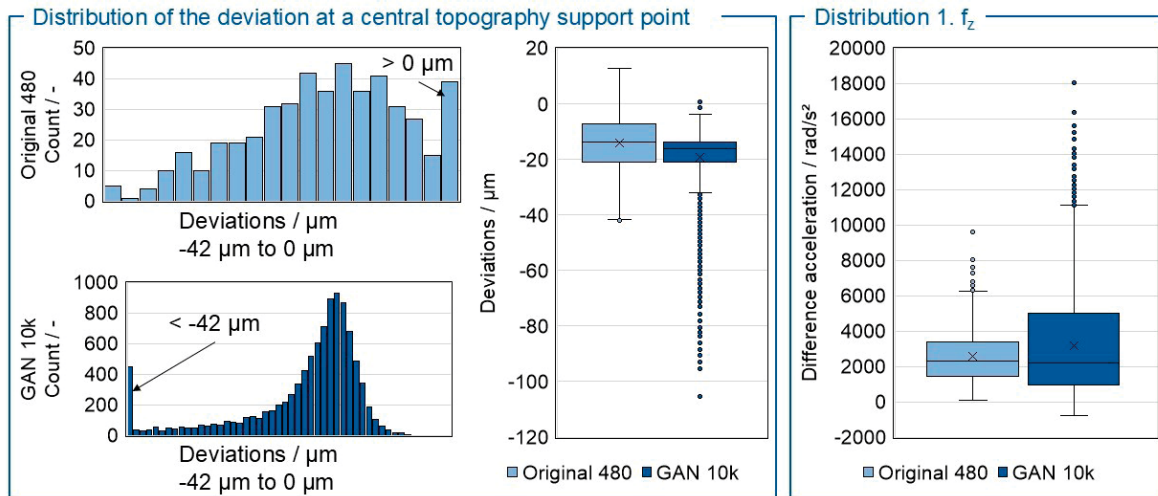
Figure 14—Comparison of the value ranges of the original and the GAN data set.

data generated by the GAN must be questioned here. The aspect of plausibility is further examined in Figure 14. Here, the distributions of the deviation values at a point in the center of the topography are shown in comparison on the left side as an example. From the distribution shapes as histograms, principal differences in the data sets can be recognized. If the original data set also shows deviations greater than $\Delta_{topo} = 0$ µm at the point under consideration, the GAN data set instead shows values smaller than $\Delta_{topo} = 42$ µm. In the boxplot diagram to the right of the histograms, this scatter is also illustrated.

The right side of Figure 14 shows the distribution of the first tooth meshing frequency as a result of values for both data sets side by side. Here it can be seen well that although the input variables seem to be concentrated in a narrower range, the result variables scatter over a wider range. After optimizing the hyper-parameters of the generator network, the hyper-parameters of the discriminator network are optimized. Hereby the R^2 value of the GAN data, related to the DNN from the section "Validating the Model with a Test Dataset," can be improved again to $R^2 = 0.995$. However, many data points still lie outside the range of values defined by the "real" data. If one removes all data points newly generated by the GAN, which would represent

an extrapolation, then the R^2 value of the remaining generated points decreases. The remaining "artificial" points are then added to the initial data set of $n = 480$ "real" points. Thus, a new DNN is generated analogous to the section "Investigating the Influence of the Number of Data Points in the Training Data Set." However, the newly generated DNN has the same prediction quality as the network consisting of the $n = 480$ "real" points alone, cf. Figure 10. The additional "artificial" data have therefore not yet improved the prediction of the DNN. However, the "artificial" data also did not worsen the prediction of the DNN either, so that in principle functionality of the GAN can be assumed. To achieve the desired added value with the GAN, further research is necessary.

Summary and Outlook

Within the scope of this work, an EMBS model for the first stage of a two-stage gearbox was built in the *Simpack* simulation environment. The high-speed demonstrator gearbox of the WZL Gear Research Circle was chosen as the application case. During modeling, attention was paid to a high-performance design of the model. Structural components are integrated as modally reduced bodies. To ensure convergence in the modal reduction, an FE study was carried out. Furthermore, a simulation study on the influence of the maximum con-

sidered eigenmodes in combination with the maximum necessary sampling frequency was carried out in the built EMBS model. The settings were optimized in such a way that a minimum calculation time could be achieved without accepting significant deviations in the calculation of the acoustic parameters. The microgeometry of the gears was imported into the model as topography via Input Functions of *Simpack*. To analyze the operating point for the following simulations, a run-up was calculated and evaluated in a waterfall diagram. With the simulation model optimized in this way, two variant calculations were carried out at the operating point. One with $n = 3,000$ variants to generate the training data set for the neural network and one with $n = 480$ variants to generate the test data set. The points in both data sets were distributed with a Latin Hypercube. Based on the preliminary work of Willecke et al. the basic structure of the DNN was chosen for the construction of the network developed in this work (Ref. 3). Subsequently, an optimization of the hyper-parameters was carried out for the network to achieve an optimal prediction quality. A coefficient of determination of $R^2 = 0.98$ was achieved for the test data set. In the prediction of the test data set, a speed advantage by a factor of $\Delta_{DNN} = 8,127,666$ was achieved. This speed advan-

tage makes it possible to predict the excitation behavior of the gearbox based on the topography in parallel with the process time of manufacturing. The objective is therefore fulfilled.

