Deep Learning Based Soft Sensors for Industrial Machinery

Benjamin Maschler[a,c]*, Sören Gansslosser[b,c], Andreas Hablizel[b], Michael Weyrich[a]

[a]*University of Stuttgart, Institute of Industrial Automation and Software Engineering, Pfaffenwaldring 47, 70550 Stuttgart, Germany*
[b]*AVAT Automation GmbH, Derendinger Straße 40, 72072 Tübingen, Germany*
[c]*These authors contributed equally to this publication.*

* Corresponding author. Tel.: +49 711 685-67295; Fax: +49 711 685-67302. *E-mail address:* benjamin.maschler@ias.uni-stuttgart.de

Abstract

A multitude of high quality, high-resolution data is a cornerstone of the digital services associated with Industry 4.0. However, a great fraction of industrial machinery in use today features only a bare minimum of sensors and retrofitting new ones is expensive if possible at all. Instead, already existing sensors' data streams could be utilized to virtually 'measure' new parameters. In this paper, a deep learning based virtual sensor for estimating a combustion parameter on a large gas engine using only the rotational speed as input is developed and evaluated. The evaluation focusses on the influence of data preprocessing compared to network type and structure regarding the estimation quality.

Keywords: Artificial Intelligence, Deep Neural Networks, Multi-Layer-Perceptron, Retrofitting, Soft Sensors, Upgrading, Virtual Sensors

1. Introduction

For at least ten years, there has been an increasingly strong trend of utilizing data driven artificial intelligence methods to improve machines, processes or products across different industrial domains [1]. This improvement could take different forms, from lowering energy or raw material consumption [2] across a better utilization of machinery [3, 4] or higher levels of automation [5] to increasing the quality of the output [6]. In recent years, optimizing emissions due to stricter environmental regulations has been an important driver as well [7, 8].

However, collecting the data necessary for such approaches is facing several challenges, among which is the longevity of industrial machinery: Even official service lifetime estimates for depreciation run from (rarely) 6 to over 30 years depending on country, type of machinery and industrial sector [9]. Furthermore, experience shows that especially in small and medium sized enterprises resilient equipment might be in daily use even longer. Thus, a lot of today's machinery has been manufactured long before the advent of today's hunger for data. It therefore oftentimes lacks at least some of the sensors to collect it.

Problem Statement: Upgrading this machinery to include more sensors for the direct measurement of the high quality, high-resolution data needed for data driven artificial intelligence methods [10] is oftentimes prohibitively expensive or even impossible due to a lack of space [11]. Yet, even in better cases, it requires disassembling machines and thereby causes long downtimes.

However, using already existing sensors' signals as an input for deep neural networks to indirectly infer the desired data rather than measure it directly could be a viable alternative [11]. These so-called soft sensors or virtual sensors could therefore potentially satisfy today's need for data while still using yesterday's hardware.

Objective: A virtual sensor to upgrade an existing industrial machine is developed and evaluated. In doing so, the influence of network type, network structure and data preprocessing on the sensor's estimation accuracy is analyzed.

Structure: In this paper, a use case for soft sensors in industrial machinery is described (see *Sec. 2*) and literature surveyed in order to identify promising approaches (see *Sec. 3*). Thereon, a methodology is developed in *Sec. 4* and thoroughly analyzed (see Results in *Sec. 5*). Finally, a conclusion is drawn in *Sec. 6*.

2. Case study description

For the last years, gas engines for combined heat and power plants (CHP) have been facing stricter environmental regulations, e.g. regarding their efficiency or the emission of potentially harmful substances – a process that will continue in the foreseeable future [12]. To fulfil these requirements, it is necessary to optimize the engine's operation based upon a monitoring of its combustion parameters. Such an optimization then strives to minimize irregular states of operation and balance each cylinder's combustion and load.

Currently, this monitoring is carried out by measuring the cylinder pressure by cylinder pressure sensors and then using it to calculate other relevant combustion parameters. However, such sensors are only used in large engines due to their additional costs as well as their space and maintenance requirements. Therefore, to facilitate the above mentioned optimization procedures in smaller engines, too, the use of virtual cylinder pressure sensors is a promising option.

2.1. Engine parameters

When operating a gas engine, both engine-global and cylinder-specific actuating variables can be used to alter the engine's behavior. Among the first are the injection volume, the air-fuel-ratio λ and the global ignition angle, whereas the latter consists of e.g. each cylinders' offset on the engine's global ignition angle which can be used to optimize the individual combustion processes in order to lower emissions or increase efficiency.

