

Using rectified linear unit and swish based artificial neural networks to describe noise transfer in a full vehicle context

Dimitrios Ernst Tsokaktsidis, Clemens Nau, Marcus Maeder, et al.

Citation: [The Journal of the Acoustical Society of America](#) **150**, 2088 (2021); doi: 10.1121/10.0005535

View online: <https://doi.org/10.1121/10.0005535>

View Table of Contents: <https://asa.scitation.org/toc/jas/150/3>

Published by the [Acoustical Society of America](#)

ARTICLES YOU MAY BE INTERESTED IN

[Automatic source localization and spectra generation from sparse beamforming maps](#)

The Journal of the Acoustical Society of America **150**, 1866 (2021); <https://doi.org/10.1121/10.0005885>

[Identifying the Russian voiceless non-palatalized fricatives /f/, /s/, and /ʃ/ from acoustic cues using machine learning](#)

The Journal of the Acoustical Society of America **150**, 1806 (2021); <https://doi.org/10.1121/10.0005950>

[Machine learning in acoustics: Theory and applications](#)

The Journal of the Acoustical Society of America **146**, 3590 (2019); <https://doi.org/10.1121/1.5133944>

[Use of the filtered-x least-mean-squares algorithm to adapt personal sound zones in a car cabin](#)

The Journal of the Acoustical Society of America **150**, 1779 (2021); <https://doi.org/10.1121/10.0005875>

[Respiration mask waveguide optimisation for maximised speech intelligibility](#)

The Journal of the Acoustical Society of America **150**, 2030 (2021); <https://doi.org/10.1121/10.0006235>

[Introduction to the special issue on machine learning in acoustics](#)

The Journal of the Acoustical Society of America **150**, 3204 (2021); <https://doi.org/10.1121/10.0006783>



**Advance your science and career
as a member of the**

ACOUSTICAL SOCIETY OF AMERICA

LEARN MORE



Using rectified linear unit and swish based artificial neural networks to describe noise transfer in a full vehicle context^{a)}

Dimitrios Ernst Tsokaktsidis,^{1,b)} Clemens Nau,² Marcus Maeder,¹ and Steffen Marburg¹

¹Chair of Vibroacoustics of Vehicles and Machines, Technical University of Munich, Boltzmannstrasse 15, Garching bei Muenchen, 85748, Germany

²NVH Concepts, Mercedes-Benz AG, Mercedesstrasse 120, Stuttgart, 70372, Germany

ABSTRACT:

Vehicle interior noise is a quality criterion of passenger cars. A considerable amount of resources is used to evaluate and design the acoustic environment with respect to given requirements. The customer's perception in the end-of-line vehicle is the main criterion. Therefore, full vehicle testing is a large part of today's sound comfort development. To increase efficiency, it is desirable to limit the hardware testing to a specific component. A later reassembly of the full vehicle is done virtually using transfer functions. These transfer functions of the substructures can be derived numerically or through measurements. However, full vehicle simulations are still challenging. Hence, transfer functions are typically measured but come with the burden of complex procedures. In this work, the authors propose a machine learning algorithm to reduce the effort for finding suitable transfer models in the automotive context. Artificial neural networks with rectified linear unit and swish activation functions are trained on full vehicle measurements. Multiple operation conditions are used for training. The networks compute spectral system responses and relative sensitivities for the input features. The performance is discussed with respect to the full vehicle validation data. The results indicate an effective procedure to reduce the costs of full-size vehicle measurements.

© 2021 Acoustical Society of America. <https://doi.org/10.1121/10.0005535>

(Received 15 January 2021; revised 7 June 2021; accepted 15 June 2021; published online 22 September 2021)

[Editor: Peter Gerstoft]

Pages: 2088–2105

I. INTRODUCTION

Passenger cars are developed to meet comfort levels that are characteristic of a specific brand and vehicle type. The interior noise as a quality aspect is influenced by many external and internal factors. The dominant contributors will vary depending on the driving condition of the vehicle. Aeroacoustic phenomena will typically dominate for faster vehicle speeds. The acoustics of slower speeds are often mainly influenced by the rolling noise or engine noise. For specific road conditions or maneuvers, such as the parking maneuver, other components, like the steering system, can become dominant. A common operating condition evaluated for interior sound quality is a constant turning motion of the steering wheel. The acoustic system response in the passenger cabin is a result of the excitation on the component level and transfer behavior of the vehicle structure, including the cabin cavity volume. The final assessments of the acoustic performance of the exciter and receiver combinations are still mainly conducted in full vehicle systems. Because hardware related evaluations are typically costly, a digital reproduction of the noise-, vibration-, and harshness- (NVH) transfer is an enabler for improved efficiency in development processes.^{1,2} The transfer path analysis (TPA) with its

three families, the classical TPA, component based TPA, and transmissibility based TPA, are commonly mentioned in this context.³ The transmissibility based approaches focus on determining the path contributors. Once the target vehicle is available, the transfer functions between the relevant locations can be acquired by exciting the system with a volume sound source.⁴ The spectral system description can be combined with an excitation on the component level to synthesize the system response. This allows the comparison of different operation conditions (OCs) or varying hardware combinations. Measuring the transfer functions is time extensive and influenced by multiple parameters. For example, the frequency range of the excitation must match the frequency range of the OC of interest. Another factor is the amount of energy that has to be applied to the system to generate a sufficient signal-to-noise ratio. Because the packaging situation of modern vehicles is typically very restrictive, the excitation locations must yield enough space to apply the desired forces to the structure.

Machine learning (ML) has recently become a powerful tool to extract information from data. Artificial neural networks (ANNs) are one of the most successful contributors in this branch.⁵ Transfer functions, as information given through excitation and system responses, represent a suitable field to investigate with pattern recognizing algorithms. ANNs can approximate every continuous function.⁶ Given nonlinear activation functions and a large number of trainable parameters, they excelled in topics such as speech or

^{a)}This paper is part of a special issue on Machine Learning in Acoustics.

^{b)}Also at: Department Development Steering Systems, Mercedes-Benz AG, Mercedesstrasse 120, 70372 Stuttgart, Germany. Electronic mail: dimitrios.tsokaktsidis@tum.de, ORCID: 0000-0003-2302-7970.

image recognition.⁷ Data-driven models are increasingly used throughout various automotive areas. Their latest improvements with respect to perception led to growing adoption in the full vehicle context. Camera based driving aid systems and speech assistants are two areas to which ML has contributed lately.^{8,9} The use of ML in the field of vehicle NVH is a topic that is increasingly attractive to the scientific community. Applications in this context are found in the fields of psychoacoustics,^{10–13} traffic noise,¹⁴ interior noise,^{10,15,16} structural health monitoring,^{11,17} approximation of material parameters,¹⁸ and as a complementary method for established simulation approaches.¹⁹ Speaking of ANNs and grouping these applications into classification and regression tasks, the following two paragraphs provide orientation. The first paragraph covers classification, and the second paragraph covers regression.

Ma *et al.*¹⁰ use an ANN model to categorize the interior noise of pure electric vehicles with respect to psychoacoustics. Their model inputs are 14 objective parameters, such as loudness or sharpness, and then assigns them to 10 classes that describe the subjective acoustic impression. The classifier performs with an average error of 9% on the validation data. Kane and Andhare¹¹ process statistical signal features and psychoacoustic indices with ANNs to identify faulty gearboxes in an end-of-line scenario. They use three-layered ANNs with sigmoid and tangens-hyperbolicus activation functions. While propagating psychoacoustic inputs, they achieve an accuracy of 98% for faulty and 99% for healthy gearboxes. Lerspalungsanti *et al.*¹² estimate a subjective perception from objective data with ANNs. They use neuron pruning to optimize their network topology. The approach is applied to evaluate different drive train assemblies as well as gear rattle noise. Stender *et al.*¹⁵ apply deep learning to detect, characterize, and predict brake noise. Convolutional neural networks (CNNs) perform the detection tasks in the frequency domain. State-of-the-art models, like inception or resnet, are modified for the use case. The proposed transfer learning process enables brake squeal detection and classification for large datasets. Qian and Hou¹⁶ investigate the interior sound quality of electric vehicles. They use an ANN with one hidden layer to describe the nonlinear relation between eight objective input values and subjective ratings from a listening test. Their estimates show less than a 5% error with a correlation coefficient of 0.997. Wang *et al.*¹⁷ develop a classifying method for engine fault diagnosis based on sound intensity measurements and ANNs. Wavelet transformations are applied to the input vector before processing the data through a network with one hidden layer. Distinguishing nine different engine states, they achieve 100% accuracy on their validation set.

He *et al.*¹³ analyze the influence of structure parameters on the loudness of the exhaust tail noise. They use a radial basis function network to find a relation between eight geometric values and the noise emission. Their structure-loudness model can be used as a target function for fast parameter studies to search for an optimal exhaust pipe design. Li *et al.*¹⁴ investigate tire-pavement interaction noise

as it is one of the dominant noise sources of passenger cars and trucks after reaching a certain speed. They compare 23 different tread patterns with a maximum difference of 17.2 dBA between the quietest and loudest. Testing different ANN configurations, they find a relationship between 12 spectral input parameters describing tire patterns and the related 81-dimensional spectral system response when driving at different speeds. Wang *et al.*¹⁸ investigate the noise reduction coefficient of porous materials. A single-layer ANN is trained on a feature vector consisting of density, thickness, and porosity. Applying a genetic algorithm to their ML approach, they find a parameter set resulting in a noise reduction coefficient of 0.62. Addressing the computational requirements of finite-element simulations, Capuano and Rimoli¹⁹ use ANNs as a substitution. They train surrogate elements to generate a direct relationship between the unit state and its forces. By presenting a modular approach with their smart finite elements, they establish a flexible method applicable to different geometries.

Despite being used in the abovementioned NVH related fields, to the best knowledge of the authors, no deep learning related TPA study on real vehicle operational steering data has been reported as of today. In an early study, Tsokaktsidis *et al.*²⁰ published an article in which a neural network with nonlinear activation was used on a frequency domain finite-element method (FEM) training set to extract the transfer behavior of a simple two-sided beam structure with variable geometry. The training set consisted of 1001 input and output pairs. The system response calculated with their network has shown a good fit compared with the conventional FEM simulation. Lee and Lee²¹ also used FEM training data to train an ANN. A frame for a car-body problem with a substructure, connected with four joints, was investigated. Comparing the relative importance for each path derived from the ML model with a reference calculation, they identified the same four structural main contributors. They state future research can focus on experimental validation and real vehicle application where noise will play a larger role in the data set. Also, different network architectures should be examined to better understand the deep learning process. In another article, Tsokaktsidis *et al.*²² used an ANN-TPA on full vehicle time domain data. They calculated the interior sound pressure level during a steering maneuver. The acceleration level data of eight tri-axle sensors on a steering system have been used as features. The interior sound pressure level was the target value. Besides opposing the ANN-calculated interior noise to a validation data set, they studied the effect of the network size with respect to the prediction performance.

In this paper, the authors present a vehicle interior noise calculation in the frequency domain that is performed by an ANN. Eventually, the machine-learned transfer behavior is compared with conventionally acquired functions from a reciprocal TPA. Figure 1 gives an overview of the paper. Structure-borne excitation on the steering system and air-borne responses in the passenger cabin contain information about the transfer behavior of the vehicle. Aiming at a

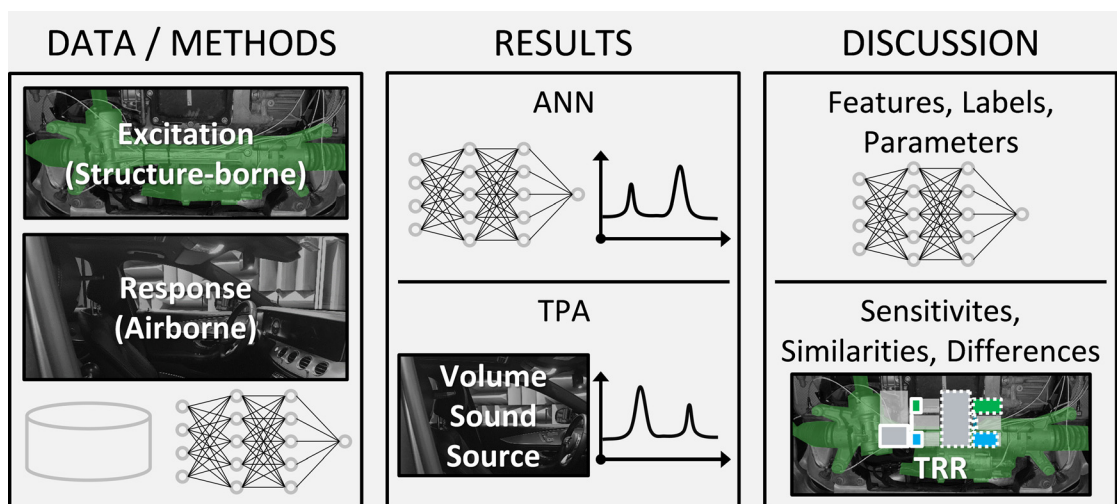


FIG. 1. (Color online) Overview of the paper structure. The data and method section provide information about a typical steering maneuver and the recorded data to give the reader an entry point. After briefly discussing general transfer behavior, the focus shifts to ANNs. In the results section, various networks and their results are presented. The final models are then used to rate the transfer behavior of the vehicle. This machine-learned behavior is compared with the reference from a volume sound source measurement. The discussion will provide further insights about ANN parameters until eventually reflecting the similarities and differences between the machine-learned and conventional approaches for transfer behavior determination.

simplified procedure to access this information, ML gets introduced in this context. The measurement effort is limited to operational measurements instead of a specific procedure (i.e., volume sound source, impact measurements, etc.). Because the presented method relies solely on accelerometers and microphones, it also weakens the typical packaging drawbacks of more complex procedures. The measured *data* will be processed with ANNs to extract the desired information. The *results* are compared with the measured data of a reciprocal full vehicle TPA. The process of finding a suitable set of features, labels, and network parameters is *discussed*. Finally, the similarities and differences for acquired sensitivities are reflected.

