

Received 30 April 2021; revised 15 September 2021; accepted 24 November 2021.  
Date of publication 6 December 2021; date of current version 4 March 2022.

Digital Object Identifier 10.1109/TETC.2021.3131371

# Artificial Intelligence for Mass Spectrometry and Nuclear Magnetic Resonance Spectroscopy Using a Novel Data Augmentation Method

FLORIAN FRICKE<sup>1</sup>, MARCELO BRANDALERO<sup>2</sup>, (Member, IEEE), SASCHA LIEHR, SIMON KERN,  
KLAS MEYER, STEFAN KOWARIK, ROBIN HIERZEGGER<sup>3</sup>, STEPHAN WESTERDICK,  
MICHAEL MAIWALD<sup>4</sup>, AND MICHAEL HÜBNER<sup>5</sup>, (Senior Member, IEEE)

Florian Fricke, Marcelo Brandalero, and Michael Hübner are with the Brandenburg University of Technology Cottbus-Senftenberg, 03046 Cottbus, Germany  
Sascha Liehr, Klas Meyer, and Michael Maiwald are with the Bundesanstalt für Materialforschung und -prüfung (BAM), 12205 Berlin, Germany  
Simon Kern is with the S-PACT GmbH, 52064 Aachen, Germany  
Stefan Kowarik and Robin Hierzegger are with the University of Graz, 8010 Graz, Austria  
Stephan Westerdick is with the Ruhr-University of Bochum, 44801 Bochum, Germany  
CORRESPONDING AUTHOR: F. FRICKE (fricke@b-tu.de)

This work was partially funded by the Bundesministerium für Wirtschaft und Energie (BMWi).

**ABSTRACT** Mass Spectrometry (MS) and Nuclear Magnetic Resonance Spectroscopy (NMR) are valuable analytical and quality control methods for most industrial chemical processes as they provide information on the concentrations of individual compounds and by-products. These processes are traditionally carried out manually and by a specialist, which takes a substantial amount of time and prevents their utilization for real-time closed-loop process control. This article presents recent advances from two projects that use Artificial Neural Networks (ANNs) to address the challenges of automation and performance-efficient realizations of MS and NMR. In the first part, a complete toolchain has been realized to develop simulated spectra and train ANNs to identify compounds in MS. In the second part, a limited number of experimental NMR spectra have been augmented by simulated spectra, to train an ANN with better prediction performance and speed than state-of-the-art analysis. These results suggest that, in the context of the digital transformation of the process industry, we are now on the threshold of a strongly simplified use of MS and MRS and the accompanying data evaluation by machine-supported procedures, and can utilize both methods much wider for reaction and process monitoring or quality control.

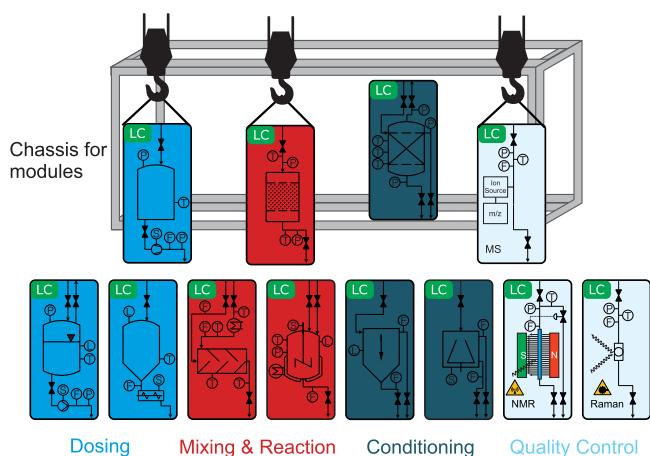
**INDEX TERMS** Industry 4.0, cyber-physical systems, artificial neural networks, mass spectrometry, nuclear magnetic resonance spectroscopy

## I. INTRODUCTION

The miniaturization of computing devices and easy access to commodity hardware are vital enablers of the fourth industrial revolution, the *Industry 4.0* paradigm. Thanks to that, it is now possible to combine intelligent sensor networks with physical processes, paving the way to exciting new applications such as flexible production, convertible factories, and customer-centered solutions [1], [2]. Modular Chemical Production (MCP), in particular, could be a gamechanging technology for the process industry for specialty chemicals.

In MCP systems, configurable physical modules can be quickly and flexibly connected to develop new processes, substantially accelerating the development of new special

chemicals and leading to faster market launches. For example, as shown in Figure 1, a new (multi-step) chemical reaction can be modularly realized by connecting standardized individual modules such as dosing units, reactors, reprocessing segments, product collectors together with quality control and monitoring units. Among the requirements for MCP is a chemical process and a quality control system that is aware of the characteristics of the ongoing process, such as the concentration of different chemicals produced. However, this information can hardly be obtained directly with chemical sensors. Instead, these sensors must be deployed and calibrated to the specific environment and process-under-analysis, and the data they collect has to be processed using complex methods



**FIGURE 1.** Depending on the process requirements, the basic framework of a modular system can be equipped with quickly reconfigurable modules for individual process steps such as dosing, reaction, reprocessing and quality control. Each module brings its own local control (LC) capability and can be interconnected.

before finally obtaining the information required for process control. This entire process requires complex calibration to the specific process (breaking the modularity principle), is substantially time-consuming, and requires a high degree of expert knowledge.

With the goal to remedy this situation, two common methods from the field of chemical structure elucidation are currently coming to the fore: Mass Spectrometry (MS) and Nuclear Magnetic Resonance Spectroscopy (NMR). Both of these offer a high amount of specific chemical information combined with considerably lower calibration requirements, making them highly promising for chemical process control strategies. The two methods cover a wide concentration range – partly as a trace method (MS) with linearity over several orders of magnitude, partly as a direct (calibration-free) method (NMR) with high linearity with respect to the composition of the sample in a technical mixture. However, for many years, both methods were considered too complex due to the cost of the hardware and the complexity of the time-consuming data analysis, rendering them unsuitable for process control. Thanks to Machine Learning (ML)-based methods that are also coming to the fore, such as the ones we present in this work, the applicability of MS and NMR can be expanded towards in-process control.

In this work, we propose and evaluate a complete flow for training and evaluating ANNs for the identification of chemical substance concentrations based on spectroscopic data. We show how this strategy can be applied to two use cases: MS and NMR, discussing and addressing the challenges of early-stage process development and limited amount of data. In summary, *our contributions can be summarized as follows:*

- We propose a flow for training ANNs towards spectroscopic analysis, where spectral information of a fluid sample is used to predict the concentrations of

individual components in the sample, allowing this analytical process to be applied for closed loop process control;

- We show how data-augmentation techniques can enable this flow to be implemented even in the early stages of the design of a new chemical process, where only limited information is available, and enable faster time-to-market;
- We evaluate the proposed flow in terms of the accuracy it achieves and the computational requirements of the resulting ANNs in two use cases: (a) A project targeting the development of a miniaturized mass spectrometer device for in-process use and (b) A project targeting the automated and early-stage chemical analysis via NMR for reducing time-to-market of new modular chemical processes.

Our results suggest that ANNs can be a key technology in enabling in-process MS and NMR to support *Industry 4.0*.

