



# Improving the surface quality of friction stir welds using reinforcement learning and Bayesian optimization

R. Hartl<sup>1</sup> · J. Hansjakob<sup>1</sup> · M. F. Zaeh<sup>1</sup>

Received: 28 February 2020 / Accepted: 30 June 2020 / Published online: 21 September 2020  
© The Author(s) 2020

## Abstract

Friction stir welding is an advanced joining technology that is particularly suitable for aluminum alloys. Various studies have shown a significant dependence of the welding quality on the welding speed and the rotational speed of the tool. Frequently, an inappropriate setting of these parameters can be detected through an examination of the resulting surface defects, such as increased flash formation or surface galling. In this work, two different learning-based algorithms were applied to improve the surface topography of friction stir welds. For this purpose, the surface topographies of 262 welds, which were performed as part of ten studies, were evaluated offline. The aim was to use reinforcement learning and Bayesian optimization approaches to determine the most appropriate settings for the welding speed and the rotational speed of the tool. The optimization problem was solved using reinforcement learning, specifically value iteration. However, the value iteration algorithm was not efficient, since all actions and states had to be iterated over, i.e., each possible parameter combination had to be evaluated, to find the best policy. Instead, it was better to solve the optimization problem directly using the Bayesian optimization. Two approaches were applied: both an approach in which the information from the other studies was not used and an approach in which the information from the other studies was used. On average, both the Bayesian optimization approaches found suitable welding parameters significantly faster than a random search algorithm, and the latter approach improved the result even further compared with the former approach. Future research will aim to show that optimization of the surface topography also leads to an increase in the ultimate tensile strength.

**Keywords** Artificial intelligence · Friction stir welding · Reinforcement learning · Bayesian optimization · Surface quality

## 1 Introduction

In friction stir welding (FSW), the mechanical properties [1] as well as the surface topography [2] are strongly affected by process parameters such as the welding speed  $v_s$  and the tool rotational speed  $n$  (r/min rate). These parameters are typically determined by trial and error, based on handbook values, and by manufacturers' recommendations [3]. This selection may neither yield optimal nor near-optimal welding performance. Furthermore, it may cause additional energy and material consumption and may also result in low-quality welds [3]. For this reason, several algorithms have already been developed to

optimize the process parameters in friction stir welding. Some of these are presented in the following section.

### 1.1 State of the art—Use of optimization algorithms in the field of FSW

Various statistical and mathematical methods have been used to investigate the influence of process parameters on mechanical properties, in particular the ultimate tensile strength, and subsequently optimize the mechanical properties [4]. In many of these investigations, either the robust parameter design (RPD) method [5] or the response surface methodology (RSM) [6] was applied:

The RPD method focuses on choosing levels of parameters in a process to ensure that the mean of the output response is at a desired target and to ensure that the variability around the target value is as small as possible [5]. Taguchi [7] proposed an approach to solve the RPD problem based on designed experiments and novel methods for analyzing the resulting

✉ R. Hartl  
roman.hartl@iwb.tum.de

<sup>1</sup> Institute for Machine Tools and Industrial Management, Technical University of Munich, Boltzmannstrasse 15, 85748 Garching, Germany

data [5]. He also simplified the use of orthogonal arrays [8]. An approach that has already been applied to FSW several times is the L9 orthogonal array. This method aims at understanding the influence of four independent factors with three steps each. With the L9 method, only nine experiments have to be performed in order to study four variables at three levels. So this design reduces 81 ( $3^4$ ) configurations to nine experimental evaluations [8].

Lakshminarayanan et al. [9] determined the optimum settings for the rotational speed  $n$ , the welding speed  $v_s$ , and the axial force  $F_z$  at FSW by adapting the Taguchi L9 orthogonal array method and maximizing the signal-to-noise (S/N) ratio. In the Taguchi method, the S/N ratio is used to determine the deviation of the quality characteristics from the desired value [9]. In order to investigate nonlinearities, each of the three process parameters was varied in three levels. Welding experiments were conducted for only nine out of the 27 possible parameter combinations. For each of the nine applied parameter combinations, three tensile tests were performed, and the mean of the ultimate tensile strength was calculated. Based on the mean values for the S/N ratio and the ultimate tensile strength, an ideal parameter set was determined. The expected ultimate tensile strength  $UTS_{exp}$ , when using this ideal parameter set, was calculated with the following formula [9]:

$$UTS_{exp} = \overline{UTS}_{n,L} + \overline{UTS}_{v_s,L} + \overline{UTS}_{F_z,L} - 2 \cdot \overline{UTS} \quad (1)$$

whereby  $\overline{UTS}_{n,L}$ ,  $\overline{UTS}_{v_s,L}$ , and  $\overline{UTS}_{F_z,L}$  are the mean ultimate tensile strengths at level  $L$  of the corresponding process parameters  $n$ ,  $v_s$ , and  $F_z$ , and  $\overline{UTS}$  is the overall mean of all 27 determined ultimate tensile strengths. Subsequently, the expected maximum ultimate tensile strength  $UTS_{exp}$  was compared with the actual ultimate tensile strength obtained by adjusting the previously determined ideal parameter set, and the deviation was 2.6%. It was also determined that the rotational speed  $n$  had an influence on the tensile strength of 41%, the welding speed  $v_s$  of 33%, and the axial force  $F_z$  of 21%. The remaining 5% were referred to as errors. Ugender et al. [10] also used the Taguchi technique and the S/N ratio to find an optimum setting for the ratio of the diameter of the shoulder  $D_s$  to the diameter of the probe  $d_p$ , the tilt angle, and the welding speed. The results showed that the  $D_s/d_p$  ratio and the welding speed are the most important factors, followed by the tilt angle, when deciding on the mechanical properties of friction stir welds of aluminum alloys. Ganapathy et al. [11], Abbas et al. [12], and Ma et al. [13] also adopted Taguchi's L9 orthogonal array design and maximized the S/N ratio to optimize FSW process parameters. Vijayan et al. [14] investigated an approach using the Taguchi-based grey relational analysis (GRA) [15] instead of the S/N ratio.

The RSM is an approach to solve the RPD problem that not only allows the use of Taguchi's robust design concept but also provides a more sound and more efficient approach to

experiment design and analysis [5]. Furthermore, the RSM is a collection of mathematical and statistical techniques for analyzing problems in which several independent variables influence a dependent variable and the goal is to optimize the dependent variable [16]. Rajakumar et al. [3] applied the RSM and established an empirical relationship between the independent variables (tool rotational speed, welding speed, axial force, shoulder diameter, probe diameter, and tool material hardness) and the dependent variable, which was the ultimate tensile strength of the joint. For this purpose, a multiple regression model was developed for the ultimate tensile strength of the weld. The model was able to predict the ultimate tensile strength of FSW joints within the 95% confidence level. Khansare et al. [17] proposed a hybrid optimization methodology based on the combination of the RSM and a genetic algorithm (GA) [18] to approximate the optimal welding speed and tool rotational speed in which a maximum ultimate tensile strength could be achieved.

Tansel et al. [19] developed a genetically optimized neural network system (GONNS) for modeling and optimizing the FSW process. The GONNS was introduced by Tansel et al. [20] by using artificial neural networks (ANNs) in combination with a GA. The GONNS models the system by using the ANNs trained with the experimental data or observations. The optimal operating conditions are estimated by using a GA [19]. Tansel et al. [19] used one GA for searching the optimal tool rotational speed and welding speed by using five ANNs representing the FSW operation. The five separate neural networks with two identical inputs (welding speed and tool rotational speed) estimated the mechanical and metallurgical properties of the friction stir welds.

## 1.2 State of the art—Evaluation of the surface of friction stir welds

Trueba et al. [21] performed an optimization experiment using a factorial design to evaluate the effect of process parameters on the weld temperature, surface and internal quality, and mechanical properties during bobbin-tool friction stir welding. To evaluate the surface appearance, a semi-quantitative visual appearance rating (VAR) was developed based on the presence and severity of visually observable defects. The rating scale ranged from nine (poorest surface quality) to zero (best surface quality), and the criteria *wormhole*, *galling*, *flash*, and *narrow bead* were included. The *wormhole* was defined as an internal void extending to the surface. It was found that high levels of tool rotational speeds and welding speeds resulted in high welding temperatures and insufficient weld metal constraint. This in turn led to *galling* and the formation of *wormholes* with a corresponding decrease in surface quality. It was taken into account that there is a relationship between rotational speed, weld temperature, surface appearance, and void formation.

According to Zuo et al. [2], the surface topography of friction stir welds plays an important role in the performance of the joints. A larger surface roughness leads to a more serious stress concentration, which will cause the occurrence of fatigue damage and the reduction of fatigue strength of the parts [2]. Important process parameters to control the surface topography of friction stir welds are the welding speed and the rotational speed of the tool [22]. Hartl et al. [23] presented key indicators for quantifying the surface topography of friction stir welds and showed that some of these can be predicted by evaluating process variables such as the process forces or temperatures [24].

To date, there have not been any investigations regarding the very promising algorithm-based optimization of the surface topography of friction stir welds or on the application of reinforcement learning (RL) [25] and Bayesian optimization (BO) in the field of FSW, which is why these modern learning-based algorithms are used in this work. The fundamentals regarding these algorithms are contained in [Appendix I](#).

## 2 Methodology

### 2.1 Approach

Previous investigations have shown that the surface quality of friction stir welds significantly depends on the welding speed  $v_s$  and the rotational speed  $n$  of the tool. The optimal setting of these parameters depends on factors such as the sheet thickness, the aluminum alloy, and the tool geometry used, for instance. Due to the complex interrelations, the ideal welding speed  $v_s$  and tool rotational speed  $n$  can often only be found through experience and trial and error. In this work, a learning-based system was developed that helps the FSW user to find optimal settings for these two parameters. Since the production of friction stir welds is time-consuming, as few parameter combinations as possible should be sampled to find suitable parameters.

The evaluation of the surface quality was conducted on the basis of surface topography indicators for friction stir welds, which were presented in Hartl et al. [23]. The task was addressed as an optimization problem:

$$\operatorname{argmin}_{(v_s, n) \in \mathcal{X} \subseteq \mathbb{R}^2} \operatorname{def}(v_s, n) \quad (2)$$

Here,  $\operatorname{def}$  is a function that indicates how defective the surface of the friction stir weld is for the given parameters. The function value  $\operatorname{def}(v_s, n)$  is smaller than the function value  $\operatorname{def}(v_s', n')$  if the parameter combination  $(v_s, n)$  leads to fewer surface defects than the parameter combination  $(v_s', n')$ . The evaluation of  $\operatorname{def}(v_s, n)$  can be equated with the explicit testing of the parameter combination  $(v_s, n)$ , i.e., the production of the friction stir weld with the given parameters, the recording of the surface topography, and the calculation of the topography

key indicators based on Hartl et al. [23]. Since this process is associated with considerable effort, the number of evaluations of  $\operatorname{def}$  should be kept as low as possible. When implementing the algorithms, it had to be taken into account that there was no information about the gradient of  $\operatorname{def}$ . Additionally,  $\operatorname{def}$  contains an error: even if the process parameters are identical for two experiments, the surface topography of these two welds will not be completely identical. Small measurement inaccuracies may also occur when recording the surface topography with the three-dimensional profilometer. However, for simplification purposes, it was assumed that  $\operatorname{def}$  has no error. To solve the optimization problem, three different approaches were considered:

- I. For the first approach, the optimization problem was modeled as a Markov decision process (MDP) and solved using the RL-based value iteration algorithm.
- II. For the second approach, the optimization problem was solved with BO. In the further discussion, the second approach will be called single-task.
- III. For the third approach, the optimization problem was also solved using Bayesian optimization. In contrast to the single-task approach, here the Gaussian process (GP) was provided with additional data that it could use to find the optimum. The GP was provided with information about the type of aluminum alloy, the sheet thickness, and the shoulder geometry used. In the further discussion, the third approach will be called multi-task.

