Visual Exploration and Analysis of Simulation and Testing Data in Motor Engineering

Patrick Louis, *Pro2Future GmbH, Graz, 8010, Austria*

Lena Cibulski, *Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, 64283, Germany Technical University of Darmstadt, Darmstadt, 64283, Germany*

Josef Suschnigg, *Pro2Future GmbH, Graz, 8010, Austria*

Edmund Marth, *Institute for electrical drives and power electronics, Johannes Kepler University, Linz, 4040, Austria*

Hubert Mitterhofer, *Linz Center of Mechatronics GmbH, Linz, 4040, Austria*

Jörn Kohlhammer, *Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, 64283, Germany*

Tobias Schreck, *Graz University of Technology, Graz, 8010, Austria*

Belgin Mutlu, *Pro2Future GmbH, Graz, 8010, Austria*

Abstract—
End-of-line tests and defect detection are vital for ensuring the reliability of electric motors. However, automated defect detection methods, e.g., data-driven approaches, face challenges due to the limited availability of real data from failed motors. Simulated data, though beneficial, lacks the complexity of real motors, impacting the performance of these methods when applied to actual observations. To tackle this challenge, we introduce a visual analysis tool designed to facilitate the analysis of measured and simulated data, presented in the form of time series data. This tool helps identify domain-invariant features and evaluate simulation data accuracy, assisting in selecting training data for reliable automated defect detection in real-world scenarios. The main contribution of this work is a design proposal based on visual design principles, specifically tailored to address the unique requirements of electric motor professionals. The visual design is validated by findings from a think-aloud study with specialized engineers.

The world we live in today is powered by systems, machines, and processes that are the results of engineering. Electric motors are at the basis of many industrial and domestic applications, converting electrical energy into various forms of mechanical energy in a highly efficient manner. Their performance affects key indicators such as consumption, lifetime or performance of the mechanical process. This places high demands on the validated design of electric motors.

To ensure that each electric motor produced functions properly and meets the target characteristics, the motors are tested at the end of the production line. This process is called end-of-line testing. Besides traditional physics-based failure prediction models, so-called expert systems, contemporary data-based approaches open up new possibilities.

The achievable quality of such data-based models for automated defect detection depends heavily on the quality and quantity of available data. However, real data from motors on the verge of failure is sparse, as each production line inherently aims to produce only defect-free motors, resulting in many more end-of-line tests being performed on correct motors than on defective motors. In order to obtain more real data from defect motors, such motors would have to be produced on purpose, which is not economically

XXXX-XXX © 2024 IEEE
Digital Object Identifier 10.1109/XXX.0000.0000000

viable. To expand the set of training data in a more efficient way, motor defects can be simulated. However, each simulation model is likely to deviate from the real system. Consequently, automated defect detection methods trained on simulations will most likely lead to a drop in performance. Marth et al. [1] introduce a method that utilizes a model trained on simulated data of healthy and faulty motors to generalize to real-world measurement data. The method relies on the crucial assumption that the domain shift between measurements and simulation results remains constant under different system operating conditions. The visualization approach proposed in this work aims to either confirm this assumption by seamlessly inspecting measured and simulated data, or, if not valid, identify features within the time series with minimal variance with respect to the two domains in the parameter space. Visualizations, rooted in fundamental principles of visual cognition, provide a universal and intuitive means for interpreting complex data by users. By representing time series data graphically, users can quickly grasp the overall structure and identify critical features.

The main contribution of this work can be summarized as a design proposal based on established principles of visual design and a thorough requirements analysis of the target domain, thus seamlessly adapting to the specific needs of professionals in the field. We propose an interactive analytical lens that enables a context-dependent identification of patterns between simulation and measurement data. The lens operates in the space of torque and rotational speed, which are fundamental motor engineering parameters, and supports finding global and local patterns in this space. In addition, we present the results of a use case scenario that gives an insight into the application of the proposed tool from the perspective of an electric motor engineer. A think-aloud study, involving three engineers with specialized knowledge, validated the visual design. Their insights proved invaluable, given their firsthand experience. While the overall tool received positive feedback, we identified areas for improvements to enhance the overall usability of the tool.

RELATED WORK

We structure previous works along our two main contributions: the analysis of motor behaviors and the visual exploration of time series data.

Methods to analyze the behaviour of electric motors are diverse, ranging from physics-based expert systems [2] to black-box models using artificial intelligence [3]. The recent research activities and the traditional expert analysis methods are summiarized in [4], [5]. A common source of information is time series data generated by current or voltage probes, acceleration sensors, or other specialized measurement devices that are sometimes complemented with simulated time series from a digital twin [6]. The time series data is commonly processed by, e.g., fast Fourier or wavelet transformations that produce characteristic frequency signatures. Black-box models or human experts then map these signatures to motor defects. In previous research [1], combining simulations and measurements enhanced the prediction capability of black-box models. Our goal is to utilize human domain expertise to analyze simulated and measured time series data, identifying defect-specific features that a data-driven model can exploit.

Various approaches exist for visually exploring time series data. Circular visualizations, such as those by Zhao et al. [7], efficiently display tree structure relationships, while highlighting cyclic patterns, as seen in the work by Ceneda et al. [8]. Hao et al. [9] dynamically allocate display space to show sub-intervals in long time series at varying resolutions based on user interest. Different data processing methods have also been applied to enhance visualization and exploration. For instance, Shi et al. [10] use reinforcement learning to offer exploratory visual analysis suggestions, while Andrienko et al. [11] abstract multi-variate time series distribution to episodes and topics for applying topic modeling techniques. Another approach involves dimensionality reduction [12] to visualize multi-variate time series. In the domain of electric motors, prior methods have simplified comparisons across numerous measurements at various construction stations during manufacturing [13], and visualized trade-offs in motor characteristics during design [14]. In contrast, our tool concentrates on end-of-line electric motor data.

The core view of our tool uses a scatterplot [15]. The two axes of the scatterplot represent the operating conditions of the electric motor as an engine map, showing rotational speed on the horizontal axis and torque on the vertical axis. Each scatterplot entry represents one measurement time series and one simulation time series. Our tool utilizes semantic zooming [16] to transform scatterplots and employs lenses, similar to the Magic Lenses framework [17], for aggregating multiple scatterplot entries.

DOMAIN CHARACTERIZATION

We start by introducing fundamental concepts of electric motors and their end-of-line testing. Based upon that, we provide a characterization of the data and tasks that engineers face when searching for domain-invariant features in cases where measurements and simulations differ from each other. We conclude with a number of requirements to be addressed by the visual tool.

