Available online at www.eccomasproceedia.org
Eccomas Proceedia UNCECOMP (2025) 93-111

ECCOMAS Proceedia

UNCECOMP 2025
6th ECCOMAS Thematic Conference on
Uncertainty Quantification in Computational Science and Engineering
M. Papadrakakis, V. Papadopoulos, G. Stefanou (eds.)
Rhodes Island, Greece, 15-18 June 2025

LEVERAGING PHYSICS-INFORMED MODELS AND PRODUCTION DATA FOR QUALITY ASSESSMENT IN RESISTANCE SPOT WELDING

Miriam B. Dodt[1,2], Bouwe Verkens[1,2], Augustin Persoons[3], Matthias G.R. Faes[4], David Moens[1,2]

[1] KU Leuven, LMSD Division, Department of Mechanical Engineering
Jan Pieter de Nayerlaan 5, 2860 Sint-Katelijne-Waver, Belgium
e-mail: {miriam.dodt,bouwe.verkens,david.moens}@kuleuven.be

[2] Flanders Make @ KU Leuven, Belgium

[3] Université de Technologie Troyes
12 rue Marie Curie, CS 42060, 10004 Troyes CEDEX , France
e-mail: augustin.persoons@utt.fr

[4] TU Dortmund University, Chair for Reliability Engineering
Leonhard-Euler-Strasse 5, 44227 Dortmund, Germany
e-mail: matthias.faes@tu-dortmund.de

Abstract. *Ensuring high-quality and reliable production is a central challenge in industrial manufacturing. Within a production line, three main sources of knowledge are typically available: the general process settings, measurements taken during production, and post-process quality assessments. While quality assessments are essential for building accurate models and enabling process control, they are often costly and limited in availability. It is therefore important to choose a modelling approach that can handle limited data availability while also accounting for uncertainty and measurement noise. Some manufacturing processes are also subject to degradation, which must be properly accounted for in any predictive or control-oriented modelling approach.*
This paper investigates how modelling electrode degradation affects the quality prediction of resistance spot welds. Using an industrial dataset, various single- and multi-fidelity approaches, with and without an explicit drift model were developed and compared to each other. Special focus was placed on whether explicitly modelling process drift improves predictive performance. Results show that multi-fidelity models outperform single-fidelity ones, especially when degradation is explicitly modelled. Including a separate degradation model leads to a significantly better fit than models that rely only on electrode metadata or ignore drift altogether.

Keywords: Multi-fidelit Modelling, Resistance Spot Welding, Process Monitoring, Degradation Model, Surrogate Modelling.

ISSN:2623-3339
Peer-review under responsibility of the organizing committee of UNCECOMP 2025.
doi: 10.7712/120225.12343.21505

1 INTRODUCTION

Resistance spot welding is a crucial production process in the manufacturing of cars and batteries [1], which is strongly influence by both random variation and the gradual degradation of consumables, such as electrode caps. [2] To iterate, a modern vehicle consists of 2000 to 5000 individual spot welds, with a single production line producing more than seven million welds a day [1]. Reliable process control and monitoring require careful consideration of subtle changes in material properties, electrode conditions, environmental factors, and the wear state of consumables. This paper explores the capabilities of utilising a degradation model on the predictability of an output quantity. In this context, we develop and compare data-driven single- and multi-fidelit models using an industrial dataset. It is further explored, how using a multi-fidelit model improves predictability by combining different types of information. Multi-fidelit modelling has already been explored in related techniques, such as laser-bed welding [3, 4] but not yet in resistance spot welding. Multi-fidelit techniques as well as digital twin approaches, relevant for reliable process control, have been frequently tested on analytical cases or small-scale experimental set-ups. [5, 6] Evaluations on industrial datasets from a running production line, like in this paper, are seldom conducted. Information considered are the weld settings, which give a general expected weld quality, as well as measurements during the process. These measurements are cheap to acquire and provide information about each individual weld. To that front, the current and voltage is measured while welding, which allows to compute the dynamic resistance, one of the key parameters of resistance welding. In an industrial context, a weld is required to fulfi various quality criterion, such as a sufficient nugget diameter and indentation as well as passing visual inspection. Further detail about the criteria are found in the next section. In this paper, the indentation or rather is inverse, the remaining sheet thickness, are used as a quality assessor.

Additional to the multi-fidelit approach, a separate model is used to model the degradation. This allows us to improve our predictability and also estimate the effect of the degradation already before welding.

1.1 Resistance spot welding

During the resistance spot welding process, two to three thin metal sheets (often steel or aluminium) are joined through a localised spot. As a firs step, the sheets are pressed together by two electrodes on opposing ends with a large force (range in kilonewton) during which a high electrical current (range in kiloampere) is sent through for a short amount of time (tens to hundreds of milliseconds). The thermal energy Q is generated following the equation of Joule heating $Q = I^2Rt$, with current I, resistance R and current fl w duration t. [7] As heat is generated proportional to the local resistance, the resistance should be high where the weld should be formed and as small as possible elsewhere. Fig. 1 shows a schematic of the different local resistances, that affect the heat generation. [8] The total resistance is split in two types, bulk resistance of the metal R_m and contact resistance between two surfaces R_c. To minimise energy loss, the contact resistance between the two sheets $R_{c,3}$ should be by far the largest, whereas the resistance between the electrodes and the metal sheets $R_{c,1}$ and $R_{c,2}$ should be as small as possible. Copper, which has good thermal and electrical conductivity, is therefore usually chosen as the electrode material.

When sufficient heat is generated between the sheets, the material begins to melt, forming a pool of molten material. After re-solidificatio this pool creates the welding nugget. During

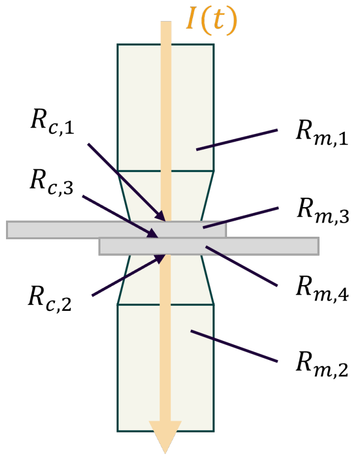


Figure 1: Schematic of contact resistance $R_{c,1-3}$ and material resistance $R_{m,1-4}$ of a two-sheet RSW-process. Also visible are the electrodes, metal sheets and current fl w.

welding, both the current $I(t)$ and voltage $V(t)$ fl w can be measured, allowing calculating the dynamic resistance $R(t)$ using Ohm's law $R(t) = V(t)/I(t)$. Fig. 2 [8] shows a typical resistance curve observed during the welding process. In phase 1, the sheets are pressed together by the electrodes to establish optimal contact. The resistance in that phase is concentrated at the micro-contacts and is initially high due to debris, dirt and oil on the workpieces surface acting as an insulator. The sharp drop in resistance occurs as the current begins to burn off these contaminants. In phase 2, most of the contaminants are removed, establishing metal-to-metal contact. As temperature rises, the material softens, increasing the contact area. While higher temperatures tend to increase resistance, the growing contact area has the opposite effect. In this phase, the effect of increasing contact area dominates, causing a further reduction of resistance. In phase 3, the resistance due to increased temperature becomes dominant. The temperature continues to raise until the onset of firs melting near the end of this phase. In phase 4, melting progresses, expanding the molten zone, which increases the cross-sectional area, reducing the resistance. The peak resistance R_{max} corresponds to the point where temperature stabilises and the growth of the weld nugget becomes the predominant cause of resistance. [8] When current fl w is continued after this point, the nugget will continue growing, until the surrounding solid material, under the large compressive electrode force, is insufficient to contain the molten material, it is being spat out, which is called expulsion. This expulsion can also be observed in the resistance curve by a sharp drop. [9]
To conclude, the dynamic resistance curve is a key feature of the welding process and its maximal value R_{max} and the respective time $t_{R_{max}}$ are two key parameters.