II. DATA AND METHODS

This section covers the investigated OCs and measurement setup. It provides a detailed overview of the used data. Furthermore, tools and mathematical formulations for describing transfer behavior are presented. The complexity

is illustrated by describing a typical approach for approximating interior noise. ML as the data-driven method is explained, beginning with the required hardware setup and data acquisition. The required preprocessing to prepare the acquired database for ANN usage is explained. This includes the network architecture and its hyperparameters. The network training is covered. Last, the prediction process is explained in the post-processing step.

A. Description of maneuver and measurement setup

A commonly evaluated OC of modern steering systems is the parking maneuver. It can be described by the steering wheel position α_{StWhl} and speed of its turning motion $\dot{\alpha}_{StWhl}$. Figure 2 shows both parameters during a steering maneuver with a target speed of 300 deg/s. The steering wheel angle [see Fig. 2(a)] of 0 deg indicates a centered position. Positive values represent a wheel position right of center, and negative values each represent a position left of center. The steering wheel angular speed [see Fig. 2(b)] reaches

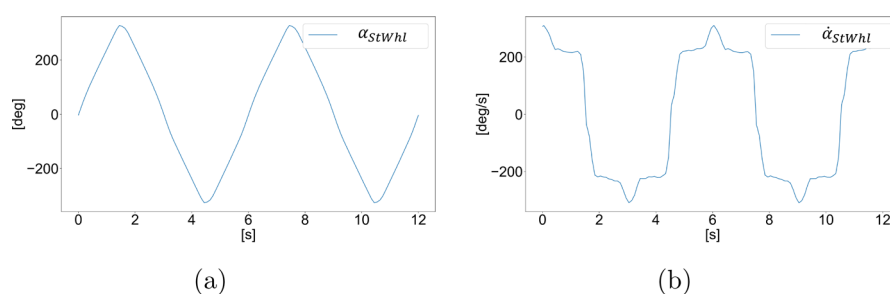


FIG. 2. (Color online) The steering maneuver over time. (a) Steering wheel angle, starting in the middle position, positive values correspond to wheel positions right of center, and negative values correspond to wheel positions left of center; (b) steering wheel angular speed, positive values correspond to the clockwise operation, and negative values correspond to the counterclockwise operation. The data were derived from the vehicles controller area network (CAN)/the electrical control unit (ECU) of the steering. The exact signal form results from the regulation of the ECU when requesting a certain OC.

roughly ± 300 deg/s. Today's steering systems are typically evaluated for different steering speeds. They range from 100 deg/s to 700 deg/s.

To assess the NVH behavior for the OC in the full vehicle system, structure-borne and airborne noise is recorded. Figure 3 shows the measurement setup for the accelerometers on the steering system. The sensors are placed close to the interfaces between the steering subsystem and main vehicle body. All sensors are applied on the component side. The coupling positions front left (FL), front right (FR), back left (BL), and back right (BR) are rigid connections to the vehicle and affixed using screws. Flexible bushings connect the left and right tie rods (TRL and TRR, respectively) to the wheel carriers. The input shaft of the steering column (STCOL) connects with the rack bar and is guided through the ball bearings. While operating, the electric drive applies a force to the rack bar. Because its behavior is characteristic for the investigated parking maneuver, an accelerometer (EMOT) is placed on it. The assurance of a specific acoustic response in the passenger cabin is typically verified for each available seat. In this paper, the sound pressure at the position of the co-driver's left ear is evaluated. An omnidirectional microphone is mounted to the headrest according to Fig. 4. A sampling rate of 48 kHz is used.

B. Data

The recorded database of the operational measurements consists of five separate conditions of which each is measured two times. The measurements are distinguished by their target steering wheel angle speeds. The lowest speed recorded targets 200 deg/s. The highest speed recorded targets 600 deg/s. The targets for the intermediate measurements are 100 deg/s apart from each other in terms of their steering wheel angle speed. Figure 5 gives an overview of the processing chain from the measurement until the data are fed into the ML algorithm for model extraction. The left path covers the structure-borne data. The right path covers the airborne data. The structure-borne raw data $a_{\text{raw}}(t)$ consists of 129 024 000 samples. These result from 24 measured channels (CH), multiplied by two measurements (M) per speed, multiplied by the number of samples (S) needed to record the steering maneuver according to Fig. 2.

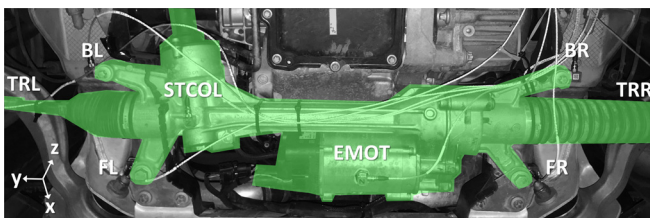


FIG. 3. (Color online) The vehicle bottom view with the structure-borne measurement setup. The accelerometers are on chosen subsystem positions. From left to right: left tie rod (TRL), back left (BL) coupling, front left (FL) coupling, steering column (STCOL), electric drive (EMOT), front right (FR) coupling, back right (BR) coupling, and right tie rod (TRR). The driving direction corresponds to the positive x axis.



FIG. 4. The vehicle passenger cabin view in the anechoic chamber. The airborne measurement setup with the omnidirectional microphone is shown. The measurement position is head high on the left side of the co-driver seat (CD_L), and the microphone is mounted to the headrest.

Generating blocks with 16 384 samples and transforming them to the frequency domain results in 7776 narrow-band spectra. Eventually, calculating third-octave bands creates $a_{\text{raw}}(\omega)$. Within the model-validation split, the 7776 third-octave spectra will be equally distributed into $a_{\text{mod}}(\omega)$ and $a_{\text{val}}(\omega)$. Both data containers consist of the same amount of samples for each OC. $a_{\text{mod}}(\omega)$ is further split into $a_{\text{train}}(\omega)$ and $a_{\text{test}}(\omega)$ using a two-to-one ratio. These containers are later used for model creation.

The processing steps for the airborne data container are identical to the structure-borne procedure with the exception that there is only 1 measured channel compared to 24. The initial 5 376 000 samples in $p_{\text{raw}}(t)$ end in 107 third-octave spectra in $p_{\text{train}}(\omega)$ and 55 third-octave spectra in $p_{\text{test}}(\omega)$. After training the ANN, $a_{\text{val}}(\omega)$ and $p_{\text{val}}(\omega)$ are used for validation.

The desired target value is centered on the vehicle interior noise. It is the criterion on which a quality decision for the steering is made. Two possible ways of representing $p(\omega)$ are linear or logarithmic. Figure 6 shows the distribution of these values throughout the dataset. On the left side, the system response is given in Pa values in a third-octave resolution [Fig. 6(a)]. The sound pressure level distribution with reference to 2×10^{-5} Pa is shown on the right side [Fig. 6(b)]. The orange indicator on both of the plots marks the median for each third-octave. The boxes end on the lower and upper quartile of the data. The whiskers mark the entire data range.

C. Transfer behavior

To achieve the goal of calculating the system response in the passenger cabin from the component data, the description of the transfer function $H_{pa}(\omega)$ between the source and receiver, according to Fig. 7, is possible. Because the source input is given as the acceleration $a(\omega)$ and the desired receiver output is the sound pressure $p(\omega)$, the transformation from input to output requires two intermediate steps.⁴ The accelerance,

$$H_{aF}(\omega) = \frac{a(\omega)}{F(\omega)}, \quad [H_{aF}(\omega)] = \frac{1}{\text{kg}}, \quad (1)$$

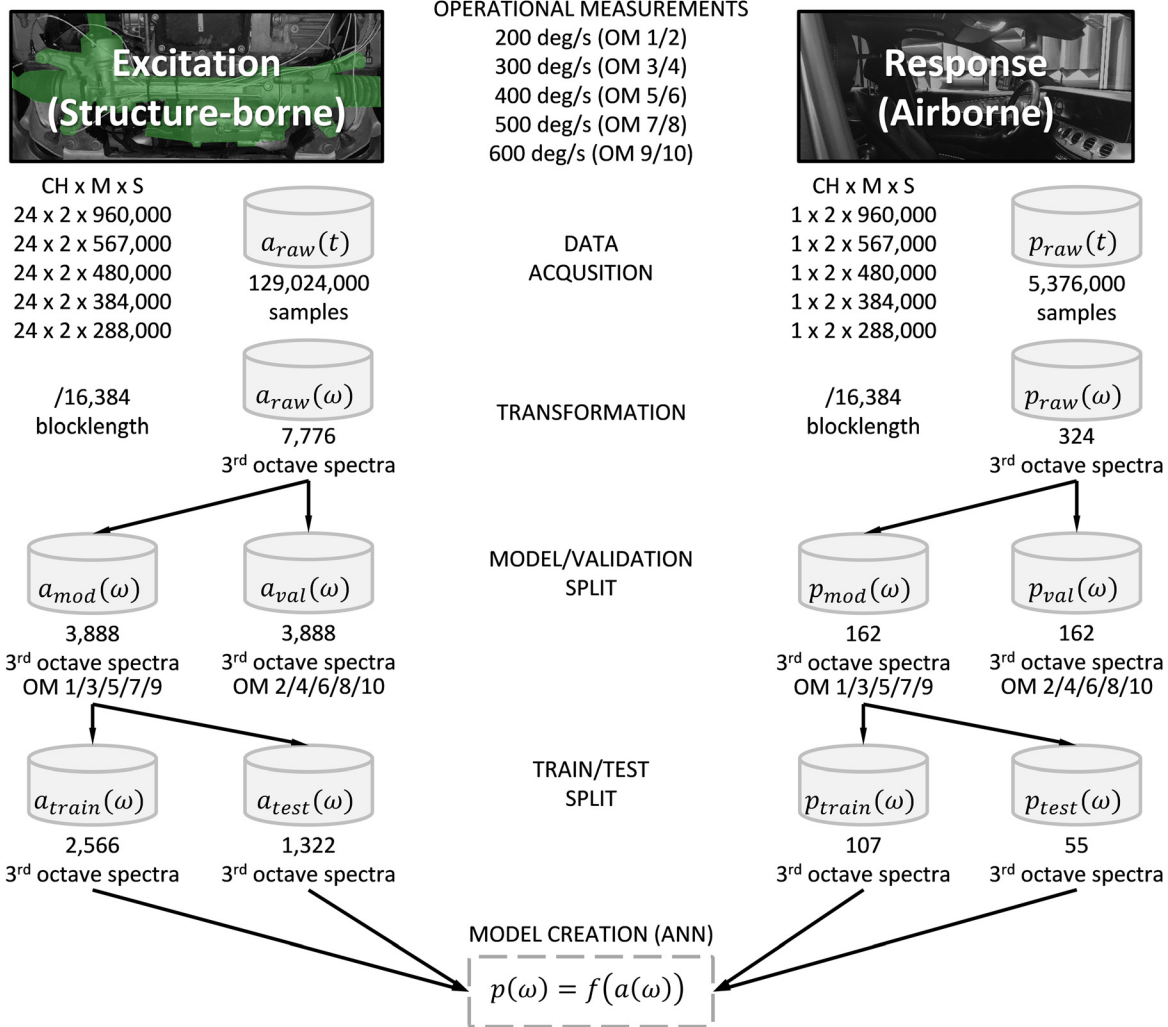


FIG. 5. (Color online) The overview of the data used, including the processing steps from the raw measured data to features and labels used for model creation. The left path shows the processing for the structure-borne data. The right path shows the processing for the airborne data. The middle column represents the general preprocessing steps from the data acquisition to model creation.