End-of-Line Testing Background

For products with critical importance in their operational use case, e.g., security-relevant devices, end-of-line testing is an essential step in the production process [18]. It validates the functionality of units near the end of the production line. Products that do not function or whose characteristics do not match the tolerance band around the target characteristics must be identified and removed from the further production or sales process. To shorten production cycles and improve reliability, automated testing increasingly replaces manual inspection performed by human operators.

We choose an end-of-line test for permanent magnet synchronous motors (PMSM) as a use case, because they are a popular choice for applications which demand high efficiency and power density. These range from the automotive sector to industrial drives and household appliances or medical devices but continue to expand into all other fields of use of electric motors. Diverse defects can occur during the manufacturing of PMSMs, many of which are related to geometric deviations of motor parts or to misplacements of magnets, windings, or angular sensors. Misplaced angular sensors produce an erroneous rotor angle φ as input to the motor's control circuit. This results in incorrect control currents, leading to a non-optimal working point with reduced motor efficiency and torque ripple, and, in the extreme case, reduced lifetime. We choose a misplaced angular sensor as the defect property under investigation, because its measurements can be easily emulated by manipulating the control software of the frequency converter of an otherwise functioning motor. Going forward, the amount by which an angular sensor is misplaced will be denoted as 'rotor angle offset' or 'angle offset'. Despite the use of misplaced angle sensors for this work, the approach is versatile and can be applied to any situation with available time series measurements and simulation data. It helps in the identification of range shifts and the investigation of time series signals within a given two-dimensional state space.

A motor in an end-of-line test is exposed to different load points to observe its response using different sensors. Because the phase current signals can be obtained relatively easily, they are often used as the primary source of information. The resulting multivariate time series typically do not show characteristics that can be directly attributed to a specific error. *Erroneous rotor angles*, for example, will lead to the phase current signals being distorted with respect to their characteristic shape (e.g., sinusoidal shapes) and their amplitudes. However, as the distortions can also originate from varying load conditions, a simple mapping of the time series acquired during the end-of-line test to a particular error is not possible.

For a model to automatically identify motor defects from the current signals, training data is essential. While real-world data from motors functioning as intended might be plenty, data from motors at or near failure is much sparser. In such cases, the training data can be augmented with synthetic data, i.e., simulated motor defects [1]. Synthetic data can be generated in a controlled way to only contain a specific error. This means that it is available in large amounts, can be properly labelled and most often does not require prior data cleansing since it is not subject to sensor noise or external disturbances.

Despite its advantages, a simulation model can only approximate the real-world behaviour of the motor. Many data-driven models (e.g., for fault detection), however, rely on the assumption that the data samples to be used for training are drawn from the same distribution. Therefore, characterizing the shift or differences between these data distributions becomes crucial for ensuring the training of an accurate and reliable model. A human-in-the-loop approach using interactive visualizations enables experts to proactively identify and understand subtle shifts or discrepancies between the simulation and real-world data distributions. Through direct interaction with visual representations, experts can identify differences in the data distributions and, if necessary, pre-process the data for model training, thereby guaranteeing a more precise and resilient representation of motor behavior.

Data

The time series data describe the behavior of the permanent magnet synchronous motor (PMSM) under different load conditions. Measurement data is acquired using a dedicated test bench (Fig. 1). The setup consists of 1) the motor under test, 2) a high-resolution encoder to measure the rotor angle, 3) a break and a torque sensor to control the load condition, and 4) a power electronics to control the motor and the brake and to measure the motor's phase currents.

Simulation data is acquired using a surrogate model of the motor that is created based on finite

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3392969

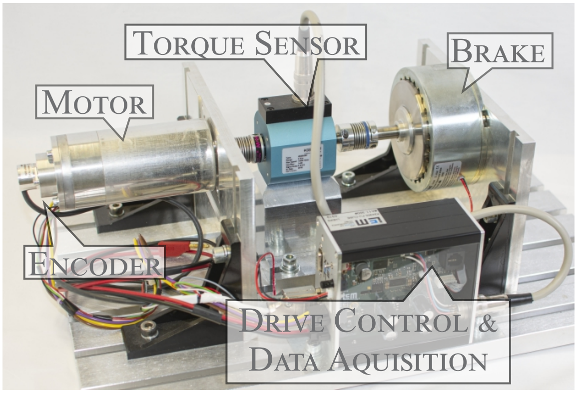

FIGURE 1. Test-bench to gather measured time series data of the motor's phase currents under certain load conditions.

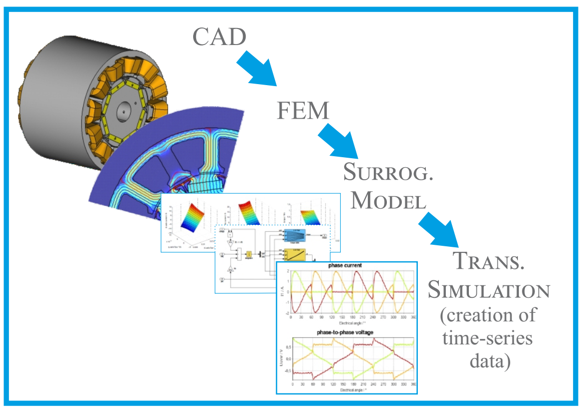

FIGURE 2. Simulation chain to gather time series data of the motor's phase currents under certain load conditions.

element simulations [19]. Using this model, transient simulations can be performed to generate time series data of the motor's phase currents under different load conditions (Fig. 2). Motor modeling and simulation were performed with *SyMSpace*, a development and optimization tool for mechatronic components and systems [20].

With the test bench and the simulation framework at hand, 1476 randomly selected load points, i.e., combinations of rotational speed and load torque, were evaluated (Table 1). For each measurement, a simulation with the same load point and error was performed. This opens up the possibility to directly compare the measurement and simulation time series.

1106 measurements do not show an erroneous rotor angle, while 370 measurements do. Equivalently, 1106 simulations do not show an error and 370 simulations do. All measured time series contain a fixed number of 408 samples with a sampling rate of 20,000 samples per second, leading to a fixed time series length of 0.02035 seconds. The simulations have the same sampling rate but contain a varying number of samples between 298 and 1197.