To calculate the actuating variables necessary for this optimization (or: balancing) process, each individual combustion process should be monitored, usually by measuring the cylinder pressure. The most relevant combustion parameters that can be calculated using a known combustion pressure curve are usually considered the maximum combustion pressure (p_{Max}), the indicated mean effective pressure (IMEP), the center of combustion (CoC), the *duration of combustion* (DoC) and the transformed fuel energy (HR).

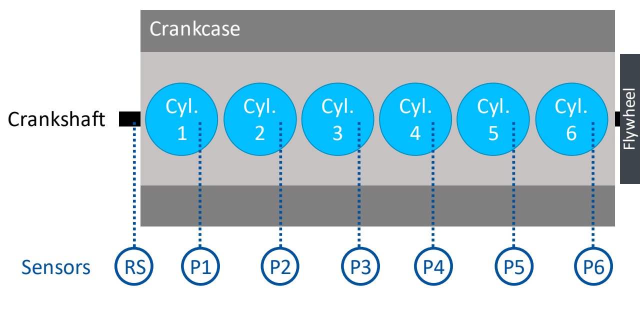

Figure 1: Schematic representation of a gas engine's crankcase indicating the position of a sensor for rotational speed (RS) and sensors for cylinder pressure (P1 to P6) as used in the data collection setup

These differ for each individual combustion process and are therefore *cylinder-specific combustion parameters*.

By averaging the cylinder-specific combustion parameters, *engine-global combustion parameters* can be derived which cannot be used to facilitate the optimization process. However, they provide a (less accurate) overview of the engine's combustion behavior and can thereby serve as an indicator of the need for optimization.

As a substitute for the cylinder pressure measurement, in this study, the *crankshaft's rotational speed* is used as the input measurement from which the DoC as an example for all other relevant cylinder-specific combustion parameters is inferred by a deep neural network.

2.2. Dataset

The dataset used was created on a mixture-supercharged six-cylinder gas engine with a displacement of 12 liters and an output of 210 kW, which is installed in a CHP. Each cylinder had a cylinder pressure sensor (P1 to P6). The rotational speed sensor (RS) was mounted on the crankshaft before cylinder 1 (see Fig. 1). Thereby, each cylinder's pressure curve as well as the engine's crankshaft's rotational speed were measured and the DoCs calculated based on the pressure curves.

Both the cylinder pressure indexing and the measurement of the rotational speed signal were performed with a sampling rate of 48 kHz. The speed signal was recorded on a gear rim with 120 teeth, resulting in a measuring point every 3° or 120 pulses per crankshaft revolution and 240 pulses per working cycle.

The data set contains data from more than 500,000 working cycles from different operating points of the engine. For that purpose, the engine-global actuating variables have been varied in a way suitable for simulating the actual operation of such an engine. In detail, the target power was varied between 25 % and 100 %, the ignition angle between 24° and 32° and an air-fuel-ratio λ between 1.52 and 1.60.

For the training and evaluation of the neural networks, the dataset was divided into three parts: a training, a validation and a test dataset with respectively 64%, 16% and 20% of the total dataset.

3. Related work

In published literature, different approaches using deep neural networks in which the rotational speed signal was used to estimate combustion parameters can be found. Most publications (e.g. [13–16]) initially estimate the cylinder pressure curve and then derive the combustion parameters (usually at least p_{Max} and its position in the crank angle range) for subsequent evaluation. There are considerably fewer publications, e.g. [17], which, as in this study, estimate and evaluate the combustion parameters directly by means of a neural network.

Furthermore, these publications differ significantly as to which factors are considered to be most relevant for the algorithms' performance.

Reference [13] uses a neural network with non-linear autoregressive with exogenous input (NARX) architecture to estimate the cylinder pressure. It was evaluated based on the relative p_{Max}-error and p_{Max}'s position in the crank angle range showing good results: Depending on the operating condition, the mean error lay between 5.3 % and 33.6 % respectively between 1.7° and 4.3°.

In [14], the same authors achieved a significant improvement of this result by using a time-delay neural network (TDNN). In the same scenarios as used in [13], the mean error decreased to just 1.14 % to 1.32 % respectively 1.65° to 3.08°.