### 2.2 Welding experiments

The welding experiments were conducted on a four-axis milling machining center MCH 250 from Gebr. Heller Maschinenfabrik GmbH, which was adapted for friction stir welding. The maximum axial force of the system was 30 kN. In the experiments, the sheets were joined in the butt joint configuration and a rigid clamping device avoided gaps between the two joining partners. All tests were performed in position-controlled operation with a  $2^\circ$  tilt angle of the tool. Two-piece tools consisting of a shoulder and a conical welding probe with a thread and three flats were used. A total of 262 welding experiments were conducted within the scope of 10 studies. In the 10 studies, the type of aluminum alloy, the tool shoulder geometry, and the sheet thickness were varied. [Table 1](#) provides an overview of the different studies. Some of the studies have already been described in more detail in previous research conducted by Hartl et al. [23, 26]. The evaluated weld seam length varied in the ten studies, but was always between 70 and 170 mm. The evaluated weld seam area started 10 mm after the plunge point and ended approximately 20 mm before the exit hole.

**Table 1** Welding experiments used for this work

Study no.	Experiment no.	Number of exp.	Number of different ...		Aluminum alloy	Shoulder geometry	Sheet thickness
			... $v_s$	... $n$			
1	001–016	16	4	4	EN AW-5754-H111	Concave	4 mm
2	017–032	16	4	4	EN AW-5754-H111	Spiral	4 mm
3	033–048	16	4	4	EN AW-6082-T6	Concave	4 mm
4	049–064	16	4	4	EN AW-5754-H111	Rings	4 mm
5	065–080	16	4	4	EN AW-6082-T6	Rings	4 mm
6	081–096	16	4	4	EN AW-6082-T6	Rings	4 mm
7	097–112	16	4	4	EN AW-5754-H111	Concave	2 mm
8	113–128	16	4	4	EN AW-5754-H111	Rings	4 mm
9	129–141	13	1	13	EN AW-6082-T6	Concave	4 mm
10	142–262	121	11	11	EN AW-6082-T6	Rings	3 mm

The welding speed  $v_s$  and the tool rotational speed  $n$  were varied in a large parameter window. As high welding speeds  $v_s$  are becoming increasingly important for industrial applications, especially in the context of electromobility [27], welding speeds of up to 1500 mm/min were employed. In order to protect the welding equipment, the minimum  $n/v_s$  ratio was limited to 1 mm<sup>-1</sup>. In studies no. 1 to 8, the welding speed  $v_s$  and the tool rotational speed  $n$  were varied in a full factorial manner in four steps, respectively. Thereby, the welding speeds  $v_s$  ranged from 500 to 1500 mm/min and the tool rotational speeds  $n$  from 1500 to 3500 min<sup>-1</sup>. In study no. 9, a total of 13 different rotational speeds  $n$  from 1500 to 3500 min<sup>-1</sup> were set at a welding speed  $v_s$  of 833 mm/min. In study no. 10, the welding speed  $v_s$  was varied in eleven steps from 500 to 1500 mm/min and the rotational speed  $n$  was varied in eleven steps from 1500 to 3500 min<sup>-1</sup> in a full factorial design.

### 2.3 Data preprocessing

The topography of the friction stir welds was recorded using a three-dimensional profilometer VR-3100 from Keyence Deutschland GmbH which was based on phase-coded structured light projection. Thereby, white LEDs projected light from two places onto the welds and the reflected light was measured by a CMOS sensor. The smallest measurable difference in the height direction normal to the sheet surface was 1 μm. The sheet surface was defined as the zero height. The distance between the individual topography points in the plane of the sheet surface was approximately 24 μm. A total of about 250,000 height information points per 10 mm weld seam length were generated. The point cloud was processed to determine the key indicators listed in Table 2 for each weld. A more detailed description of the key indicators is given in Hartl et al. [23].

Table 3 shows the ideal value for each of these eight key indicators as well as the best and the worst values obtained for the 262 welding experiments performed. The value of -2.80 mm for the largest seam underfill was notably high. This value was caused by a lack of fill occurring in some experiments in study no. 2 (see also Hartl et al. [23]). The maximum value for the peak material volume of 37.36 ml/m<sup>2</sup> was also remarkably high. This high value could be explained by flash that reached into the weld. The values displayed in Table 3 were therefore all considered plausible.

The eight topography indicators obtained for the 262 welds were then scaled to values between 0 and 1. The ideal value for each topography key indicator was scaled to a value of 0 and the worst occurring value for each topography key indicator was scaled to a value of 1. The ideal value for the ratio  $r_{arc}$  is 1 [23]. The largest deviation from this ideal value was 0.95 at an  $r_{arc}$  of 1.95, which is why that deviation was scaled as 1. The eight scaled values  $N$  for the eight topography indicators were then averaged for each weld according to:

$$def = \frac{N_{f_m} + N_{u_m} + N_{S_f} + N_{S_u} + N_{S_d} + N_{r_{arc}} + N_{V_{mp}} + N_{S_w}}{8} \quad (3)$$

and a scaled and averaged key indicator  $def$  was obtained that took into account all eight topography indicators defined before (see Table 2). In Eq. 3, all defined topography key indicators were weighted equally. If a quality characteristic would be particularly relevant in the application, for example, the flash height, this could be weighted more prominently in Eq. 3. The perfect friction stir weld surface would therefore have the value  $def$  of 0. The best actual weld of all 262 conducted experiments was experiment no. 53, which had the value  $def$  of 0.021. The worst obtained value for  $def$  was 0.603 for experiment no. 123. Figure 1 shows the evaluated areas of these two welds as color and topography images. The images were generated using the three-dimensional profilometer. Experiment no. 53, on the one hand, contained no surface defects. Neither pronounced flash

**Table 2** Key indicators derived from the three-dimensional point cloud to quantify the features of the weld surface

Surface feature	Key indicator 1	Key indicator 2
Flash formation	Mean flash height $f_m$	Standard deviation of flash height $S_f$
Seam underfill	Mean seam underfill $u_m$	Standard deviation of seam underfill $S_u$
Surface galling	Peak material volume $V_{mp}$ [28]	n/a
Arc texture formation	Standard deviation of the difference between the local minima and the following local maxima along the weld center line $S_d$	Ratio between the measured number and the theoretical number of local extrema along the weld center line $r_{arc}$
Weld seam width	Standard deviation of the weld seam width $S_w$	n/a

formation, nor surface galling, nor cracks were visible on the surface of the weld. The topography image in Fig. 1 shows that the seam underfill was also low and regular. Experiment no. 123, on the other hand, showed a very strong flash formation and pronounced surface galling. The defined key figure *def* was therefore assessed as suitable for documenting the surface quality in a scalar quantity.

### 3 Results

#### 3.1 Reinforcement learning

In order to solve the optimization problem using RL, it first had to be formulated as an MDP. Two formulations, labeled as formulation 1 and formulation 2, were implemented, which differed in the state transition function and the possible actions. In both formulations, the state transition function was deterministic. For a clearer presentation, the state transition function  $p$  was represented deterministically with the two functions  $T(s, a): S \times A \rightarrow S$  and  $r(s, a): S \times A \rightarrow \mathbb{R}$ . These two functions described in which state  $s'$  the environment resulted and which reward  $r$  the agent received when the agent executed the action  $a$  in the state  $s$ . The state transition function could be derived from both functions as follows:

$$p(s', r | s, a) = \begin{cases} 1 & \text{if } s' = T(s, a) \text{ and } r = r(s, a) \\ 0 & \text{else} \end{cases} \tag{4}$$

Both formulations were solved using the value iteration algorithm. The parameter  $\delta$  (see Algorithm A2) was set to  $10^{-8}$ .

In both formulations, the states  $S$  were the parameter combinations of welding speed  $v_s$  and tool rotational speed  $n$ . For example, from study nos. 1 to 8, each of which included 16 different parameter combinations,  $S$  was assigned:

$$S = (v_s, n) = \{500, 833, 1167, 1500\} \times \{1500, 2167, 2833, 3500\} \tag{5}$$

The reward function  $r$  should lead the policy  $\pi$  (see Appendix I) into a minimum as fast as possible, which is why the following reward function was chosen for both formulations:

$$r(s, a) = -def(T(s, a)) - b_s \tag{6}$$

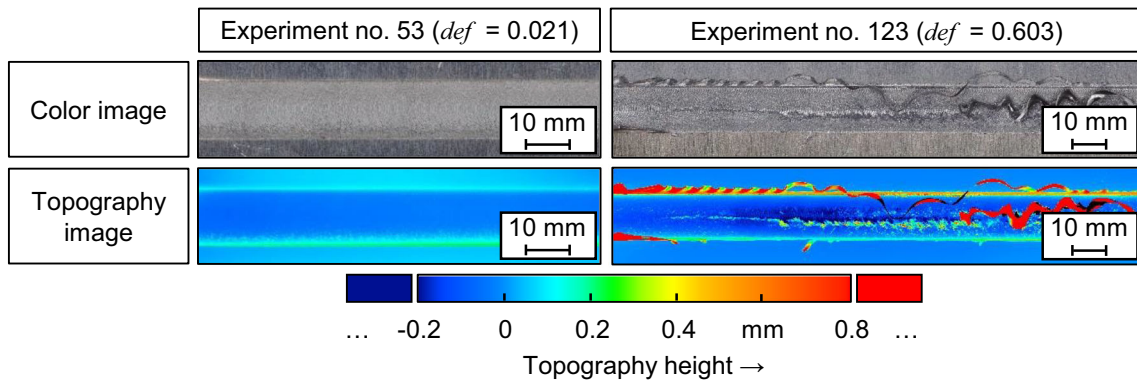
where *def* was the function to be optimized and  $b_s \in \mathbb{R}^+$  was a constant. This reward function gave the agent a higher reward  $r$ , the smaller the value of *def* was, i.e., the less defective the topography of the friction stir weld was. In addition, the agent received a penalty of  $b_s$  for each step, because each step was coupled with an explicit evaluation of *def* and thus connected with effort. The value for  $b_s$  was set to 1. Since the value of *def* was between 0 and 1, the reward  $r$  for each step was between  $-2$  and  $-1$ .

In formulation 1, only the following four different actions were allowed:

$$A = \{\text{increase welding speed } v_s, \text{ reduce welding speed } v_s, \text{ increase rotational speed } n, \text{ reduce rotational speed } n\} \tag{7}$$

**Table 3** Ideal values for each of the eight key indicators as well as the best and worst values obtained for one of the 262 welding experiments performed

Key indicator	$f_m$	$u_m$	$S_f$	$S_u$	$S_d$	$r_{arc}$	$V_{mp}$	$S_w$
Unit	mm	mm	mm	mm	mm	-	ml/m <sup>2</sup>	mm
Ideal value	0	0	0	0	0	1	0	0
Best occurred value	0.09	-0.02	0	0	0	1.00	0.27	0.05
Worst occurred value	3.79	-2.80	0.90	0.75	0.37	0.21/1.95	37.36	4.31



**Fig. 1** Color images and topography images of the evaluated weld surfaces with the lowest ( $def=0.021$ ) and highest ( $def=0.603$ ) value for the scaled and averaged surface topography indicator  $def$

For example, for study nos. 1 to 8, the state transition function  $T$  for formulation 1 was defined as:

$$\begin{aligned}
 T(s, \text{increase welding speed } v_s) &= \begin{cases} (v_s + 333, n) & \text{if } v_s < 1500 \\ s & \text{else} \end{cases} \\
 T(s, \text{reduce welding speed } v_s) &= \begin{cases} (v_s - 333, n) & \text{if } v_s > 500 \\ s & \text{else} \end{cases} \\
 T(s, \text{increase rotational speed } n) &= \begin{cases} (v_s, n + 667) & \text{if } n < 3500 \\ s & \text{else} \end{cases} \\
 T(s, \text{reduce rotational speed } n) &= \begin{cases} (v_s, n - 667) & \text{if } n > 1500 \\ s & \text{else} \end{cases}
 \end{aligned} \tag{8}$$

Figure 2 shows the 16 different states with the respective possible actions for formulation 1 for study nos. 1 to 8. The states and possible actions for study nos. 9 and 10 were analogous.