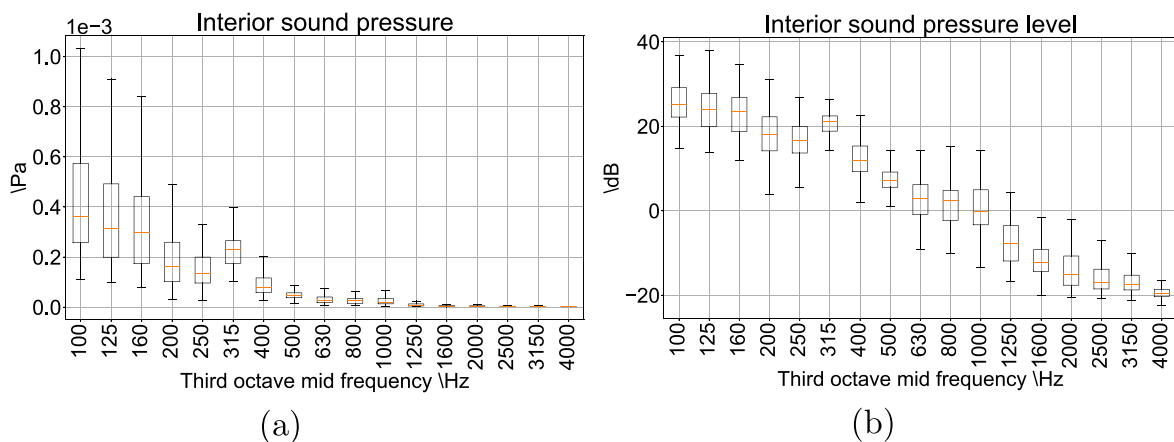


FIG. 6. (Color online) The box and whisker plot for $p_{raw}(\omega)$ in the linear and logarithmic presentation. (a) shows the Pa values (linear), and (b) shows the dB values (logarithmic).

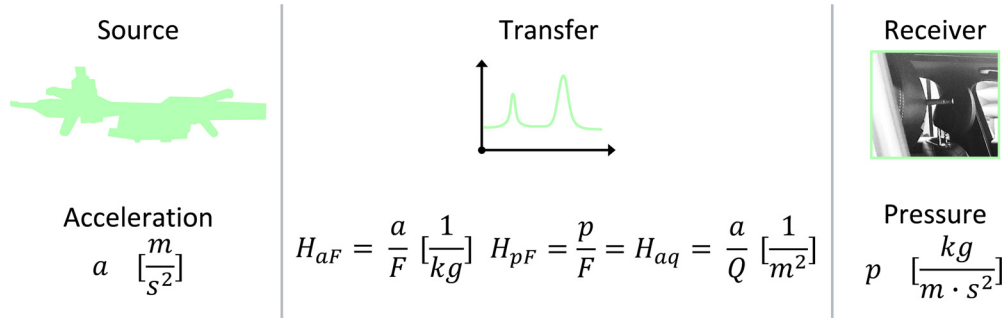


FIG. 7. (Color online) The source, transfer, and receiver relations to pair the component acceleration and passenger cabin sound pressure.

as the ratio between the acceleration a and force F has to be calculated for each measurement position. It can be experimentally determined through the impact hammer measurements. Here, the square brackets indicate the unit. Given F , the sound pressure p is calculated using

$$H_{pF}(\omega) = \frac{p(\omega)}{F(\omega)} = H_{aq}(\omega) = \frac{a(\omega)}{Q(\omega)},$$

$$[H_{pF}(\omega)] = [H_{aq}(\omega)] = \frac{1}{\text{m}^2}, \tag{2}$$

which are equivalent based on reciprocity for linear, time-invariant systems.⁴ H_{pF} requires exciting each measurement position on the steering system with an impact hammer while measuring the sound pressure in the passenger cabin. Depending on the vehicle package, this process is difficult to perform as the desired measurement positions for the operational measurements will typically be unreachable for the engineer conducting the experiments. Measuring H_{aq} as the equivalent transfer function is performed by applying a volume excitation $Q(\omega)$ to the cavity in the passenger cabin at the desired response position while simultaneously measuring the accelerations at all desired sensor positions. For n structure-borne excitations,

$$p(\omega) = \sum_{u=1}^n a_u(\omega) \cdot H_{pa,u}(\omega) \tag{3}$$

can be used to compute the pressure signal as a linear combination of the accelerations a_u and their respective transfer functions $H_{pa,u}$.

D. Data-driven model

ML approaches like ANNs are driven by data and trained for a specific purpose. To approximate the spectral system response in a vehicle passenger cabin using acceleration inputs from the steering system, the database has to reflect the desired relation between the input and output. The required hardware setup is equivalent to the operational measurements configuration from Figs. 3 and 4.

In contrast to the measurement procedures involving an impact hammer or a volume sound source to determine the transfer behavior, no additional exciters or special hardware

are needed when applying the proposed ML approach. Figure 8 illustrates the data processing to generate a spectral vehicle surrogate with respect to the NVH transfer from acceleration $a(t)$ on the component level to sound pressure $p(\omega)$ in the passenger cabin (also recall Fig. 5). Because the ML algorithm is desired to extract spectral information, the raw data are transformed to the frequency domain by applying a discrete Fourier transformation (DFT). In a second step, the spectral data $a_n(\omega)$ for each sensor n are concatenated to create an input vector for the ANN. Multiplying all of the inputs $a_n(\omega_m)$ with the weights $w_{n,m,k}$ connected to neuron k in the hidden layer results in the neuron input. The sum for these values will then be transformed by the activation function to generate the neuron output y_k . Researchers have introduced many weight initialization methods and activation functions. Glorot uniform initialization is widely used; see Glorot and Bengio²³ for a deeper insight. The most common nonlinear activation,

$$f(x) = \max(0, x), \tag{4}$$

is the rectified linear unit (ReLU) function.^{7,24} Another possible nonlinear activation is the swish function

$$f(x) = x \cdot \text{sigmoid}(x); \tag{5}$$

see Ramachandran *et al.*²⁵ Moving from the hidden layer to the output layer, the neuron activations y_k are multiplied with the weights $w_{k,j}$. Creating another weighted sum and feeding the value into the identity function generates the network output y_j in the output layer. During the training process, this network prediction will be compared with the desired target data $p(\omega)$. A mean squared error function,

$$\text{MSE} = \frac{1}{j} \sum_{s=1}^j (p(\omega) - y_s)^2, \tag{6}$$

serves as a metric for the backpropagation. The dimensions of the layers can vary largely depending on the task. In the illustrated case, the input layer, hidden layer, and output layer have the dimensions

$$d_{\text{IN}} = n \cdot m, \tag{7a}$$

$$d_{\text{HL}} = k, \tag{7b}$$

$$d_{\text{OUT}} = j. \tag{7c}$$

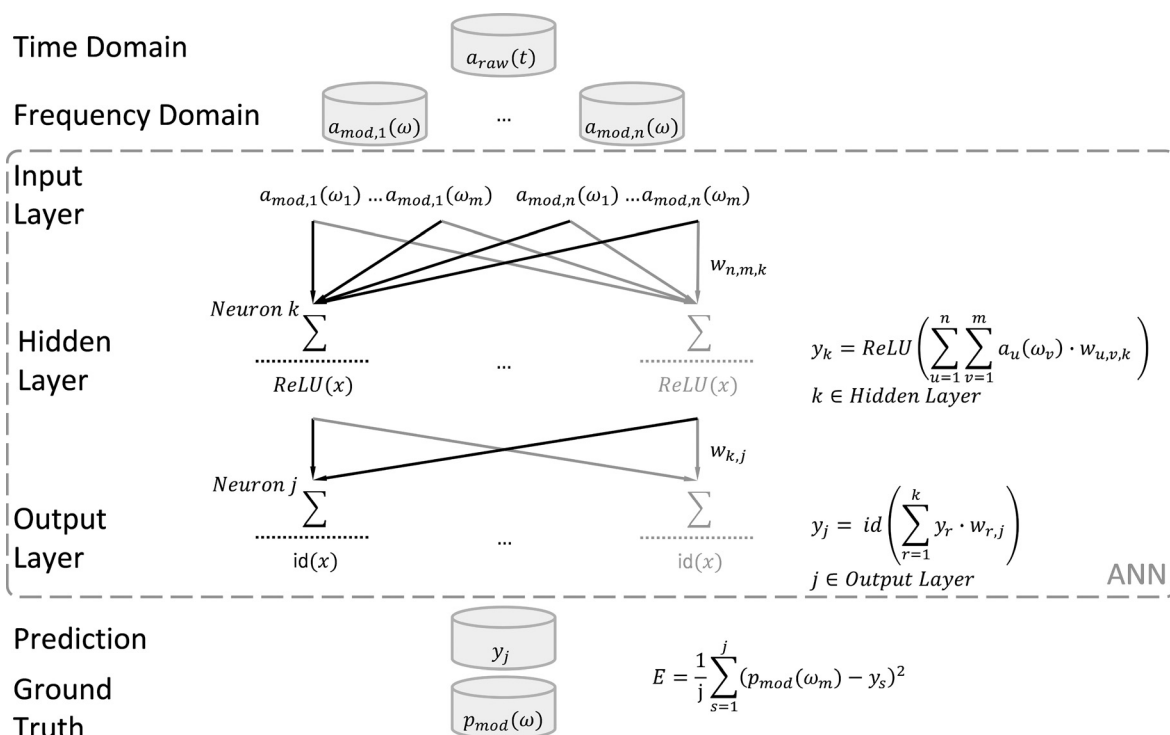


FIG. 8. The data processing with ANNs for structure-borne excitation and acoustic response. The features consist of acceleration, and the labels consist of sound pressure. The network is trained on the spectral data. The hidden layer uses ReLU activations. The output layer uses identity activations. A mean squared error function is applied for backpropagation.

Hyperparameters, such as the number of layers, dimension d_{HL} , or the activation function used, are often determined through an iterative process. As an example, a grid search will vary d_{HL} while evaluating the network performance. The mean absolute percentage error (MAPE) and mean absolute error (MAE) are common evaluation metrics.^{26,27} The MAPE will allow us to relatively compare the training performance even when the target value ranges are different. A MAE metric, on the other hand, will express errors in the target unit of the model. An example, speaking of dB values in an automotive NVH environment, is those values provide a very concrete understanding for the performance of the model. In addition to optimizing the hyperparameters, the composition of the input and target data decides about the performance of the ML approach. In the presented explanation of a parking maneuver involving eight triaxial accelerometers of which all possibly contribute to the interior sound pressure, the concatenated input vector dimension can easily reach the order of $\mathcal{O}(10^3)$ or higher. It mainly depends on the frequency resolution. Considering an output vector that also has thousands of dimensions, the ANN training becomes complex.

III. RESULTS

Full vehicle data, an established set of methods from the NVH sector, and state-of-the-art ML techniques are combined in this work. One of the most common frameworks for working with ANNs is Python. Finding a suitable network architecture while adapting the network parameters is a highly

iterative process. After implementing a new architecture, the evaluation and interpretation take considerable effort due to the ANNs being black box models. This section is divided into two parts. In part one, the network architectures, training results, and validation results for different ANNs are provided. This leads to two networks that are compared to each other in terms of a prediction for the interior noise target value. Eventually, a reciprocal volume sound source measurement is used to create a relative sensitivity ranking for each translatory structure-borne degree of freedom. These are then compared to the machine-learned ranking to assess how strongly ANNs correlate to the physical domain.

A. Architectures, training, and validation

To find network architectures, grid searches have been used throughout the training process. Recalling Fig. 5, the features $a_{mod}(\omega)$ and labels $p_{mod}(\omega)$ are the basis for model fitting. According to Fig. 8, the features vector is assembled with a third-octave frequency resolution. A frequency range between 89.2 and 4467 Hz was used because this will cover the majority of today's operational noise phenomena of steering systems. The exact upper and lower limits are defined by the third-octave resolution. The structure-borne setup consists of 24 channels (8 sensors \times 3 degrees of freedom). Having 17 third-octave amplitudes per channel in the mentioned frequency range results in an input vector d_{IN} of size 408. The feature amplitudes used for training are given in m/s^2 . The output vector d_{OUT} has size 17, consisting of the third-octave amplitudes of $p_{mod}(\omega)$. The labels have been processed as linear and logarithmic values. In total, eight network

TABLE I. Overview of trained networks with feature, label, and OC combinations.