Fig. 3 shows the cross-section of a welding nugget, formed between two sheets with thickness t_1 and t_2 respectively. The nugget diameter w_n is visible, as well as the indentations $t_{i,1}$ and $t_{i,2}$ left by the electrode cap on both sides of the sheet. The weld quality is determined by its yield strength, which can only be tested destructively. It however directly corresponds to the size of the welding nugget w_n, which can also be estimated in a non-intrusive manner. This parameter alone is however not sufficient alone to determine the quality. Other factors such as the visual performance or indentation are relevant also. Visual performance refers to welds that appear "burnt" or show other visible defects. These defects do not directly affect the weld's

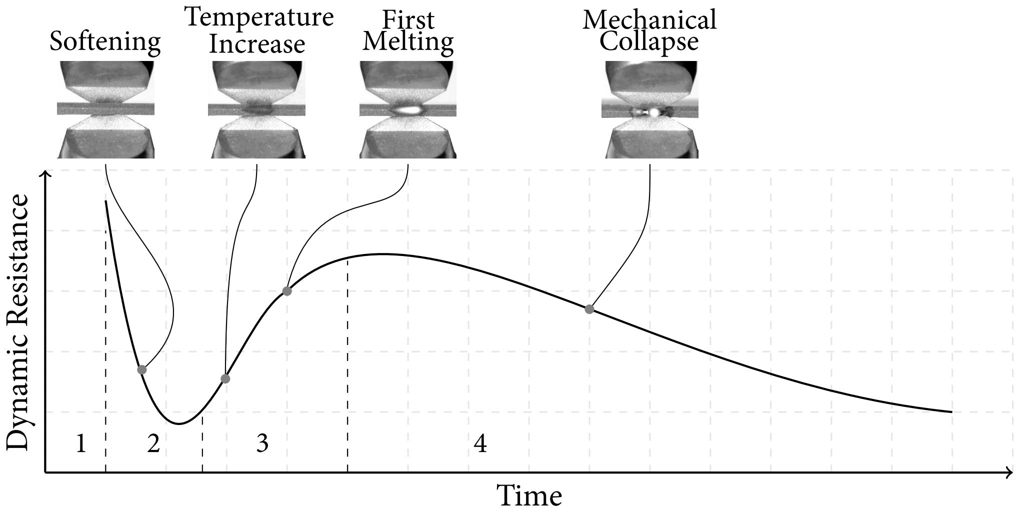


Figure 2: Theoretical dynamic resistance curve interpretationand characterisation. The images on top were created with a Phantom VEO 640 ultra high speed camera and illustrate the evolution of the weld nugget at the given phase. [8]

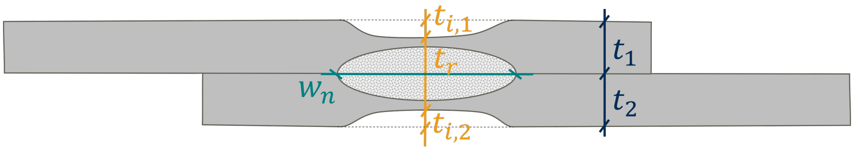


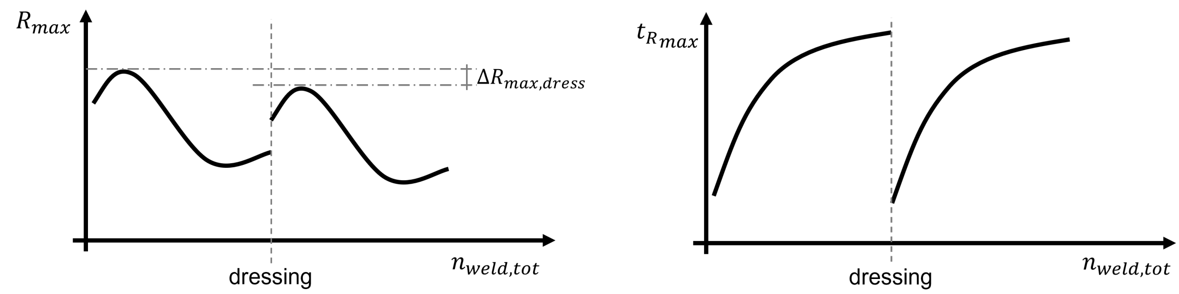
Figure 3: Schematic of the cross section of a weld. Indicated are sheet thickness ($t = t_1 + t_2 = t_{i,1} + t_r + t_{i,2}$), indentation ($t_i = t_{i,1} + t_{i,2}$) and remaining sheet thickness t_r as well as the welding nugget with width w_n.

strength but rather the surface finis and are visually unappealing to the costumer. During welding, the sheets melt and their contact point and the surrounding material softens. As the sheets are constantly pressed together with large force, the material deforms and the electrode leaves an indentation after cooling. The indentation is an indicator for the amount of material molten during welding and the corresponding yield strength. If the indentation is too small ($< 2-3\%$), there is the danger of a stick weld: the two sheets are not unified rather just glued together, the weld's strength will not be sufficient. If the indentation is too large ($> 60\%$ of the thinnest sheet), the strength of the boundary between weld and sheet is low and can tear easily. The listed values are based on the standards DVS 2960 [10] and EN ISO 6520 [11] and contain a safety factor.

While the electrodes are manufactured to reduce their resistance as much as possible, heat generation inside the electrodes and especially at their contact to the metal sheets cannot be avoided completely. Over time, this leads to change of the electrode shape and thus the contact area and contact resistance. Additionally, dirt, oil and dust on the metal sheet can adhered to the electrode, further increasing the resistance and heat generation at the wrong contact surface. To decrease costs and increase electrode lifetime, the electrode is equipped with an electrode cap, made out of the same material. Instead of exchanging the full electrode after degradation, it is sufficient to exchange only the cap. Additionally, it is possible to dress the electrode a number of times before exchanging it. This entails taking away a small layer of material to restore the shape and surface quality.

1.2 Effect of consumables on the process

The degradation of the surface quality and the change of the contact area affect the dynamic resistance observed during the welding process. Based on expert knowledge and physical considerations, one can expect the maximal resistance R_{max} to increase the more often an electrode cap has welded. The time at which R_{max} is observed, $t_{R_{max}}$, also increases with each additional weld. The dynamic resistance describes the total resistance, including all seven resistances shown in fig 1. As mentioned before, the dressing resets the surface finis and shape of the electrode, therefore resetting $t_{R_{max}}$. However, the thickness of the electrode cap reduces each time material is taken away. Although slowly, this introduces an additional facette of the drift, leading to a reduction in maximal observed resistance in each new dress cycle. In other words, while the time the maximal resistance is observed returns to its initial value after dressing, the maximal resistance is lower than in the previous dress cycle. Fig. 4 shows a schematic of the expected drift behaviour.



(a) Maximal observed dynamic resistance.

(b) Time of maximal resistance.

Figure 4: Schematic of the effect of electrode degradation on the observed values of R_{max} (a) and $t_{R_{max}}$ (b). Includes degradation within one dress cycle and between dresses. $n_{weld,tot}$ is the number of total welds performed.

It is important to note that welding is a highly sensitive and noisy process, influence by the individual characteristics of each weld, local material variations, and other factors. Additionally, process control is not perfectly repeatable, and measurement errors in current and voltage are to be expected. Since dynamic resistance is computed from these measurements, it is inherently subject to the same uncertainties and variations.
Another important consideration is the condition of the electrode. Each weld affects the electrode's state, and electrode degradation is influence not only by the weld history but also by the specifi machine in use. As a result, the drift behaviour can vary significantl between different electrodes.