In formulation 2, in contrast to formulation 1, the agent was able to change from any state to any other state. This followed:

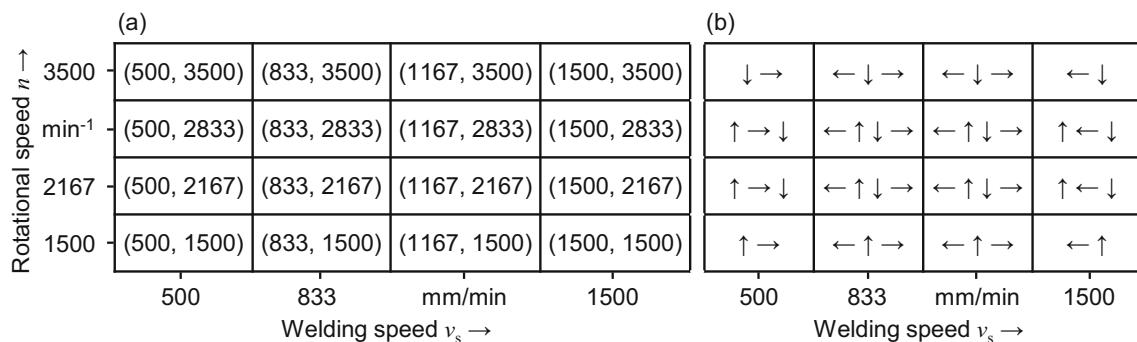
$$A = S \tag{9}$$

The state transition function was thus simplified to:

$$T(s, a) = a \tag{10}$$

In both formulations, the discount factor  $\gamma$  was set to 1 (see Appendix I). This expressed the fact that both immediate and future evaluations of  $def$  were equally unwanted [25]. Additionally, the number of iterations no longer depended on the choice of  $\gamma$ .

To ensure that the algorithm still terminated, the minimum of  $def$  was regarded as the terminal state [25]. This meant that the value of the value function in the terminal state was always zero. The algorithm terminated because the agent could get from any state to the terminal state in finitely many steps and the agent received a negative reward in every state except the terminal state. So, in order to get the least negative reward possible, the agent



**Fig. 2** a Possible states  $s$  and b possible actions  $a$  in the different states  $s$  for study nos. 1 to 8 when using formulation 1

had to get to the terminal state as quickly as possible. In Algorithm A1, the slightly modified value iteration algorithm is described. The changes, compared with Algorithm A2, are underlined. The changes ensured that the value functions of the terminal states were always zero. In addition, the algorithm was adapted for the deterministic formulation.

**Algorithm A1** Adapted value iteration algorithm

```

1 initialize  $V^0$  arbitrarily (e. g.  $\forall s \in S: V^0(s) = 0$ )
2 set for each terminal state  $s \in S: V^0(s) = 0$ 
3 loop:
4 |  $\Delta \leftarrow 0$ 
5 | for each  $s \in S$  that is not a terminal state:
6 | |  $v \leftarrow V^0(s)$ 
7 | |  $V^0(s) \leftarrow \max_{a \in A} (r(s, a) + \gamma \cdot V^0(T(s, a)))$ 
8 | |  $\Delta \leftarrow \max(\Delta, |v - V^0(s)|)$ 
9 until  $\Delta < \delta$ 
    
```

Figure 3 a shows the values for the scaled and averaged surface topography indicator *def* (see Eq. 3) for the 16 different states in study no. 1. It is evident that the best value for *def*, 0.041, in study no. 1 was obtained in the state (500, 1500). This was in good agreement with the result from Hartl et al. [23], wherein this parameter combination led to the best-rated result in study no. 1 according to the visual inspection. Figure 3 b shows the initialization of the value function  $V^0(s)$  with zeros (compare line 1 in Algorithm A1).

In both formulations, the algorithm evaluated the function *def* for all 16 parameter combinations. Figure 4 illustrates the results for the value iteration algorithm for study no. 1 when using formulation 1. Thereby, the strategy for each state was the direction in which the sum of future rewards was maximized (see Eq. 19 in Appendix I). The algorithm terminated after five iterations. Figure 4 a shows the values of the value function  $V^0(s)$  for the first and the fifth (last) iteration. Figure 4 b shows the values of the value function  $V^0(s)$  plus the reward, if the agent changes from another state to this state for these two iterations. Figure 4 c demonstrates the direction to the neighboring state that yields the highest improvement

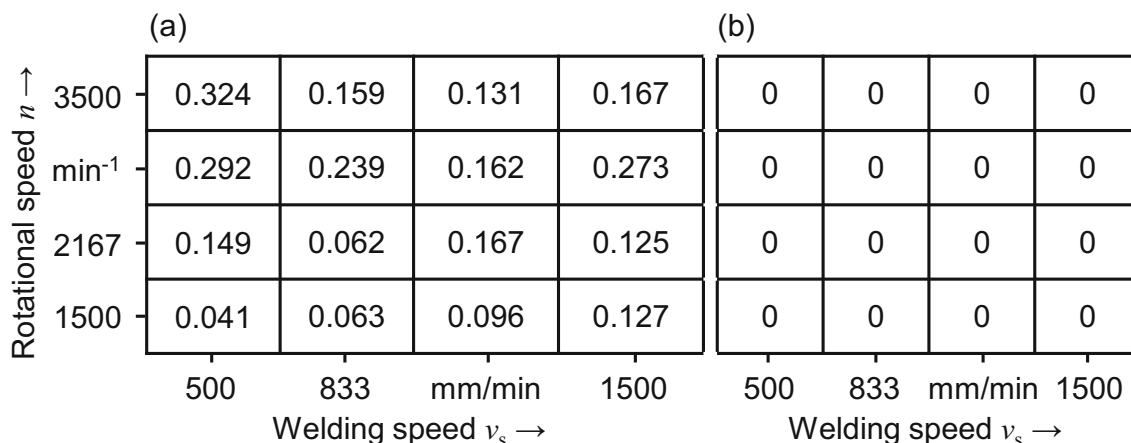
according to Fig. 4b. For example, for state (1500, 3500) in iteration 1, that would be the reduction of the welding speed  $v_s$ . The terminal state was reached for (500, 1500), so the value function  $V^0(s)$  for this state is zero from the beginning. The results for the other studies (see Table 1) led to the same findings.

Figure 5 displays the results for the value iteration algorithm for study no. 1 when using formulation 2 analogous to Fig. 4. For formulation 2, the strategy in each state  $s$  was the action (500, 1500). The labeling of the boxes in Figs. 4 and 5 is analogous to the labeling in Fig. 2a.

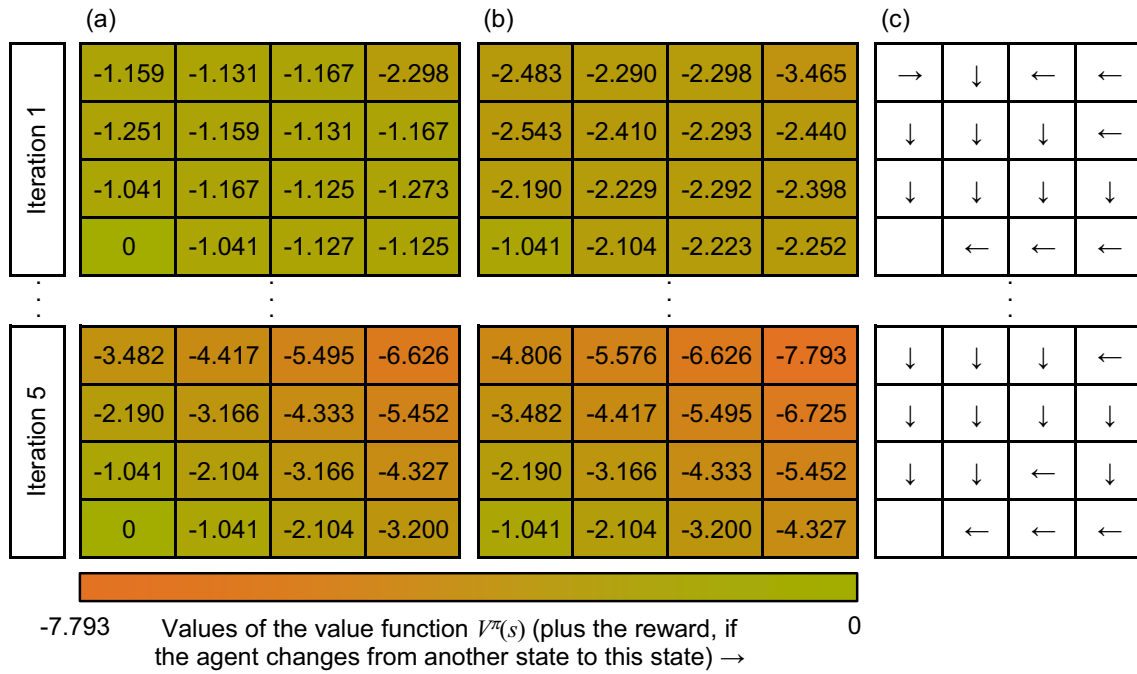
The results showed that the optimization problem (see Eq. 2) can be solved using RL. However, the value iteration algorithm was not efficient. Since all states  $s \in S$  and all actions  $a \in A$  were iterated over, the function *def* had to be evaluated for each process parameter combination. There are algorithms for RL that are more efficient in this respect. For example, the value function could be approximated using a Gaussian process to reduce the number of evaluations [29].

**3.2 Bayesian optimization**

**Optimization of the surface quality** In the second approach (single-task), the optimization problem was solved using Bayesian optimization (BO). Thereby, only the data from the respective study were used. Additional information from other studies or information such as the type of aluminum alloy were not utilized. Since no information was available at the beginning of the optimization for the first selection of a parameter set, a random parameter set had to be selected in the first trial. A parameter set consisted of the welding speed  $v_s$  and the tool rotational speed  $n$ . To ensure that the results were independent of the selected starting point, all parameter combinations were used once as the starting point and then the means and standard deviations of the required number of steps to find good parameter sets were subsequently calculated.