	ANN 1	ANN 2	ANN 3	ANN 4	ANN 5	ANN 6	ANN 7	ANN 8
Features	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$	$a(\omega)$ $\frac{m}{s^2}$
Labels	$p(\omega)$ Pa	$p(\omega)$ Pa	$p(\omega)$ Pa	$p(\omega)$ Pa	$p(\omega)$ Pa	$p(\omega)$ Pa	$p(\omega)$ dB	$p(\omega)$ dB
Data container	OC: all	OC: 200	OC: 300	OC: 400	OC: 500	OC: 600	OC: all	OC: all

architectures have been investigated with the feature, label, and OC combinations given in Table I. Details on the parameter variation for each ANN with more insights on the results can be found in the Appendix in Table III.

As a result of ANN 1–ANN 6 suffering from dying ReLU (see the Appendix for details), further results focus on ANN 7 and ANN 8. Figure 9 shows the loss curves for both ANNs during the training. Figure 9(a) shows training and validation loss for ANN 7, which is ReLU based. The loss curves for the swish based ANN 8 are shown in Fig. 9(b). Both networks are trained on the entire $a_{mod}(\omega)$ and $p_{mod}(\omega)$ data. Hence, the networks reflect the entire range of OCs at once.

To rate the prediction performance of each trained ANN, $a_{val}(\omega)$ and $p_{val}(\omega)$ are processed for every OC separately. The MAE values for the comparison of the network output versus the ground truth are presented in Table II. The values are supported by Figs. 10–13, which will be explicitly addressed in the following.

The predictions with the ReLU based ANN 7 are performed for all five OCs. The spectral responses of ANN 7 [Figs. 10(a)–10(e)] generate MAE values ranging from 2.94 to 5.37 dB for the different steering wheel angular speeds. Figure 11 results from predicting the third-octave band responses for a steering cycle over time. The values calculated by ANN 7 for a steering speed of 200 deg/s [Fig. 11(a) left] are opposed to the corresponding ground truth values [Fig. 11(a) right]. The faster steering wheel angular speeds are compared similarly in Figs. 11(b)–11(e).

Comparable to ANN 7, Figs. 12 and 13 show the predictions of the swish based ANN 8. The averaged third-octave predictions are given in Fig. 12. Depending on the

steering wheel angular speed, the MAE values range from 3.02 to 5.22 dB. The time-discrete responses for an entire steering cycle from slowest to fastest steering speed are shown in Figs. 13(a)–13(e). The predictions are shown on the left, and the ground truths are shown on the right.

B. Physical domain versus ML domain

A direct comparison for the time-discrete responses of ANN 7 and ANN 8 against the measured ground truth for the 400 deg/s validation measurement is shown in Fig. 14. The ReLU prediction of ANN 7 is displayed in Fig. 14(a), and the swish prediction of ANN 8 is displayed in Fig. 14(b). The ground truth spectrum in Fig. 14(c) is marked with numbers 1–4, indicating the moments in time when there was a steering direction change, and *A* and *B* denote the areas of high amplitude for the 315 Hz and 1000 Hz third-octave bands, respectively.

Having performed a reciprocal volume sound source measurement, a relative sensitivity ranking based on Eq. (2) was calculated for each structure-borne measurement position. The airborne noise reference position in the passenger cabin is CD_L . This physically based ranking is opposed to a ranking that was derived via ANN 7. To create the ANN ranking, the median values for each feature in $a_{mod}(\omega)$ were processed through the network. To determine the influence of each measured channel, the median values were processed 24 more times, subsequently zeroing the input for each measured channel at a time. To clarify with an example, in the first iteration, the input of the TRL in the *x*-direction is set to zero. The ANN response is then compared to the response acquired with the full input vector. Normalizing and opposing the results of the measured and

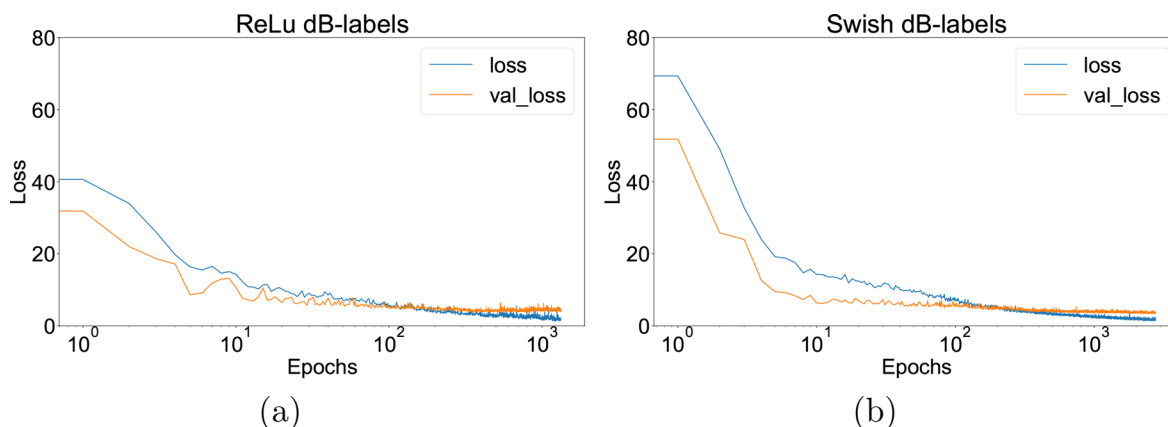


FIG. 9. (Color online) The loss curves for a ReLU-based and swish based networks trained on $a_{mod}(\omega)$ and $p_{mod}(\omega)$. The labels are in dB. (a) shows the loss curves of the ReLU-based ANN, and (b) shows the loss curves of the swish based ANN.

TABLE II. Prediction performances for ANN 7 and ANN 8 for all OCs. The mean absolute dB deviation for calculated third-octave bands is shown.

	OC: 200 MAE dB	OC: 300 MAE dB	OC: 400 MAE dB	OC: 500 MAE dB	OC: 600 MAE dB
ANN 7 (see Fig. 10)	4.0	4.0	2.94	3.22	5.37
ANN 8 (see Fig. 12)	4.49	3.02	3.04	4.41	5.22

machine-learned sensitivities results in Fig. 15. In the top part, the sensitivities are overlaid on a picture of the vehicle underbody. For each measurement position, there are 2×4 bars displayed. Using the TRR as an example, the left four bars are solidly contoured. The vertical, gray-filled bar represents the sum of the x -, y -, and z -sensitivity for this measurement position according to ANN 7. The growth direction is from the bottom to top. The related translatoric sensitivities increase horizontally from left to right with the colors given at the coordinate system. The dashed contoured bars follow the same logic for the measured sensitivities. For the TRR, the ANN and measurement rate the x - and z -direction to be more sensitive than the y -direction. Figure 15(b) gives a direct overview of all relative

sensitivities from Fig. 15(a), excluding the sensitivity sums. The ANN weighs the coupling points, especially the BL, FL, and FR, as being stronger compared to the measured values. The z -direction is emphasized for the BL and FL. The electric drive (EMOT) has low to medium importance for both methods. Again, the ANN shows the strongest influence on the z -direction, whereas the measurement highlights the x -direction. The TRL and STCOL are evaluated comparably. For the TRR, both methods show the highest importance in the x - and z -directions. The measurement roughly shows a sensitivity that is three times as high. Figure 15(c) shows the average coherence for the measured relative sensitivities. The values reflect the frequency range between 200 and 2000 Hz as a result of the technical limitations of

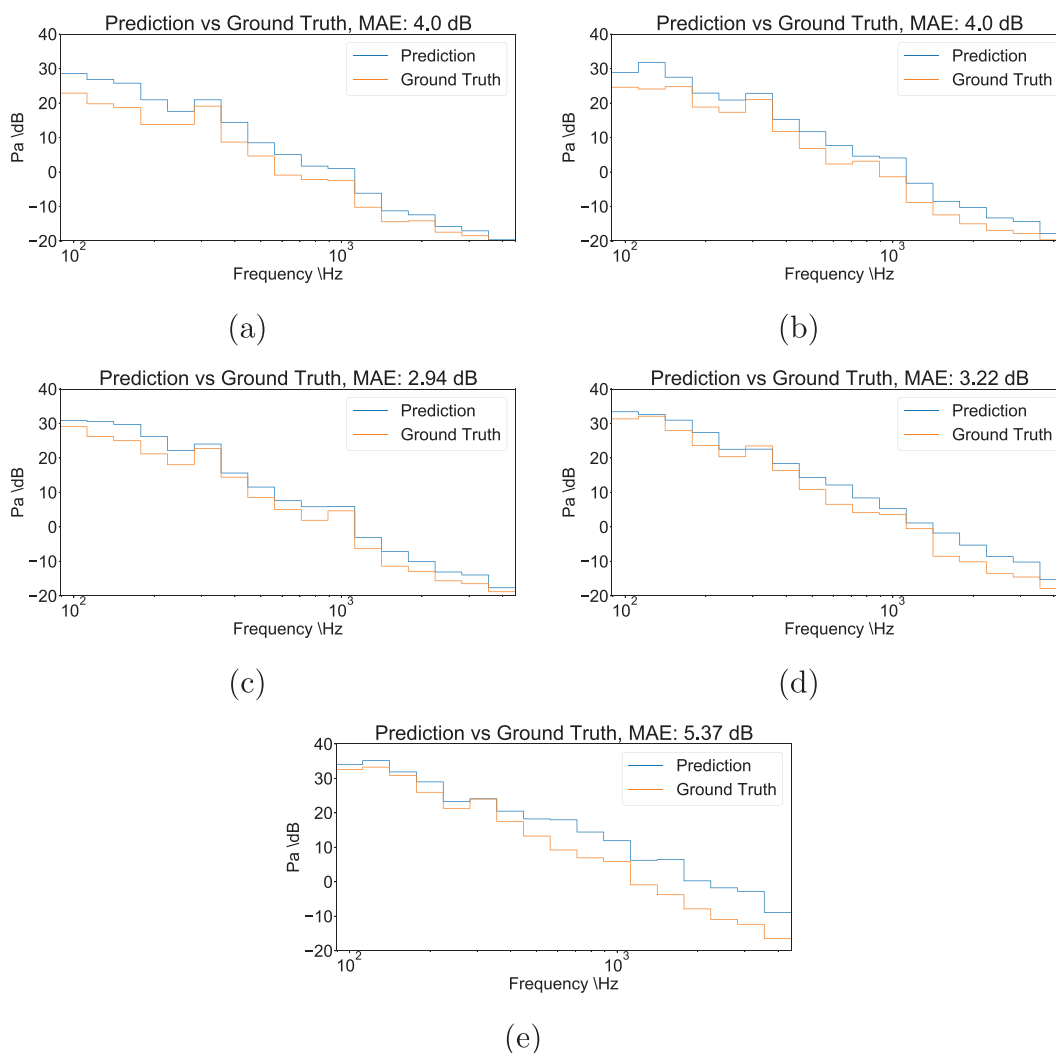


FIG. 10. (Color online) ANN 7 (ReLU) predictions versus the ground truth. The ANN was trained on dB values. (a) 200 deg/s, (b) 300 deg/s, (c) 400 deg/s, (d) 500 deg/s, and (e) 600 deg/s are the steering wheel angular speeds.