TABLE 1. Operation parameters of the motor for measurements and simulations. A motor with a rotor angle offset between $28^{\circ,el}$ and $32^{\circ,el}$ is considered defect-free. $^{\circ,el}$ designates degrees in the electrical reference frame, where in case of the motor under test $\pm10^{\circ,el}$ means a deviation of $\pm2.5^{\circ}$ mechanically. This range for the correct rotor angle offset was determined to be useful for this specific motor setup by domain experts. Other electric motors might necessitate a change of the correct value range.

The time series have been sampled at $50\mu s$.

Parameter	Min.	Max.	Unit	Description
n	1003	4030	rpm	Rotational speed
T	0.1019	0.6696	Nm	Load torque
φ_{offset}	20	39	$^{\circ,el}$	Rotor angle offset

Task Analysis

We identified the following key tasks and analysis questions based on readings about the target domain background as well as feedback and existing research from domain experts in electrical engineering [4], [21]. Asking three experts about their current workflows and discussing an exemplary usage scenario in multiple sessions provided us with a fundamental understanding of the targeted problem. This resulted in three key Tasks (T) for the analysis of time series data during end-of-line testing:

- **T1 - Simulation Validation**: Determine how well the simulation model approximates the real-world measurements. If the simulated time series are not representative, they will not provide meaningful information that can be used for defect detection. This leads to the question what level of precision is required. An additional benefit of comparing measurements to simulations is the ability to determine under which load conditions the simulation model works well or poorly.
- **T2 - Visual Feature Exploration**: Identify time series characteristics that meaningfully relate to certain motor conditions (e.g., full operability or defects). The characteristics need to be domain-invariant, i.e., they should help distinguish between defective and non-defective motors in both the measurement and the simulation domain.
- **T3 - Process Improvement**: Identify load conditions where correct and defective motors cannot be clearly distinguished. This could hint at load points

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3392969

that should be tested in more detail to increase the expressiveness of visual features with regard to defect detection. Also identify load conditions where real-world measurements are inconsistent compared to expectation based on the simulation. This could suggest a problem at the end-of-line testing station.

DATA PRE-PROCESSING

Given the heterogeneous data sources, certain pre-processing steps are necessary to ensure that simulated and measured time series can be shown in a consistent way. This is critical for comparisons across different load conditions.

Measuring the phase currents of the PMSM results in three time-shifted but otherwise identical time series at each load point. To avoid redundancy, only one time series representing one of the phases is visualized. This also matches the cardinality of the single simulation time series.

Due to measuring time-shifted phase currents, the measurement time series and the simulation time series do not overlap perfectly. To correct this, we perform a cross-correlation analysis on each pair of measurement and simulation. For different time shifts, the sample correlation coefficient [22] between measurement and simulation is calculated. The simulation is then shifted by the time step corresponding to the largest coefficient.

As mentioned previously, the number of samples in simulated time series varies, while the number of samples for each measurement time series is constant. To keep comparisons consistent, the length of longer simulations is trimmed to the length of measurements.

When aggregating multiple time series, the rotational speeds at which they were captured varies. Because the period length of the phase current pattern depends on the rotational speed, this leads to a problem when calculating the average time series (Figure 3a). This problem is solved by converting the x-axis for each time series from time to rotation angle before the calculation. This leads to an x-axis independent from the rotational speed of the data which allows the aggregation of multiple time series (Figure 3b).

Finally, because the measurements were done independently they might still have a relative time shift. By choosing one measurement as a reference, this time shift can also be eliminated by a cross-correlation analysis. Afterwards, multiple measurements or simulations can be seamlessly aggregated and compared.

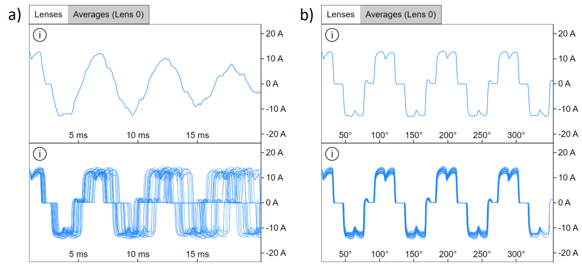

FIGURE 3. An example of an aggregation of the phase current of multiple measurements over time (left) and over the rotation angle (right). The top graph shows the average of all measurements, the bottom graph shows a superposition of all averaged measurements.

ITERATIVE DESIGN PROCESS

The visual analytics tool was developed iteratively over five versions. Discussions by experts in both the visualization domain and the electric motor domain led to updates between each version (Figure 4). While we considered novelty important, our focus was on providing an effective solution for our experts' domain problem by leveraging proven visualization and interaction techniques. This required a deep understanding of the unique challenges, constraints, and user expectations within the targeted domain to provide a solution that is relevant to the real-world problem and tailored to its requirements. The purpose of our iterative approach was to refine the tool in response to expert feedback and to align it to the specific requirements of the electric motor domain. The feedback and new features added between successive iterations are discussed in this section.

Design Goals

In Section DOMAIN CHARACTERIZATION, we outlined three main tasks in the electric motor domain. To enhance the effectiveness, clarity, and interpretability of data visualization for these tasks, we prioritize core design goals. Despite a broad range of design objectives in existing literature ([23], [24], [25]), our primary focus is on meeting the specific requirements of domain experts.

- **DG1 - Visual Clarity:** When looking at or interacting with a part of the visualization it should always be clear what information is shown. Avoid re-using names or colors for different parts of the visualization. When needed, add tool-tips or descriptions to aid understanding. For the electric motor domain, make sure that measurement data and simulation