## II. BACKGROUND

*Modular Chemical Production* is one of the interesting new technologies that are being enabled as a tangible translation of the digital transformation of the process industry for specialty chemicals. In such production systems, configurable physical modules can be flexibly connected to quickly develop new processes, substantially accelerating the development of new special chemicals and leading to faster market launches. For example, as shown in Figure 1, a (multi-step) chemical reaction can be realized modularly by connecting standardized individual modules, e.g., using dosing units, reactors, reprocessing and conditioning segments, as well as product collectors together with quality control and monitoring units. Among the requirements for modular chemical production is a *chemical process* and a *quality control system* based on special chemical sensors. These sensors collect data that can be used to assess the concentration of selected chemical components and the presence of undesired by-products, using this information in a feedback loop to adjust the production system accordingly. Today, standard chemical sensors are based on optical spectroscopic methods such as near-infrared, infrared, or Raman spectroscopy, all of which have to be manually calibrated to the specific chemical components under analysis. This process requires a considerable degree of expertise and often causes expenditures in the amount of an analytical instrument’s purchase price. Because of this dependency, these sensing methods lose the much desired modularity requirement.

In the hope of remedying this situation, two common methods from *chemical structure elucidation* are currently coming to the fore, as they are also extremely interesting as quantitative methods: Mass Spectrometry (MS) and Nuclear Magnetic Resonance Spectroscopy (NMR). MS and NMR offer a high amount of specific chemical information combined with considerably lower calibration requirements, making them highly promising for such modular chemical process control while behaving highly linear. These methods have always been considered complicated, because of the hardware, which is now

available and affordable as a benchtop system for both methods, and because of the data analysis, which has so far been considered complicated, but which is now becoming manageable through machine-based methods. However, thanks to Machine Learning (ML)-based methods such as the ones we present in this work, the applicability of MS and NMR can be expanded toward in-process control.

In the next subsections, we describe the applications of MS and NMR, and finally present the challenges in applying them to real-world chemical production processes.

### A. MASS SPECTROMETRY (MS)

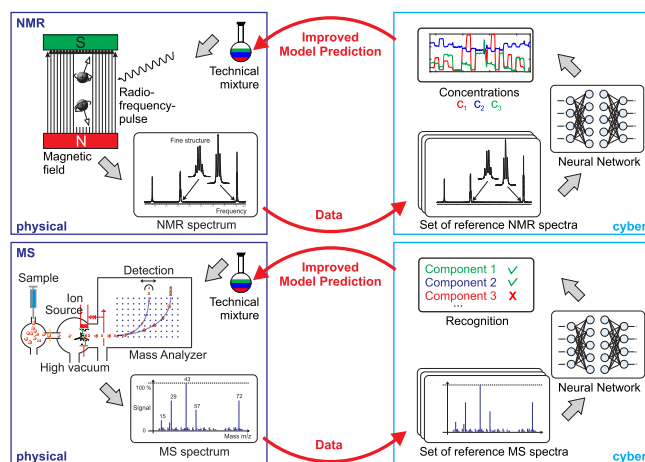
The MS process consists of two phases: sample Ionization and Ion-Separation. The Ionization phase consists of ionizing the sample, so the different substances that compose it may fragment or decay into their own ions, depending on the molecules contained in the sample. To avoid interaction between ions, this process should ideally be carried out in a vacuum chamber. The Ion-Separation phase consists of separating the different ions or fragments released by the sample. To accomplish this, a dynamic electrical field is generated, pushing the ions at different velocities based on their mass and the influence of the electric field on their electrical charge. The spectrometer records the ion mass-to-charge relation. By monitoring that, and knowing the effect of the ionization on different substances, it is possible to estimate the sample composition with high precision.

In such analyses, the *resolution* indicates the capacity to differentiate and catalogue the ions according to their mass. A high resolution indicates that two elements with similar mass can still be differentiated, while in the case of low resolution this is not the case, both elements would be catalogued as the same one.

### B. NMR SPECTROSCOPY

Nuclear Magnetic Resonance (NMR) spectroscopy is today one of the most important instrumental analytical methods in the natural sciences and medicine [3]. Unlike MS, which requires ionizing a sample in order to acquire information (thus being a *destructive* analysis process), NMR is able to obtain unambiguous structural information on the compounds in a sample in a *non-destructive* way. It has become indispensable in chemistry, biochemistry, and pharmacy (high-resolution NMR spectroscopy of liquids), physics and materials research (solid-state NMR spectroscopy), and medicine (magnetic resonance imaging) [3], [4].

A radio frequency pulse leads to a momentary change in the population between nuclei in the ground state and in the excited state (see Figure 2). The resulting change in overall magnetization can be detected with a radio frequency coil as a decaying receiver signal and digitally recorded. The NMR spectrum is produced by Fourier transformation (FT). Atomic nuclei magnetically differently shielded by electron currents lead to individual signals. The characteristic fine splittings in the spectrum allow the unambiguous clarification of structural assignments in the molecule. The most important NMR-active nuclei are the isotopes  $^1\text{H}$  and  $^{13}\text{C}$ , as they are found in most



**FIGURE 2. Cyber-Physical Systems for MS (bottom) and NMR (top) applications.** MS: The ion separation by a static magnetic field results in individual flight paths for fragments and molecules of the sample, which were generated in an ionization chamber. A detector collects intensities for various mass-to-charge ratios. NMR: The effect is based on interactions of electromagnetic radiation in the radio frequency range (typically 300–1200 MHz) with atomic nuclei (nuclear spins) that occur in strong external magnetic fields. Resonances are recorded with a radio frequency antenna yielding spectra for individual atom groups in a chemical molecule. Both, MS and NMR spectra can be analyzed by neural networks to predict quantitative information such as the concentrations of individual components in a complex mixture within seconds.

molecules of interest. Other important nuclei are  $^{19}\text{F}$ ,  $^{31}\text{P}$ ,  $^{17}\text{O}$ ,  $^{15}\text{N}$  and some metals.

High-resolution NMR spectroscopy is particularly suitable for the non-invasive study of the properties and behavior of complex multicomponent mixtures under technical conditions [5]. It is one of the few non-invasive analytical methods that are able to provide quantitative and, at the same time, detailed chemical-structural information on complex reacting mixtures, e.g., on the type of reactions taking place or on by-products. In contrast to routine NMR spectroscopy, which usually aims at qualitative information, sample preparation in which the analytes are dissolved or diluted in deuterated solvents is not suitable for technical mixtures – especially in online operation. Nevertheless, if the NMR parameters and experimental conditions are carefully selected, it is also possible to analyze undiluted technical mixtures quantitatively under increased pressure and at high temperature, even those with high electrolyte concentrations or extreme pH values.

NMR spectroscopy is a direct analytical comparison method; in other words, it exhibits a direct correlation between *the signal area in the spectrum* and *the number of observed nuclei in the active sample region*, allowing for a calibration-free relative quantification compared to the *counting* of nuclear spins. In combination with a reference compound of known quantity or a single-point calibration, e.g., on a pure reactant, absolute quantification becomes feasible. This makes NMR spectroscopy a very promising method for online reaction monitoring applications. [5]

With the availability of compact benchtop NMR instruments, a very wide range of applications for NMR spectroscopy has opened up, and the instruments can be used in a fume hood or in a production environment. Work is presently underway to fully integrate an NMR analyzer and the associated data evaluation in the sense of quality control and process control unit [6]. The application of NMR spectroscopy as online method in such a manner is currently in technical readiness level 7 (TRL 7 – i.e., system prototype demonstration in an operational environment) [7].

### C. CHALLENGES OF CHEMICAL INFORMATION AND PROCESS MODELS

MS and NMR spectroscopy are indispensable tools for analysis in different science fields, which have undergone enormous technical development in recent years. Just a few years ago, both methods were too sophisticated and were extremely sensitive to their laboratory environment. Thus, they required costly, large-scale research equipment and staff with a high level of expertise.