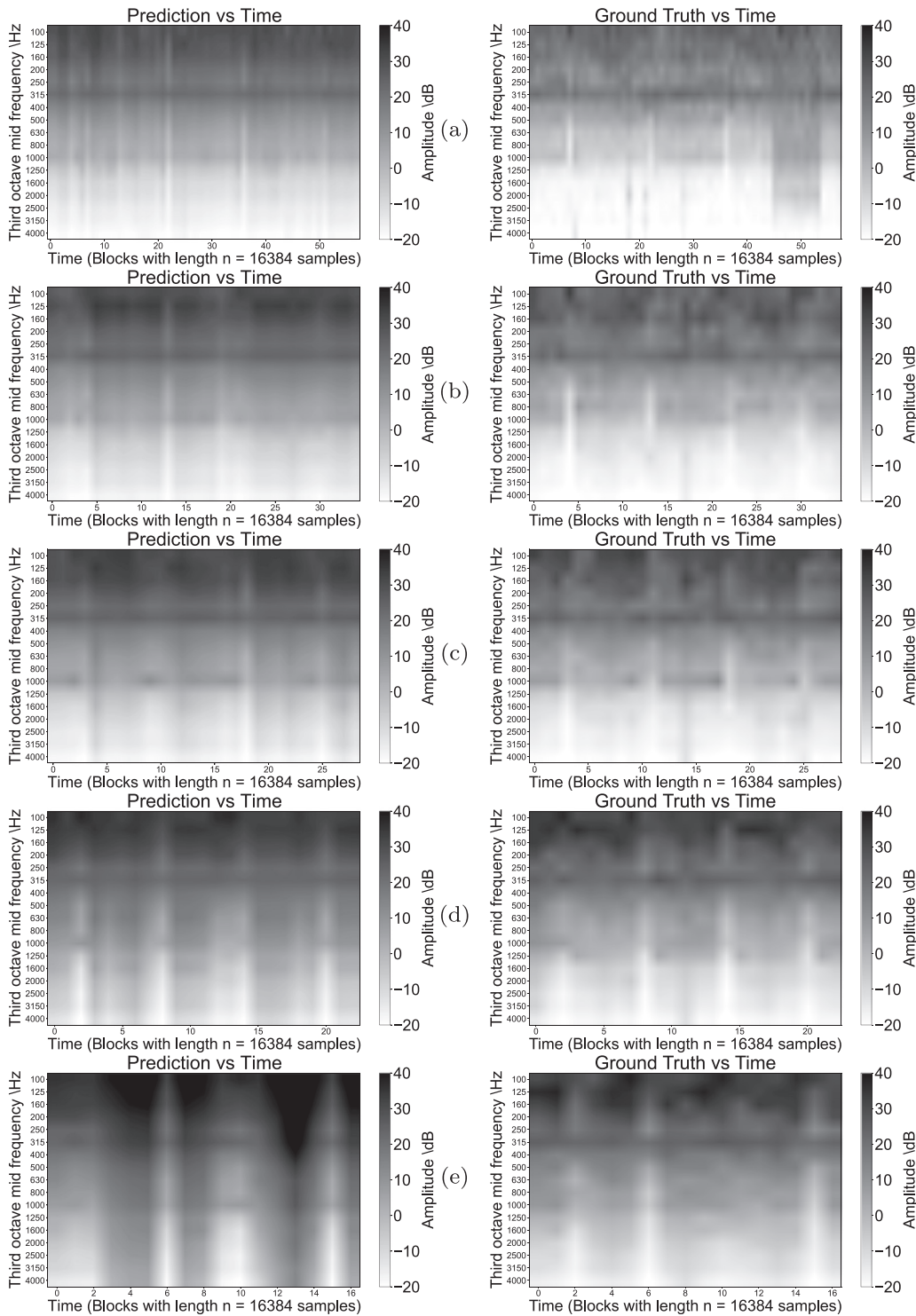


FIG. 11. The ANN 7 (ReLU) predictions versus the ground truth (time discrete). The ANN was trained on dB values. (a) 200 deg/s, (b) 300 deg/s, (c) 400 deg/s, (d) 500 deg/s, and (e) 600 deg/s are the steering wheel angular speeds.

the volume sound source that was used. With the exception of the electric drive in the z -direction and TRL in the x -direction, a coherence of more than 70% is reached.

IV. DISCUSSION

Training ANNs to describe noise propagation of a steering system in a full vehicle environment requires a specific

conditioning of the ML procedure. Using accelerometers on the vehicle component and a microphone on the desired response position, this paper proposes to train ANNs with operational data to acquire a vehicle surrogate. The network architecture, as well as the feature and label combination, has been shown to largely influence the outcome. The dying ReLU phenomenon and its countermeasures are discussed in the first subsection to sensitize future research. Due to

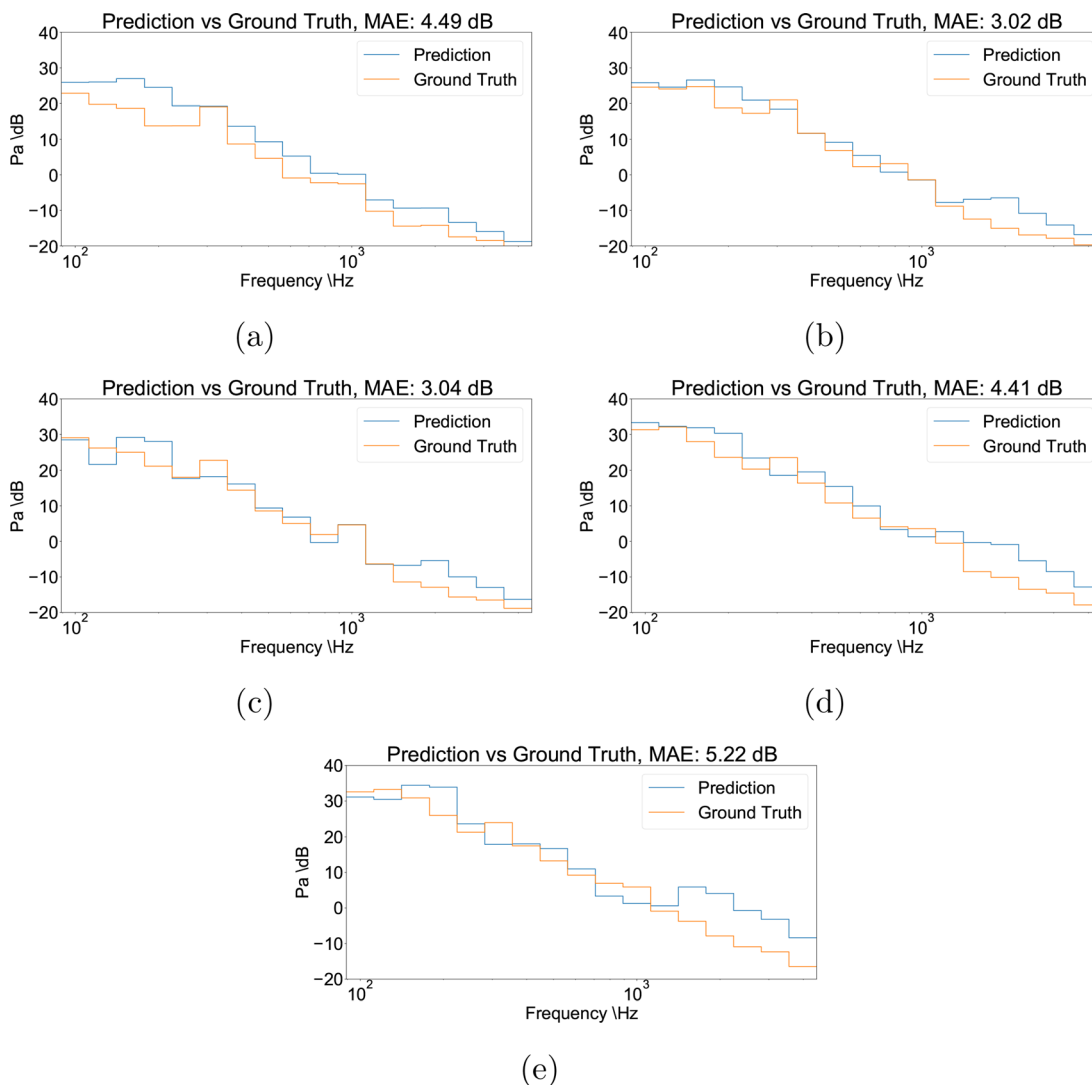


FIG. 12. (Color online) The ANN 8 (swish) predictions versus the ground truth. The ANN was trained on dB values. (a) 200 deg/s, (b) 300 deg/s, (c) 400 deg/s, (d) 500 deg/s, and (e) 600 deg/s are the steering wheel angular speeds.

explainability being a major factor for today’s ML approaches, a comparison with conventionally measured sensitivities was shown. Thus, the second subsection of the discussion aims at the physical comparison of the approach. To account for the practical usage besides the sensitivities, the final vehicle surrogates have been used to calculate the spectral system response for a steering maneuver. The usability in an automotive environment is discussed.

A. Dying ReLU

The results for eight different ANNs have been presented. ANN 1–ANN 6 from Table III are ReLU based networks because it is the most common nonlinear activation as of today. ANN 1 is trained for five different operational conditions at once. ANN 2–ANN 6 are trained for each individual operational condition. The ReLU based average third-octave band results from ANN 1 in Fig. 16 show promising results, especially for the 300 and 400 deg/s target steering wheel angular speeds. MAE values of 1.31 and

1.05 dB with a good fit for the third-octave band trend are displayed. Training ANN 2–ANN 6 for each specific OC decreases the error for the average third-octave band amplitude predictions even further. The lowest MAE of 0.42 dB is reached for a 400 deg/s target speed (Fig. 17). Calculating the time-discrete system response with ANN 2–ANN 6 shows constant values for each frequency regardless of the input (Fig. 18). As the ReLU function from Eq. (4) becomes zero for negative values, it is possible for the network to learn to ignore all of the inputs. The dying ReLU phenomenon hinders the neurons to recover because the gradient for negative values is zero as well.^{28,29} Despite ignoring the actual input vector, the network predicts a learned average value from the dataset. This happens by updating the weights for the bias of the last layer. The weights of this bias layer of the ANN are exactly equal to the predicted amplitudes. As the last layer is using the identity function, it is not affected by the dying ReLU phenomenon. It throughputs the last bias layer values. The same observation can be made when investigating ANN 1. This behavior also

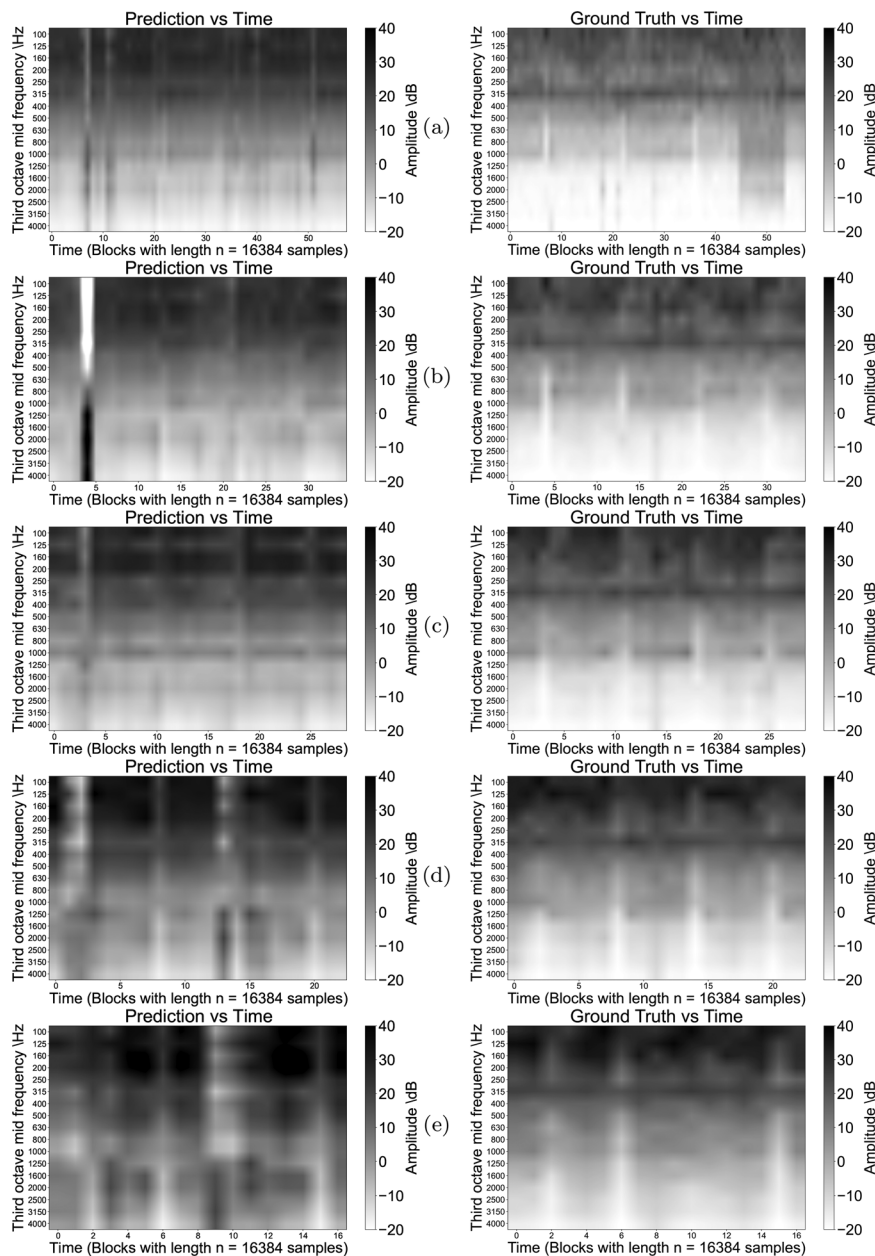


FIG. 13. The ANN 8 (swish) predictions versus the ground truth (time discrete). The ANN was trained on dB values. (a) 200 deg/s, (b) 300 deg/s, (c) 400 deg/s, (d) 500 deg/s, and (e) 600 deg/s are the steering wheel angular speeds.

explains the better MAE values for ANN 2–ANN 6, which are only trained for a single steering wheel angular speed. The spread of the values that has to be covered within one OC is smaller than the spread covering all of the OCs.