1.3 Motivation

To summarise, resistance spot welding is a fast manufacturing process, which is affected by large uncertainties as well as a periodic degradation. In this paper, it is explored how to model such a process to enable reliable and robust predictions for use in process control, monitoring, and, eventually, in building a digital twin. One important aspect to consider is, that the measured output quantity is subject to measurement error and is therefore noisy. The chosen modelling approach needs to take this into account. In cases where the noise is unbiased, regression methods are well suited to filte out the noise.

To address the challenges of noise, reliability and fast evaluation, grey-box and multi-fidelit modelling are used. Grey-box modelling unifie fast-to-evaluate data-based regression models, the so-called black-box models, as well as physics-informed white-box models. [12, 13] White-box models are usually numerical or analytical models, which makes them slow to evaluate for complex systems but more accurate in areas without sufficient data. [13] For processes with large variability, they are often not sufficiently accurate, because it is difficult to adjust them to each process individually due to lack of knowledge and cost restraints. To mitigate this, a multi-fidelit approach is used, which consists of a number of low- and high-fidelit models combined through a model management strategy. [14] Low-fidelit refers to models, that are less accurate, e.g. due to their simplicity, lack of complexity but also non-converged models. This makes them more cost efficient. High-fidelit models have a good accuracy but also significan costs attached. Combining both types allows for a resulting model with larger accuracy while limiting the costs. [6]
In our case, the low-fidelit model uses weld settings to predict the expected output, specificall the remaining sheet thickness t_r. However, it lacks any information about individual welds and real-time process variations, so its predictions can only reflec the average expectation. The high-fidelit model on the other hand incorporates process data, such as the dynamic resistance, voltage and current curve, which gives access to the actual conditions of individual welds. This makes the underlying model more precise and reliable but also requires a black-box model instead of a numerical model. To train this black-box model, measurements of the remaining sheet thickness are required, which are costly and resource-intensive to collect in large quantities, as pointed out earlier.

Traditionally, a parallel approach is used when designing a model management strategy, meaning that at each step, it is decided anew, which model to evaluate. [14] High-fidelit evaluations are only performed when necessary and seen as the ground truth. The low-fidelit model gives a general fit which is then enhanced by the more reliable data. In the case of a noisy high-fidelit model, special caution is necessary to not overfi the model. Alternatively, a nested multi-fidelit scheme can be utilised. Giannoukou et al. [15] have shown, that it is very well suited to address even large noise floors allowing to significantl reduce the number of necessary high-fidelit data points without curbing the accuracy. The idea behind is, that the low-fidelit model captures the general relationship between inputs and outputs and the high-fidelit model can then be used to focus on the discrepancy between the latter's prediction and the observed output. This simplifie the relationship the high-fidelit model needs to capture accurately, leading to better predictions. Furthermore, the approach is simple and allows for any number of sub-models, as long as there is a clear hierarchy. It is not necessary for the sub-models to be of the same type nor be identical in their inputs. The next section gives a more detailed introduction to the scheme used.

Additional to the multi-fidelit model, a degradation model is trained based on observed change in the resistance curve. This degradation model is additionally included in the low-fidelit model, as it is possible to evaluate it before welding. Section 3.1 details how it was built.

1.4 General setup of a nested multi-fidelity model

As shown in previous work [16], this approach is well suited for process control of fast production processes and is able to deal well with noisy high-fidelit data [15, 16]. Here, as the name suggests, the next lower-fidelit (LF) model $\mathcal{M}_{LF}$ is included in the high-fidelit (HF) model $\mathcal{M}_{HF}$. The high-fidelit model $\mathcal{M}_{HF}$ can then be expressed as the sum of the scaled low-fidelit model $\mathcal{M}_{LF}$ and a discrepancy term δ

$$\mathcal{M}_{HF} = \rho \cdot \mathcal{M}_{LF} + \delta. \tag{1}$$

The constant scaling coefficient ρ can be estimated through

$$\hat{\rho} = \mathrm{E}\left[\frac{\mathcal{Y}_{HF}(\mathcal{X}_{HF})}{\mathcal{M}_{LF}(\mathcal{X}_{HF})}\right] \tag{2}$$

the expected scale of the high-fidelit output $\mathcal{Y}_{HF}$ to the low-fidelit model, evaluated at the high-fidelit input $\mathcal{X}_{HF}$ [15]. The estimation is expressed through the hat-accent. To compute the discrepancy $\hat{\delta}$, eq. 1 can be rearranged to

$$\hat{\delta} = \mathcal{Y}_{HF}(\mathcal{X}_{HF}) - \hat{\rho} \cdot \mathcal{M}_{LF}(\mathcal{X}_{HF}). \tag{3}$$

In the next step, a surrogate model $\hat{\mathcal{M}}_{\hat{\delta}}$ is trained on $\hat{\delta}$. The fina output prediction $\hat{Y}$ is computed following

$$\hat{Y} = \hat{\mathcal{M}}_{\hat{\delta}}(x) + \hat{\rho}\mathcal{M}_{LF}(x) \tag{4}$$

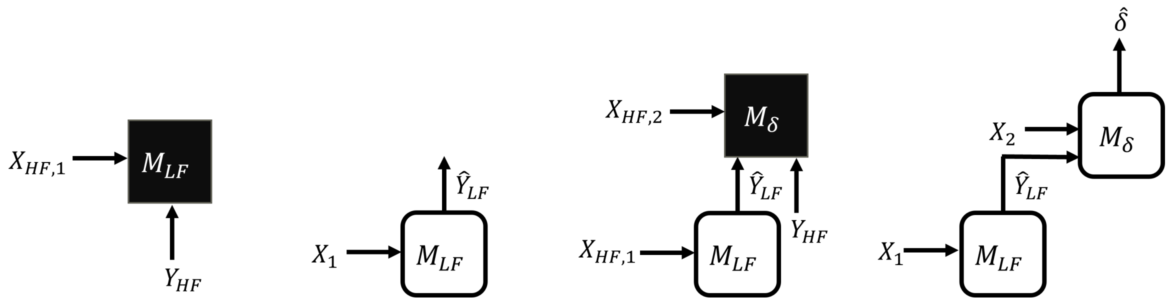
for any input x. Fig. 5 illustrates the described procedure of building a multi-fidelit model.

The black squares describe the model, that is being trained. Arrows pointing from the left into the squares indicate the inputs used for training and evaluation. For example in fig 5a, a regression model is trained on $X_{HF,1}$ and Y_{HF}. The arrows pointing to the right inside the box indicate input data, the arrows pointing upwards into the model indicate output data relevant for training only. Fig. 5b shows the finishe LF-model with X_1 as input and $\hat{Y}_{LF}$ as its output, indicated by the arrows pointing outwards the box.

The operator is free to choose the architecture and information richness for the models. In this case, $X_{HF,1}$ and $X_{HF,2}$ both correspond to the high-fidelit observations Y_{HF} but contain different information. The operator can also decide to use the same input but change the model architecture, depending on their problem. In this paper, the goal is to explore the predictability before welding, so without process information; $X_{HF,1}$ therefore contains only information known before performing any welds, while $X_{HF,2}$ consists of observations unique to each weld. Further, each model is shown as a data-based regression model. Alternatively it is possible to build a surrogate model as the low-fidelit model (fig 5a) on the inputs and outputs of a corresponding numerical model or use the predictions of the numerical model directly as the LF-model.