Nowadays, these methods have improved in stability and are easier to apply so that a sample can be taken and analyzed offline by instructed staff. This approach is useful for basic research or quality assurance but cannot be used for direct process control, as these analysis must still be performed manually. This not only requires an available specialist but can also be very time-consuming creating long delays in process feedback. In principle, the application of NMR and MS with compact online instruments in industrial scale production is feasible and was already demonstrated [7]. However, the real-time data analysis with fully automated feedback control remained challenging without affordable instruments and simplified automated data analysis. Here, we show how this area can be advanced considerably by using Artificial Intelligence (AI) and Machine Learning.

Sensors that react directly or indirectly to changes in the target components are needed in order to achieve chemically specific measurements. The measurements must be as precise as possible (i.e., undisturbed by other components such as the solvent) as well as with sufficient sensitivity – such as MS. *Direct analytical methods*, such as NMR spectroscopy, would be the optimum where the signal obtained is independent of matrix effects and is causally linked to a concentration of a component in a multicomponent mixture.

The improvement of production parameters, space-time yield, or energy efficiency can only be achieved by data- or knowledge-based models from the raw data via a more detailed description of the material on several scale levels. Nowadays, typically multivariate tools and algorithms are used for automated feature extraction, such as the extraction of chemical information from the data sources using Principle Component Analysis (PCA), Partial Least Squares (PLS), or Latent Discriminant Analysis (LDA). This is currently done manually and is time-consuming. However, access to consistently compiled data alone is not sufficient for this purpose, because the understanding of the process is often hidden as so-called *expert*

*knowledge* in the design of the process and its control and is usually undocumented.

To find an acceptable way to control *chemical* processes so that sensors and actuators can better meet the requirements of digital transformation and the associated tasks in the future, they must be equipped with *intelligent* functions that also accumulate the expert knowledge, e.g., in the form of process models. For a holistic process-analytical approach and such decision support systems, the semantic level (context information brought together with sensor data) must be considered beyond the pure interface definition (syntax, such as data acquisition, data connectivity, and data integrity). Efficient communication with all the components in the system is thus key and will require novel architectures [8].

### III. ML-ASSISTED MS AND NMR

In this work, we show realizations that lead to a novel calibration concept for spectroscopic and spectrometric sensors using ML. These approaches lead to a considerably shorter procedure for modeling, require less calibration data than traditional ones, and can be efficiently implemented in Field Programmable Gate Arrays (FPGAs) [9]. The methods described here could significantly enhance the real-time release and continuous production as well as traditional batch chemistry for fine chemicals.

One of the critical challenges in this field, however, is the acquisition of the massive amounts of data needed for the training of ANNs. Such data is usually not available for chemical processes due to the limited number of measurement devices and their typical intervals. Thus, the term *Big Data* as a prerequisite for data-driven evaluation procedures is not appropriate for the process industry because even with quantitatively large datasets, information is typically only available for campaigns about a few batches with a series of measurement data that do not exhibit sufficient variance for a data-driven evaluation – not comparable with the datasets from large Internet groups. In the process technology environment, the term *Smart Data* is more appropriate. Smart Data includes, among other things, the clever selection of data for analysis and the combination of data-driven procedures and expert knowledge for their analysis. *In this work, we address this problem by using simulated variants of the measured data (i.e., synthetic MS and NMR spectra [10]), a form of data augmentation.* Past works [11], [12] have used similar procedures, although not in the context of MS nor NMR.

#### A. A MINIATURIZED IN-PROCESS MASS SPECTROMETER

Our project aimed at the development of a miniaturized mass spectrometer for use within chemical processes. One objective was to evaluate the measured spectra of a specific measurement task using artificial neural networks on an embedded system. In addition to a low energy requirement, a small size of the embedded system was also required, as the system was ultimately to be installed together with the miniaturized mass spectrometer in a 19 inch housing. With regard to the structure

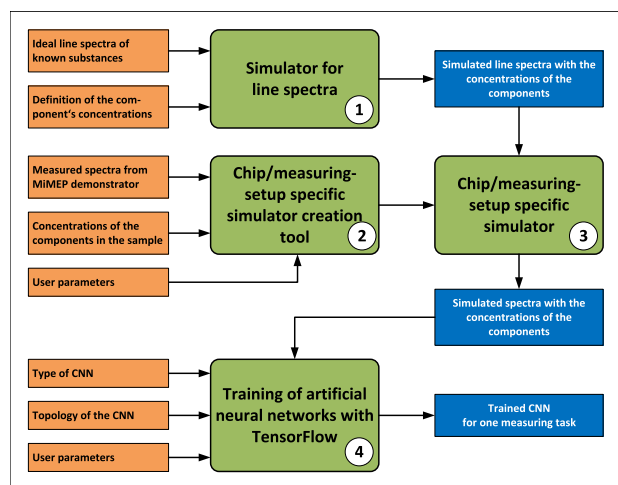


FIGURE 3. Toolchain overview.

of the network, initially only the structure of the input and output data was specified: a measured spectrum should be able to be used directly as input for the neural network. In order to increase flexibility and to keep the number of required networks small, it was determined that missing values would be interpolated when the resolution was changed. This was necessary due to the fact that the stepsize as well as the range on the  $m/z$ -axis can be set for the mass spectrometer prototype. The percentages of the individual substances in the sample were to be used as output of the network. This determined that a network can only be used for a measurement task defined in advance and that in practical application measures are required to check the plausibility of the input data. This is due to the fact that in the case of inputs containing unknown compounds or completely different substances, no meaningful output can be expected. To assist in the development of new chemical processes and have a toolflow that can also be applied to other systems, we developed a complete flow that includes the following tools:

- an ideal *line spectra simulator* (Step 1 in Figure 3) that can generate ideal mass spectrometers for arbitrary compounds based on already-known spectra;
- a *simulator of the portable mass spectrometer* (Step 3 in Figure 3) that takes into account all the nonidealities of the real measuring device and can produce a non-ideal mass spectrum matching those characteristics (given an ideal spectrum);
- a tool for the *automatic generation of the mass spectrometer simulator* (Step 2 in Figure 3) based on measured data from the real measuring device;
- a tool for the *automatic training of ANNs* (Step 4 in Figure 3) that can identify the chemical compounds from the non-ideal mass spectra.

#### 1) EVALUATION METHODOLOGY / TOOLFLOW

A major limitation in the development of a new type of miniaturized mass spectrometer is the parallel development of the actual measurement system and the software for analyzing

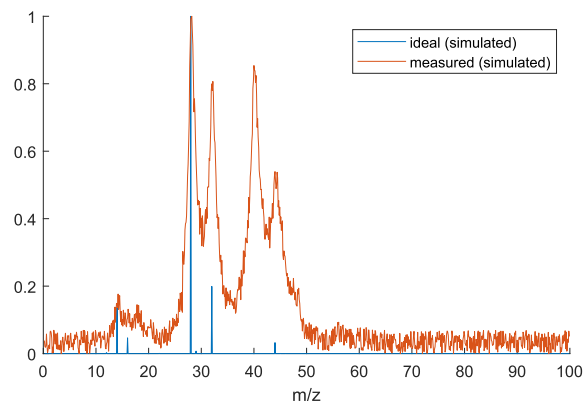


FIGURE 4. Ideal spectrum from Tool 1 (blue) and the simulated spectrum from Tool 3 (orange).

the measurement data. Since it was not possible to generate a sufficient number of labelled measurement series for the training of an ANN by means of the prototype, a program had to be developed that generates a sufficient number of simulated measurement series, which is then suitable for the training of an ANN for the evaluation of the measured spectra of the real measurement system.

The program (Figure 3) works as follows.