B. Measures: Swish and label range

To overcome the dying neurons, two measures have been implemented before starting the grid search of ANN 7 and ANN 8. The first measure makes the swish activation function available for the parameter variation. Swish features a nonzero gradient for each input value and is, therefore, more robust against vanishing gradient phenomena like the dying ReLU. The second measure was changing the value range for the labels. Originally training on Pa values,

the targets were close to zero. For Glorot uniform initialized layer weights, zero initialized bias weights, and typical structure-borne accelerations as inputs, this results in mostly negative bias weights for all of the layers except the output layer at the end of training. Hence, the ReLU activations will output zero. Using dB labels for training seems to stabilize the backpropagation process. The neurons remain active. Both, ANN 7 (ReLU) and ANN 8 (swish) have larger error values than ANN 1 (dying ReLU), which was also trained on all OCs. Calculating the time-discrete system responses shows that both networks are processing the inputs to the output layer. The calculated system response is changing throughout the steering cycle (Figs. 11 and 13). The time-discrete predictions show comparable characteristics as the ground truth (Fig. 14). Discrete phenomena, like

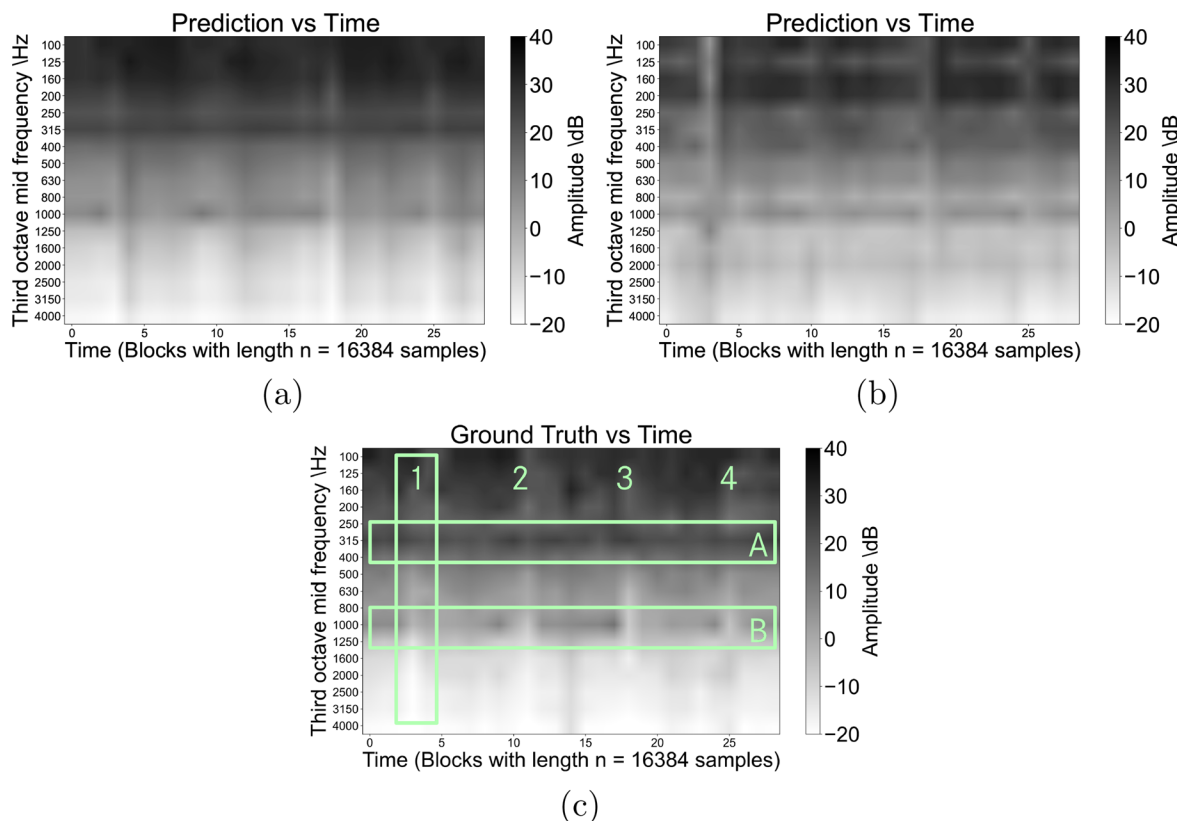


FIG. 14. (Color online) The comparison of (a) ANN 7 and (b) ANN 8 predictions for the 400 deg/s steering wheel angular speed versus (c) the ground truth. The change of the rotation direction for the ground truth is marked as 1–4. The constant high amplitude bands for the ground truth are marked with A and B.

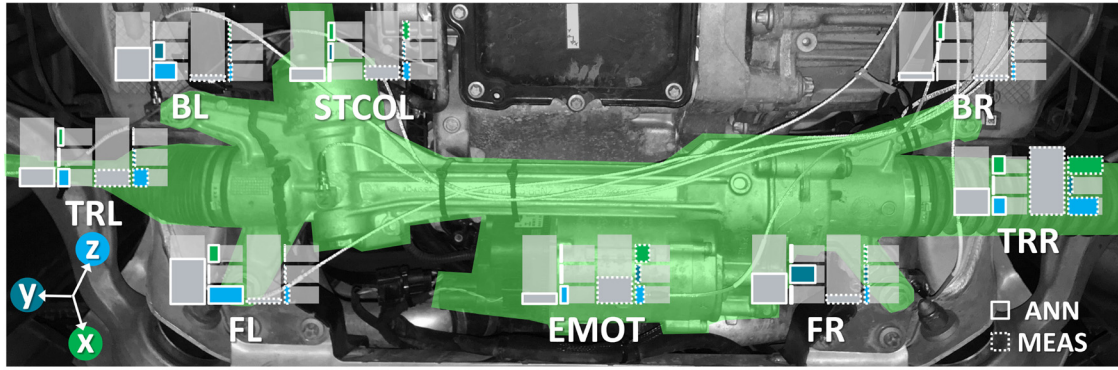
the moment of steering direction change, as well as constant phenomena, such as the continuous high amplitudes at 315 and 1000 Hz, are predicted correctly. Comparing ANN 7 and ANN 8, the objective overall performance of the ReLU based network is slightly better compared to swish with respect to the MAE metric. The predictions in Figs. 11 and 13 subjectively confirm the objective values. Especially for 500 and 600 deg/s, the swish prediction contains more noise. The architecture of the swish network is smaller in terms of the trainable parameters. It is to be investigated whether a larger swish based architecture can outperform a ReLU based approach. Another topic to investigate in the future is the influence of unbalanced training data. As the faster OCs are of shorter duration until completion of one steering cycle, these speeds are underrepresented in the dataset. This might have led to poorer learning for these conditions. To improve precision, a larger and more balanced dataset should be tested for training.

C. Sensitivities and practical usage

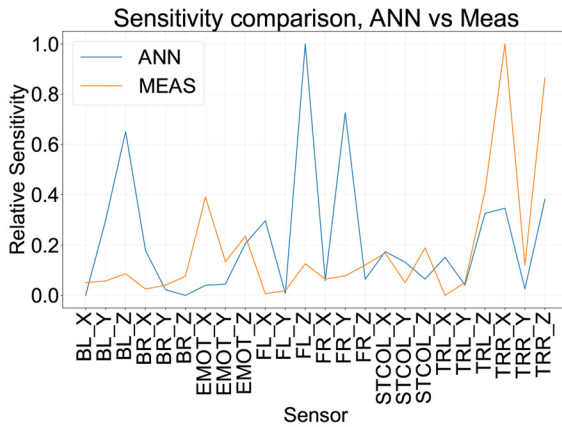
Because ANN 7 is slightly superior compared to the others, it has been additionally compared to a reciprocal volume sound source measurement. The sensitivities of ANN 7 and reciprocal TPA have similarities and differences at the same time. Generally speaking, the coupling points are more important for the ANN, whereas the tie rods dominate the ML approach. Despite showing this overall difference, both approaches agree

on the TRR being valuable for the interior noise. Showing good overall coherence, the measured TPA seems trustworthy. As mentioned before, the ANN shows higher sensitivity for the coupling points and, therefore, reduces the impact rating of the tie rod in terms of the noise propagation. Keeping in mind that the TPA was only performed for a frequency range from 200 to 2000 Hz and also recalling the predictions over time from Fig. 14, a possible explanation could be that the ANN is extracting additional information between 2000 and 4467 Hz. a_{mod} and p_{mod} have spectral information available that the conventional approaches data do not contain. This is due to the volume sound source (conventional) being limited in the frequency range it can produce excitation. Since the ANN is working with a_{mod} and p_{mod} it has chances to extract spectral links from the data in higher frequencies that are not contained in the conventional approaches data.

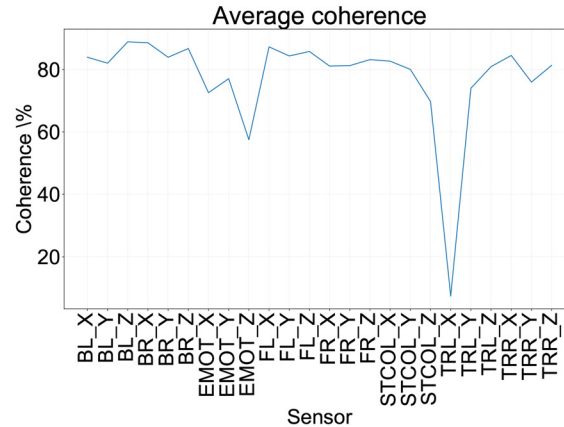
Discussing the practical use of the approach for industrial application, the spectral responses over time and sensitivity extraction both show potential. For a structure-borne interior noise calculation, one of the main drawbacks is the time needed to assess the full vehicle transfer behavior. The ANN-calculated system responses over time subjectively enable us to evaluate a steering maneuver. Whether the objective values are sufficient enough depends on the requirements. Using the ANN method is by far faster than the conventional approach. In addition to the faster execution, the ANN sensitivities indicate the counter spectral limitations of typical exciters like impact hammers or volume



(a)



(b)



(c)

FIG. 15. (Color online) The comparison of the relative sensitivity for each structure-borne channel according to the ANN and reciprocal measurement. (a) shows a graphical visualization. Gray bars with solid outlines are the sum of the x -, y -, z -sensitivities for each position according to the ANN. The growth direction is from bottom to top. The colored bars with the solid outline are the x -, y -, z -sensitivities according to the ANN. The growth is from left to right. The bars with the dashed outlines follow the same logic, representing the relative sensitivities for the measurements. (b) shows the corresponding numerical values for the ANN sensitivities and measured sensitivities. (c) provides the average coherence from 200 to 2000 Hz.

TABLE III. Architectures and training performances for different ReLU and swish based ANNs. The hyperparameter variation is done via the grid search. Backpropagation for ANN 1–ANN 6 based on Pa labels. The backpropagations for ANN 7 and ANN 8 is based on dB labels.