**Fig. 3** a Values for the scaled and averaged surface topography indicator *def* for study no. 1. b Initialization of the value function  $V^0(s)$

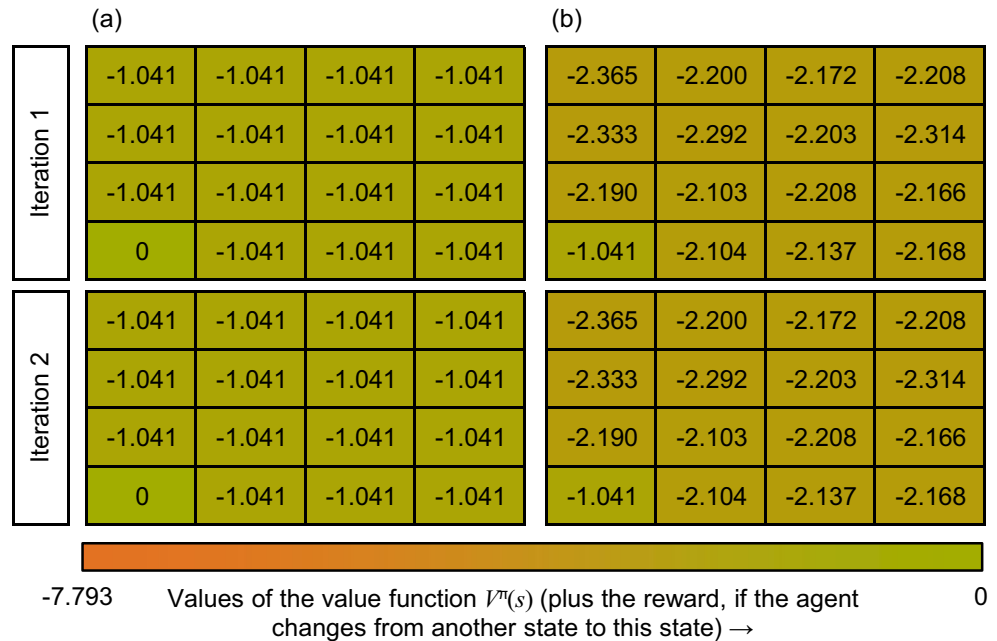


**Fig. 4** Visualization of the algorithm for the 16 different states in study no. 1 using formulation 1 for the five iterations. **a** Values of the value function  $V^{\pi}(s)$ . **b** Values of the value function  $V^{\pi}(s)$  plus the reward, if the agent changes from another state to this state. **c** Optimal action according to **b**

In order to avoid overfitting the hyperparameters, it was assumed that they follow certain distributions. For the hyperparameter length-scale  $l$  of the Matérn 5/2 kernel, a uniform distribution was assumed, since the Gaussian process degenerated strongly at values below 0.1 and those were thus excluded. A logarithmic normal distribution was assumed for the variance  $\sigma$  to determine the approximate interval in which  $\sigma$  should be located. The exact choice of the

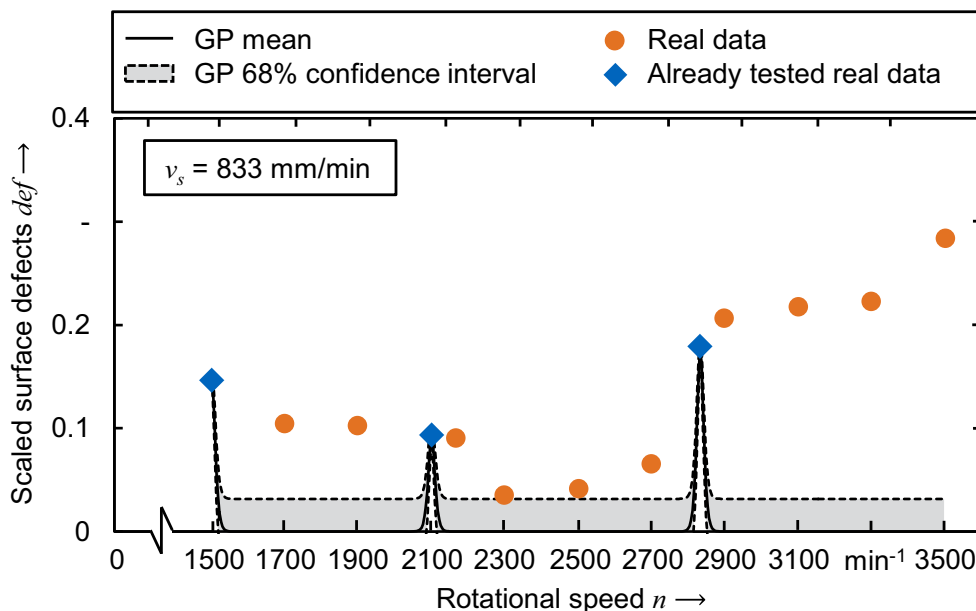
distribution parameters was not significant. The parameters in Table 4 in Appendix II were found by trial and error using the data sets from study nos. 1 to 10, so that the distributions cover approximately the range of the hyperparameters that do not degenerate the GP. Figure 6 shows a degenerate GP for study no. 9 that had to be avoided. The cause of the degeneration was that the hyperparameter  $l$  was selected to be 0.01 and

**Fig. 5** Values of the value function  $V^{\pi}(s)$ , which is the sum of rewards if the agent follows the policy in this state, for the 16 different states in study no. 1 using formulation 2 for the two iterations. **a** Sum of rewards if the agent follows the policy in this state. **b** Sum of rewards if the agent changes from another state to this state and then follows the policy





**Fig. 6** Example of a degenerated GP due to unsuitable hyperparameters



therefore too small. Consequently, the mean in Fig. 6 is always zero and only spikes at known points. In addition, the hyperparameter variance  $\sigma$  of 0.001 was too small.

In the third approach (multi-task), the optimization problem was also solved using the BO, but in this case, the GP received the data sets from the nine other studies as additional information, respectively. It was suspected that this would allow the GP to better estimate the function to be minimized and to find the optimum more quickly. In order for the GP to be able to estimate which other studies were similar, it was given additional features as input variables which influence the setting of the process parameters. These features were the type of aluminum alloy, the sheet thickness, and the shoulder geometry. Since they were identical in study nos. 3 ↔ 9, 4 ↔ 8, and 5 ↔ 6 (see Table 1), the study number was also used as an additional input variable. This resulted in the following kernel for the GP:

$$\begin{aligned}
 g(v_s, n, m_f, t_f, s_f, i) : \mathbb{R} \times \mathbb{R} \times M_f \times T_f \times S_f \times \{1, 2, \dots, 10\} &\rightarrow \mathbb{R} \\
 M_f &= \{\text{EN AW-5754-H111, EN AW-6082-T6}\}, \\
 T_f &= \{2 \text{ mm}, 3 \text{ mm}, 4 \text{ mm}\}, \\
 S_f &= \{\text{concave, spiral, rings}\},
 \end{aligned}
 \tag{11}$$

where  $M_f$ ,  $T_f$ , and  $S_f$  represent the quantities of the aluminum alloys  $m_f$ , sheet thicknesses  $t_f$ , and shoulder geometries  $s_f$  used. Additionally,  $i$  is the number of the study and  $\{1, 2, \dots, 10\}$  is the set of natural numbers less than or equal to 10, since data from 10 studies were used in total. For the welding speed  $v_s$ , the tool rotational speed  $n$ , and the sheet thickness  $t_f$ , the Matérn 5/2 kernel (see Appendix I) was used. Since the type of aluminum alloy, the shoulder geometry, and the study number were

categorical values, the coregionalization kernel (see Appendix I) had to be used for those. In that way, their covariance could be learned via hyperparameter optimization. The covariance function  $k$  was defined as follows:

$$\begin{aligned}
 &k\left(\left(v_s, n, m_f, t_f, s_f, i\right), \left(v'_s, n', m'_f, t'_f, s'_f, i'\right)\right) \\
 &= k_E\left(\left(v_s, n, t_f\right), \left(v'_s, n', t'_f\right)\right) \cdot k_M\left(m_f, m'_f\right) \cdot k_S\left(s_f, s'_f\right) \cdot k_I\left(i, i'\right)
 \end{aligned}
 \tag{12}$$

where  $k_E$  is the Matérn 5/2 kernel,  $k_M$  and  $k_S$  are coregionalization kernels for two or three categories, and  $k_I$  is a coregionalization kernel for ten categories. Since the GP was given information from the nine other studies, the first parameter set to be tested no longer had to be chosen randomly as in the single-task approach.

In order to prevent overfitting the hyperparameters, it was also assumed for the multi-task approach that the hyperparameters follow certain distributions. A uniform distribution was assumed for the hyperparameter length-scale  $l$  of the Matérn 5/2 kernel. A logarithmic normal distribution was assumed for hyperparameters, which can only have positive values, and a normal distribution for all remaining hyperparameters. Table 5 in Appendix II lists the hyperparameters for the multi-task approach and their distribution analogous to Table 4.

The number of expected steps until the random search algorithm finds a suitable parameter setting has been calculated as described below: Let  $Z$  be a random variable that indicates after how many steps a random search finds one of the  $o$  optima for the first time. Thus,  $p_z$  was the probability that a random search has not found an optimum in the previous  $z$  steps and finds an optimum in the  $(z + 1)$ -th step:

$$p_z(Z = z + 1) = \frac{(q - o)!}{(q - o - z)!} \cdot \frac{(q - z)!}{q!} \cdot \frac{o}{q - z}$$

In the first  $z$  steps, none of the  $o$  optima were found.
An optimum is found in the  $(z + 1)$ -th step.

(13)

whereby  $q$  is the number of possible parameter combinations and  $Z \in \{1, 2, \dots, q - o + 1\}$ , whereby  $\{1, 2, \dots, q - o + 1\}$  is the set of natural numbers less than or equal to  $(q - o + 1)$ , since there are at most  $(q - o)$  parameter sets that are not an optimum and one of the optima is found in the  $(q - o + 1)$ -th step at the latest. Thus, the expected number of necessary steps to find an optimum using random search was:

$$\mathbb{E}[Z] = \sum_{z=1}^{q-o+1} z \cdot p_z(z) = \sum_{z=0}^{q-o} (z+1) \cdot p_z(Z = z+1) \quad (14)$$

Table 6 in Appendix II shows the number of steps required to find a suitable parameter set for the ten different studies. Initially, the best 20% of each study were defined as suitable. This resulted in the number of  $o_{20}$  parameter sets for each study, which led to an optimum result. If the random search (RS) algorithm was used, the calculated expected value until one of the  $o_{20}$  good parameter sets was found was 4.3 steps on average. When using the single-task approach and the probability of improvement (PI) acquisition function, an average of 3.3 steps, and when using the multi-task approach, an average of 2.4 steps were required to find a parameter set that was among the best 20%. The average results for the expected improvement (EI) acquisition function were almost identical with the results when using the PI acquisition function. Since the application of the single-task approach also depended on which parameter set was tested first, the mean value and standard deviation are given for each study in Table 6. The BO was started once with each of the parameter sets. Since the determination of suitable settings for the welding speed  $v_s$  and the tool rotational speed  $n$  with both the single-task and the multi-task approaches succeeded faster than an RS algorithm would suggest, the use of the BO was considered suitable. It was expected that the BO would lead to suitable parameters faster than RS, since the BO models the function to be optimized. This made it possible to estimate which parameter setting should be tested next. With RS, the parameter settings were evaluated in random order. In addition, the information given in Table 6 shows that, on average, the multi-task approach led to better results than the single-task approach. The information from the other data sets could thus be successfully used to determine suitable welding parameters faster. Since the similarity to the other data sets was modeled, it was also possible to derive from which of the other studies the information could be better transferred in the multi-task approach.

Figure 7 illustrates the number of steps required for the different studies and approaches. It becomes obvious again that the multi-task approach led to the best results. Only in study nos. 7 and 9 the results were the worst for the multi-task approach. The reason was assumed to be that study no. 7 was the only study in which a sheet thickness of 2 mm was used (see Table 1), which is why the information from the other studies in this case even had a negative effect on the result compared with the single-task approach. Study no. 9 also differed from the other studies. In this study, the welding speed  $v_s$  was not changed and only the rotational speed  $n$  was varied. Due to the greater difference as compared with the other studies, this information could therefore not be used advantageously.

Table 7 in Appendix II shows the number of steps required to find suitable welding parameters for the various approaches in analogy to Table 6. Now, only the best 5% of each study were defined as suitable. This resulted in the number of  $o_5$  good welding parameter sets for each study. This specification could be used for applications that require a very high surface quality, for example for visible welding seams. Since the number of optimal parameter sets was now lower, the number of steps necessary to find good parameter sets was higher on average. It once again became clear that the use of the BO reduced the number of necessary steps compared with RS. The average number of necessary steps was again lower with the multi-task approach than with the single-task approach. This time, the differences between the two different acquisition functions were slightly bigger and the PI acquisition function was on average better for the single-task approach, whereby for the multi-task approach, the EI acquisition function performed marginally better.

Figure 8 illustrates the steps necessary to find a parameter set that was among the best 5%. Compared with Fig. 7, it is particularly noticeable that in study no. 2, the multi-task approach required the most steps to find a good parameter set. This was assumed to be due to the fact that study no. 2 was the only study in which a spiral shoulder geometry was used (see Table 1) and the information from the other data sets was rather confusing than useful for the multi-task BO algorithm.