Thus, for these two publications, the type of network used greatly affected the algorithms‘ performance.

In [16], a recurrent neural network is proposed that, in addition to the rotational speed signal, also receives information about the air-fuel-ratio λ, the ignition angle and the boost pressure of the turbocharger. Here, the focus lay on the structure of the neural network, which greatly affected the algorithm's performance.

In [15], a neural network with radial basis functions (RBF) and, thus, without recurrence is used to estimate cylinder pressure curves. In contrast to other studies that have used an RBF network, the authors of [15] do not use the raw rotational speed signal, but transform it into the frequency domain and processes only the first 20 harmonics. In addition, they use the 21st-50th harmonics of the structure-borne sound signal. Thus, the decisive factor in this work is the preprocessing of the available data. The mean errors achieved thereby are 3.4 % for p_{Max} and 1.5° for its position in the crank angle range.

In contrast to the publications mentioned so far, [17] estimates the combustion parameters directly from the crankshaft's rotational speed and acceleration signal using a multi-layer perceptron (MLP). Here, the mean error lays between 4.1 % and 8.0 % respectively between 1.38° and 9.1°.

4. Approach

Based on prior research (see *Sec. 3*), it is highly likely that estimating the combustion parameters from the rotational speed signal by deep neural networks is possible. However, as described above, the identified studies all differ significantly in their methodology. In addition, all of them except for [13] and [14] which were from the same authors, used different datasets for evaluation greatly limiting comparability of the results.

Therefore, this study is designed to be able to test all three major factors (preprocessing of the input data, type of neural network, structure of neural network). Due to time restraints, the different approaches were tested on estimating engine-global parameters first. Based on their performance therein, the most promising ones were then adapted to estimate cylinder-specific combustion parameters.

4.1. Data preprocessing

In a first step of data preprocessing, the rotational speed signal is transformed from the time into the crank angle domain, because any evaluation of the combustion parameters takes place in this domain as well. Furthermore, implausible data is filtered from the dataset. An example of the resulting data is depicted in Fig. 2.

To analyze the *impact of different preprocessing approaches*, differently preprocessed variations of the same dataset are created thereon:

A *first variant* focuses on the much too high fluctuations of the rotational speed signal, which render it physically not plausible: The combustion increases the pressure within a cylinder, which results in a torque transmission to the crankshaft. This then leads to a more or less continuous acceleration respectively deceleration of the crankshaft – the more cylinders, the smoother the synchronization. Therefore, a considerably smoother signal is to be expected. Due to vibrations, jitter and a relatively small number of sampling points (240) per working cycle of 720° overlaying this signal with noise, it is smoothed with a moving average. An example of the resulting data is depicted in Fig. 3.

Due to the fact that the rotational speed signal is a periodic signal, following [15], a Fast Fourier Transformation (FFT) is conducted on the raw and smoothed (see Fig. 4) rotational speed datasets as a *second and third variant*. It can be seen that after the 25th harmonic there is hardly any power left in the signal, so these frequencies were not considered in the investigation. As expected, the FFT of a filtered signal differs from a raw signal only in its power at higher frequencies. The difference becomes visible between the 20th and the 25th harmonic.

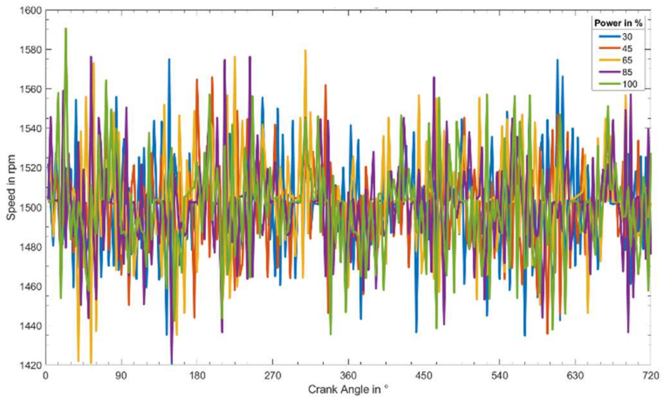

Figure 2: Raw rotational speed signal in the crank angle domain over one working cycle under different loads