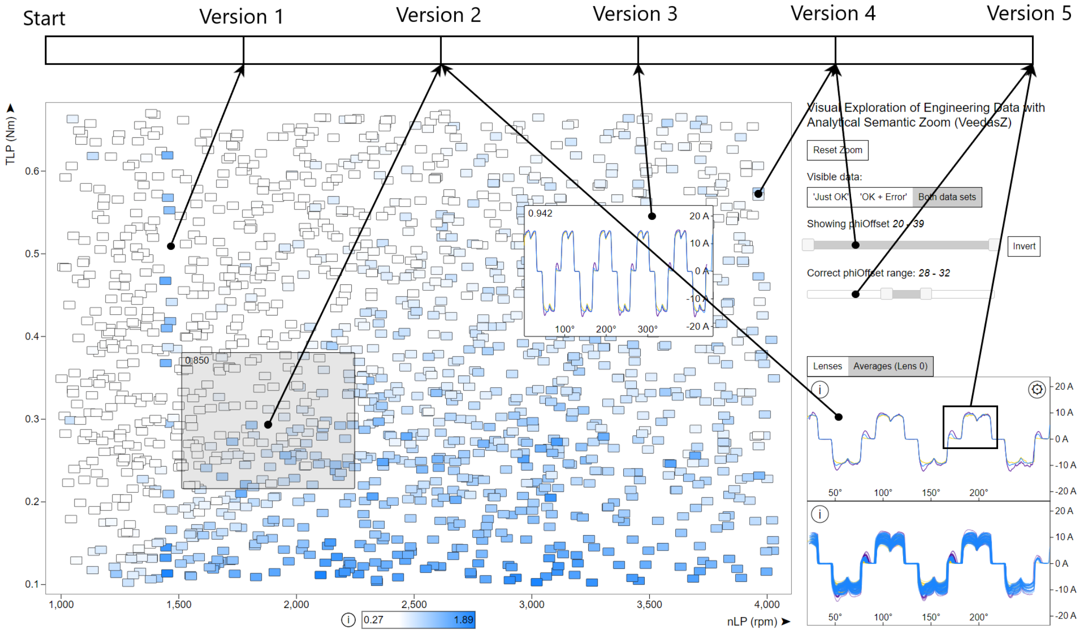

FIGURE 4. A history of the added changes during development of the visualization. The scatterplot with semantic zoom functionality and the general layout was added in version 1. Scatterplot lenses and the connected average measurement calculation were added in version 2. Version 3 added the functionality to show the average in the lens. The color gradient in the scatterplot and the range slider to hide entries based on their angle offset were added in version 4. Version 5 added the split of the lens averages based on the angle offset of the electric motor, the ability to calculate average simulations and the range slider to set the correct angle offset range.

data can be distinguished and compared (cf. **T1**).

- **DG2 - High Level of Interactivity:** This visualization functions as a way to get knowledge from domain experts before trying to train or adjust a simulation model for further use. Before this human knowledge is integrated, it can not be known which part or aspect of the visualized data is most important for the analysis. Therefore the visualization should have a high level of interactive functionality to allow the electric motor domain experts to separately analyze measurement and simulation data from defective and non-defective motors (cf. **T2**).
- **DG3 - Flexible Exploration:** Related to the last point, it is important to provide a high level of freedom with the implemented exploration functionalities. Because the usage scenario is not known in advance, assumptions about how the analysis process will work should be avoided. Instead, the separation of measurement data and simulation data (cf. **T1**) and the separation of defective and non-defective motor data (cf. **T2**) should be accessible at any step of an analysis.
- **DG4 - Scalability:** The visualized data consists of time series pairs created at different motor load points. The number of the pairs is only dependent on how many experiments were done and there are no upper or lower limits for the number of pairs in a data set. The visualization should be able to display large numbers of time series pairs while still allowing comparisons within and between motor load points (cf. **T3**).

Initial Prototype

The data is arranged in a two-dimensional scatter plot with rotational speed (in rotations per minute) on the horizontal axis and load torque (in Newton-metres) on the vertical axis. This encoding resembles engine maps that are commonly used in engineering (**DG1**). Each scatter plot entry represents a measurement and a simulation at a particular load point. The scatterplot can be transformed through semantic zooming (**DG2**). Depending on the zoom level, the rectangle is transformed over a small multiple to a detailed line chart visualization of both time series (Figure 5).

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3392969

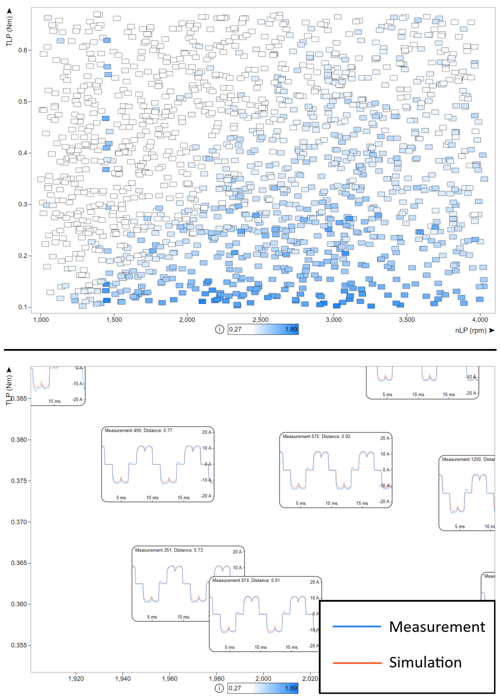

FIGURE 5. Two views of the scatterplot implementation in the visualization. The top view shows the lowest zoom level with all data visible and the entries colored according to the color scale. The legend below the scatterplot describes the distance between measurement and simulation within each scatterplot entry. The bottom view shows a high zoom level with only a few entries visible as line plots due to semantic zooming.

Introduction of Lenses

The idea to average time series over scatterplot regions was first introduced and implemented as a single scatterplot lens, which could be moved and resized (**DG2**). A scatterplot lens is a rectangular region of the scatterplot which aggregates the measurement time series contained within it to calculate a single average measurement (**DG4**). The average distance between measurement and simulation of entries within the lens is shown in the top left corner of the lens (**DG1**). The average measurement time series and all measurement time series in the lens in superimposition are also shown in a separate window in the bottom left of the layout. (**DG3**).

Enhancement of Lenses

Next, the functionality to create and delete multiple independent lenses was added to increase the flexibility of their usage (**DG2**). Additionally, the average measurement time series of entries in a lens could now be shown directly in the lens to allow comparisons directly in the scatterplot (**DG3**).

Comparison Functionality

Following this, the data alignment procedures described in the chapter DATA PRE-PROCESSING were implemented, leading to more exact averages and visualizations (**DG1**). Another addition was the color scale used to color the scatterplot entries. For each entry, the color scale value is calculated as the Euclidean distance between the measurement and the simulation (low -> white, high -> blue). This allows an overview on how well the simulation model reflects the real-world measurements under different load conditions (**DG4**).

To allow for separate analysis of time series from non-defective motors and time series from defective motors, a slider to hide or show scatter plot items depending on the defect property (in this data set: the angle offset φ_{offset}) was also added (**DG3**).

Feature Exploration

For data-driven defect detection, it is important to identify features which distinguish non-defective motors from defective motors. Therefore, the average calculation within a lens was changed to distinguish between low angle offsets, correct angle offsets and high angle offsets. The average measurement time series is then calculated for each group and three average time series are displayed in the separate window and in the lens (**DG2**). To allow the same analysis for simulation time series, the functionality to aggregate simulations in the scatterplot lenses was added. The displayed average can then be switched between measurement and simulation on demand (**DG3**, Figure 6). Another related addition was a second range slider, used to set the range of the angle offset which is considered correct (**DG2**). This is necessary because the correct range of the angle offset can be context-dependent, making it hard to specify in advance.