1.5 Overview

In this paper, the drift behaviour is studied on an industrial spot welding dataset. A degradation model is developed, which is then included in prediction models regarding the weld's indentation estimation. A single fidelit model is being compared with a nested multi-fidelit model,



(a) Training of $\mathcal{M}_{LF}$ (b) Trained $\mathcal{M}_{LF}$ (c) Training of $\mathcal{M}_{mf}$ (d) Trained $\mathcal{M}_{mf}$

Figure 5: The nested multi-fidelit model contains a low-fidelit model (b), which is trained (a) on the relation between $\mathcal{S}$ and Q. Using its predicted output $\hat{Q}_{LF}$, a discrepancy model can be trained (c). Evaluating it (d) on $\mathcal{S}$ and $\mathcal{P}$ gives the estimated output discrepancy $\hat{\delta}$ between $\hat{Q}_{LF}$ and Q.

whose low-fidelit model only takes information known before welding and whose high-fidelit model contains all the features observed during welding itself. Pimenov et al. [17] provide a comprehensive overview of recent advances in AI-based methods to model tool degradation, such as Neural Networks, Fuzzy logic and image recognition. They conclude that decision trees and ensemble methods are particularly effective, offering high predictive accuracy without the need for extensive hyperparameter tuning or large volumes of training data, as typically required by neural networks. In this study, the degradation can be analysed independently and exhibits a relatively simple structure, eliminating the need for complex models. Nevertheless, a bagging ensemble of decision trees [18] was chosen for both the low-fidelit and multi-fidelit models due to its robustness and performance. Both (sub-)models make use of the same architecture. For each model class, three variations are compared to each other: considering no drift at all, considering the electrode health/degradation state, namely the relative lifetime/total number of welds performed and the number of welds performed within the current dress cycle and lastly, considering the degradation model.
In the following section, the dataset used is described, followed by a more in-depth model overview in section 3. In section 4 and 5 the results are presented and discussed. This paper then closes with its conclusion in section 6.

2 DATASET

The dataset to study the degradation and build the models on, was kindly provided by DAF Trucks Vlaanderen from one of their production lines. Included are observations of welds performed at 8 different spots on the same part, two each are welded with the same electrode and welding robot. One observation consists of information from four main sources:

- General welding settings $\mathcal{S}$
  Such as the current profile welding duration or material type and thickness
- Electrode information $\mathcal{E}$:
  Here: relative lifetime n_{rel}, weld counter in current dress n_{dress}, total welds number n_{welds}
- Process data logged during each process $\mathcal{P}$
  Current and voltage curve, $I(t)$ and $V(t)$ respectively
- Quality assessment Q
  Here: Remaining sheet thickness (see fig 3, t_r), binary failure evaluation

Table 1: Overview over the dataset used to train the degradation and process model. The limited availability of Q constraints the dataset size.

Information source	Information on welds in dataset No. locations	No. machines	No. samples	No. features
Process settings $\mathcal{S}$	8	/	8	12
Electrode information $\mathcal{E}$				
for degradation model	2	1	934	2
for process model	8	4	140	2
Process data $\mathcal{P}$				
for degradation model	2	1	934	15
for process model	8	4	140	15
Quality assessment Q	8	4	140	1

The general settings $\mathcal{S}$ determine the expected welding result and are consistent per spot. It entails all the information necessary to perform a successful weld. Most relevant are here the current profile welding duration and thickness. Material variations between welds are considered to be small, and therefore have minor impact on the analysis. The information on the electrode caps $\mathcal{E}$ is highly relevant to determine the effect of electrode degradation and will be used to build a drift model. The process data $\mathcal{P}$ gives live-information about each individual spot and is relevant to predict the individual weld quality. During the process, the current and voltage over the time is logged and relevant features can be computed from them. Relevant features entail the mean value of current and voltage but also the dynamic resistance $R(t) = \frac{V(t)}{I(t)}$ and power $P(t) = V(t) \cdot I(t)$ curve. Especially the dynamic resistance is important to determine the welds quality, as discussed above. Additional to parameters such as mean and maximal value, principal component analysis was performed on the curves of resistance and current to detect changes in its features and gain further information from the curve.

The dataset was created over two weeks. Tab. 1 gives an overview over the utilised data. Welding settings are predefined for the eight different spots, they contain eight distinct information. Basic logging data is readily available at minimal cost, as they are automatically recorded in any industrial resistance spot welding system due to process steering. In contrast, estimating the quality of each weld is significantl more expensive, since it requires manual, non-destructive testing by an operator. Although such testing is performed regularly to ensure product quality, the number of available measurements remains limited, restricting the available amount of samples. The dataset used in this study includes a total of 140 welds with corresponding ultrasonic measurements. To study the behaviour of degradation, no quality assessment is necessary, thus more samples can be used. During testing, several failure criteria are evaluated, including the remaining sheet thickness and whether it falls within an acceptable threshold. In the case of the dataset, the remaining sheet thickness t_r is used as Q.

3 MODELLING APPROACHES

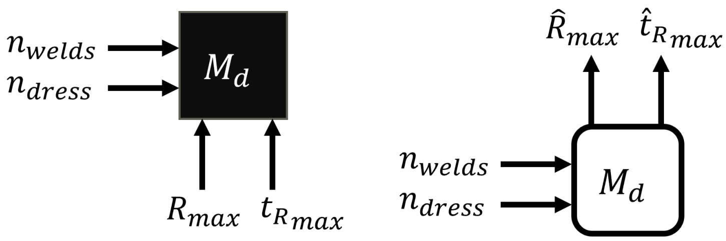
In this chapter, the various modelling approaches are presented in further detail. This paper follows two goals: studying the effect of a degradation model on predictability, especially before actually welding and studying in what way multi-fidelit modelling can improve the results. The former is relevant for process control, monitoring, knowing when to exchange/dress the electrode. The focus here lies in separating the information known before welding and the information gained during the welding process to build a reliable and robust model, that can focus on the discrepancy of individual welds from their expected line. For this, a single-fidelit model without information on neither the degradation nor the electrode state is used as the base-model. This model is enhanced once with the degradation model and once with the electrode information. The same is done for the multi-fidelit model. Tab. 2 gives a short overview over the six different models. In this chapter, the different models are explained in more detail, starting with the degradation model.

Table 2: Overview models

Name	Model Type	Information
$\mathcal{M}_{sf}$	Single-fidelit model	no information on electrode state
$\mathcal{M}_{sf,e}$	Single-fidelit model	with information on electrode state
$\mathcal{M}_{sf,d}$	Single-fidelit model	with drift model
$\mathcal{M}_{mf}$	Multi-fidelit model	no information on electrode state
$\mathcal{M}_{mf,e}$	Multi-fidelit model	with information on electrode state
$\mathcal{M}_{mf,d}$	Multi-fidelit model	with drift model

3.1 Degradation model

As discussed earlier, it is known, that the degradation of the electrode cap has an influenc on the maximal resistance and the time thereof. The idea is then, build a surrogate model based on the dynamic between information on the electrode, namely the number of welds performed with the electrode and number of welds in current dress, and the dynamic resistance curve data. The model training and evaluation is shown in fig 6. The black square of the training model (fig 6a) indicates that the model is a data-based black-box model. The arrows pointing from the left inside the box are the inputs used to train the model and the arrows pointing from the bottom show the outputs observed for training. After training (fig 6b) only the inputs are necessary to evaluate the model. Its predicted output is symbolised by arrows pointing away from the top of the square.