	ANN 1	ANN 2	ANN 3	ANN 4	ANN 5	ANN 6	ANN 7	ANN 8
	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ Pa	$a(\omega) \frac{m}{s^2} p(\omega)$ dB	$a(\omega) \frac{m}{s^2} p(\omega)$
	OC: all	OC: 200	OC: 300	OC: 400	OC: 500	OC: 600	OC: all	dB OC: all
In layer d_{IN}	408	408	408	408	408	408	408	408
Layer 1 d_{HL1}	2008	1050	954	720	450	40	520	119
Activation 1	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	Swish
Layer 2 d_{HL2}	669	350	318	480	50	60	607	81
Activation 2	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	ReLU	Swish
Layer 3 d_{HL3}	223	—	328	—	50	60	693	—
Activation 3	ReLU	—	ReLU	—	ReLU	ReLU	ReLU	—
Layer 4 d_{HL4}	—	—	—	—	450	40	780	—
Activation 4	—	—	—	—	ReLU	ReLU	ReLU	—
Dropout	0.15	0.15	0.15	0.15	0.15	0.15	0.15	0.15
Out layer d_{OUT}	17	17	17	17	17	17	17	17
Activation out	id	id	id	id	id	id	id	id
Parameters	2 318 511	803 267	804 101	648 737	239 767	25 617	1 504 868	59 785
Validation	41.00	21.18	29.78	23.07	38.23	31.85	29.92	31.89
MAPE %								

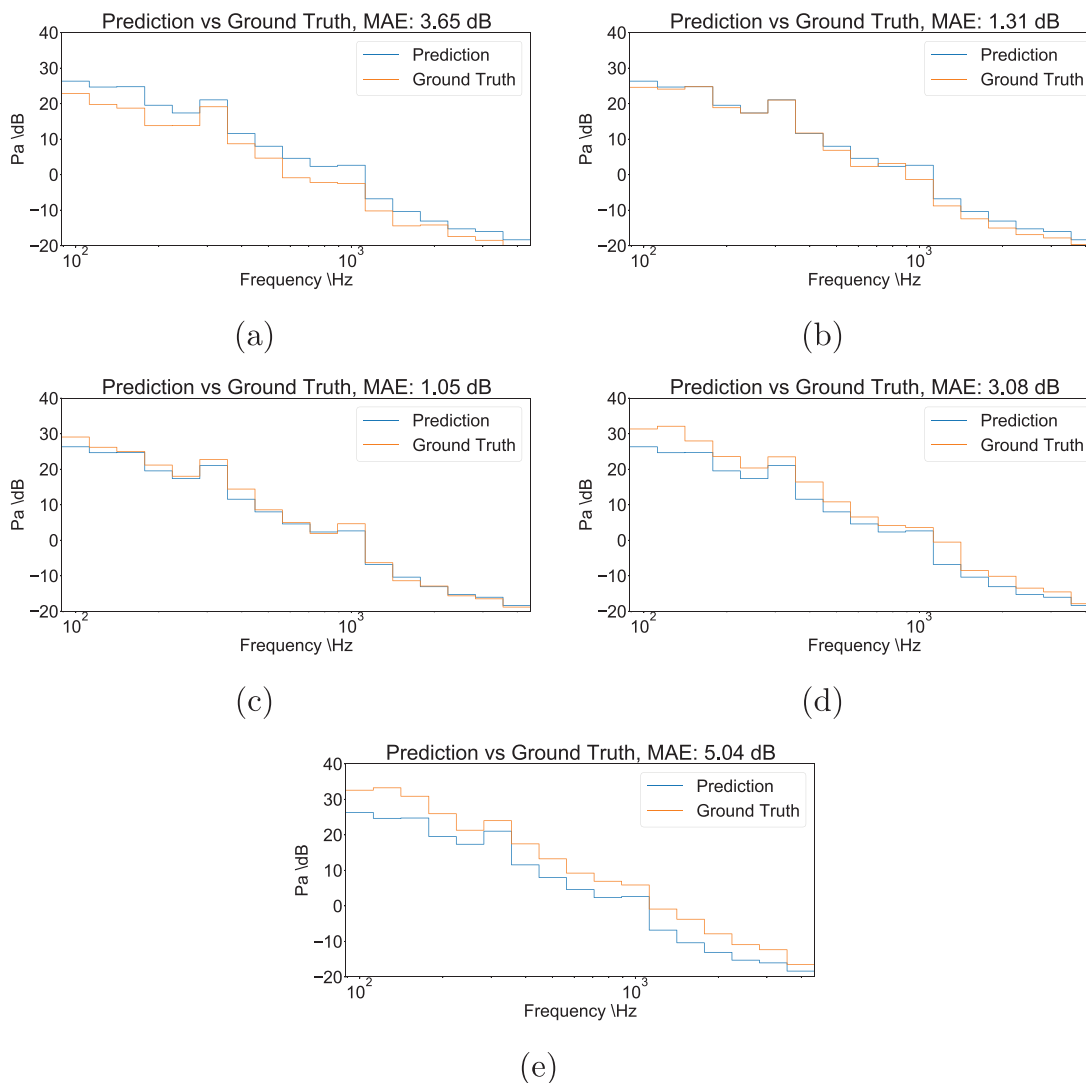


FIG. 16. (Color online) The ANN 1 (ReLU) predictions versus the ground truth for (a) 200 deg/s, (b) 300 deg/s, (c) 400 deg/s, (d) 500 deg/s, and (e) 600 deg/s steering wheel angle target speeds.

sound sources. If this assumption proves to be true, it would reduce hardware costs by a substantial amount. Further investigations with a wider range of exciters are needed to strengthen the hypothesis.

V. CONCLUSION

Passenger cars are complex mechanical systems with a large number of components, some of which are producing enough excitation to be acoustically perceived in the passenger cabin. Determining the transfer behavior between the excitation and response is an enabler for the digital development processes. ML as a data-driven method for identifying transfer functions shows a high potential to be used for frequency domain interior noise predictions. Exclusively using operational measurements decreases costs compared with the existing experimental methods. Iteratively searching ANN architectures with nonlinear activation functions leads to a model that is capable of predicting the interior noises for different OCs. The network’s internal behavior is not

intuitively assessable due to the large number of trainable parameters. Further comparisons with conventional physical approaches throughout larger frequency ranges are necessary to enhance the understanding of the capabilities of the ML approach.

To conclude, the findings of this paper are as follows:

- NVH transfer behavior in the frequency domain can be extracted from operational measurements with ANNs. A system description to pair component excitation with vehicle influence is acquirable without performing common procedures (impact measurements, volume sound source, etc.). For steering systems, multiple OCs can be trained into one final network.
- The extraction of the transfer behavior with ML is not limited by the often narrow range of the exciter. For example, the volume sound source used was able to range up to 2 kHz, whereas the accelerators and microphone recorded with a sampling rate of 48 kHz. Hence, the analyzable range is much higher when using the natural excitation of

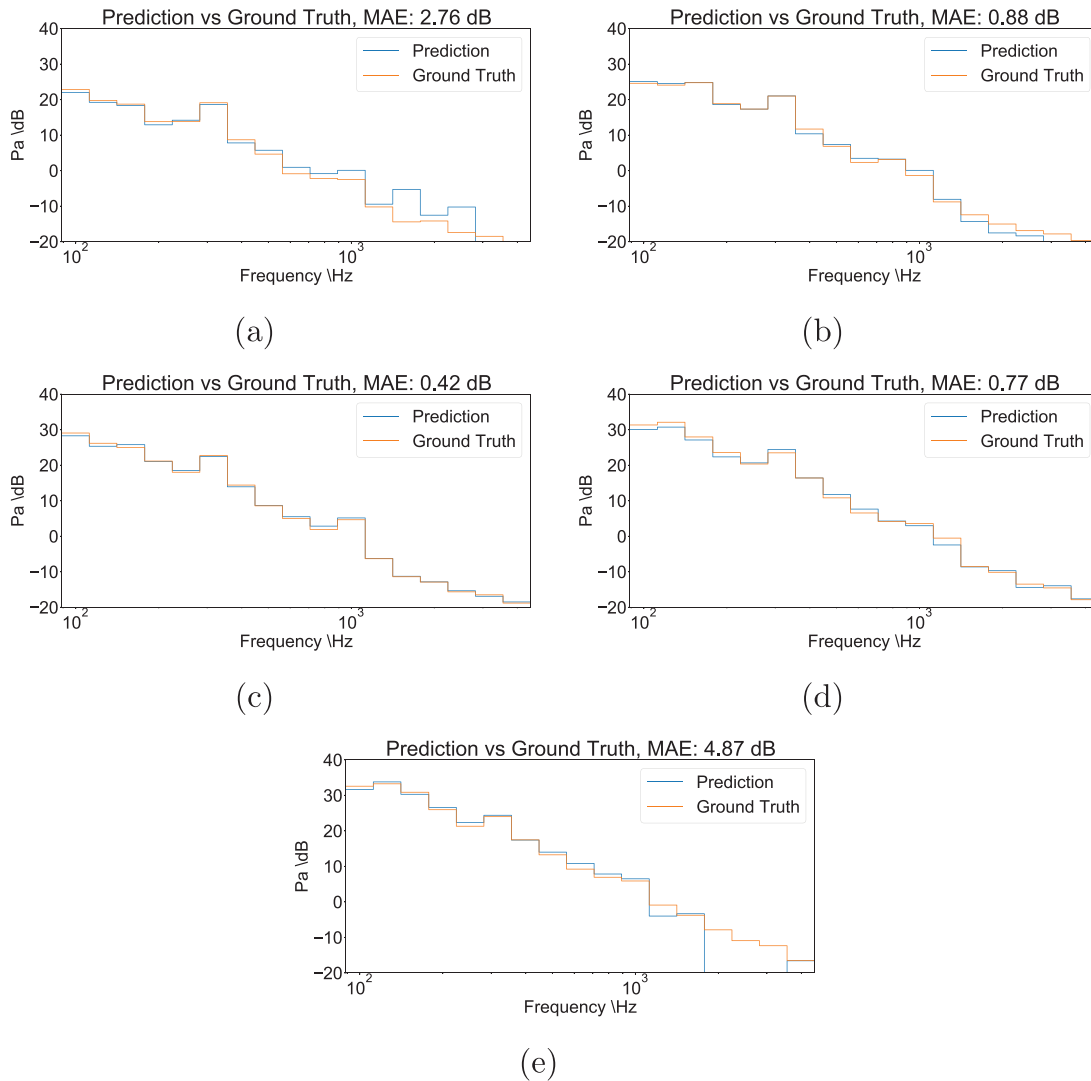


FIG. 17. (Color online) (a) The ANN 2 (ReLU) prediction for 200 deg/s, (b) ANN 3 (ReLU) prediction for 300 deg/s, (c) ANN 4 (ReLU) prediction for 400 deg/s, (d) ANN 5 (ReLU) prediction for 500 deg/s, and (e) ANN 6 (ReLU) prediction for 600 deg/s.

the system for a specific OC. An ANN trained on operational measurements can be used to derive sensitivities for structure-borne measurement positions. The results shown partly correlate with the results from a volume sound source measurement. Further investigations would be interesting to evaluate whether the ANN does extract correct sensitivities over the entire frequency range recorded through the accelerometers and microphone.

- The composition of the feature and label vectors is highly relevant to the training outcome. For airborne targets, the label range plays a key role in terms of the training success. If no normalization is applied to the labels in a NVH surrogate modelling task for airborne targets, then the user has to pay close attention to the training process. It is advisable to track the weight updating, since Pa values are close to zero, which might interfere with the backpropagation. Also, for ML use with airborne targets, countermeasures to vanishing gradient phenomena should be applied.
- Finding a suitable network architecture and its hyperparameters is an iterative process worth improving to save

time. Generally speaking, the internal behavior of ANNs for NVH propagation tasks should be closely analyzed. The large number of parameters and self-optimizing process lacks convenience in the post-processing.

From an automotive point of view, development processes are continuously shortened. Providing a method that reduces hardware needs and also time needs provides major benefits for the NVH investigations. Emphasizing the large efforts of the conventional TPA, getting a feeling for transfer behavior solely based on operational measurements seems to be a promising tool. Further, mentioning the often inconvenient handling of typical exciters with their limited frequency ranges in a full vehicle context makes it even more desirable to establish new ways with data based methods.

APPENDIX

Table III lists eight network configurations. All architectures shown are found by a grid search, which is limited to four hidden layers. For ANN 1–ANN 6, only ReLU

activations are implemented in the grid. ANN 7 and ANN 8 also feature swish as the possible activation. ANN 1 inputs a vector concatenating all structure-borne accelerations $a(\omega)$ from Fig. 3. Their amplitudes in the frequency domain are fed into the ANN without distinguishing between the OCs. The target values $p(\omega)$ are given in Pa. The training is conducted for data in a frequency range between 89.2 and 4467 Hz, where a third-octave resolution is used. Having 8 structure-borne sensors with 3 degrees of freedom and 17 third-octave bands per signal results in 408 input features. The best performing ReLU based architecture for all OCs trained on Pa values is three-layered. The hidden layers have the dimensions d_{HL1} , d_{HL2} , and d_{HL3} . The output layer has the dimension d_{OUT} with an identity activation. Its output reflects the acoustic system response $p(\omega)$. The entire network has 2 318 511 trainable parameters. The final validation MAPE of the network training reaches 41%. From Table III, ANN 2–ANN 6 are each trained on data that is

only comprised of excitations and responses of a single OC. The speeds are rising from 200 to 600 deg/s. The best performing architecture for each speed with regard to the training validation MAPE is listed. Values between 21.18% and 38.23% are found. ANN 7 and ANN 8 are trained on dB labels. All of the OC are included in the training data for these networks. ANN 7 is the best performing ReLU architecture with a training validation MAPE of 29.92%. ANN 8 is the swish based network with the lowest MAPE for these input and output configurations. The value reaches 31.89%.