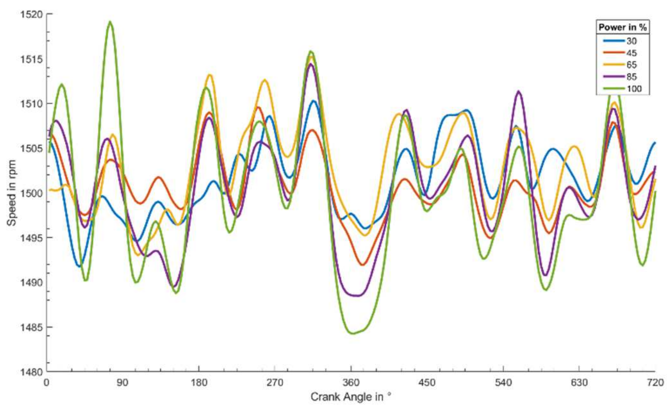

Figure 3: Smoothed rotational speed signal in the crank angle domain over one working cycle under different loads

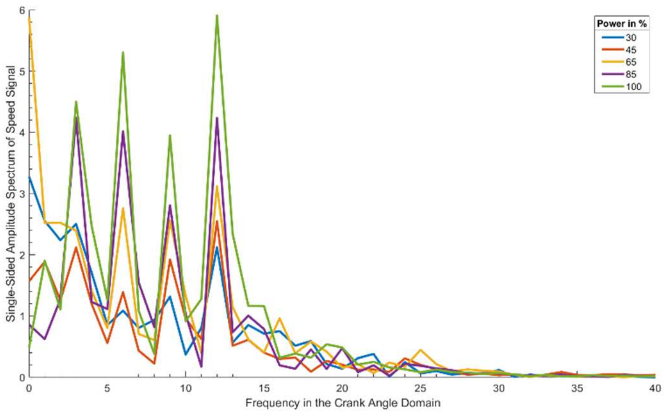

Figure 4: FFT of the raw rotational speed signal in the crank angle domain over one working cycle under different loads

Furthermore, FFTs of raw and smoothed rotational speed signals over two working cycles were created as a *fourth and fifth variant* to test the effect of longer sample times. These then contained the first 50 harmonics.

Additionally, a Principal Component Analysis (PCA) was performed on the smoothed rotational speed signal as a *sixth variant*. Of these, the first 25 components already representing over 90% of the signal were used.

4.2. Algorithms

To analyze the *impact of different network types*, two promising ones from the literature presented in Sec. 3 are selected to be tested: a TDNN as used by [14] and a MLP as in [17].

However, in our setup the TDNN network does neither yield good nor reliably reproducible results. Furthermore, this does not change substantially when the network is slightly modified. Contrastingly, the MLP network shows promising results.

Therefore, we selected the MLP for further optimization over the course of our study and did not continue with the TDNN.

Next, to analyze the *impact of different network structures*, a hyperparameter optimization in form of a grid search is carried out on the MLP. The parameters subject to the grid search are:

- Number of hidden layers: 1, 2 or 3
- Activation function per hidden layer: Sigmoid or Rectified Linear Unit (ReLU)
- Number of neurons in each hidden layer: 50, 75, 100 or 125

Initially included in the grid search, different dropout strategies [18] were discarded as they always had a negative effect on the estimation accuracy.

All networks make use of batch normalization [19] and have a single, linear output neuron to estimate the DoC in °. The dataset variant used is FFT (raw, one working cycle).

The results of this grid search listed by lowest and highest root mean squared error (RMSE) per number of layers is depicted in Table 1. It reveals that the estimation accuracy is largely independent of the used neural network's structure: There is a difference of just under 20% between the best and the worst result.

5. Case study results

Since adjustments to the network structure hardly result in any significant improvements and only the network type MLP delivers useful results at all (see Sec. 4.2), the focus is now laid entirely on the *impact of different preprocessing methods*.

5.1. Engine-global combustion parameters

Estimating the engine-global mean DoC using an algorithm as described in Sec. 4.2 and the dataset variants created in Sec. 4.1 yields RMSEs as depicted in Table 2:

The smoothed dataset delivers the best estimates with an RMSE of just 0.94°, closely followed by the raw dataset.

The performance using the FFT data (variants 3 to 6) is striking. While there is hardly a difference between raw and smoothed FFT datasets (variants 3 and 5 or 4 and 6), explainable by the merely minute differences in the selected harmonics, the difference to non-FFT preprocessing is substantial. FFT preprocessing leads to an RSME up to 50% higher. Interestingly, extending the length of the rotational speed signal used from one to two working cycles improves the estimation accuracy by about 13%. Apparently, the increase in signal length adds valuable information.