An essential prerequisite for the generation of the DNN, which is used for the prediction, is a sufficiently large training data set. Generating sufficiently large data sets can take a significant amount of time. To be able to generate a sufficiently large training data set with fewer data points, the usability of GAN was investigated. It could be shown that a GAN is in principle capable of generating data points similar to those contained in the reference data set. However, stable convergence could not be achieved with the GAN yet. Likewise, the generated data could not yet improve the prediction network. Further investigations are necessary until the GANs can be used for the described application. For this, further hyper-parameter studies are essential on the one hand; on the other hand, the learning and evaluation method of the GAN can still be optimized so that nongradient-based methods are used.

PTE

Marius Willecke, M. Sc. studied Mechanical Engineering at RWTH Aachen University 2019, graduating with distinctions. Since 2019, he has served as a research assistant at WZL of RWTH Aachen University. His primary interests include end-of-line testing development of tooth contact simulation software and simulation of the acoustic behavior of gearboxes.

Dr.-Ing. Jens Brimmers is the head of the gear department at the Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University since June 2019. He graduated from RWTH Aachen University with master's degrees in mechanical engineering and business administration. His Ph.D. thesis focused on beveloid gears and topological tooth flank modification.

Prof. Dr.-Ing. Christian Brecher has been, since January 2004, Ordinary Professor for Machine Tools at the Laboratory for Machine Tools and Production Engineering (WZL) of the RWTH Aachen, as well as Director of the Department for Production Machines at the Fraunhofer Institute for Production Technology IPT.

Bibliography

1. Hemmelmann J. E., 2007, "Simulation des lastfreien und belasteten Zahneingriffs zur Analyse der Drehübertragung von Zahnradgetrieben," Diss., Werkzeugmaschinenlabor (WZL) der RWTH Aachen, RWTH Aachen University, Aachen. Eng. Title: Simulation of load-free and loaded gear meshing for analysis of rotary transmission of gear trains.
2. Berger V., Wilbertz A., and Meyer I., 2010, "Körperschall Im End-Of-Line-Test Von Doppelkupplungsgetrieben," ATZproduction(3), pp. 24–27. Eng. Title: Structure-borne noise in the end-of-line test of dual clutch transmissions.
3. Willecke M., Brimmers J., and Brecher C., 2023, "Surrogate model-based prediction of transmission error characteristics based on generalized topography deviations," Forschung im Ingenieurwesen.
4. Brownlee J., 2019, "A Gentle Introduction to Generative Adversarial Networks (GANs)," *https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/*.
5. Mescheder L., Geiger A., and Nowozin S., 2018, "Which Training Methods for GANs do actually Converge?," International Conference on Machine Learning.
6. Goodfellow I., 2016, "NIPS 2016 Tutorial: Generative Adversarial Networks," NeurIPS|2016: Thirtieth Conference on Neural Information Processing Systems.
7. Brimmers J., 2020, "Funktionsorientierte Auslegung topologischer Zahnflankenmodifikationen für Beveloidverzahnungen," Diss., RWTH Aachen University, Aachen. Eng. Title: Function-oriented design of topological tooth flank modifications for beveloid gears.
8. Willecke M., Brimmers J., and Brecher C., 2022, "Evaluating the Capability of Acoustic Characteristic Values for Accessing the Gearbox Quality," 13th Aachen Acoustics Colloquium 2022, 1st ed., Aachen Acoustics Colloquium GbR, Aachen, pp. 125–134.
9. SIMPACK GmbH, 2020, "*SIMPACK* 2020x Documentation," Firmenschr., SIMPACK GmbH.
10. Dassault Systèmes, 2011, "*Abaqus 6.11: Abaqus*/CAE User's Manual," Bed.-Anl., Providence, RI, USA.
11. Shannon C. E., 1948, "A Mathematical Theory of Communication," Bell Sys. Tech. J. (27), pp. 623–656.
12. Willecke M., Brimmers J., and Brecher C., 2021, "Virtual Single Flank Testing—Applications for Industry 4.0," Procedia CIRP, 104, pp. 476–481.
13. Ketkar N., 2017. Deep Learning with Python: A Hands-On Introduction, Apress L. P, Berkeley, CA.