To rate the prediction performance of each trained ANN, an unseen dataset is processed for every OC. The MAE values for the comparison of the network output versus the ground truth are presented in Table IV.

Figures 16(a)–16(e) represent the responses for ANN 1. It is trained on a dataset containing all five OCs. The sound pressure third-octave band amplitude spectra are chronologically sorted from the slowest to fastest steering wheel

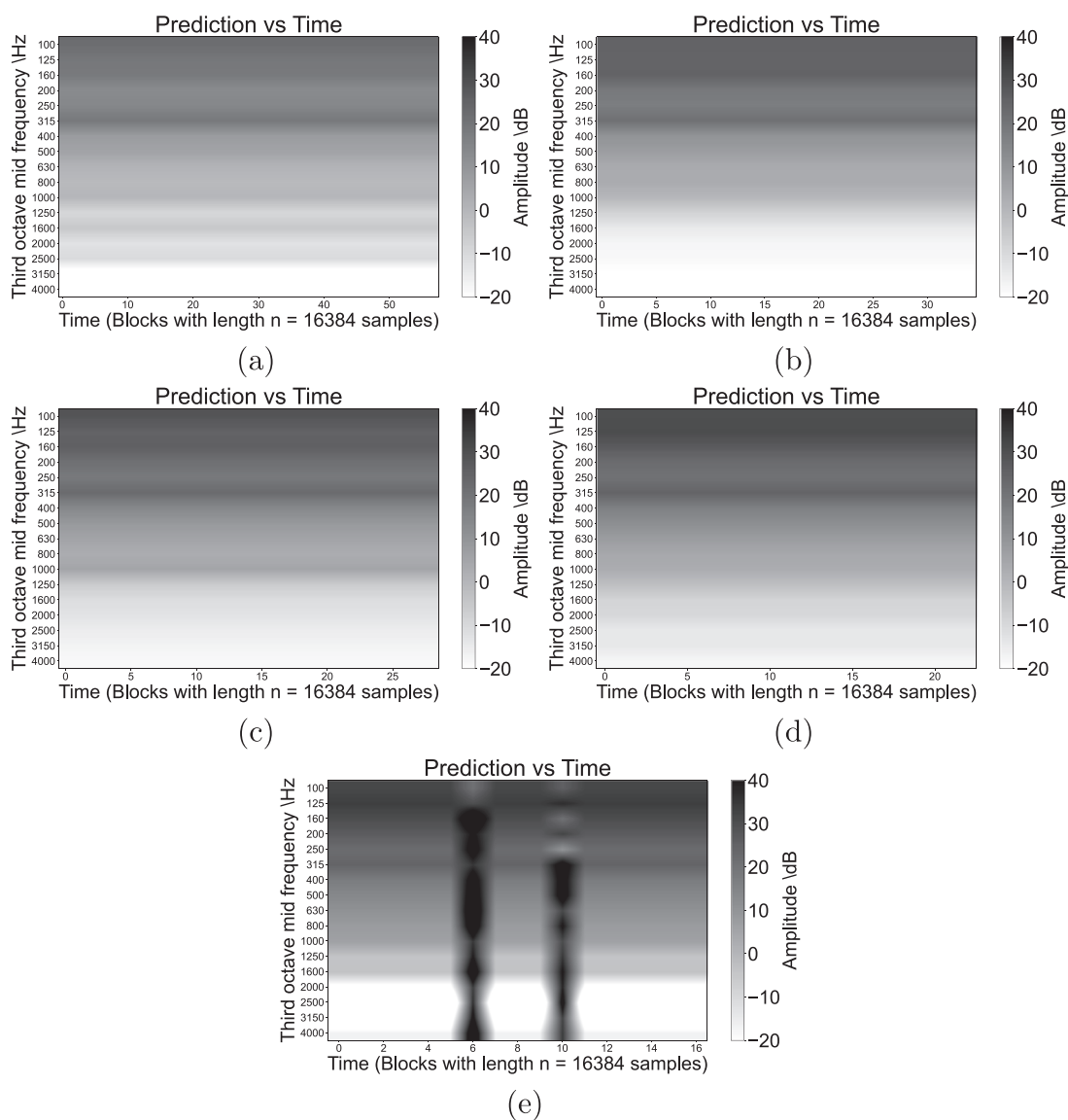


FIG. 18. The time-discrete third-octave band prediction of (a) ANN 2 (ReLU), 200 deg/s; (b) ANN 3 (ReLU), 300 deg/s; (c) ANN 4 (ReLU), 400 deg/s; (d) ANN 5 (ReLU), 500 deg/s; and (e) ANN 6 (ReLU), 600 deg/s.

TABLE IV. Prediction performances for ANN 1–ANN 8 for all OC. Mean absolute dB deviation for calculated third-octave bands is shown.

	OC: 200 MAE dB	OC: 300 MAE dB	OC: 400 MAE dB	OC: 500 MAE dB	OC: 600 MAE dB
ANN 1 (see Fig. 16)	3.65	1.31	1.05	3.08	5.04
ANN 2 [see Fig. 17(a)]	2.76	—	—	—	—
ANN 3 [see Fig. 17(b)]	—	0.88	—	—	—
ANN 4 [see Fig. 17(c)]	—	—	0.42	—	—
ANN 5 [see Fig. 17(d)]	—	—	—	0.77	—
ANN 6 [see Fig. 17(e)]	—	—	—	—	4.87
ANN 7 (see Fig. 10)	4.0	4.0	2.94	3.22	5.37
ANN 8 (see Fig. 12)	4.49	3.02	3.04	4.41	5.22

angular speed. One MAE value for each state from 200 to 600 deg/s is calculated.

Figure 17 contains the responses of ANN 2–ANN 6. The processed structure-borne accelerations of a maneuver with 200 deg/s result in the prediction in Fig. 17(a). The following amplitude spectra [Figs. 17(b)–17(e)] are subsequently showing the faster speeds.

Figure 18 shows the time-discrete third-octave band prediction of ANN 2–ANN 6 for their corresponding steering wheel angle target speeds. Each third-octave block is produced through 16 384 samples from each structure-borne measurement position. Linear interpolation is applied for the time and frequency axes. The prediction in Fig. 18(a) shows the airborne response for a maneuver with a 200 deg/s target speed. The dB values remain constant throughout the steering cycle for each frequency. The maneuvers in Figs. 18(b)–18(d) show comparable behavior. Processing the data of the 600 deg/s maneuver results in Fig. 18(e). It shows changes for the response of block six and block ten.

¹D. C. Barton and J. D. Fieldhouse, *Automotive Chassis Engineering* (Springer International, Cham, 2018).

²M. Ersoy and S. Gies, eds., *Fahrwerkhandbuch (Chassis Handbook)* (Springer Fachmedien Wiesbaden, Wiesbaden, 2017).

³M. V. van der Seijs, D. de Klerk, and D. J. Rixen, “General framework for transfer path analysis: History, theory and classification of techniques,” *Mech. Syst. Signal Process.* **68–69**, 217–244 (2016).

⁴P. Zeller, ed., *Handbuch Fahrzeugakustik (Handbook Vehicle Acoustics)* (Vieweg and Teubner, Wiesbaden, 2012).

⁵J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural Networks* **61**, 85–117 (2015).

⁶G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Math. Control, Signals, Syst.* **2**(4), 303–314 (1989).

⁷Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature* **521**(7553), 436–444 (2015).

⁸A. Brunetti, D. Buongiorno, G. F. Trotta, and V. Bevilacqua, “Computer vision and deep learning techniques for pedestrian detection and tracking: A survey,” *Neurocomputing* **300**, 17–33 (2018).

⁹Y. Zheng, Y. Liu, and J. Hansen, “Navigation-orientated natural spoken language understanding for intelligent vehicle dialogue,” in *2017 IEEE Intelligent Vehicles Symposium (IV)* (IEEE, New York, 2017).

¹⁰C. Ma, C. Chen, Q. Liu, H. Gao, Q. Li, H. Gao, and Y. Shen, “Sound quality evaluation of the interior noise of pure electric vehicle based on neural network model,” *IEEE Trans. Ind. Electron.* **64**(12), 9442–9450 (2017).

¹¹P. V. Kane and A. B. Andhare, “Application of psychoacoustics for gear fault diagnosis using artificial neural network,” *J. Low Freq. Noise, Vib. Active Control* **35**(3), 207–220 (2016).

¹²S. Lerspalungsanti, A. Albers, S. Ott, and T. Düser, “Human ride comfort prediction of drive train using modeling method based on artificial neural networks,” *Int. J. Automot. Technol.* **16**(1), 153–166 (2015).

¹³Z. He, Y. Qiu, E. Li, H. Wang, Y. Huang, Y. Shen, and Y. Du, “Analysis on loudness of exhaust noise and improvement of exhaust system based on structure-loudness model,” *Appl. Acoust.* **150**, 104–112 (2019).

¹⁴T. Li, R. Burdisso, and C. Sandu, “An artificial neural network model to predict tread pattern-related tire noise,” in *SAE Technical Paper Series* (SAE International, Warrendale, PA, 2017).

¹⁵M. Stender, M. Tiedemann, D. Spieler, D. Schoepflin, N. Hoffmann, and S. Oberst, “Deep learning for brake squeal: Brake noise detection, characterization and prediction,” *Mech. Syst. Signal Process.* **149**, 107181 (2021).

¹⁶K. Qian and Z. Hou, “Intelligent evaluation of the interior sound quality of electric vehicles,” *Appl. Acoust.* **173**, 107684 (2021).

¹⁷Y. S. Wang, N. N. Liu, H. Guo, and X. L. Wang, “An engine-fault-diagnosis system based on sound intensity analysis and wavelet packet pre-processing neural network,” *Eng. Appl. Artif. Intell.* **94**, 103765 (2020).

¹⁸L. Wang, S. Chen, D. Wang, Y. Jiang, and J. Chen, “Sound absorption optimization of porous materials using bp neural network and genetic algorithm,” in *SAE Technical Paper Series* (SAE International, Warrendale, PA, 2016).

¹⁹G. Capuano and J. Rimoli, “Smart finite elements: A novel machine learning application,” *Comput. Methods Appl. Mech. Eng.* **345**, 363–381 (2019).

²⁰D. E. Tsokaktsidis, T. von Wysocki, F. Gauterin, and S. Marburg, “Artificial neural network predicts noise transfer as a function of excitation and geometry,” in *Proceedings of the 23rd International Congress on Acoustics, Integrating 4th EAA Euroregion*, edited by M. Ochmann (RWTH, Aachen), pp. 4392–4396.

²¹D. Lee and J. W. Lee, “Operational transfer path analysis based on deep neural network: Numerical validation,” *J. Mech. Sci. Technol.* **34**(3), 1023–1033 (2020).

²²D. E. Tsokaktsidis, C. Nau, and S. Marburg, “Time domain full vehicle interior noise calculation from component level data by machine learning,” in *SAE Technical Paper Series* (SAE International, Warrendale, PA, 2020).

²³X. Glorot and Y. Bengio, “Understanding the difficulty of training deep feedforward neural networks,” in *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics* (PMLR, 2010), Vol. 9, pp. 249–256, available at <http://proceedings.mlr.press/v9/glorot10a.html> (Last viewed July 6, 2021).

²⁴X. Glorot, A. Bordes, and Y. Bengio, “Deep sparse rectifier neural networks,” in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (2011), pp. 315–323.

²⁵P. Ramachandran, B. Zoph, and Q. Le, “Searching for activation functions,” *arXiv:1710.05941* (2017) (Last viewed July 6, 2021).

²⁶A. de Myttenaere, B. Golden, B. L. Grand, and F. Rossi, “Mean absolute percentage error for regression models,” *Neurocomputing* **192**, 38–48 (2016).

²⁷Y. Dodge, *The Concise Encyclopedia of Statistics* (Springer, New York, 2008).

²⁸D. Pedamonti, “Comparison of non-linear activation functions for deep neural networks on MNIST classification task,” *arXiv:1804.02763* (2018) (Last viewed July 6, 2021).

²⁹L. Lu, Y. Shin, Y. Su, and G. Karniadakis, “Dying ReLU and initialization: Theory and numerical examples,” *arXiv:1903.06733* (2019).