In summary, the iterative design process has led to the further development of our visualization tool, culminating in a version that effectively meets the needs of the electric motor sector. The following sections look at the impact of these refinements on the usefulness and applicability of our visualization tool.

USAGE SCENARIO

This section describes a possible usage scenario of our visualization tool from the perspective of an electric motor engineer. In particular, the usage scenario aims

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3392969

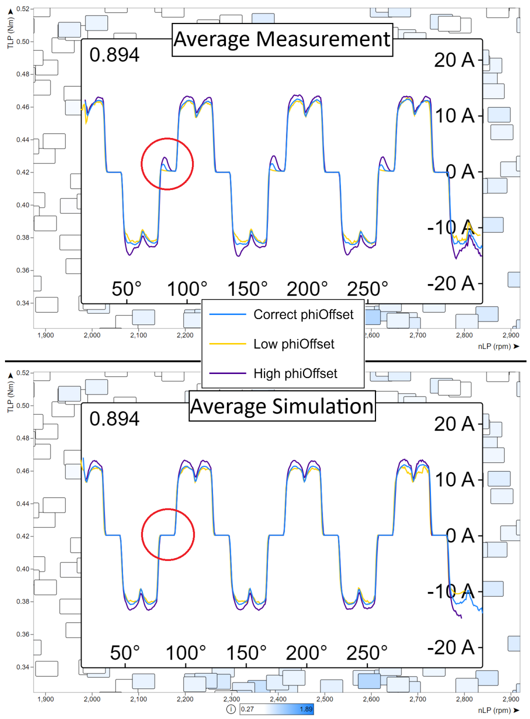

FIGURE 6. Two images showing the same scatterplot lens. In the top image, the lens calculates the average measurement of its enclosed entries. In the bottom image, the lens calculates the average simulation. The horizontal axis measures the motor rotation in degrees, the vertical axis measures the electrical current. This comparison helps find domain-invariant features and identify differences between the domains. An example of a difference in shape between domains is the overshooting highlighted in the measurement (top), which is not present in the simulation (bottom).

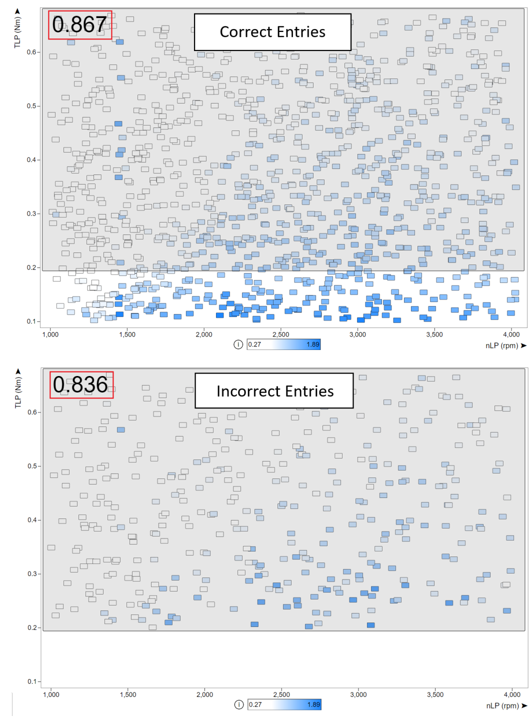

FIGURE 7. A comparison of two large lenses spanning the scatterplot. The legend below the scatterplot describes the distance between measurement and simulation within each scatterplot entry. In the top image, the scatterplot only shows correct entries (entries with a correct rotor angle). In the bottom image, the scatterplot only shows incorrect entries. The lenses are sized such that they cover all incorrect entries because correct entries and incorrect entries cover different load point ranges. For this data set, the average distance between measurement and simulation in the lenses (shown in the top left corner) is very similar in both cases. This shows that the simulation model works equally well for correct and incorrect motor conditions.

to identify similarities and differences between measurement data and simulation data. This gives insights into the quality of the simulation model.

The first analysis target the engineer is interested in is how accurately the simulation model depicts the real measured data (**T1 - Simulation Validation**). For this, the engineer looks at the color of the scatterplot entries. In this data set, the entries in the low torque range consistently have a higher distance than entries in the high torque range, which is represented by a darker color (Figure 5, top).
To investigate the reason for this difference, the engineer zooms into the lower torque range until the individual time series become visible. From this, they can see that the high distance stems from overshooting in the measurement time series which is more pronounced at low torque loads (Figure 8a).

A second interesting observation from the entry color is a small number of entries with high distance between measurement and simulation located at approximately 1500 rotations per minute. Again, the engineer zooms closer to these entries and discovers an anomalous current path in the time series (Figure 8b). This discovery leads them to assume that the testing bench had some issue at this specific rotational speed (**T3 - Process Improvement**).

To further investigate the accuracy of the simulation model, the engineer next wants to determine if the angle offset has an influence on the simulations (**T1 - Simulation Validation**). Therefore, they create a lens which covers the complete data set and focus

This article has been accepted for publication in IEEE Computer Graphics and Applications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/MCG.2024.3392969

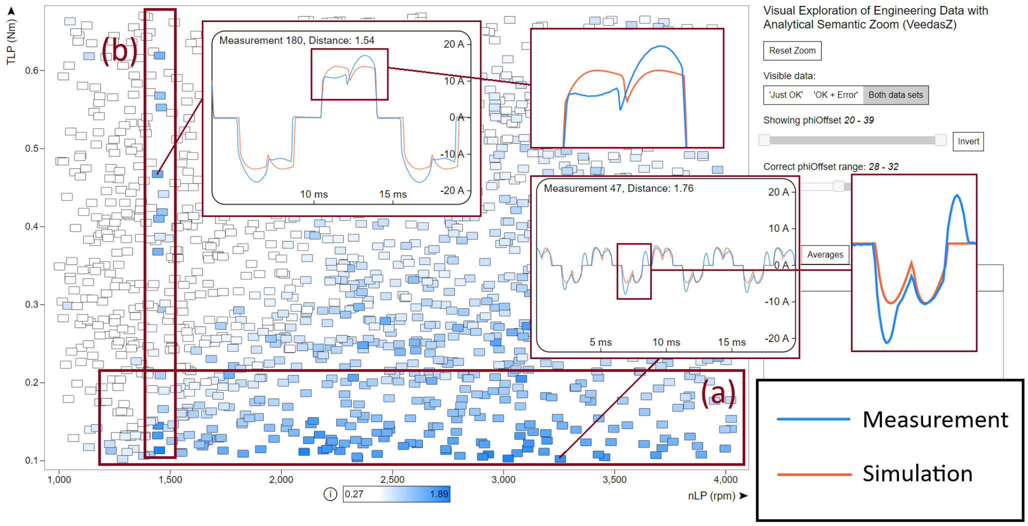

FIGURE 8. An illustration of two anomalous regions of the scatterplot based on the color scale showing the distance between measurement and simulation. The legend below the scatterplot describes the distance between measurement and simulation within each scatterplot entry. Region (a) shows higher distances at low torque, region (b) shows higher distances at a specific rotational speed close to 1500 rpm. Overlaid is a representative plot for each region. The horizontal axis of the plots measures time, the vertical axis measures the electric current.