(a) Training (b) Evaluation

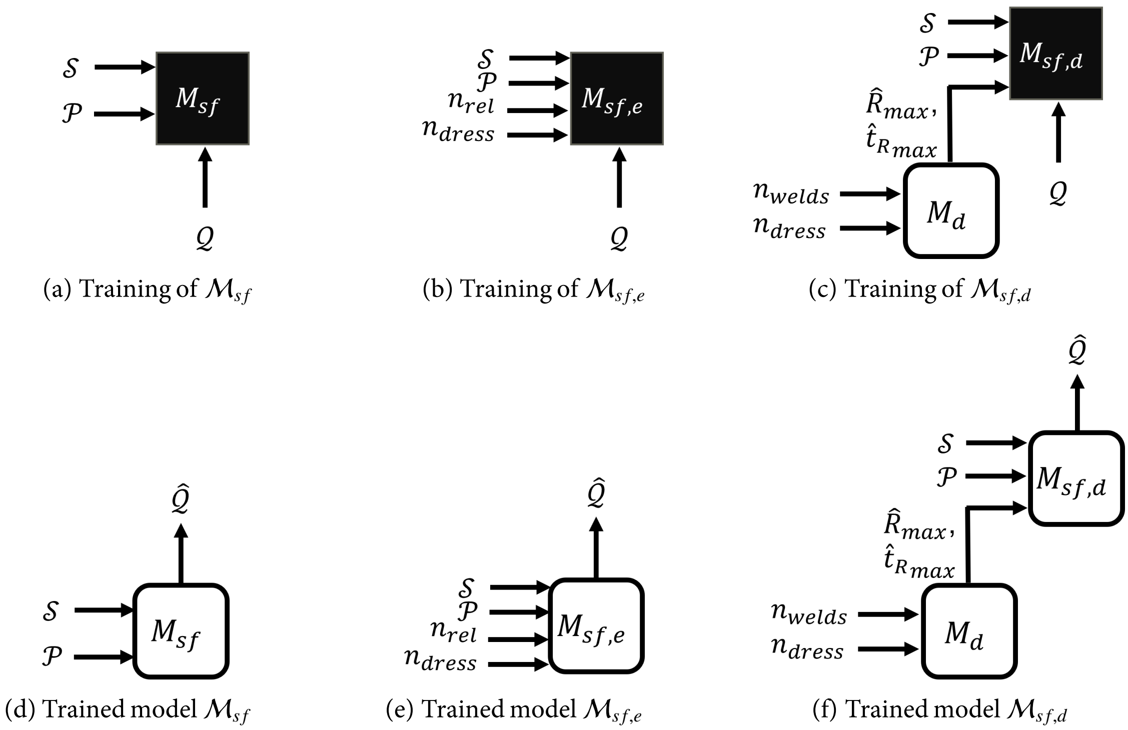
Figure 6: The degradation model is a black-box model trained (a) on the input-outputrelation of $\mathcal{E}$ (n_{welds} and n_{dress}) and R_{dyn}-key parameters (R_{max} and $t_{R_{max}}$).

The degradation between electrodes varies, some clearly show the expected behaviour, while it is not as obvious for others. To build a reliable drift model, 934 data points belonging to two spots welded with the same electrode were chosen. The drift behaviour of those two spots is the same and clearly identifiabl Therefore, only two out of the 8 spots are used for that. A gradual mostly monotonic change was observed, so a low order polynomial of degree 5 was chosen for the regression model with dress cycle counter and total weld counter as inputs and R_{max} and $t_{R_{max}}$ as outputs.

3.2 Single-fidelity model

For the single-fidelit (SF) model $\mathcal{M}_{sf}$ a bagging ensemble tree [18] was chosen with 200 learning cycles which selects the best split predictor based on the interaction-curvature-test. Here the split predictor is chosen to minimise the p-value of the chi-square tests of independence between each predictor and response as well as each pair of predictor and their response. [19]

Fig. 7 shows a schematic of the different single-fidelit models used as well their input and outputs. The base-multi-fidelit model $\mathcal{M}_{sf}$ (fig 7a and 7d), which contains no further information on the electrodes, is trained on a 27-dimensional input consisting of the welding settings $\mathcal{S}$ as well as observations during welding $\mathcal{P}$. The corresponding output for training is the measured remaining sheet thickness $\mathcal{Q}$ (fig.7a) $\mathcal{M}_{sf,e}$ additionally contains information on the relative lifetime n_{rel} of the electrode and the weld counter within the current dressing n_{dress} (fig 7b and 7e). In the degradation-informed model $\mathcal{M}_{sf,d}$, the prediction of the degradation model $\hat{R}_{max}$ and $\hat{t}_{R_{max}}$ is used as additional inputs to the single-fidelit model. An overview of the different methods and their input is provided by tab. 3.



(a) Training of $\mathcal{M}_{sf}$ (b) Training of $\mathcal{M}_{sf,e}$ (c) Training of $\mathcal{M}_{sf,d}$

(d) Trained model $\mathcal{M}_{sf}$ (e) Trained model $\mathcal{M}_{sf,e}$ (f) Trained model $\mathcal{M}_{sf,d}$

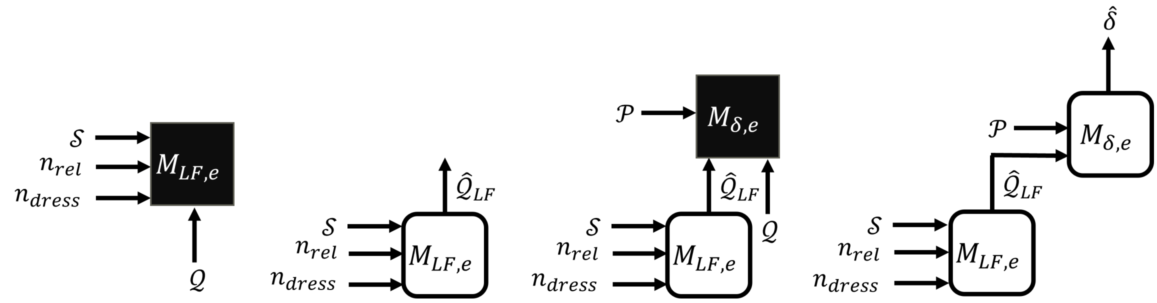
Figure 7: The single-fidelit model has $\mathcal{S}$ and $\mathcal{P}$ as inputs. It is trained on output $\mathcal{Q}$ (a), of which it gives a prediction $\hat{\mathcal{Q}}$ after training is complete (d). $\mathcal{M}_{sf,e}$ additionally takes electrode information as input (b,e) and $\mathcal{M}_{sf,d}$ the output of the degradation model $\mathcal{M}_d$ (c,f).

3.3 Nested multi-fidelity model

Additionally to the single-fidelit model, a nested multi-fidelit model is used as well. In this case, it is a random forest model, build with the same hyperparameters as the single-fidelit models and the high-fidelit ones. The low-fidelit surrogate model was trained on 140 samples using a 12-dimensional input features space $\mathcal{S}$ (excluding drift information) and a one-dimensional output measure $\mathcal{Q}$. The base model was enhanced in two manners: a) the drift model evaluation of R_{max} and $t_{R_{max}}$ or b) the information on relative lifetime and number of welds in current dressing were added to the input features space. An overview is given in tab. 3.

As the model is trained on the measured remaining sheet thickness, the scaling factor $\hat{\rho}$ can be set to 1. In the next step, the discrepancy $\hat{\delta}$ of each LF-model prediction and actual measured value is computed. This difference becomes the new output features to build the discrepancy model $\mathcal{M}_{\hat{\delta}}$ on. The 15-dimensional input-space for the high-fidelit model contains $\mathcal{P}$, the information gained during the welding process itself. The expected output can then be computed by adding the low-fidelit prediction to the discrepancy model prediction. Fig. 8 and fig 9 described how the multi-fidelit models with electrode information and degradation model respectively are build. For the base multi-fidelit model, it is referred to fig 5. It is visible, that the difference in these three models lies not in the high- but only in the low-fidelit model and its predictions. For the schematic of the degradation model $\mathcal{M}_d$, the reader is referred to fig 6.

In conclusion, the multi-fidelit model consists of a high-fidelit model, that is informed about each individual process and has access to evaluations of the low-fidelit model, which gives only the generally expected behaviour. Each of these models is build from a bagging ensemble tree model.