Figure 9 shows the sequence of the tested parameter sets in study no. 1 based on the multi-task approach. The sequence was identical for both acquisition functions: PI and EI. Figure 9 b shows the achieved values for the surface topography  $def$  in each step. From the two figures, it becomes clear that the optimum parameter set (500, 1500) was already found in the first step. This was also in agreement with the result from Hartl et al. [23], where for study no. 1 the parameter set (500, 1500) led to the best result regarding the visual inspection. Figure 10 shows a color image and a topography image of the evaluated welding surface from this experiment. Since the approach tested all 16 parameter settings and have only tested each parameter set once (see Algorithm A3), the quality of the surface topography decreased in the subsequent steps (see Fig. 9b).

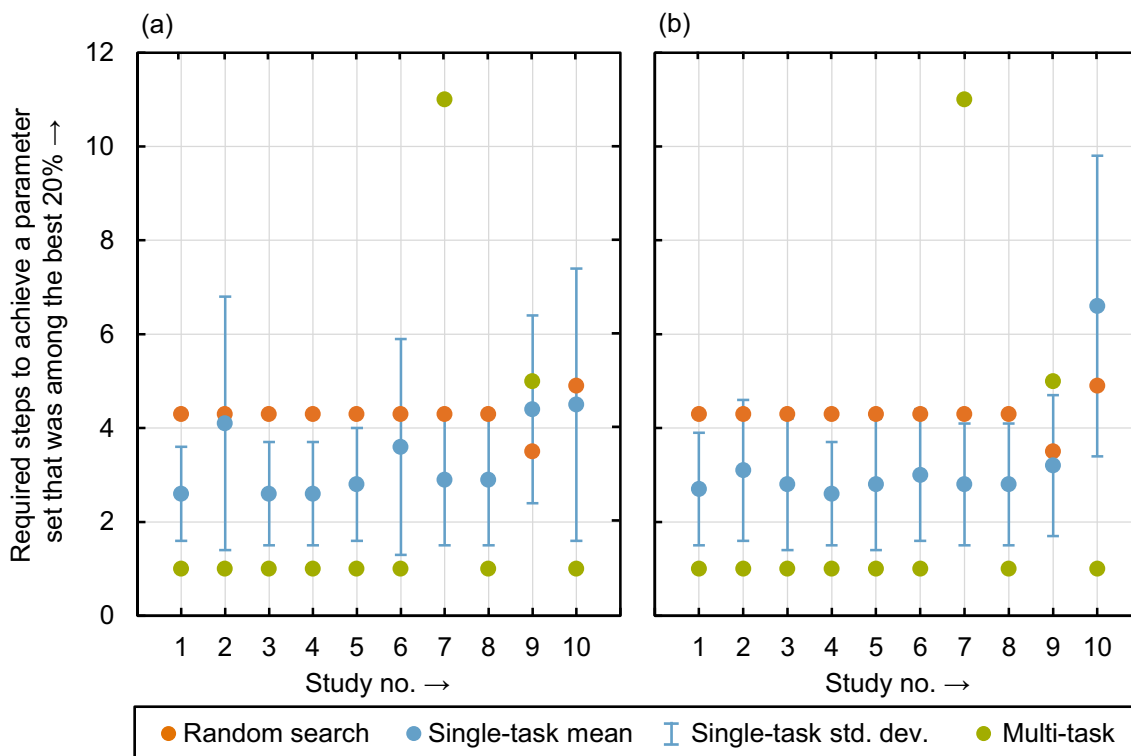


Fig. 7 Number of steps required to achieve a parameter set that was among the best 20% by using the acquisition functions **a** probability of improvement and **b** expected improvement

Figure 11 visualizes the GP for study no. 1 when using the BO multi-task approach with the mean function and the 68% confidence interval. The first three steps have already been performed, so the points (500, 1500), (500, 2167), and (833,

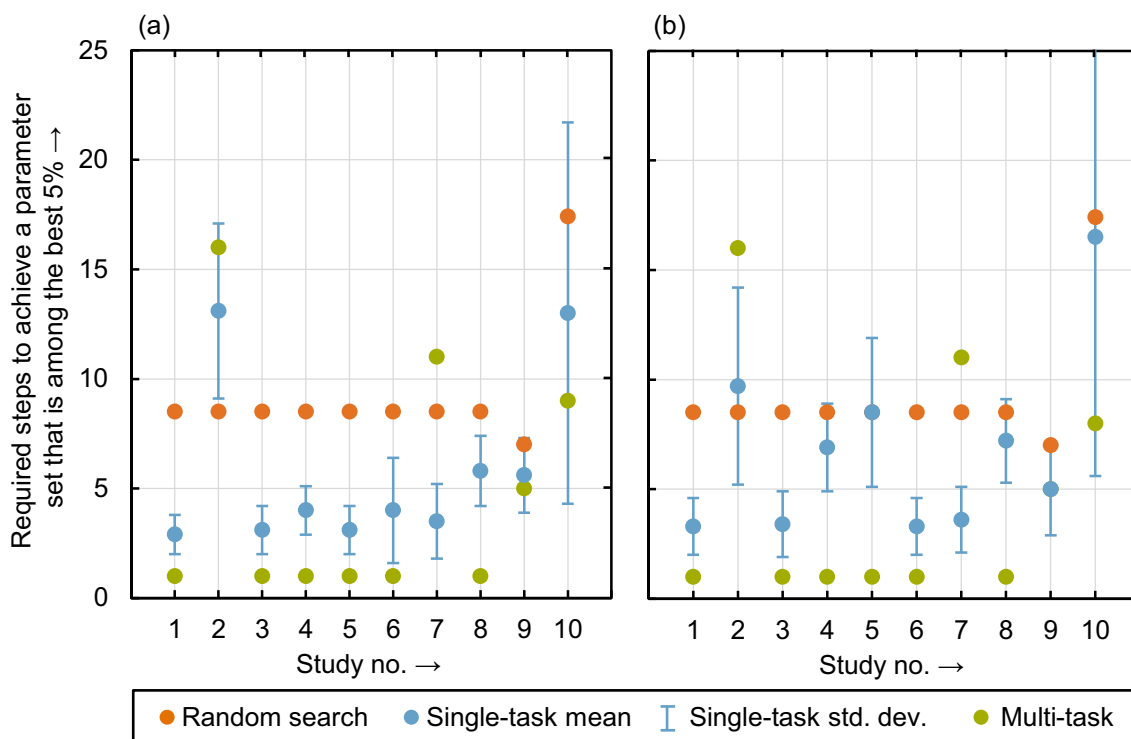


Fig. 8 Number of steps required to achieve a parameter set that is among the best 5% by using the acquisition functions **a** probability of improvement and **b** expected improvement

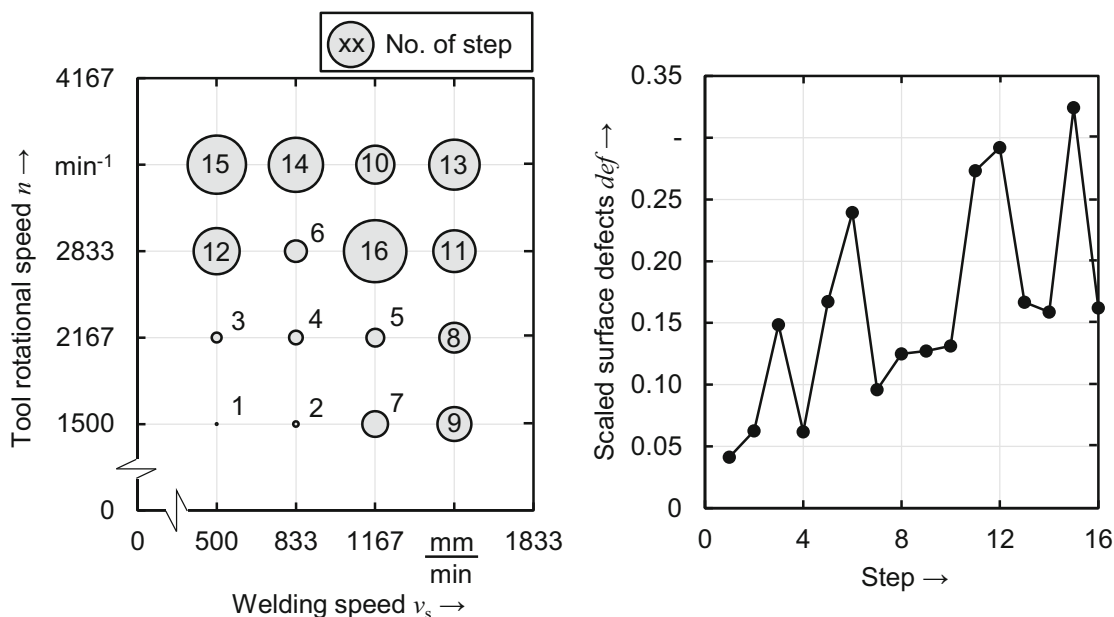


Fig. 9 a Sequence of the selected parameter sets for study no. 1 when using the multi-task approach for the two acquisition functions tested. b Corresponding values for the scaled surface defect  $def$

1500) in Fig. 11 are already known. Since the acquisition function for the point (833, 2167) was the highest compared with the other points not yet evaluated, this point was evaluated next for both the PI and the EI acquisition function.

Figure 12 is analogous to Fig. 11 when using the BO single-task approach and the first three steps have already been performed. Thereby, it was specified that in the first step, the parameter set (500, 1500) was tested. It became clear that the GP mean function and the associated 68% confidence interval could estimate the real data considerably less accurately as compared with the multi-task approach, which is illustrated in Fig. 11.

In the investigations on the multi-task approach presented so far, all nine other studies shown in Table 1 were used to find suitable parameters. In the study nos. 3 ↔ 9, 4 ↔ 8, and 5 ↔ 6, the same aluminum alloy, the same sheet thickness, and

the same shoulder geometry were used. The disadvantage of these studies, hereinafter referred to as duplicates, was that the GP needed additional hyperparameters (see Eq. 11), which made their learning more complex. Furthermore, it was not possible to investigate how well the other data sets, which had no duplicates, could be used for the data sets that had duplicates. In further investigations using the multi-task approach, the duplicates were therefore not used and the study number  $i$  (see Eq. 11) was no longer required for the differentiation, which also simplified the covariance function. Table 8 in Appendix II shows which data sets were used to calculate suitable parameter sets, whereas Table 9 displays a comparison of the results with and without the duplicates. On average, the results could be improved without the duplicates. The assumed reason for this was that fewer hyperparameters had to be learned if the duplicates were