Based on the known ideal (e.g., without measurement noise) line spectra of the substances contained in the mixture, ideal spectra of the different substance mixtures with arbitrary concentrations are generated by linear superposition (Tool 1). To adjust this ideal spectra to the real (noisy) behaviour, characteristics of the measurement system such as the deformation of the peaks to a curve, the frequency-dependent attenuation, the drift and the noise model are to be obtained from real measurements using MATLAB tools (Tool 2). In a further part of the simulation tool, these ideal spectra are converted into a continuous spectrum with the desired resolution using the characteristics of the real measuring system (Tool 3). The output of Tool 1 and Tool 3 for one specific substance mixture is shown in Figure 4. Please note, that the peak in the simulated continuous spectrum which has no counterpart in the linespectrum is caused by the utilized ignition gas.

An important factor in this flow is the necessary number of measured values for creating the simulation program, as it directly corresponds to the effort that has to be taken when creating a new network for an improved system or a new measurement task. Our approach to determining the required number of measured samples needed to create a functioning simulator is considered in the upcoming section. With the simulator created in this way, a sufficient number of simulated and labelled measurement series can be generated in minutes to train an artificial neural network. To handle the big amounts of data (measured samples, simulated samples, trained networks, ...) a MongoDB database is used to store the data of all tools in the presented toolflow. In addition to the actual data, all objects stored in the database also store metadata that make it possible to trace the basis on which the

respective data was generated. This has been done to comprehend which measurements have been used to train the simulators and which data has been used to train a specific network.

## 2) DESIGN AND TRAINING OF ANN MODELS

The dataset used for the training consisted of 100000 simulated spectra, labelled with the percentage of different substances in the sample. This dataset was divided into 80% for the training and 20% for the testing. The process of designing and training the ANN is primarily defined by the TensorFlow Framework, which is used in this project (Tool 4). To enable the integration of the database and to allow the automated evaluation of different network topologies using the same training data, additional tools have been developed. The tools that assist in the definition phase allow the definition of one or more network topologies and the training- and validation datasets to use without modifying the source code. The whole training process can then run without user interaction. Backend tools help with the evaluation of the trained networks with different training datasets, the selection of the best-performing networks, based on selectable quality criteria and the export of analysis data to spreadsheet applications or data analysis tools, e.g., MATLAB or Pandas. This allows a deeper investigation of the data used for training and validation. Furthermore, the comparison of measured and simulated data is eased. Additionally, a tool to export the desired ANN for use on embedded platforms is available.

Due to the large design space available in the creation of ANN architectures, we performed a preliminary investigation considering a broad set of ANN topologies, before selecting to focus on a subset of these that gives a most promising direction. These included Multi-Layer Perceptron (MLP) networks, the ResNet and Highway network architectures [13], and Convolutional Neural Networks (CNN). The preliminary investigations showed that CNNs represent a good compromise between performance and effort in training and inference. Therefore, a CNN architecture was selected for the given task.

Since there were only very few measurements from the real system with known substance composition at this point in the project, these could not yet be used to assess the quality. So we initially set at a mean error of no more than 0.005 on the validation data, which was at this point also simulated, as target for the network. As we used the mean absolute error (MAE) as loss function for training the networks, the target means an absolute deviation of not more than 0.5 % for all output substances. Based on this target, the topology of the network was developed by starting with only one convolutional layer and one MLP layer for the output. Based on this we exploratively added more convolutional layers and adjusted the parameters of these layers until a satisfactory result could be achieved. The topology of the network created in this phase of the project is the basis for all further considerations and is shown in Table 1.

The simulated measurement values enabled the creation of ANNs that fulfilled the above-mentioned requirements. As

**TABLE 1. Structure of the ANN used for mass spectrum analysis.**

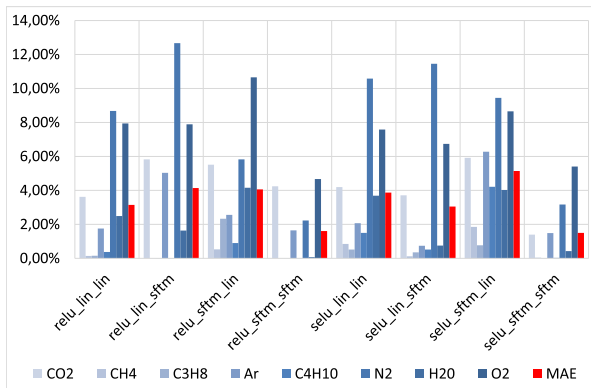
Layer	Type	Filters	Kernel	Stride	Activation
<b>1 (Input)</b>	Input				
<b>2</b>	Reshape				
<b>3</b>	Conv1D	25	20	1	SELU
<b>4</b>	Conv1D	25	20	3	SELU
<b>5</b>	Conv1D	25	15	2	SELU
<b>6</b>	Conv1D	15	15	4	Softmax
<b>7</b>	Flatten				
<b>8 (Output)</b>	Dense				Softmax

stated earlier, the initial validations have been done using simulated values, which lead to promising results. The initial version, which used linear activation functions for layer 6 and 8 has a mean absolute error of 0.14% on the validation dataset.

When evaluating the ANN with real measurement series, the quality of the outputs was significantly lower than on the datasets used for training. The MAE for the abovementioned network, using the linear activation function in the output layer increased to 3.15%. This behaviour was to be expected due to the prototype status of the measurement equipment and the resulting fluctuations in the quality of the measurement results. Therefore, in further iterations and in parallel to the further development of the mass spectrometer prototype, we started investigating the influence of the activation functions for the layers in the ANN and the required inputs for the simulator generation tool, to enable good generalization.

*Impact of Different Activation Functions.* The investigation on the impact of the utilized activation functions has been done using simulated values for training and an initial evaluation. Afterwards, labelled samples, which have been measured with the mass spectrometer prototype were used for evaluating the quality of the different network versions. We distinguished between the functions in the hidden layers of the network and the functions in the output layer. For the hidden layers we examined the activation functions RELU and SELU, for the output layers Softmax and Linear. A distinction was made between the final convolutional layer and the actual output layer. In this experiment, 8 different networks were trained, 4 each using either the RELU or the SELU function in the hidden layers and for each variant all possible combinations of softmax and linear for the final convolutional layer and the output layer. In the analysis of the trained networks using the simulated datasets, the differences between the different ANNs were quite small.

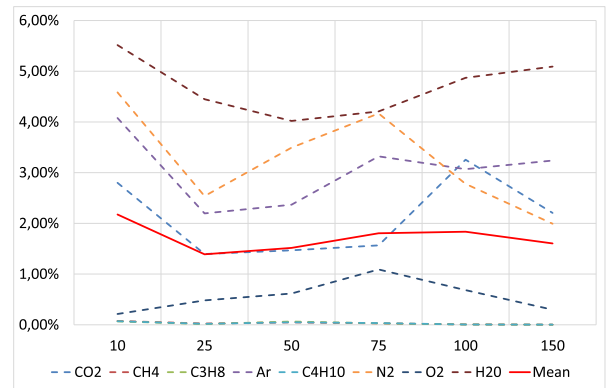
All networks using the RELU activation function have an MAE between 0.145% and 0.156%. The MAE of the SELU networks is higher on average with the network using linear activation in layer 6 and softmax activation in layer 8 being a clear outlier with an MAE of 1.1%. The other SELU networks have an MAE between 0.28% and 0.21%. When using the measured data for the evaluation of the networks, the medium performance of the networks decreases and the differences between them become significantly larger. An



**FIGURE 5.** Comparison of the mean absolute error on measured data for ANN, using different activation functions.

overview is depicted in Figure 5. The blue bars show the error for specific substances in the compound, whereas the red bar shows the mean absolute error for all substances. The labels on the horizontal axis denote the activation functions used, Softmax is abbreviated "sftm" and "lin" means Linear. The MAE is now in the range between 1.5% and 5.14% and the best performing network is now the version which uses the SELU activation and softmax in layer 6 and 8. It was expected, that the Softmax function in the output layer is beneficial especially for nets whose output values add up to 1, which is the case here. This expectation was confirmed, as the use of the Softmax function in the output layers resulted in a massive improvement of the mean error, for both the SELU (1.5%) and the RELU (1.61%) version. All other versions have an MAE between 3.05% and 5.14%, which is significantly worse. The use of the SELU function in the hidden layers also resulted in a slight improvement in quality for the best performing networks, but this was significantly less (about 0.1%) than the effect of the softmax function.