Similarly to FFT, the PCA preprocessing tries to reduce the amount of irrelevant information (or: noise) within the signal. This hardly decreases the estimation accuracy at all, showing a great difference to using FFT. However, the cause of this difference was not analyzed any further.

The scatterplot depicted in Fig. 5 shows the mean DoC estimation performance on preprocessing variant 2 in detail:

Table 1: Results of the grid search

No. of hidden layers		1	2	3
Best DoC estimation result		1,45°	1,41°	1,41°
	1st layer	125 neurons, sigmoid	100 neurons, sigmoid	100 neurons, ReLU
Network yielding best result	2nd layer	-	125 neurons, sigmoid	100 neurons, sigmoid
	3rd layer	-	-	150 neurons, sigmoid
Worst DoC estimation result		1,69°	1,48°	1,48°
	1st layer	75 neurons, ReLU	100 neurons, sigmoid	125 neurons, ReLU
Network yielding worst result	2nd layer	-	100 neurons, ReLU	75 neurons, ReLU
	3rd layer	-	-	100 neurons, ReLU

Table 2: Influence of different preprocessing approaches on results of *engine-global* mean DoC estimation

No. of preprocessing variant	Preprocessing variant	RMSE [°]
1	Raw, one working cycle	1,02
2	Smoothed, one working cycle	0,94
3	FFT (raw, one working cycle)	1,42
4	FFT (raw, two working cycles)	1,23
5	FFT (smoothed, one working cycle)	1,41
6	FFT (smoothed, two working cycles)	1,23
7	PCA (smoothed, one working cycle)	0,95

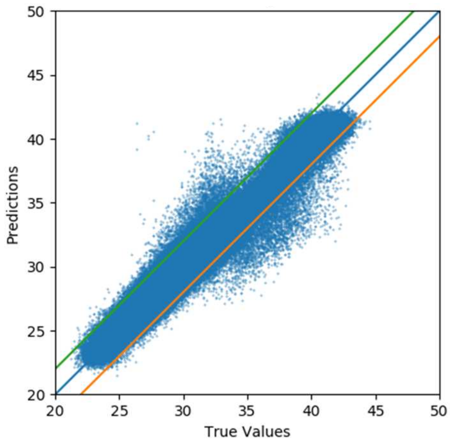

Figure 5: Scatterplot showing the predicted values and true values of the *engine-global* mean DoC estimate (blue line: target; orange and green lines: 2° deviation from target)

98% of all estimates lie between the green and orange line, indicating a deviation of 2° from the blue target line.

Based on these results of the engine-global estimation quality investigations, the following findings can be derived for the cylinder-specific investigations:

- The FFT datasets do not need to be examined any further as they do not yield good results.
- An in-depth investigation of an extended rotational speed signal's effect on the estimation accuracy should be carried out as it improved the FFT results substantially.

This results in the creation of the preprocessed input datasets as listed in Table 3 to be further examined.

Table 3: Different preprocessing methods for estimating *cylinder-specific* DoC values

No. of preprocessing variant	Preprocessing variant
I	Raw, one working cycle
II	Raw, two working cycles
III	Smoothed, one working cycle
IV	Smoothed, two working cycles
V	PCA (smoothed, one working cycle)
VI	PCA (smoothed, two working cycles)

5.2. Cylinder-specific combustion parameters

Estimating the cylinder-specific DoC using an algorithm as described in Sec. 4.2 and dataset variants as listed in Table 3 yields RMSEs as depicted in Fig. 6. The cylinder number was neither used as an input variable nor were different instances of the algorithm trained for the different cylinders.

It can be clearly seen that not only the preprocessing method, but also the cylinder number greatly influences the estimation accuracy. Generally speaking, the estimation is better, the smaller the cylinder number.

This might be due to the rotational speed sensor sitting on the crankshaft next to cylinder 1, meaning that all other cylinders are increasingly further away (see Fig. 1). This might cause the signal to become distorted as the crankshaft warps because of torsion. This effect increases with an increasing distance to the sensor.

Furthermore, it can be seen that the preprocessing method has an even greater influence here than on the engine-global mean DoC: The now cylinder-specific DoCs' RMSEs increase from 9% to between 8% and 15% and from 0.08° to between 0.12° and 0.26°.