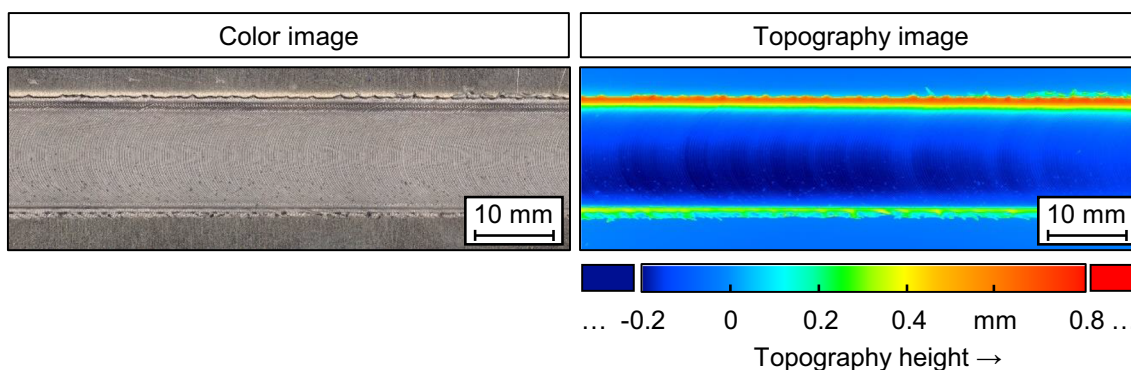
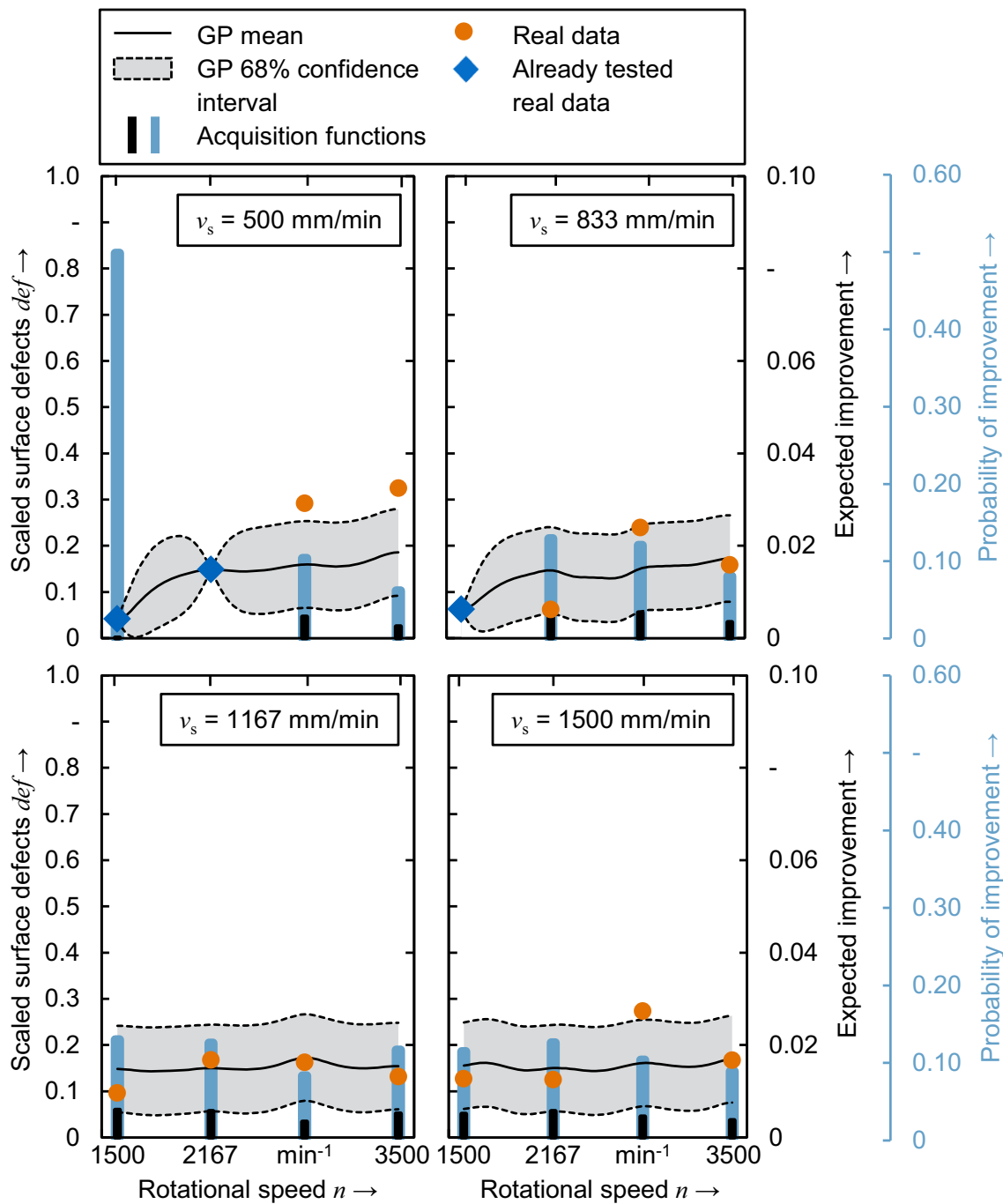


Fig. 10 Color image and topography image of the best weld surface achieved in study no. 1 in which the parameter set was found in the first step using the Bayesian optimization multi-task approach

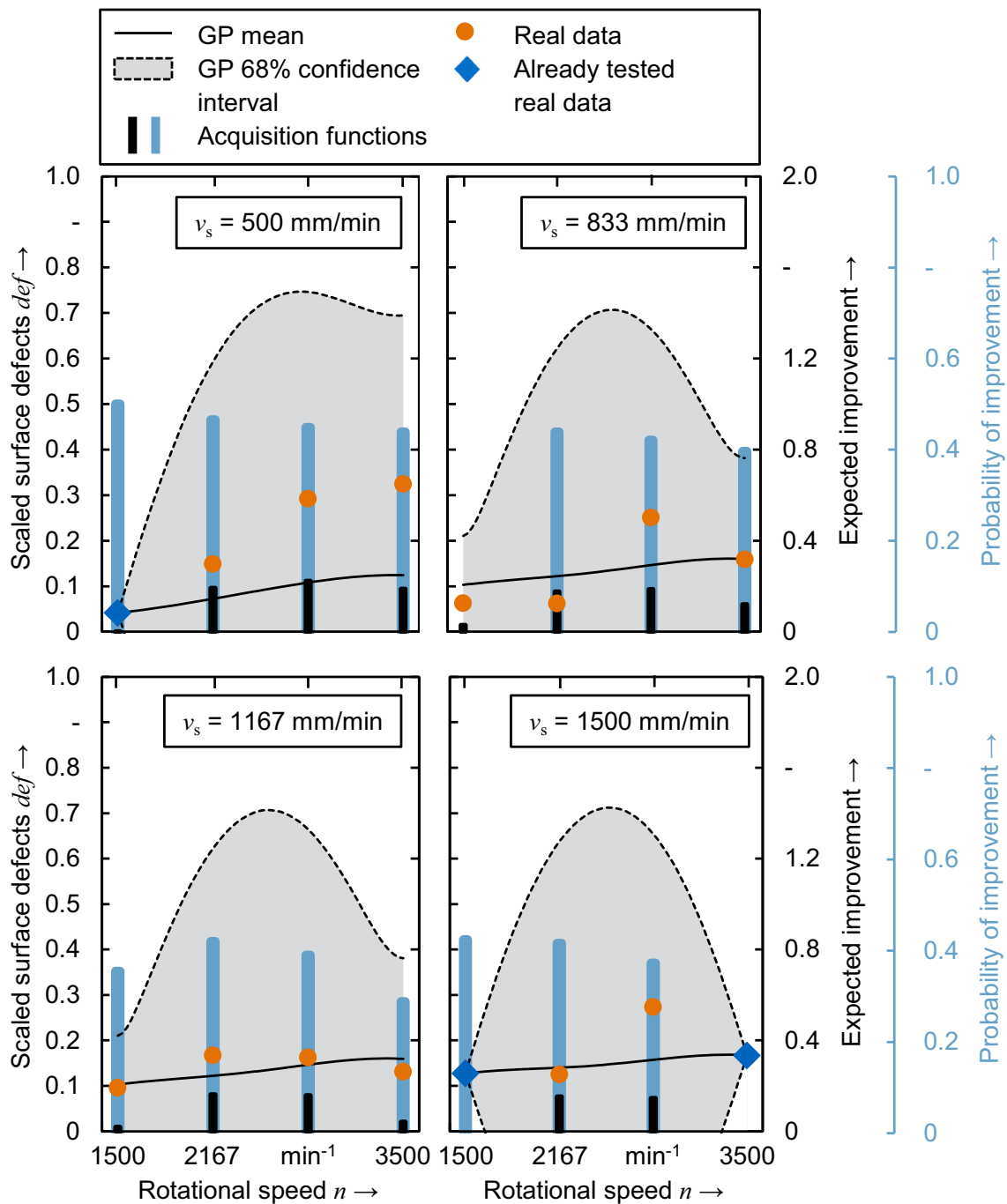


**Fig. 11** Visualization of the Gaussian process and the acquisition functions after three tested parameter sets in study no. 1 after applying the multi-task approach

omitted. Especially in study nos. 2 and 7, which had no duplicates, the results could be improved. In the studies that had duplicates, the results were worse if the duplicates were not used. Only in study no. 9 did the results improve, even though there was a duplicate. This was probably due to the fact that study no. 3 showed a poor welding result for the welding speed of 833 mm/min

and a rotational speed of 2167 min<sup>-1</sup>, whereas study no. 9 showed the best welding results in this parameter range.

Overall, it could be shown that Bayesian optimization can be used very efficiently to find suitable process parameters for friction stir welding. The Bayesian optimization was better suited than reinforcement learning for the aim of

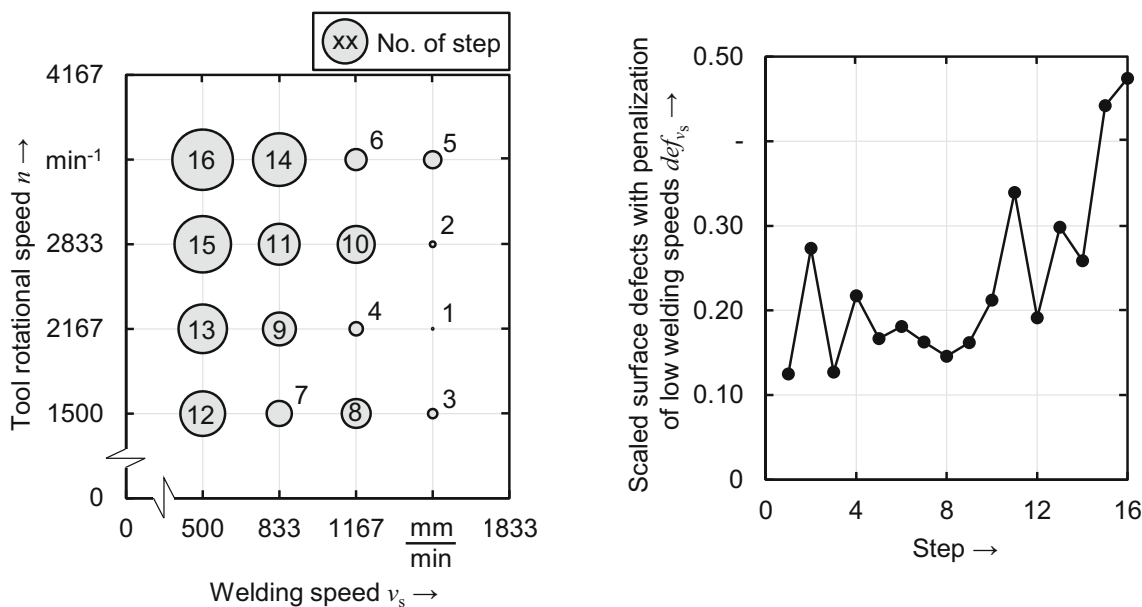


**Fig. 12** Visualization of the Gaussian process and the acquisition functions after three tested parameter sets in study no. 1 after applying the single-task approach

this project to optimize the surface quality of FSW seems as efficiently as possible. For the optimization using reinforcement learning, a strategy had to be found that maximizes the expected value of an infinite sum of random variables and the function  $def$  had to be evaluated for each process parameter combination. Using the Bayesian optimization, the function  $def$  could be directly optimized by

searching for values for the welding speed and the tool rotational speed that minimize the function  $def$ .

In this work, only the welding speed and the tool rotational speed were varied. This can be extended to other process parameters in further work. For example, the tilt angle or the immersion depth of the tool could be implemented additionally to optimize the surface quality. The Bayesian



**Fig. 13** a Sequence of the selected parameter sets using the multi-task approach for the two acquisition functions tested with penalization of low welding speeds. b Corresponding values for  $def_{vs}$

optimization method can also be transferred to force-controlled processes in order to find a suitable setting for the axial force of the tool.