on the average distance between measurement and simulation which is displayed in the top left corner of the lens. By using the range slider, the engineer then switches between showing only entries with a correct angle offset and showing only entries with an incorrect angle offset. For this data set, the average distance for these two cases are very similar (Figure 7). Therefore, the engineer can conclude that the simulation model works consistently for correct and incorrect angle offsets.

Next, the engineer find possible time series characteristics which point to a measurement or simulation having a correct/incorrect angle offset (**T2 - Visual Feature Exploration**). To this end, the engineer creates a lens and focuses on the average measurement time series inside the lens. At a specific scatterplot region in this data set, they notices that the average time series for high angle offset entries has a larger overshooting than for correct offset entries (Figure 6, top), making this overshooting shape a input candidate for data-driven defect detection methods. To be useful for training such models, the shape must be domain-invariant, i.e., appear in both measurement and simulation. To verify this, the engineer switches the lens view to show the average simulation. In the average simulation, the overshooting does not appear in any of the three time series (Figure 6, bottom). Therefore, the overshooting shape in the measured phase current is not domain-invariant and can not help with defect detection methods. However, the remaining parts of the time series are the same in both measurement and simulation, making them good candidates for such methods.

Think-Aloud Study with Electric Motor Experts

To assess the effectiveness of our visualization, we conducted a think-aloud study involving three electric motor experts, two of whom are co-authors. They possess specialized knowledge in the domain. The study utilized the described dataset collected from end-of-line testing.

Procedure

Participants began by watching an introductory video detailing the functionalities of the tool. Following this, they undertook a series of open-ended tasks designed to estimate their perception of the tool's utility. Participants were asked to utilize all provided visual features and articulate the information they observed. The tasks were:

1) Explore the scatterplot visualization and explain what you see in the scatterplot.
2) Reason why the distance between measurement and simulation is larger for smaller torque values.

3) Change the range slider to only show entries with an incorrect angle offset. Explain what you see now in the scatterplot.
4) Create two new Lenses. Move one Lens to any area with high torque and move one Lens to any area with low torque. Right-click the Lenses to show the average time series inside the Lenses. Explain what you see in each Lens.
5) When comparing the two lenses, what do you notice? Explain what you see.
6) When comparing the time series within a lens, reason what (if any) influence does the distance between measurement and simulation have?
7) When comparing the time series within a lens, reason what (if any) influence does the rotational speed and the torque (i.e., the position of the Lens in the scatterplot) have?

While solving the tasks, the participants were encouraged to talk about their thought process and to mention any other observations not explicitly mentioned in the tasks. After completing all tasks, the participants were asked to fill out a positive System Usability Scale (SUS) [26], [27] using a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree)[28]. The SUS measures the overall usability of the visualization tool.

Results

The scatterplot arrangement and the color scale were well understood by all participants during the initial exploration. One participant mentioned that "the distance color scale can be used to identify where there are differences between measurement and simulation (P1)".

To answer the next question, the participants zoomed into the low torque area of the scatterplot to see the individual time series. They identified that "through semantic zooming it can be seen that overshooting and undershooting of measurements in the low torque area leads to higher measurement-simulation distance (P1, P2)". One participant noted that this overshooting also happens in some high torque measurements but because the absolute amplitude is larger at higher torque and the absolute value of the overshooting is the same, the resulting difference between measurement and simulation is not as large. One participant suggested an indicator of which percentage of the whole data set is visible when some entries are hidden via the range slider. It was also noted by one participant that "entries with a low angle offset have on average the lowest distance between measurement and simulation (P3)".

Next, the participants created the lenses and experimented with the averages within the lens. They noted that the overall average behaved as expected when moving the lens, with an increase in the covered rotation angle for higher rotational speed and an increase in the current amplitude for higher torque values. It was also noted that the overshooting in the measurement happens most prominently in entries with a high angle offset. Overall, the functionality of the lenses and the information they convey was understood quickly by the participants. The lenses were described as a novel and useful addition to the visualization ("Lenses and Semantic zooming are useful and have the potential to show even more information (P1)").

All participants agreed that the distance between measurement and simulation does not influence the average time series within a lens. Instead, one participant described it as "two separate layers of information". These two layers could describe the first two goals of the visualization, with the distance between measurement and simulation being useful for Simulation Validation and the average time series being useful for Feature Exploration. When comparing the lens averages at different positions in the scatterplot, only the previously mentioned increases in covered rotation angle and amplitude were noticed. The averages were seen as largely consistent across the scatterplot which suggests that the end-of-line test setup and the simulation model work equally for a variety of motor load conditions.

The results of the SUS can be seen in Table 2. The low average scores for question 2 ('I found this Tool to be simple') and question 4 ('I think that I could use this Tool without the support of an expert') suggest that the visualization might be challenging to understand for first-time users, requiring a familiarization step. The relatively small quantity of the collected data prevented us from drawing valid statistical conclusion. This study thus needs a follow-up involving more participants.

DISCUSSION

The visualization tool described in this paper can be used to explore time series data created from end-of-line testing for electric motors. It uses semantic zooming and scatterplot lenses to allow an analysis of both real-world and simulation data as well as comparisons between these domains. In the following, we discuss how the visualization achieves the goals laid out at the beginning:

- **T1 - Simulation Validation:** How well the approximation of real-word data by the simulation model works can be seen through the color scale representing the distance between measurement and

TABLE 2. Results of the system usability scale (SUS).