*Impact of Different Sample Sizes.* We also investigated how many samples are required to parameterize the simulator for a new measurement task or a new model of the mass spectrometer. This is desirable, as a smaller number of required samples reduces the effort in terms of the time needed, but furthermore less known samples (e.g., reference gas) are necessary. However, a sufficient number of measurement series must be used to be able to extract the characteristics reliably. For this reason, 6 different neural networks were trained, which were generated on the basis of simulators parameterized with 10, 25, 50, 75, 100 and 150 measurement series per mixture. In each case, 14 different mixtures were used. The quality of the ANN generated in this way was subsequently determined with both simulated values and real measured values in the same way as in the investigation of the activation functions. The analysis based on the simulated values showed a comparable performance for all networks, even the one with only 10 samples per mixture. The MAE for all networks is quite close and lies in the range between 0.2% and 0.22%. When evaluating the networks with measured data, the differences are higher.



**FIGURE 6.** Comparison of the mean absolute error on measured data for ANN, trained with simulators, parametrized with a different number of samples.

An overview of the mean absolute error, which occurs during the evaluation of the ANN with real measured data for the different substances in a mixture, is shown in Figure 6. The MAE is between 1.39% and 2.18%, which is the value for the network created with 10 samples. This shows, that the use of only 10 measurement series per mixture is not sufficient, as this network delivers a significantly lower quality. However, it is interesting to note that the best network in the context of this investigation is the network based on 25 measurement series (MAE: 1.39%), the others are qualitatively inferior in the evaluation. In particular, the nets with 75 and 100 measurement series are clearly inferior (MAE: 1.81% and 1.83%), the nets with 50 (MAE: 1.52%) and 150 (MAE: 1.6%) measurement series have a comparable, albeit inferior, performance to the version with 25 values. The fact that the measurement series considered for the training were randomly selected may be one reason. However, the behaviour can be reproduced at least qualitatively.

As a conclusion, it can be stated that a significantly larger number of measurement series for the parameterization of the simulator does not necessarily lead to a better artificial neural network. However, further investigations are appropriate and, in our view, sensible.

### 3) EVALUATION OF ANN MODELS

To evaluate the networks with measured data, we mixed gases with known spectra by using mass flow controllers, allowing us to create mixtures with controlled concentrations of compounds. The mass spectra of the mixtures have been measured with the MMS prototype and the ANN has been used to infer the compound concentrations. The ANN, which has been used for this evaluation has the structure and the activation functions shown in Table 1 and has been trained with simulated training data, generated by a simulator parameterized with real measurements from 14 different mixtures and around 200 samples per mixture. This ANN reached an average MAE of 0.27% on the validation dataset. The average MAE on the measured dataset is 1.5%. Figure 7 shows, that for most of the compounds the mean absolute error is smaller than 3%. Evaluations showed that this could partly

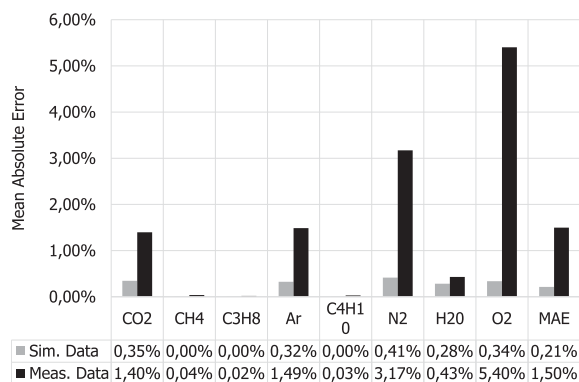


FIGURE 7. Results of the MMS when identifying the compounds in a simulated (gray) and a real sample (black).

be explained with impurities and changes in the configuration of the prototype.

In future versions, we plan to introduce additional inputs to the ANN, which provide parameters of the system and the environment to evaluate their impact on the quality of the results. Another reason could be that the simulator only considers a static system state, fluctuations of certain parameters, such as the displacement of the peaks do not affect the simulated values. One substance in Figure 7 ( $O_2$ ) shows bigger deviations of over 5%. The error for Oxygen ( $O_2$ ) and the detection of Water ( $H_2O$ ), which should not be part of the compound, can be explained by the aforementioned impurities.  $H_2O$  was no purposed compound, but air humidity caused a signal in the reference measurement, which was not considered while creating the simulator. Therefore, the ANN is able to detect water, but the reference gas should not contain water. So, it is possible that  $O_2$  is partially recognized as  $H_2O$ .

Apart from the software, an embedded system, that is able to run the ANN and to graphically present the analysis result has been developed. The system is based on the Nvidia Tegra platform. These systems combine a Multicore ARM processor and an embedded GPU in a System-on-Module (SoM). For the prototype of our embedded system, the Nvidia Jetson Nano development Kit has been used. This system has been combined with a 7 inch monitor and integrated into a 3D-printed case. The system uses a full-featured Linux distribution and is able to run the TensorFlow ANN on the CPU and the embedded GPU. It allows fast iterations as the adoption of all software components and the ANN can be done like on a desktop system. Performance measurements show that the system could be replaced by a smaller and cheaper system in the future.

*Performance and Energy Consumption on Embedded Platforms.* Table 2 shows the performance, power and energy consumption of executing the trained ANN model in two platforms (NVIDIA's Jetson Nano & Jetson TX2), each featuring both a CPU and GPU execution unit. The evaluation has been done using a dataset of 21600 samples. NVIDIA Jetson Nano features a Quad-Core ARM Cortex-A57 and an NVIDIA Maxwell GPU with 128 CUDA cores, while the Jetson TX2 features the same ARM core but also an NVIDIA

TABLE 2. Results of the execution of the complete dataset on the two Jetson platforms.

Metric	Platform			
	Jetson Nano		Jetson TX2	
	CPU	GPU	CPU	GPU
Execution Time/s	30.19	6.34	21.64	3.03
Power Consumption/W	5.03	4.77	5.92	6.68
Energy Consumption/J	151.86	30.24	128.11	20.24

Dual-Core Denver 2 CPU and an NVIDIA Pascal GPU with 256 CUDA cores.

The results show that the execution time can be improved by  $4.8\times$ – $7.1\times$  by leveraging the embedded GPUs compared to the CPUs. Likewise, energy consumption can be improved by  $5.0\times$ – $6.3\times$  when executing in GPU compared to CPU, although the power consumption levels are similar and in the range of 5 W. The reason for that is the high regularity of ANN computations: they can easily be modelled in a Single-Instruction-Multiple-Threads (SIMT) format, which GPUs excel at. Finally, doubling the number of CUDA cores from 128 to 256 (Jetson Nano versus Jetson TX2) improves performance and energy consumption by  $2.1\times$  and  $1.5\times$ , respectively, highlighting that computing such ANNs in GPUs scales well with the size of the hardware resources.