The overall accuracy with RMSEs between 1.43° and 1.76° is worse than on the engine-global mean DoC with an RSME of 0.94°. This might be due to the much higher variability of the cylinder-specific datasets, which would have to be reproduced by the neural networks.

Because of the promising results shown by the two working cycle FFT datasets on engine-global mean DoC, the impact of an extended signal length was examined on the cylinder-specific DoCs as well (see Fig. 7).

It can be seen that extending the rotational speed signal length from one to two working cycles does indeed increase the estimation accuracy by about 0.03° to 0.1°. This increase is larger on cylinders number 1 and 6 and smaller on cylinder number 2.

6. Conclusion and Transfer

In this paper, a virtual sensor for a gas engine's duration of combustion (DoC) was developed based on deep neural network. Its estimation accuracy was evaluated with a special focus on the influence of network type, network structure and data preprocessing.

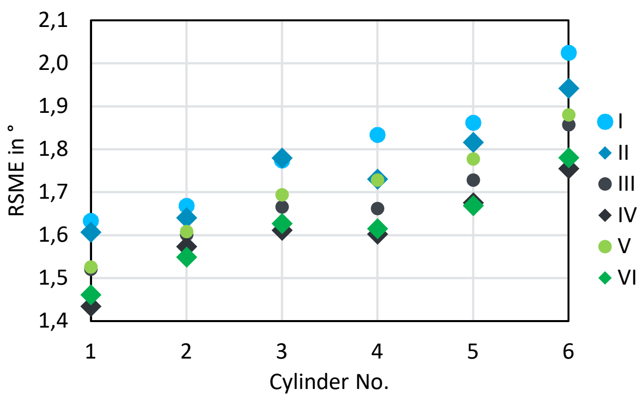

Figure 6: Comparison of the effect of the different preprocessing methods (see Table 3) on the estimation of *cylinder-specific* DoCs

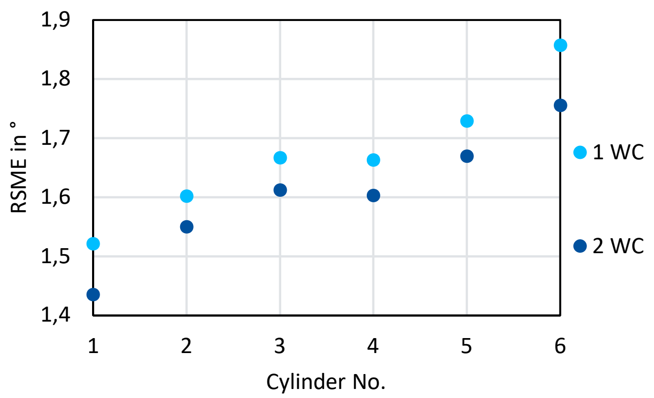

Figure 7: Comparison of the effect of the rotational speed signal length on the estimation of *cylinder-specific* DoCs using the best results of one working cycle (1 WC) and two working cycles (2 WC)

The literature study conducted revealed that all three factors do have an influence, but it remained unclear as to which was (most) relevant.

A careful analysis showed that in this scenario, only one of the two network types considered did work at all. Therefore, only this multi-layer perceptron approach was further examined.

A grid search to assess the influence of the networks structure resulted in just small differences between the worst and the best structure.

Key result: Our analysis revealed that input and method of data preprocessing had the most significant effect on estimation accuracy. Surprisingly, a Fast Fourier Transformation lowered the algorithm's output quality whereas a simple smoothing of the measured rotational speed signal delivered the best results on the engine's mean DoC.

The cylinder-specific DoC was best estimated by a dataset on which a Principal Component Analysis had been performed. Still, the estimation accuracy was lower than on the engine-global mean value, most likely due to torsion causing a distortion of the measured signal.

The overall performance of the virtual sensor developed was at par with literature values or better. Future work should further focus on the effects of data preprocessing.

References

[1] B. Maschler, D. White, and M. Weyrich, "Anwendungsfälle und Methoden der künstlichen Intelligenz in der anwendungsorientierten Forschung im Kontext von Industrie 4.0," University of Stuttgart, 2020, DOI: 10.18419/opus-10740.

[2] S. Yin, W. Ji, and L. Wang, "A machine learning based energy efficient trajectory planning approach for industrial robots," *Procedia CIRP*, vol. 81, pp. 429–434, 2019, DOI: 10.1016/j.procir.2019.03.074.