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
Participant 1	7.5	7.5	7.5	2.5	10	10	10	7.5	10	10	82.5
Participant 2	7.5	2.5	7.5	7.5	7.5	10	10	7.5	5	10	75
Participant 3	10	10	10	7.5	7.5	7.5	7.5	7.5	7.5	7.5	82.5
Average	8.33	6.66	8.33	5.83	8.33	9.17	9.17	7.5	7.5	9.17	80

simulation in each scatterplot entry. While the color scale only shows the correctness of the simulation in each entry individually, an overview of the simulation model performance can be gained by aggregating the color of all entries. Many bright entries signify a better approximation of the real-world data while many dark entries signify a worse approximation. Additionally, the distance between measurement and simulation can also help find problems at the test bench if the simulation data stays similar across a region of the scatterplot while some measurements show unusual patterns. These methods can be useful when there is no precise match between measurement and simulation in the torque-rotational speed parameter space. In such cases, comparing data from both domains within a specific range can be valuable. Additional functionalities like statistical similarity measures across a limited parameter space could enhance the current implementation.

- **T2 - Visual Feature Exploration:** The most effective way to find time series characteristics which identify defective and non-defective motors is by using scatterplot lenses. Scatterplot lenses help explain the differences between them in terms of time series patterns. The average time series within each lens is divided between non-defective and defective motors, therefore it is possible to see which features appear only in one of the time series. Such features can be seen as input candidates for data-driven defect detection methods. By switching the average time series between measurements and simulation it can also be seen if these features are domain-invariant.
- **T3 - Process Improvement:** To find load conditions where there is no clear distinction between defective and non-defective motors, the most useful functionality is the scatterplot lens. Because the average time series are divided between defective and non-defective motors, it is easy to see when the time series are very similar. By moving a lens to different regions of the scatterplot, load conditions with no clear distinctions between correct and incorrect motors can be found.

The analysis of the functionality of this visualization and the results of the think-aloud study show that the visualization achieves the defined goals.

While originally developed for engineers, our tool facilitates the dynamic exploration and analysis of time series data through intuitive visualizations and interactive features. Beyond the detection of defects in electric motors, its applicability extends to various fields of engineering and beyond. It can be easily adapted to domains requiring pattern recognition or anomaly detection such as financial analysis, manufacturing quality control, or condition monitoring [29].

CONCLUSION

We developed a visual analysis tool tailored for electrical motor engineering, specifically to identify key requirements in end-of-line testing. It integrates measurements of phase currents from motor tests with simulated data from a motor model, focusing on the rotor angle offset (φ_{offset}) which is a motor parameter that can influence the performance. The tool effectively assists in comparisons between measurements and simulations at different load conditions to determine if the domain shift is constant, as required for the approach described by Marth et al. [1].

The tool employs scatterplots based on motor parameters like rotational speed and torque, enabling comparison between real and simulated data with semantic zooming and detailed line graphs. Scatterplot lenses aggregate multiple time series, aiding accurate simulation model evaluation, particularly in φ_{offset} relationships. This approach enhances end-of-line testing and data-driven defect detection, streamlining data generation processes for engineers.

The performed expert study laid the groundwork by establishing the validity of the proposed visual design. Our future work will include comparative evaluations to further improve the research results. This approach will not only address concerns regarding the generalizability of findings but also offer additional insights into enhancing usability. Additionally, we aim to automate end-of-line test documentation for customers and developers, integrating annotation features to construct a knowledge base from extensive test and simulation data. Our vision is to develop a system that suggests significant data segments to experts, thereby enhancing the analysis workflow for users of varying expertise levels.

ACKNOWLEDGMENT

This work was supported by the Austrian FFG-COMET-K1 Center Pro²Future (Products and Production Systems of the Future), Contract No. 881844, and by the COMET-K2 "Center for Symbiotic Mechatronics" of the Linz Center of Mechatronics (LCM) funded by the Austrian federal government and the federal state of Upper Austria.

REFERENCES

1. E. Marth, P. Zorn, F. Schmid, S. Masoudian, K. Koutini, and W. Amrhein, "Simulation-assisted Training of Neural Networks for Condition Monitoring of Electrical Drives: Approach and Proof of Concept," in *IKMT 2022; 13. GMM/ETG-Symposium*, pp. 1–7, VDE, 2022.
2. M. Pacas and S. Villwock, "Development of an expert system for identification, commissioning and monitoring of drives," in *2008 13th International Power Electronics and Motion Control Conference*, pp. 2248–2253, 2008.
3. S. Zhang, S. Zhang, B. Wang, and T. G. Habetler, "Deep Learning Algorithms for Bearing Fault Diagnostics—A Comprehensive Review," *IEEE Access*, vol. 8, pp. 29857–29881, 2020.
4. S. B. Lee, G. C. Stone, J. Antonino-Daviu, K. N. Gyftakis, E. G. Strangas, P. Maussion, and C. A. Platero, "Condition monitoring of industrial electric machines: State of the art and future challenges," *IEEE Industrial Electronics Magazine*, vol. 14, no. 4, pp. 158–167, 2020.
5. M. Kande, A. J. Isaksson, R. Thottappillil, and N. Taylor, "Rotating electrical machine condition monitoring automation—A review," *Machines*, vol. 5, no. 4, p. 24, 2017.
6. A. Rassõlkin, V. Rjabtšikov, T. Vaimann, A. Kallaste, V. Kuts, and A. Partyshev, "Digital twin of an electrical motor based on empirical performance model," in *2020 XI International Conference on Electrical Power Drive Systems (ICEPDS)*, pp. 1–4, IEEE, 2020.
7. Zhao, Jian and Chevalier, Fanny and Balakrishnan, Ravin, "Kronominer: using multi-foci navigation for the visual exploration of time-series data," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1737–1746, 2011.
8. Ceneda, Davide and Gschwandtner, Theresia and Miksch, Silvia and Tominski, Christian, "Guided visual exploration of cyclical patterns in time-series," in *Proceedings of the IEEE Symposium on Visualization in Data Science (VDS). IEEE Computer Society*, 2018.
9. Hao, Ming C and Dayal, Umeshwar and Keim, Daniel A and Schreck, Tobias, "Multi-resolution techniques for visual exploration of large time-series data," in *EUROVIS 2007*, pp. 27–34, 2007.
10. Shi, Yang and Chen, Bingchang and Chen, Ying and Jin, Zhuochen and Xu, Ke and Jiao, Xiaohan and Gao, Tian and Cao, Nan, "Supporting guided exploratory visual analysis on time series data with reinforcement learning," *IEEE Transactions on Visualization and Computer Graphics*, 2023.
11. Andrienko, Natalia and Andrienko, Gennady and Shirato, Gota, "Episodes and Topics in Multivariate Temporal Data," in *Computer Graphics Forum*, vol. 42, p. e14926, Wiley Online Library, 2023.
12. Fujiwara, Takanori and Sakamoto, Naohisa and Nonaka, Jorji and Yamamoto, Keiji and Ma, Kwan-Liu and others, "A visual analytics framework for reviewing multivariate time-series data with dimensionality reduction," *IEEE transactions on visualization and computer graphics*, vol. 27, no. 2, pp. 1601–1611, 2020.
13. J. Eirich, G. Koutroulis, B. Mutlu, D. Jäckle, R. Kern, T. Schreck, and J. Bernard, "ManEx: The Visual Analysis of Measurements for the Assessment of Errors in Electrical Engines," *IEEE computer graphics and applications*, vol. 42, no. 2, pp. 68–80, 2022.
14. R. Silva, A. Salimi, M. Li, A. R. Freitas, F. G. Guimarães, and D. A. Lowther, "Visualization and analysis of tradeoffs in many-objective optimization: A case study on the interior permanent magnet motor design," *IEEE Transactions on Magnetics*, vol. 52, no. 3, pp. 1–4, 2015.
15. M. Friendly and D. Denis, "The early origins and development of the scatterplot," *Journal of the History of the Behavioral Sciences*, vol. 41, no. 2, pp. 103–130, 2005.
16. T. Büring, J. Gerken, and H. Reiterer, "User interaction with scatterplots on small screens-a comparative evaluation of geometric-semantic zoom and fisheye distortion," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 829–836, 2006.
17. E. A. Bier, M. C. Stone, K. Pier, W. Buxton, and T. D. DeRose, "Toolglass and magic lenses: the see-through interface," in *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pp. 73–80, 1993.
18. W. Klippel, "End-of-line testing," *Assembly Line-Theory and Practice, Prof. Waldemar Grzechca (Ed.)*, pp. 181–206, 2011.
19. G. Weidenholzer, S. Silber, G. Jungmayr, G. Bramerdorfer, H. Grabner, and W. Amrhein, "A flux-based PMSM motor model using RBF interpolation for time-stepping simulations," in *2013 International Elec-*