**Optimization of the surface quality with consideration of the welding speed** Since the welding speed  $v_s$  has a direct influence on the process productivity [30] in any welding operation in an industrial context, the objective behind the selection of suitable welding parameters is to maximize the welding speed  $v_s$  while ensuring an acceptable welding quality [31]. The growing market for the use of friction stir welding in the electromobility sector also requires sufficiently high welding speeds in order to enable economic production [27]. Richter [27] recommends aiming to attain a welding speed  $v_s$  of at least 1000 mm/min. Therefore, in the investigations described in this section, slower welding speeds were penalized with  $p_{lvp}$  according to the following

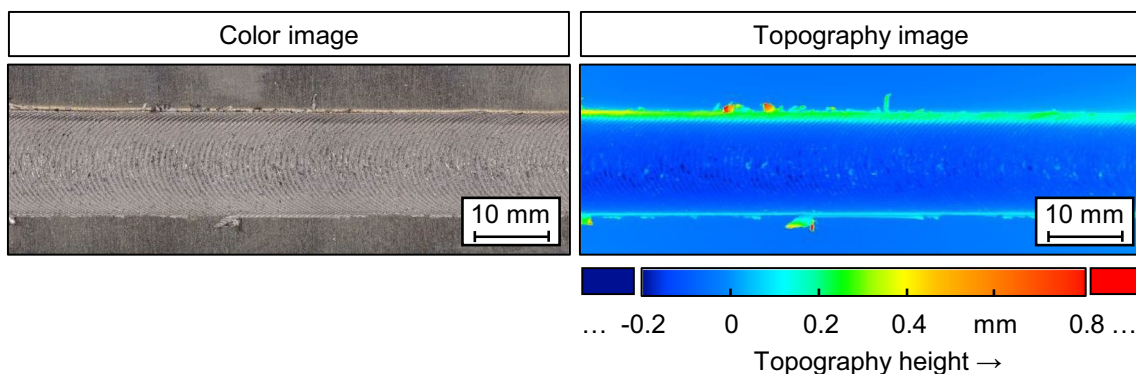
formula:

$$p_{lvp} = lvp \cdot \frac{v_{smax} - v_s}{v_{smax} - v_{smin}} \tag{15}$$

with  $lvp$  being a factor indicating the magnitude of low welding speed penalty,  $v_{smax}$  being the maximum welding speed in a study, and  $v_{smin}$  being the minimal welding speed in a study. The higher the selected  $lvp$  factor, the higher welding speeds are preferred. The value for  $p_{lvp}$  was then added to the values for  $def$ , which were calculated according to Eq. 16 to generate the new variable  $def_{vs}$ , which took into account both the surface quality and the welding speed  $v_s$ :

$$def_{vs} = def + p_{lvp} \tag{16}$$

In this project, the maximum welding speed in all ten studies was 1500 mm/min and the minimum welding speed was 500 mm/min. Additionally,  $lvp$  was chosen to be 0.15. For a



**Fig. 14** Color image and topography image with the best weld in study no. 1, taking into account the weld surface quality and the welding speed; the parameter setting was achieved in the first step using the Bayesian optimization multi-task approach

welding speed  $v_s$  of 833 mm/min, for example, this resulted in a  $p_{lvp}$  of 0.10, which was then added to all values for  $def$  where a welding speed  $v_s$  of 833 mm/min was applied. Figure 13 a shows the sequence of tested parameter sets when using the BO multi-task approach with all ten studies used (including the duplicates) and punishing low welding speeds  $v_s$  with  $p_{lvp}$ . It becomes clear that parameter settings with a welding speed  $v_s$  of 1500 mm/min were preferred then. Figure 13 b shows the corresponding values for  $def_{vs}$ . The optimum for  $def_{vs}$  of 0.12 for study no. 1 was already reached in the first step again, and thus, the parameter set (1500, 2167) was used.

Figure 14 shows a color image and a topography image of the evaluated welding surface for the parameter set (1500, 2167). It becomes clear that the surface has a small irregular flash formation and slight surface galling compared with the result shown in Fig. 10. However, the welding speed  $v_s$  was three times as high.

A possibility was thus found which allows the FSW user to weight the two criteria surface quality and welding speed individually adjusted to the requirements by setting the parameter  $lvp$  and to consider this in the learning-based automated search for suitable process parameters.

## 4 Conclusions and future research

A total of 262 friction stir welds were performed within 10 studies. Subsequently, with reinforcement learning and Bayesian optimization, two learning-based algorithms were tested for their applicability in optimizing the surface topography by adjusting the welding speed and tool rotational speed. The following conclusions were drawn:

- The optimization problem could be solved by means of reinforcement learning, but not efficiently. Furthermore, it was complicated to solve the problem with reinforcement learning, because a policy had to be found that maximizes the expected value of an infinite sum of random variables. Instead, it was better to solve the optimization problem directly by using the Bayesian optimization.
- The Bayesian optimization found suitable settings for the process parameters significantly faster than random search, both without (single-task approach) and with (multi-task approach) the aid of the data sets from the other studies.
- In the multi-task approach, the information from the other studies could be successfully used to find suitable welding parameters even faster compared with the single-task approach.
- By penalizing low welding speeds, both the surface topography and the welding speed could be considered for the optimization.

In a future project, the aim will be to show that the optimization of the surface topography in general leads to an increase

in the ultimate tensile strength of the friction stir welded joint. The transfer of the algorithms developed in this work for the inline optimization of the weld seam surface is also the subject of future investigations.

**Acknowledgments** The IGF-research project no. 19389 N of the “Research Association on Welding and Allied Processes of the DVS” has been funded by the AiF within the framework for the promotion of industrial community research (IGF) of the Federal Ministry for Economic Affairs and Energy because of a decision of the German Bundestag.

**Funding** Open Access funding provided by Projekt DEAL.

## Appendix I

### Reinforcement learning

Reinforcement learning (RL) describes the concept of goal-oriented learning by an agent through the interaction with its environment. At each discrete time step, the agent performs an action, which influences the state of its environment and which rewards the agent. The agent seeks to maximize its rewards, whereby the agent is guided towards the goal. In general, the charm of RL is that the agent learns autonomously to achieve the goal, while the user only needs to assess, by using rewards, how well the goal was achieved [32].

The Markov decision processes (MDPs) have become the standard formalism to mathematically describe sequential decision-making tasks in stochastic environments, such as RL problems [33]. An MDP is defined by [25]:

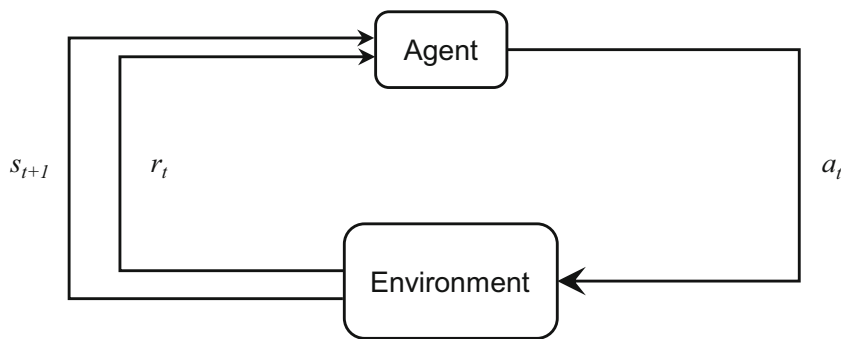
- $S$  is the set of states  $s$ , in which the environment can be,
- $A$  is the set of actions  $a$ , which the agent can perform in the environment, and
- $p(s', r | s, a): S \times R \times S \times A \rightarrow [0,1]$  is the state transition function. Thereby,  $p$  describes the probability distribution of the state  $s'$  and the reward  $r$  after the agent has executed the action  $a$  in the state  $s$ .

At the beginning (time step  $t = 1$ ), the agent is in an initial state  $s_1$ . At every time step  $t$ , the environment is in the state  $s_t$  and the agent executes an action  $a_t$ . According to the state transition function, the agent is in the state  $s_{t+1}$  at the next time  $t + 1$  and gets the reward  $r_t$ . Figure 15 illustrates this cycle.

By using RL, a possibly stochastic policy  $\pi(a | s): A \times S \rightarrow [0,1]$  is searched for, according to which the agent can act in the environment. The policy is supposed to maximize the sum of all expected weighted future rewards  $\mathbb{E}_{r_1, r_2, \dots} [\sum_{t=1}^{\infty} \gamma^{t-1} r_t]$ , where  $\mathbb{E}_{r_1, r_2, \dots}$  is the expected value for the random variables  $r_1, r_2, \dots$ , and  $\pi(a | s)$  indicates the probability with which the agent executes the action  $a$  in the state  $s$ . To keep the sum of all future rewards finite,  $r_t$  is weighted by  $\gamma^{t-1}$ , where  $\gamma \in [0, 1]$ .



**Fig. 15** The agent-environment interaction in an MDP [25]



A  $\gamma$  smaller than one can also express that immediate rewards are more important to the agent than future rewards [25].

Value iteration is an algorithm to find a policy  $\pi$ . It uses an auxiliary function, the value function  $V^\pi(s): S \rightarrow \mathbb{R}$ , which represents the sum of all expected weighted future rewards when the agent follows the policy  $\pi$  [25]:

$$V^\pi(s) = \mathbb{E}_{a_t, r_t, s_{t+1}, a_{t+1}, r_{t+1}, \dots} [\sum_{k=0}^{\infty} \gamma^k \cdot r_{t+k} \mid s_t = s] \quad (17)$$

The value function  $V^\pi$  can in turn be defined by itself [25]:

$$V^\pi(s) = \mathbb{E}_{a_t, r_t, s_{t+1}} [r_t + \gamma \cdot V^\pi(s_{t+1}) \mid s_t = s] \quad (18)$$

The policy is deterministic and returns the action  $a$  that maximizes the expected weighted future reward and can be defined using the value function  $V^\pi$  [25]:

$$\pi(s) = \operatorname{argmax}_{a \in A} \mathbb{E}_{r_t, s_{t+1}} [r_t + \gamma \cdot V^\pi(s_{t+1}) \mid s_t = s, a_t = a] \quad (19)$$

The value function is represented by a look-up table. Since the value function and thus also the values of the look-up table are unknown, the values are learned with the help of the value iteration algorithm. First, the value function is initialized arbitrarily; then, it is updated until the change is only marginal (i.e., smaller than  $\delta$ ) [25].

The value iteration algorithm is described in Algorithm A2.

**Algorithm A2** Value iteration algorithm [34]

```

1 initialize  $V^\pi$  arbitrarily (e. g.  $\forall s \in S: V^\pi(s) = 0$ )
2 Loop:
3    $\Delta \leftarrow 0$ 
4   for each  $s \in S$ :
5      $v \leftarrow V^\pi(s)$ 
6      $V^\pi(s) \leftarrow \max_{a \in A} \mathbb{E}_{r_t, s_{t+1}} [r_t + \gamma \cdot V^\pi(s_{t+1}) \mid s_t = s, a_t = a]$ 
7      $\Delta \leftarrow \max\{\Delta, |v - V^\pi(s)|\}$ 
8 until  $\Delta < \delta$ 
    
```

The value iteration algorithm converges to the optimal value function, i.e., to the optimal policy as well [25]. But both  $S$  and  $A$  must be finite, because the algorithm iterates over all states  $s$  and actions  $a$ .

**Gaussian processes**

A Gaussian process (GP) is a stochastic process, which, among other things, can be used to model probability distributions over functions. Therefore, the GPs can be used as a model for regression analysis [35].

Figure 16 shows a GP with the mean of the GP, the 95% confidence interval, and three samples, whereby four data points  $(x, f(x)) \in \{(-5.0, 2.5), (-2.0, 2.0), (3.0, -0.5), (3.5, 0)\}$  of the unknown function  $f(x)$  are known.