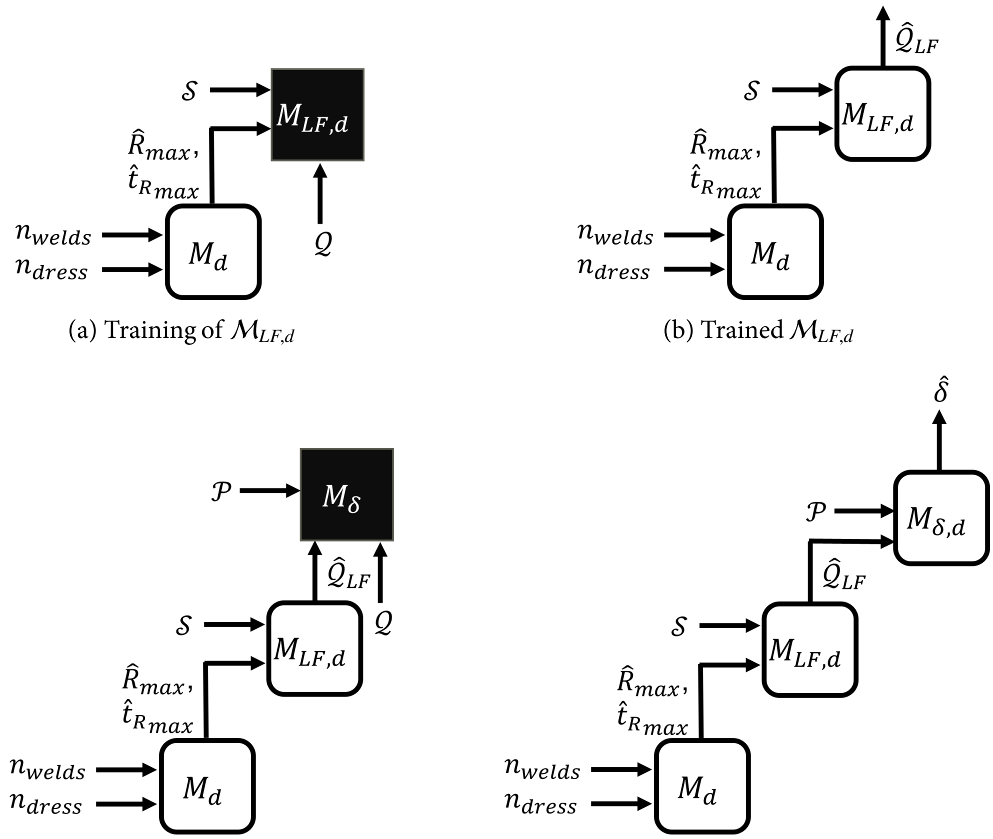


(a) Training of $\mathcal{M}_{LF,e}$ (b) Trained $\mathcal{M}_{LF,e}$ (c) Training of $\mathcal{M}_{mf,e}$ (d) Trained $\mathcal{M}_{mf,e}$

Figure 8: $\mathcal{M}_{mf,e}$ contains a low-fidelit model (b), which is trained (a) on the relation between settings $\mathcal{S}$, electrode information n_{rel}, n_{dress} and observed output $\mathcal{Q}$. Using its predicted output $\hat{\mathcal{Q}}_{LF}$, a discrepancy model can be trained (c). Evaluating it (d) on $\mathcal{S}$, $\mathcal{P}$ and $\mathcal{E}$ gives the estimated output discrepancy $\hat{\delta}$.

3.4 Overview used models

In total, six models, three single- and three multi-fidelit models (tab. 2), are compared to each other in this paper. Each of these models (apart from the degradation model) uses a random forest surrogate model with the same training hyperparameters. The main difference is, that the multi-fidelit model firs builds an expectation model based on information before the welding and refine it with measurements during the process. The single-fi elity model utilises everything at once. Further information which information on the models can be found in tab. 3.



(a) Training of $\mathcal{M}_{LF,d}$

(b) Trained $\mathcal{M}_{LF,d}$

(c) Training of $\mathcal{M}_{mf,d}$

(d) Trained $\mathcal{M}_{mf,d}$

Figure 9: $\mathcal{M}_{mf,d}$ contains a low-fidelit model (b), which is trained (a) on the relation between $\mathcal{S}$, degradation model $\mathcal{M}_d$ result and Q. Using its prediction $\hat{Q}_{LF}$, a discrepancy model can be trained (c). Evaluating it gives the estimated output discrepancy $\hat{\delta}$ between $\hat{Q}_{LF}$ and Q (d).

4 RESULTS

This section presents the key results of the different models. Table 4 summarises the R-squared values of the different models. The best performing model with this metric is, as expected, $\mathcal{M}_{mf,d}$, the multi-fidelit model, that considers the degradation model in its low-fidelit model with an R-squared value of 0.89. The other multi-fidelit models show a similar performance, all with R-squared values above 0.8. Based on the table alone, the performance of the single fidelit models appears to be comparable to all of the low-fidelit models. However, when inspecting the actual prediction in fig 10, fig 11a and fig 12a, significan differences with the model fi can be observed. These figure compare the actual measured remaining sheet thickness to the predicted values. Without noise in the process and measurements, a perfect prediction would have points aligned with the red dotted line. As they are however noisy, a scatter around the line would be the expected optimal result. $\mathcal{M}_{sf}$ in fig 10 demonstrates generally good predictive performance, except for three significantl overestimated and six underestimated samples. The discrepancy on those 9 samples appear consistently across models.

While the discrepancy model clearly improves the performance, this is not reflecte in the R-squared values of the discrepancy model. The reason lies in the outliers but might also point to a lack of information to further improve the model performance.

The degradation-uninformed low-fidelit model $\mathcal{M}_{mf,LF}$ predicts a limited number of dis-

Table 3: Overview about which information is contained in which model. A fille circle indicates that the information is included directly, an empty circle, that it is included implicitly and a dash, that it is not included.

Name	Sub-model	Has information on				Dimension	
		Settings $\mathcal{S}$	Measurements $\mathcal{P}$	Degradation model $\mathcal{M}_d$	Electrode information $\mathcal{E}$	Input	Output
$\mathcal{M}_{sf}$		●	●	-	-	27	1
$\mathcal{M}_{sf,e}$		●	●	-	●	29	1
$\mathcal{M}_{sf,d}$		●	●	●	-	29	1
$\mathcal{M}_{mf}$	LFM	●	-	-	-	12	1
	MFM	○	●	-	-	15	1
$\mathcal{M}_{mf,e}$	LFM	●	-	-	●	14	1
	MFM	○	●	-	○	15	1
$\mathcal{M}_{mf,d}$	LFM	●	-	●	-	14	1
	MFM	○	●	○	-	15	1

tinctive values, see fig 11a. This behaviour was expected, as the model exclusively considers information on the weld settings. The process is highly noisy and affected by drift but the model contains no information on neither. Additionally, there are only 8 different settings, so an average output value per setting is computed.
The updated multi-fidelit model $\mathcal{M}_{mf}$ (Fig. 11b) demonstrates a significantl improved fit with an R-squared value of 0.83, compared to 0.62 for the single-fidelit model. The outliers are still prevalent however, although slightly closer to the measured result.

The drift-informed low-fidelit model $\mathcal{M}_{mf,d,LF}$ (Fig. 12a) suggests that drift plays a crucial role in output predictions. While its visual fi appears worse than that of the single-fidelit model, it achieves similar R-squared values and slightly improves outlier predictions. The corresponding multi-fidelit model performs best out of the six fina models with an R-squared value of 0.89 and a better visual fit

Explicit figure for the other three models are omitted, as their fi closely resembles that of one of the figure shown. The single-fidelit models all perform similarly and the multi-fidelity-mode informed about the electrode state, instead of the explicit drift model, shows similar behaviour to the latter. While it is possible, to include the drift model as well as the information on the electrode state, it does not provide additional insights and slightly decreases overall performance.