tric Machines & Drives Conference, pp. 1418–1423, IEEE, 2013.

20. S. Silber, W. Koppelstätter, G. Weidenholzer, G. Segon, and G. Bramerdorfer, “Reducing development time of electric machines with SyMSpace,” in *2018 8th international electric drives production conference (EDPC)*, pp. 1–5, IEEE, 2018.
21. S. Nandi, H. Toliyat, and X. Li, “Condition Monitoring and Fault Diagnosis of Electrical Motors—A Review,” *IEEE Transactions on Energy Conversion*, vol. 20, pp. 719–729, dec 2005.
22. A. G. Asuero, A. Sayago, and A. González, “The correlation coefficient: An overview,” *Critical reviews in analytical chemistry*, vol. 36, no. 1, pp. 41–59, 2006.
23. E. R. Tufte, *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press, 2 ed., 2001.
24. S. Few, *Information Dashboard Design: Displaying Data for At-a-glance Monitoring*. Analytics Press, 2013.
25. A. Cairo, *The Functional Art: An Introduction to Information Graphics and Visualization*. USA: New Riders Publishing, 1st ed., 2012.
26. J. Brooke, “Sus: a “quick and dirty’usability,” *Usability evaluation in industry*, vol. 189, no. 3, pp. 189–194, 1996.
27. J. Sauro, “Measuring usability with the system usability scale (SUS),” Feb 2011.
28. A. Joshi, S. Kale, S. Chandel, and D. K. Pal, “Likert scale: Explored and explained,” *British journal of applied science & technology*, vol. 7, no. 4, pp. 396–403, 2015.
29. J. Suschnigg, B. Mutlu, G. Koutroulis, V. Sabol, S. Thalmann, and T. Schreck, “Visual exploration of anomalies in cyclic time series data with matrix and glyph representations,” *Big Data Research*, vol. 26, p. 100251, 2021.

Patrick Louis is an assistant researcher at Pro²Future. His research interests include data visualization, interface design and anomaly detection. He received the bachelor’s degree in computer science from TU Graz. Contact him at patrick.louis@pro2future.at.

Lena Cibulski is a researcher at Fraunhofer IGD and a PhD candidate at TU Darmstadt, Germany. Her research interests are at the intersection between information visualization and multi-attribute decision-making, with a particular focus on engineering applications. She holds a master’s degree in computer science from Otto-von-Guericke University Magdeburg. Contact her at lena.cibulski@igd.fraunhofer.de.

Josef Suschnigg is a researcher at Pro²Future who received his Master’s degree in Software Engineering and Management from TU Graz. His research interests are in Visual Analytics and interactive data analytics, mainly for industrial time series data. Currently, he is a computer science PhD candidate. Contact him at josef.suschnigg@pro2future.at.

Edmund Marth is Senior Scientist at the Institute of Electrical Drives and Power Electronics at JKU Linz. While his PhD was on magnetically passively stabilized rotor systems, actual research interests include robust and sustainable design, surrogate modeling, and condition monitoring of electric machines. Contact him at edmund.marth@jku.at

Hubert Mitterhofer is the head of the Area Electric Drives at the Linz Center of Mechatronics, Austria. His research interests include electric drives, magnetic bearings, and renewable energies. He received his PhD degree in mechatronics from JKU Linz. Contact him at hubert.mitterhofer@lcm.at.

Jörn Kohlhammer is head of the competence center for information visualization and visual analytics at Fraunhofer IGD, and honorary professor for user-centered visual analytics at the TU Darmstadt. His research interests are focused on decision-centered visualization and its application in various application areas. Contact him at joern.kohlhammer@igd.fraunhofer.de.

Tobias Schreck is Professor and head of the Institute of Computer Graphics and Knowledge Visualization at TU Graz. His research interests are in Visual Analytics, Information Visualization, and Applied 3D Object Retrieval. He received his PhD degree in Computer Science from University of Konstanz, Germany. Contact him at tobias.schreck@cgv.tugraz.at.

Belgin Mutlu has a Master’s degree in Computer Engineering and a PhD in Computer Science, both from TU Graz. She is head of the research area Cognitive Decision Making at Pro²Future. Her research focuses on visual data analytics, human-computer interaction, recommender systems, personalisation, and machine learning. Contact her at belgin.mutlu@pro2future.at