## B. MACHINE-ASSISTED MODEL BUILDING OF NMR SPECTRA

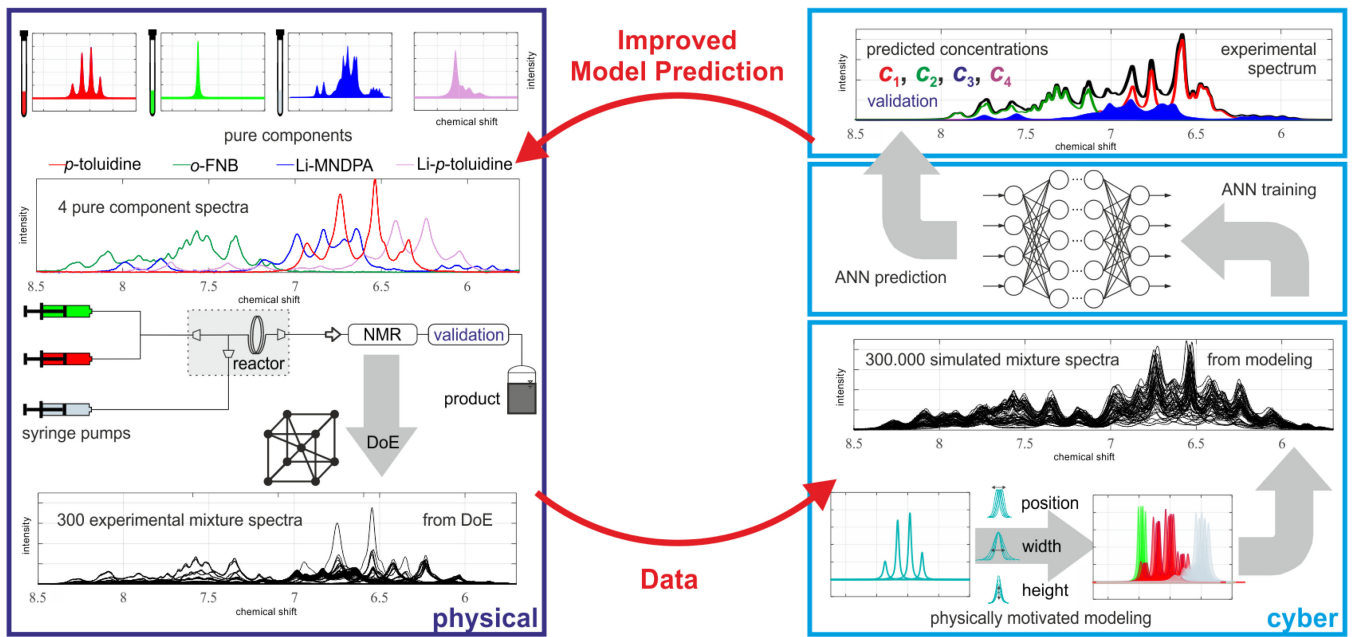
As already mentioned, data-driven methods for predicting quantitative information from complex process analytical data require very large training datasets together with contextual information. These are typically not available if a process development is still in its early stages or the chemicals involved are still being modified. Usually only a few hundred or a few thousand specific datasets are available. Data for the validation of online methods obtained by sampling and comparative analysis are even smaller. We have therefore based our study on a dataset which corresponds to this typical framework: different reaction conditions for an organic lithiation reaction were generated with the help of laboratory equipment and measured simultaneously online using two methods: medium-resolution and high-resolution NMR spectroscopy resulting in a set of 300 spectra as raw data basis with four compound concentrations as the four labels of interest.

The experimental setup is based upon a laboratory scale flow reactor, the chemical example reaction used, NMR spectra acquisition, data enhancement, ANN model building and ANN concentration prediction is depicted in Figure 8.

### 1) SIMULATION OF SYNTHETIC SPECTRA FOR TRAINING

Similar to the MS example, we again used an NMR line spectra simulator to generate a large number of synthetic training data covering the full concentration range of interest. The simulator uses NMR spectra of the pure components, which have been fitted with an *Indirect Hard Modelling*





**FIGURE 8.** The synthesis of nitro-4'-methylphenylamine (MNDPA) by aromatic substitution of p-toluidine and 1-fluoro-2-nitrobenzene (o-FNB) is a relevant example for the pharmaceutical industry [7]. p-toluidine was activated by a proton exchange with the organolithium compound lithium bis(trimethylsilyl)amide (Li-HMDS), giving four relevant components in all mixtures, as shown in the upper left figure. The flow reactor (middle left) was operated along a DoE yielding representative mixture spectra (bottom left). These were enhanced to 300.000 spectra on basis of a physically motivated simulation method (bottom right). ANN model building and ANN concentration prediction of the abovementioned is depicted in the upper right. On the left side is the physical part of the cyber-physical system, on the right side the virtual part. The data predicted by ANN were validated against a reference method used in parallel with low field NMR spectroscopy.

(IHM) approach that can also be used for an analysis of the spectra [14]. IHM is a description of the NMR spectra of mixtures in which each component produces one or more signals in the total spectrum, which more or less overlap. Based on a physical assumption (hard model), each component can be described as a pure component, which is done with a series of Lorentz-Gauss functions. With IHM, these pure components can be found in the total spectrum of a mixture by fitting algorithms and their intensities and thus concentrations can be determined, although individual signals are allowed to shift or broaden. Because of the high linearity in NMR spectroscopy, this method works particularly well. Linear combinations of the parametric models of pure component spectra can then be calculated to generate NMR spectra for arbitrary values of the four compound concentrations. Note, that the IHM approach has advantages compared to just a linear combination of the experimental pure component spectra. First, noise in the experimental spectra would be inaccurately scaled and added in the linear combination of experimental spectra. Second, the mixing of compounds in solution may shift single NMR peaks within a pure component spectrum. A linear combination of experimental pure component spectra would neglect this effect, but it is included in our spectra simulator through shifting and broadening of peaks in our parametric model. Overall, the approach allows the initial training dataset to be arbitrarily sized and distributed along different prediction variables.

## 2) DESIGN AND TRAINING OF ANN MODELS

We regard two different types of small ANN models here. First, we study a network consisting of a convolutional and a dense layer acting on single NMR spectra. In a second approach we make use of the fact that the NMR spectra are recorded in real-time and use a long short-term memory (LSTM) neural network to analyse the time series. LSTM networks have the advantage that information from previous timesteps is used to make the prediction, thereby broadening the database. For the convolutional network acting on single timesteps we first conducted an optimization of the neural network architecture to arrive at a small but accurate neural network, capable of running on limited hardware in cyber-physical systems. We optimized the architecture through trying out fully connected, convolutional and locally connected neural networks with different depths. After finding that fully connected layers need to have significantly more trainable parameters than convolutional layers we further optimized the number of layers, average or max-pooling layers, size of the pooling or strides, number of filters, the number of fully connected layers after the convolutional layer and their type of activation function. We finally arrived at a comparatively simple model of just a single, locally connected 1-D convolutional layer and a single flatten layer connected to the four output variables of the four concentrations. Training this neural network, we found that after 50 epochs the performance on the experimental validation dataset is best. However, we continued training for 400 epochs, despite the risk of overfitting to

synthetic data because we want the neural network to perform well over the full range of concentrations, not just the ones available in our experimental low field NMR test dataset. Note, that one limitation of the ANN methodology is that the resulting model can only reproduce those changes that lie within the training label range, so application is limited to parameter ranges within the training data. For analyzing the time sequence of NMR spectra, we used an LSTM model as a form of recurrent neural network using feedback connections from previous time-steps. As our time dependent experimental data consists of a time series of several steady state plateaus with different concentrations, we repeated random training spectra one to twenty times to emulate plateaus with jumps between them. The LSTM model was then trained with this augmented training dataset for 100 epochs and the network with the best performance on the experimental validation dataset was selected.