Table 4: R2-values of different models compared

	Low-fidelit model	Discrepancy model	Final model
$\mathcal{M}_{sf}$	-	-	0.6150
$\mathcal{M}_{sf,e}$	-	-	0.6266
$\mathcal{M}_{sf,d}$	-	-	0.6298
$\mathcal{M}_{mf}$	0.6598	0.0800	0.8296
$\mathcal{M}_{mf,e}$	0.6462	0.1423	0.8862
$\mathcal{M}_{mf,d}$	0.6470	0.1228	0.8923

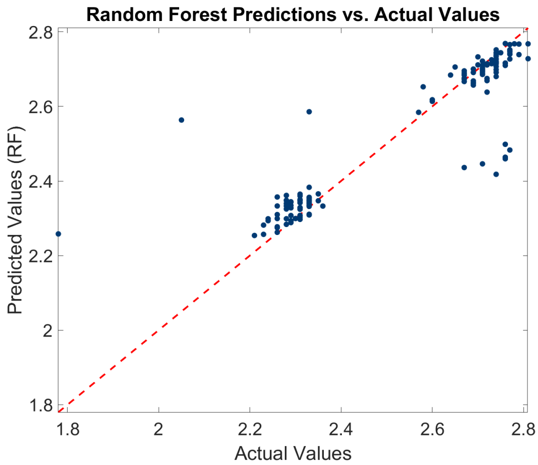
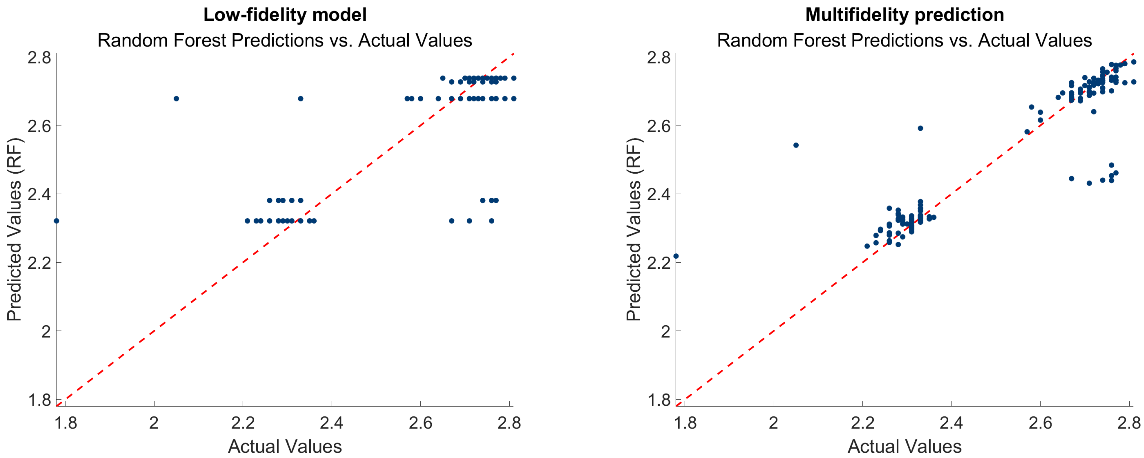


Figure 10: Remaining sheet thickness: measured vs predicted with a single-fidelit model without drift



(a) Low-fidelit model of $\mathcal{M}_{mf}$

(b) Final $\mathcal{M}_{mf}$

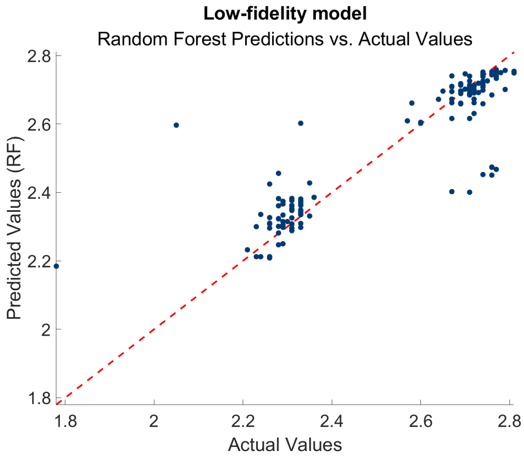
Figure 11: Remaining sheet thickness: measured vs predicted with a low-fidelit model and the fina multi-fidelit model without considering the drift model nor electrode information in the low-fidelit model.

5 DISCUSSION

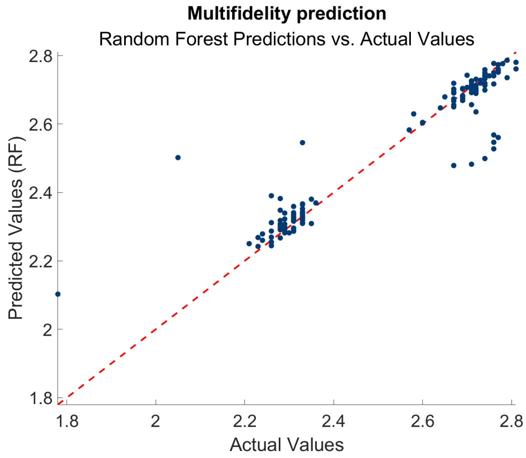
In this section, we are focussing on the meaning of the results with respect to process monitoring and reliable failure detection and prediction.

5.1 Failure detection and outliers

One aspect to evaluate additionally to the general fi of the prediction is, whether failure was correctly identified There are two failure modes with regards to the remaining sheet thickness. The firs one, is a too light indentation, which can point to a stick-weld. This means, that the two sheets are not properly joined but rather stuck together with an insufficient yield strength. To avoid this failure, the minimum indentation requirement was set to 3 % of the total thickness, based on internal guidelines and DVS 2960 [10]. There were seven welds that failed in this manner, all of which, were correctly identifie by each method. In the second failure mode the indentation is too heavy, too little material of one sheet is remaining and one can expect a shear failure, also with insufficient yield strength. In the case of the dataset, the ultrasonic device



(a) Low-fidelit model of $\mathcal{M}_{mf,d}$



(b) Final $\mathcal{M}_{mf,d}$

Figure 12: Remaining sheet thickness: measured vs predicted with a low-fidelit model and fina multi-fidelit model when considering the drift model in the low-fidelit model.

detected three failed welds. When evaluating the criterion of maximal 40 % indentation of the thinner sheet however, all welds passed. Usually the ultrasonic device uses the same threshold to detect failure. So while on paper all models were able to identify failure correctly, it still became obvious, that, especially heavy indentation, was often times not well estimated.

One possible reason for the overestimation of remaining sheet thickness are outliers, which can affect the modelling process in two distinctive ways. Firstly, outliers in the input features; when looking only at the input features that were measured during the process, 18 observations can be detected, where at least one process measurement is more than 1.5 interquartile ranges above or below the 25-75-quartile. Normally, outliers in the input features are easy to handle, especially, if, as in this case, their corresponding output values are within the expected range. Ensemble methods like random forests are especially robust against them, as they rely on averaging over multiple decision trees. This reduces the impact of a limited number of outliers. However, this filterin can also affect rare event estimation.
Secondly, the model may struggle to accurately predict samples where the output is an outlier but the input is not. This is particularly problematic when such cases are rare (i.e. in this dataset), as the model has insufficient examples to learn the mapping for these extreme cases. Since random forests rely on majority voting or averaging across trees, they inherently bias predictions toward the dominant trend in the training data. As a result, they may fail to capture rare deviations, leading to poor predictions for these output outliers. In the dataset, three such samples exist. One sample is more than 1.2 interquartile ranges below the 25-quartile and two additional samples are 0.2 interquartile ranges below. These three outliers are also identifie when looking at the scaled median absolute deviation (MAD) larger than three. Looking instead at three standard deviations from the mean, only one outlier remains. The usage of mean and median to detect outliers needs to be taken with a grain of salt, as we obviously do not have a normal but multimodal distribution here. If it is not possible to augment the dataset to include more rare events, alternative strategies such as weighted training or anomaly detection can be used to address the problem.