[3] T. Müller, N. Jazdi, J.-P. Schmidt, and M. Weyrich, "Cyber-Physical Production Systems: enhancement with a self-organized reconfiguration management," *Procedia CIRP*, 2020 (accepted).

[4] A. Mayr *et al.*, "Machine Learning in Production – Potentials, Challenges and Exemplary Applications," *Procedia CIRP*, vol. 86, pp. 49–54, 2019, DOI: 10.1016/j.procir.2020.01.035.

[5] B. Maschler, S. Kamm, N. Jazdi, and M. Weyrich, "Distributed Cooperative Deep Transfer Learning for Industrial Image Recognition," *Procedia CIRP*, 2020 (accepted).

[6] B. Lindemann, C. Karadogan, N. Jazdi, M. Liewald, and M. Weyrich, "Cloud-based Control Approach in Discrete Manufacturing Using a Self-Learning Architecture," *IFAC-PapersOnLine*, vol. 51, no. 10, pp. 163–168, 2018, DOI: 10.1016/j.ifacol.2018.06.255.

[7] Z. Tang and Z. Zhang, "The multi-objective optimization of combustion system operations based on deep data-driven models," *Energy*, vol. 182, pp. 37–47, 2019, DOI: 10.1016/j.energy.2019.06.051.

[8] P. Tan *et al.,* "Dynamic modeling of NOX emission in a 660 MW coal-fired boiler with long short-term memory," *Energy*, vol. 176, pp. 429–436, 2019, DOI: 10.1016/j.energy.2019.04.020.

[9] A. A. Erumban, "Lifetimes of machinery and equipment: evidence from Dutch manufacturing," *Rev Income Wealth*, vol. 54, no. 2, pp. 237–268, 2008, DOI: 10.1111/j.1475-4991.2008.00272.x.

[10] H. Kagermann, "Change Through Digitization—Value Creation in the Age of Industry 4.0," in *Management of permanent change*, H. Albach, H. Meffert, A. Pinkwart, and R. Reichwald, Eds., Wiesbaden, s.l.: Springer Fachmedien Wiesbaden, 2015, pp. 23–45.

[11] F. A.A. Souza, R. Araújo, and J. Mendes, "Review of soft sensor methods for regression applications," *Chemometrics and Intelligent Laboratory Systems*, vol. 152, pp. 69–79, 2016, DOI: 10.1016/j.chemolab.2015.12.011.

[12] "Directive 2015/2193 of the European Parliament and of the Council on the limitation of emissions of certain pollutants into the air from medium combustion plants," in *Official Journal of the European Journal*, 2015, pp. 1–19.

[13] C. Bennett, J. F. Dunne, S. Trimby, and D. Richardson, "Engine cylinder pressure reconstruction using crank kinematics and recurrently-trained neural networks," *Mechanical Systems and Signal Processing*, vol. 85, pp. 126–145, 2017, DOI: 10.1016/j.ymssp.2016.07.015.

[14] S. Trimby, J. F. Dunne, C. Bennett, and D. Richardson, "Unified approach to engine cylinder pressure reconstruction using time-delay neural networks with crank kinematics or block vibration measurements," *International Journal of Engine Research*, vol. 18, no. 3, pp. 256–272, 2017, DOI: 10.1177/1468087416655013.

[15] R. Johnsson, "Cylinder pressure reconstruction based on complex radial basis function networks from vibration and speed signals," *Mechanical Systems and Signal Processing*, vol. 20, no. 8, pp. 1923–1940, 2006, DOI: 10.1016/j.ymssp.2005.09.003.

[16] S. Saraswati and S. Chand, "Reconstruction of cylinder pressure for SI engine using recurrent neural network," *Neural Comput & Applic*, vol. 19, no. 6, pp. 935–944, 2010, DOI: 10.1007/s00521-010-0420-6.

[17] F. Taglialatela, M. Lavorgna, E. Mancaruso, and B. M. Vaglieco, "Determination of combustion parameters using engine crankshaft speed," *Mechanical Systems and Signal Processing*, vol. 38, no. 2, pp. 628–633, 2013, DOI: 10.1016/j.ymssp.2012.12.009.

[18] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, http://jmlr.org/papers/v15/srivastava14a.html, 2014.

[19] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Proceedings of the 32nd International Conference on Machine Learning, PMLR*, vol. 37, pp. 448–456, 2015.