### 3) EVALUATION OF ANN MODELS

For predicting compound concentrations, both the optimized convolutional ANN, LSTM ANN and state-of-the-art model predictions from IHM were benchmarked against accurate high field NMR reference data. The best performing convolutional ANN consists of only three layers with 10532 trainable parameters in total, organized as follows: 1) locally connected 1-D convolutional layer (four filters, strides and kernel size 9); Flattening Layer; Fully connected layer with four output neurons for the prediction of four concentration values. This convolutional ANN performs better than an IHM analysis as its results have a 5% lower mean square error. Beyond being more accurate, the ANN analysis takes only 0.9 ms for predicting a single spectrum batch on an Intel i7-8565U CPU (1.8 GHz) with Keras and TensorFlow and is therefore more than 1000 times faster than an IHM analysis. For the LSTM model with a five timesteps range, we found that a layer with 32 units has similar accuracy as one with 256 units. Therefore, we eventually used the smaller model with 221 956 trainable parameters, consisting of the LSTM layer with 32 units and a fully connected output layer with 4 units. The performance of the LSTM is worse than IHM or the convolutional ANN, with a mean square error that is roughly twice as large. However, due to the time averaging of the LSTM architecture for the steady state operation the temporal fluctuation at the concentration points is lower with a 20% reduced standard deviation. The prediction time of the LSTM network is still very low at 1.05 ms. The fast prediction times are relevant for online monitoring applications and may in the future enable not only an analysis of the compound concentrations but also real-time estimates of error margins. An additional advantage of the LSTM network analysis of time series is its ability to predict not only concentrations but also their rate of change as well as parameters of chemical kinetics. While the present dataset with steady state concentration plateaus does not lend itself to such an analysis, we expect this to be of relevance for applications, e.g., in feedback control loops. The LSTM performance may also be improved in the future with a different architecture such as merging our two approaches by combining a locally

connected convolutional layer as feature selector and input for an LSTM layer. To conclude, the presented results show that it is generally possible to apply ANN approaches to low-field NMR data with limited experimental data.

## IV. DISCUSSION: EMBEDDED PROCESS CONTROL

Upcoming architectures for ML-assisted Embedded Process Control will require meeting several important requirements such as hard real-time capability, high data throughput, and high energy efficiency. In particular, with the growing number of IoT devices, the new paradigm of *edge computing* has emerged as a solution to ever-increasing high data bandwidth and low-latency requirements [15]. Artificial Neural Networks as used in this work are known to have high computational and memory requirements. Therefore, edge architectures raise a difficult challenge of executing important workloads with high performance, while achieving the extremely low power consumption requirements of embedded systems (e.g., for explosion safety) – compatible with current 2- and 4-wire communication. In addition, there are further computational tasks to be performed, e.g., to meet I4.0 requirements such as soft sensors, predictive maintenance or cryptology, as mentioned earlier. This requires flexibility and expandability to support certified updates and patches to keep the system safe. It may sound like a contradiction, but current systems already meet the requirements analysis in terms of functional safety (SIL conformity).

Considering all these requirements, Field-Programmable Gate Arrays (FPGAs) represent an interesting architectural choice. FPGAs enable a high degree of customization, allowing the implementation of flexible application-specific solutions that can be updated over time, such as an ANN model that must be updated to improve the functionality, but which also must be quickly developed for a fast time-to-market. However, FPGAs, as of today, still also require expert knowledge from hardware designers to be used, despite several advances from Electronic Design Automation (EDA) tools in the past few years [16]. As such, the entire process of designing applications that leverage FPGA architectures is still expensive and unpractical.

Compared to manually developed hardware or FPGA designs created with HLS software, overlay architectures reduce the development effort significantly while slightly reducing the flexibility. An overlay architecture that has been developed for a class of applications and only needs to be parameterized for the intended use further requires no or only little knowledge in the area of hardware development. This enables software developers to use the advantages of application specific hardware.

The VCGRA architecture, for example, enables the creation of parameterizable CGRA-Architectures, that can be suited to a specific application by configuring the connections inside the CGRA and the capabilities of the processing elements [17]. It has been designed to be used on top of commercial FPGA and comes with a complete toolflow for designing the FPGA configuration and the interfaces in hard- and software. Additionally, the tools for converting an application into a configuration for the VCGRA are provided.

Using this architecture for the efficient implementation of ANNs is still a work in progress; however, the parallelisation opportunities and the feasibility to tailor the processing elements to specific operations and number formats make it a promising approach. The results show its potential for the efficient execution of ANN workloads.

Another interesting possibility is the use of soft-core GPUs on FPGAs, which can be optimized on-demand for different targets depending on the application requirements [18], [19]. Such an architecture has also shown promising results in the acceleration of fundamental kernels in ANN processing, like Matrix Multiplication, achieving an average  $4.2\times$  speedup for different workloads over an embedded ARM core with NEON support [20]. Further specializing increases the speedup numbers by  $100\times$  [19].

## V. CONCLUSION

Industrial production plants in the process industry, such as those found in the chemical, pharmaceutical or food production sectors, can no longer be operated economically and safely without automation and sensor technology. With the transition to Industry 4.0, the existing automation nodes are becoming independently acting cyber-physical systems. This increases the number and complexity of electronic components: Data is collected at the edge of the network via sensors, pre-processed close to the sensor and encrypted before being transmitted in the network (*Internet of Things – IoT*) and in part processed in powerful computer systems (*cloud*). Both at the *edge* of the network near the sensors and actuators and in the central data processing infrastructures (*cloud*), ever higher computing power is required, for example for increasingly complex evaluation algorithms, machine-based learning algorithms and artificial intelligence, but also for cryptology to enable secure data transfer in the *IoT*. As a vision, each automation component will register in a distributed network with its capabilities, workflows and computing power. On the basis of a risk-based assignment (*orchestration*), the ideally customized automation networks will be assembled. The tasks of higher automation levels are thus predominantly directed towards orchestration and no longer towards the dedicated control of the underlying automation components. It then no longer matters where and how these components will be connected.

In this paper we present new methods, to deploy *artificial intelligence* in our case *neural networks*, in measurement systems from the chemical industry. We hope that our approach will facilitate and accelerate the integration of smart sensors into the digital infrastructure. With that we could move to the threshold of a possible use of MS and NMR for reaction and process monitoring, extending the toolbox of chemically specific process analytical methods for process industry. We are already discussing to use the proposed AI methods in other analyzers like, e.g., Near Infrared Spectroscopy or for combinations of analytical methods.

Concepts based fundamentally on cyber-physical systems will change production processes and, e.g., enable new topologies of plant modules which are not economic

today or are not even possible for control or safety reasons. Such fully integrated and intelligently networked systems and processes require reliable and powerful automation nodes that are operated within the modular plant. This *edge computing* leads to the implementation of suitable platforms with processor architectures, which meet the requirements such as hard real-time capability, high data throughput and computing efficiency.

Once the feasibility of machine-supported learning and evaluation systems has been demonstrated, it has to be investigated how these systems can be automatically and reliably adapted to perturbations or changes in parameters within the life cycle of a production. Especially the advantage of a fast computing time will make it possible to use machine-supported analysis methods to increase functional safety, e.g., by not only evaluating process data but also checking them against potential hazard indicators from the data.

The precision of the results show that there is still research required to optimize the methods by investigating the used neural networks and training methods. However, the results are very promising and the advantages to use AI in this domain are very visible. The next steps in the research will be the optimization of the used AI algorithm to improve the precision and simultaneously the optimization for the integration into embedded systems for edge computing.

## ACKNOWLEDGMENTS

The authors would like to thank their colleagues Lukas Wandler and Martin Bornemann-Pfeiffer for their support in conducting the experiments and fruitful discussions.