5.2 Conclusion for process monitoring

In sec. 4, it became clear, that multi-fidelit modelling improves predictability and model fit Moreover, including additionally the degradation model further improves the fi and has a positive effect on outliers.

The drift-informed low-fidelit model $\mathcal{M}_{sf,d}$ already results in a strong predictability, pointing out how significan R_{max} and $t_{R_{max}}$ are in determining the output quality. The drift model is therefore able to capture a large portion of the process behaviour and indicate expected quality of future welds already before welding. This holds significan value for process control by enabling inverse quantificatio to determine optimal input parameters to counter the drift. Further, optimal moments when to dress and exchange the electrode can be define as well. Improving the model will also allow to detect unexpected behaviour from which further conclusions can be drawn. For actual process monitoring however, it is recommended to still include the live-process parameters, as they are cheap to include and hold additional valuable information.

While these results are already very promising and interesting, it needs to be pointed out that the degradation model was trained not including every spot available and that the degradation behaviour can vary significantl between different electrodes and robots. The observed drift behaviour largely depends on the electrode, the welding history and appropriate welding settings. So there is no one-model-fits-al and one should build their own drift model and check, where it is applicable.

6 CONCLUSION

To enable reliable process control and good quality estimation in industrial manufacturing processes, a model is necessary, that can handle limited quality assessment availability, noisy as well as degradation within the productionline. This paper investigates how modelling electrode degradation affects the quality prediction of resistance spot welds. To that end, various black-box and multi-fidelit models, with and without an explicit drift model were compared with each other. The results can be summarised in three key findings

- The multi-fidelit model outperforms the single-fidelit model.

- Models including a separate degradation model show better fi than not considering process drift or only considering the electrode information directly.

- The fi of the low-fidelit model including the degradation model $\mathcal{M}_{mf,d,LF}$ comparable to single-fidelit models and multi-fidelit models.

The degradation-informed multi-fidelit model $\mathcal{M}_{mf,d}$ and its corresponding low-fidelit model $\mathcal{M}_{mf,d,LF}$ both serve two important but different tasks. The multi-fidelit model is relevant for process monitoring, it can be used to detect abnormal behaviour and to predict each weld result more accurately. The low-fidelit model on the other hand can be used for general process steering, decisions about the electrode dressing and exchange, as it provides already many insights before the process. As such, it will need special consideration with respect to the available measurements. In this case, it can prove beneficial to include a numerical modelling scheme or perform separate experiments as well.

In the future, inverse quantificatio schemes or model updating can be applied to determine not only optimal moments to dress and exchange the electrode but also optimal control parameters to counter the effect of the drift. The latter is a nested optimisation scheme, as changing the welding parameters obviously also affects not only the maximal resistance and time thereof, but also the speed of degradation.

ACKNOWLEDGEMENTS

The authors would like to thank DAF Trucks Vlaanderen for the opportunity to work with their data and give special thanks to Gerry Volders for his support in working with and understanding the dataset as well as his expert knowledge on not only the welding process in general but also about the specific at DAF.

REFERENCES

[1] K. Zhou and P. Yao, "Overview of recent advances of process analysis and quality control in resistance spot welding," *Mechanical Systems and Signal Processing*, vol. 124, pp. 170–198, Jun. 2019.

[2] C. Mathiszik, D. Köberlin, S. Heilmann, J. Zschetzsche, and U. Füssel, "General Approach for Inline Electrode Wear Monitoring at Resistance Spot Welding", *Processes*, vol. 9, Apr. 2021, doi: 10.3390/pr9040685.

[3] A. Olleak and Z. Xi, "Calibration and Validation Framework for Selective Laser Melting Process Based on Multi-Fidelity Models and Limited Experiment Data", *Journal of Mechanical Design*, vol. 142, no. 081701, Feb. 2020, doi: 10.1115/1.4045744.

[4] P. Nath, M. Sato, P. Karve, and S. Mahadevan, "Multi-fidelit Modeling for Uncertainty Quantificatio in Laser Powder Bed Fusion Additive Manufacturing", *Integr Mater Manuf Innov*, vol. 11, no. 2, pp. 256–275, Jun. 2022, doi: 10.1007/s40192-022-00260-9.

[5] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihn, "Digital Twin in manufacturing: A categorical literature review and classification" *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018, doi: 10.1016/j.ifacol.2018.08.474.

[6] M. G. Fernández-Godino, "Review of multi-fidelit models", *ACSE*, vol. 1, no. 4, pp. 351–400, 2023, doi: 10.3934/acse.2023015.

[7] A. O'Brien and American Welding Society, Eds., *Welding handbook". Vol. 3: Welding processes, part 2* / Annette O'Brien, ed, 9. ed., vol. 3. Miami, Fla: *American Welding Soc*, 2007.

[8] L. Bogaerts, A. Dejans, M. G. R. Faes, and D. Moens, "A machine learning approach for efficient and robust resistance spot welding monitoring", *Weld World*, vol. 67, no. 8, pp. 1923–1935, Aug. 2023, doi: 10.1007/s40194-023-01519-1.

[9] Q. Fan, G. Xu, and X. Gu, "Expulsion characterization of stainless steel resistance spot welding based on dynamic resistance signal", *Journal of Materials Processing Technology*, vol. 236, pp. 235–240, Oct. 2016, doi: 10.1016/j.jmatprotec.2016.05.026.

[10] ”Unregelmäßigkeiten bei Widerstandspunktschweißverbindungen an Stahlwerkstoffen”, *DVS 2960*, Aug 2015

[11] ”Welding and allied processes — Classificatio of geometric imperfections in metallic materials”, EN ISO 6520-2:2013, 2013.

[12] D.J. Pitchforth, T.J. Rogers, U.T. Tygesen, and E.J. Cross. ”Grey-box models for wave loading prediction”. *Mechanical Systems and Signal Processing*, 159, October 2021.

[13] H. J. A. F. Tulleken, ”Grey-box modelling and identificatio using physical knowledge and bayesian techniques”, *Automatica*, vol. 29, no. 2, pp. 285–308, Mar. 1993, doi: 10.1016/0005-1098(93)90124-C.

[14] B. Peherstorfer, K. Willcox, and M. Gunzburger, “Survey of multi-fidelit Methods in Uncertainty Propagation, Inference, and Optimization,” *SIAM Review*, vol. 60, no. 3, pp. 550–591, Jan. 2018.

[15] K. Giannoukou, S. Marelli, and B. Sudret, ‘Uncertainty-aware multi-fidelit surrogate modeling with noisy data’, Oct. 29, 2024, doi: 10.48550/arXiv.2401.06447.

[16] M.B. Dodt, L. Bogaerts, A. Persoons, M.G.R. Faes, D. Moens, D. ”A multifidelit framework for grey-box modelling of automated welding processes.” *Proceedings of ISMA2024 International Conference on Noise and Vibration Engineering USD2024 International Conference on Uncertainty in Structural Dynamics*, 4142-4150. Leuven, Belgium, Sep 9-11 2024. ISBN: 978-90-82893-17-5

[17] D. Y. Pimenov, A. Bustillo, S. Wojciechowski, V. S. Sharma, M. K. Gupta, and M. Kuntoğlu, ”Artificia intelligence systems for tool condition monitoring in machining: analysis and critical review”, *J Intell Manuf*, vol. 34, no. 5, pp. 2079–2121, Jun. 2023, doi: 10.1007/s10845-022-01923-2.

[18] L. Breiman, ”Bagging predictors”. *Machine Learning*. 24 (2): 123–140. 1996. doi:10.1007/BF00058655.

[19] W.Y. Loh, “Regression Trees with Unbiased Variable Selection and Interaction Detection.” *Statistica Sinica*, Vol. 12, pp. 361–386, 2002.