## REFERENCES

- [1] J. Wan, M. Yi, D. Li, C. Zhang, S. Wang, and K. Zhou, "Mobile services for customization manufacturing systems: An example of industry 4.0," *IEEE Access*, vol. 4, pp. 8977–8986, 2016.
- [2] N. Mohamed, J. Al-Jaroodi, and S. Lazarova-Molnar, "Leveraging the capabilities of industry 4.0 for improving energy efficiency in smart factories," *IEEE Access*, vol. 7, pp. 18008–18020, 2019.
- [3] T. D. W. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, 3rd ed. Amsterdam, The Netherlands: Elsevier, 2016.
- [4] R. K. E. Harris and R. E. E. Wasylshen, *Encyclopedia of NMR*, 2nd ed. Hoboken, NJ, USA: Wiley, 2012.
- [5] M. Maiwald, H. H. Fischer, Y.-K. Kim, K. Albert, and H. Hasse, "Quantitative high-resolution on-line NMR spectroscopy in reaction and process monitoring," *J. Magn. Reson.*, vol. 166, no. 2, pp. 135–146, 2004.
- [6] K. Meyer, S. Kern, N. Zientek, G. Guthausen, and M. Maiwald, "Process control with compact NMR," *TrAC Trends Analytical Chem.*, vol. 83, pp. 39–52, 2016.
- [7] S. Kern *et al.*, "Flexible automation with compact NMR spectroscopy for continuous production of pharmaceuticals," *Anal. Bioanal. Chem.*, vol. 411, no. 14, pp. 3037–3046, 2019.
- [8] F. Kästner, D. Kuschnerus, C. Spiegel, B. Janssen, and M. Hubner, "Design of an efficient communication architecture for cyber-physical production systems," in *Proc. IEEE 14th Int. Conf. Autom. Sci. Eng.*, 2018, pp. 829–835.
- [9] F. Kästner, O. N. Guzman, B. Janssen, J. Hoffmann, and M. Huebner, "Analysis of hardware implementations to accelerate convolutional and recurrent neuronal networks," *Int. J. Advances Softw.*, vol. 10, no. 3/4, pp. 308–323, 2017.
- [10] S. Kern *et al.*, "Artificial neural networks for quantitative online NMR spectroscopy," *Anal. Bioanal. Chem.*, vol. 412, pp. 4447–4459, 2020.
- [11] C. M. Bishop, "Neural networks and their applications," *Rev. Sci. Instrum.*, vol. 65, no. 6, pp. 1803–1832, 1994.

- [12] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, and S. J. Gibson, "Deep convolutional neural networks for Raman spectrum recognition: A unified solution," *Analyst*, vol. 142, no. 21, pp. 4067–4074, 2017.
- [13] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway networks," 2015, *arXiv:1505.00387*.
- [14] D. Chen, Z. Wang, D. Guo, V. Orekhov, and X. Qu, "Review and prospect: Deep learning in nuclear magnetic resonance spectroscopy," *Chem. Eur. J.*, vol. 26, no. 46, pp. 10391–10401, 2020.
- [15] T. Adegbiya, A. Rogacs, C. Patel, and A. Gordon-Ross, "Microprocessor optimizations for the Internet of Things: A survey," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 1, pp. 7–20, Jan. 2018.
- [16] R. Nane *et al.*, "A survey and evaluation of FPGA high-level synthesis tools," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 35, no. 10, pp. 1591–1604, Oct. 2016.
- [17] F. Fricke, A. Werner, K. Shahin, F. Werner, and M. Hübner, "Automatic tool-flow for mapping applications to an application-specific CGRA architecture," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops*, 2019, pp. 147–154.
- [18] M. Brandalero, H. G. M. Hernandez, M. Veleski, M. A. Kadi, P. Rech, and M. Hübner, "(Special topic submission) enabling domain-specific architectures with an open-source soft-core GPGPU," in *Proc. IEEE Int. Parallel Distrib. Process. Symp. Workshops*, 2020, pp. 36–43.
- [19] R. Ma *et al.*, "Specializing FGPU for persistent deep learning," in *Proc. 29th Int. Conf. Field Programmable Log. Appl.*, 2019, pp. 326–333.
- [20] M. A. Kadi, B. Janssen, J. Yudi, and M. Huebner, "General-purpose computing with soft GPUs on FPGAs," *ACM Trans. Reconfigurable Technol. Syst.*, vol. 11, pp. 5:1–5:22, Jan. 2018.



**KLAS MEYER** received the diploma in chemistry and PhD degrees from Humboldt University Berlin, Berlin, Germany, and he joined Bundesanstalt für Materialforschung und -prüfung (BAM) from 2012–2017 focusing on high-pressure NMR spectroscopy for gas analysis and for process industry. He joined Magritek GmbH in Aachen as application scientist and came back to BAM in 2018 for quantitative and online NMR spectroscopy and automation.



**STEFAN KOWARIK** is professor with Physical and Theoretical Chemistry Department, Karl-Franzens University Graz. He works on real-time X-ray analytics.



**ROBIN HIERZEGGER** is currently working at applications of neural networks in chemistry with the University of Graz.



**STEPHAN WESTERDICK** received the BSc degree in electrical engineering and the MSc degree from Ruhr University Bochum, Bochum, Germany, in 2012 and 2014, respectively. He was a student research assistant with the Institute of Digital Communications Systems and the Institute of Electronic Circuits during his studies. He has focused on high frequency and optical systems.



**MICHAEL MAIWALD** received the graduate degree from Ruhr-University Bochum, Bochum, Germany, in 1994, and the habilitation degree from Technical University Kaiserslautern, Kaiserslautern, Germany, in 2012. He is since 2008 head of Division Process Analytical Technology, Bundesanstalt für Materialforschung und -prüfung (BAM), Berlin, Germany. Previously, he was group leader with Merck KGaA, Darmstadt, Germany. His research interests include online NMR and optical spectroscopy, process analytical technology, chemometrics, and process automation.



**MICHAEL HÜBNER** (Senior Member, IEEE) received the Dr-Ing and habilitation degrees from the Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, in 2007 and 2011, respectively. He is a full professor with the Brandenburg University of Technology, Cottbus, Germany, and the chair of Computer Engineering Group. His research interests include reconfigurable computing with application in automotive systems.



**FLORIAN FRICKE** received the BEng degree from the South Westphalia University of Applied Sciences (FH-SWF), Iserlohn, Germany, in 2012, and the MSc degree from the Ruhr-University of Bochum (RUB), Bochum, Germany, in 2015. Currently, he is currently working toward the PhD degree at the Brandenburg University of Technology, Cottbus, Germany working on reconfigurable systems like CGRAs and FPGA overlay architectures and their applicability for edge computing on embedded systems.



**MARCELO BRANDALERO** (Member, IEEE) received the doctor's degree in computer science from the Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, in 2019, and is currently a senior research scientist with the Brandenburg University of Technology Cottbus-Senftenberg, in Germany. His research interests cover computer architecture with emphasis on reconfigurable architectures.



**SASCHA LIEHR** received the PhD degree from Technische Universität Berlin, Berlin, Germany. He studied with Technische Universität Ilmenau. 2007–2020, he joined Bundesanstalt für Materialforschung und -prüfung (BAM) as research scientist. Since 2020, he is optical systems engineer with DiGOS Potsdam.



**SIMON KERN** received the master's degree in bioprocess engineering from Technische Universität Hamburg-Harburg, Hamburg, Germany. He worked 2015–2019 as PhD student with Bundesanstalt für Materialforschung und -prüfung (BAM) and Technische Universität Berlin focusing on process control with low-field NMR spectroscopy. Since 2019, he is working as an application scientist with S-PACT GmbH.