Let

$$\mathcal{D} = \{(x_1, y_1), \dots, (x_q, y_q)\} \subset \mathcal{X} \times \mathbb{R} \quad (20)$$

be a finite set of data points [35]. Then, a GP is a stochastic process  $(T_x)_{x \in \mathcal{X}}$  in which each finite subset of points  $\{T_{x_1}, T_{x_2}, \dots, T_{x_n}\}$  follows a multivariate normal distribution [35]:

$$\begin{pmatrix} T_{x_1} \\ \vdots \\ T_{x_n} \end{pmatrix} \sim \mathcal{N}(\mathbf{m}, \mathbf{K}), \text{ where} \quad (21)$$

$$\mathbf{m} = \begin{pmatrix} m(x_1) \\ \vdots \\ m(x_n) \end{pmatrix} \text{ and}$$

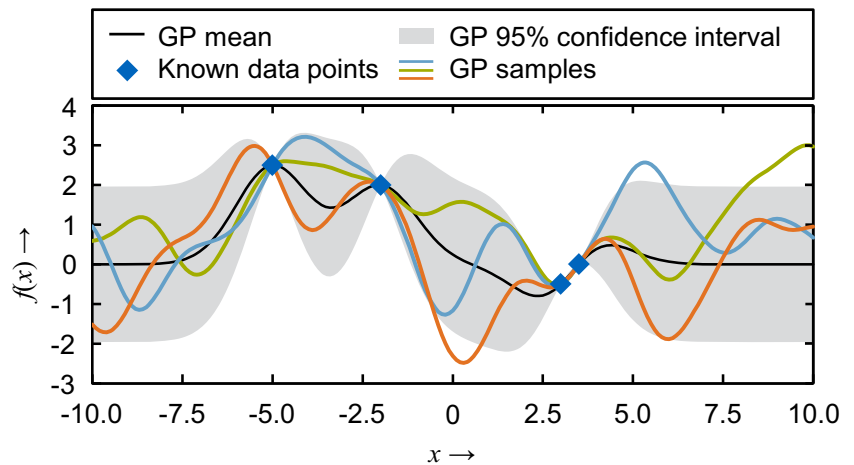
$$\mathbf{K} = \begin{pmatrix} k(x_1, x_1) & \cdots & k(x_n, x_1) \\ \vdots & \ddots & \vdots \\ k(x_1, x_n) & \cdots & k(x_n, x_n) \end{pmatrix}$$

Thereby  $\mathcal{N}$  denotes the normal distribution, and  $m(x): \mathcal{X} \rightarrow \mathbb{R}$  is the mean function which describes the expected value of the random variable  $T_x$  [35]:

$$m(x) = \begin{cases} y & \text{if } (x, y) \in \mathcal{D} \\ 0 & \text{else} \end{cases} \quad (22)$$

and  $k(x, x'): \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  is the covariance function, which can be represented by a kernel function. The covariance function describes the relationship between the two random variables  $T_x$  and  $T_{x'}$ , respectively, the similarity between the two points  $x$  and

**Fig. 16** Example of a Gaussian process [35]



$x'$ . In particular, an unknown point  $x_{n+1}$  with the corresponding random variable  $T_{x_{n+1}}$  is normally distributed, too [35]:

$$\begin{aligned} T_{x_{n+1}} | T_{x_1}, \dots, T_{x_n} &\sim \mathcal{N}(\mu, \sigma) \\ \mu &= \mathbf{k}^T \mathbf{K}^{-1} \mathbf{m}, \quad \sigma = k(x_{n+1}, x_{n+1}) - \mathbf{k}^T \mathbf{K}^{-1} \mathbf{k}, \quad \mathbf{k} = \begin{pmatrix} k(x_{n+1}, x_1) \\ \vdots \\ k(x_{n+1}, x_n) \end{pmatrix} \end{aligned} \tag{23}$$

where  $T_{x_{n+1}} | T_{x_1}, \dots, T_{x_n}$  is the random variable  $T_{x_{n+1}}$  in condition on  $T_{x_1}, \dots, T_{x_n}$ .

A kernel  $k(x, x') : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$  describes the similarity between the two points  $x$  and  $x'$ . The Matérn 5/2 kernel  $k_E$  is isotropic, i.e., the kernel depends only on the distance between the points  $x$  and  $x'$  [35]. The bigger the distance between the two points  $x$  and  $x'$ , the more different they are from each other and the smaller  $k_E(x, x')$  is [35, 36]:

$$k_E(x, x') = \sigma^2 \left( 1 + \frac{\sqrt{5} \|x-x'\|_2}{l} + \frac{5 \|x-x'\|_2^2}{3 \cdot l^2} \right) \cdot \exp \left( -\frac{\sqrt{5} \|x-x'\|_2}{l} \right) \tag{24}$$

where  $\| \cdot \|_2$  is the Euclidean norm. The Matérn 5/2 kernel has the two hyperparameters length-scale  $l \in \mathbb{R}^+$  and variance  $\sigma \in \mathbb{R}^+$ . Thereby, the length-scale  $l$  scales the distance and the variance  $\sigma$  scales the value of the function [37]. Along with the radial basis function (RBF) kernel [35], the Matérn 5/2 kernel is a common choice for kernel functions in the Euclidean space [35].

The coregionalization kernel  $k_c(x, x') : \{1, 2, \dots, n\} \times \{1, 2, \dots, n\} \rightarrow \mathbb{R}$  is used to model the similarity of the output dimensions of a function with a multi-dimensional output [38] and is defined as [36, 39]:

$$k_c(x, x') = (\mathbf{W}\mathbf{W}^T + \text{diag}(\boldsymbol{\kappa}))_{x,x'} \tag{25}$$

where

- $n$  is the number of output dimensions,
- $\{1, 2, \dots, n\}$  is the set of natural numbers less than or equal to  $n$ ,
- $\text{diag}(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$  is a function that maps a vector with  $n$  dimensions to an  $n \times n$  diagonal matrix,
- $(\dots)_{x,x'}$  is the  $x'$ -th entry of the  $x$ -th row of the corresponding matrix,
- $\mathbf{W} \in \mathbb{R}^{n \times m}$  and  $\boldsymbol{\kappa} \in \mathbb{R}^n$  are hyperparameters and are learned by hyperparameter optimization, and
- $m$  is an arbitrary natural number, but usually smaller than the variable  $n$ .

With the aid of the coregionalization kernel, functions with multi-dimensional output can be modeled using a GP [38]. Assuming  $f : \mathcal{X} \rightarrow \mathbb{R}^n$  is any function, then  $f$  can be represented by  $f'(x, i) : \mathcal{X} \times \{1, 2, \dots, n\} \rightarrow \mathbb{R}$  with  $f'(x, i) = f(x)_i$ . This means that the index of the output of  $f$  is seen as an input for the function  $f'$  [38]. The similarity between the individual output dimensions of  $f$  can be modeled with the coregionalization kernel [38]. This results in a combined kernel [38]:

$$k_{\text{multi}} \left( (x, i), (x', i') \right) = k(x, x') \cdot k_c(i, i') \tag{26}$$

where  $k$  is the kernel which measures the similarity between the points  $x$  and  $x'$ , and  $k_c$  is the coregionalization kernel which measures the similarity between the different output dimensions [38].

A GP can have hyperparameters such as the length-scale  $l$  and the variance  $\sigma$  parameter of the Matérn 5/2 kernel. These hyperparameters are usually determined using a maximum likelihood estimate (MLE) [40], i.e., the hyperparameters  $\mathbf{H}$  that best explain the data  $\mathcal{D}$  are searched [35, 37, 40]:

$$\arg \max_{\mathbf{H}} p(\mathcal{D} | \mathbf{H}) \tag{27}$$

If the GP has many hyperparameters and is optimized with too few data points, the hyperparameters may overfit on the data points, i.e., they only explain the data that was used to optimize, but not new data [35]. This can be prevented by assuming that the hyperparameters follow a certain distribution  $p(\mathbf{H})$  (e.g., a uniform, normal, logarithmic normal, or gamma distribution) [40]. This expresses which hyperparameters are likely and which are not. The hyperparameters are then determined using a maximum a posteriori (MAP) estimate [41]. In contrast to the MLE, the MAP estimation searches for the hyperparameters  $\mathbf{H}$  that are most likely for the given data  $\mathcal{D}$  [35, 40]:

$$\arg \max_{\mathbf{H}} p(\mathbf{H} | \mathcal{D}) = \arg \max_{\mathbf{H}} p(\mathcal{D} | \mathbf{H}) \cdot p(\mathbf{H}) \tag{28}$$

### Bayesian optimization

The Bayesian optimization (BO) is a class of machine learning-based optimization methods focused on solving the problem [40]:

$$\min_{x \in \mathcal{X}} f(x) \tag{29}$$

The BO is suited, when the set  $\mathcal{X}$  and the objective function  $f$  have the following properties [40]:

- It is expensive to evaluate the function  $f$ .
- The function  $f$  is a black box, i.e., only information about the input and output of the function is known. It is unknown how the function calculates the output.
- There is no information about the gradient of the function  $f$ .
- The output of  $f$  may contain an error.
- The number of dimensions of  $\mathcal{X}$  is in the order of 20 or less.
- The function  $f$  is continuous.

The BO is an iterative algorithm which evaluates the function  $f$  at a certain point in each iteration. In order for the algorithm to know at which point  $f$  should be evaluated next, the function  $f$  is first modeled using a GP and the already known points  $P$ . The strategy according to which the BO then selects the point to be evaluated next is determined by the maximum of the acquisition function  $a(x): \mathcal{X} \rightarrow \mathbb{R}$  [40]

One acquisition function is the probability of improvement (PI). This acquisition function indicates how likely it is that a

point will improve the previous optimum. The probability of improvement is defined as [42]:

$$a_{PI}(x') = P_{T_{x'}} [ T_{x'} \leq f_{\min} | T_{x_1}, \dots, T_{x_n} ] \tag{30}$$

where  $f_{\min}$  is the best point evaluated so far (in this work the goal was to minimize the surface defects),  $T_{x'}$  is the GP at point  $x'$ , and  $T_{x_1}, \dots, T_{x_n}$  is the GP at the points evaluated up to now. The disadvantage of this acquisition function is that it only indicates the probability of an improvement, but not the magnitude of improvement [42].

Another acquisition function that also includes the magnitude of improvement is the expected improvement (EI). EI provides a good balance between exploration and exploitation [43]. The EI acquisition function is defined as [44]:

$$a_{EI}(x') = \mathbb{E}_{T_{x'}} [ \max \{ 0, f_{\min} - T_{x'} \} | T_{x_1}, \dots, T_{x_n} ] \tag{31}$$

The BO can also be used to optimize a function  $f: \mathcal{X} \rightarrow \mathbb{R}^n$  with multi-dimensional output [45]. The BO algorithm is described in Algorithm A3. At the beginning,  $m \in \mathbb{N}$  random samples are taken, first to investigate the function to be optimized, and second so that there is at least one point with which the function  $f$  can be modeled in the first iteration.

#### Algorithm A3 Bayesian optimization algorithm [40]

```

1 take  $m \in \mathbb{N}$  random samples from  $\mathcal{X}$ 
2 be  $P = \{(x_1, f(x_1)), \dots, (x_m, f(x_m))\} \subseteq \mathcal{X} \times \mathbb{R}$  the number of known data points
3 loop:
4 | model  $f$  with a GP and the currently known data points  $P$ 
5 |  $x' = \operatorname{argmax}_{x \in \mathcal{X} \setminus \{x_1, \dots, x_m\}} a(x)$ 
6 |  $P \leftarrow P \cup \{(x', f(x'))\}$ 
7 as long as wanted
8 the searched optimum is the optimum of  $P$ 
    
```

In Algorithm A3,  $a(x)$  is any acquisition function and  $\mathcal{X} \setminus \{x_1, \dots\}$  is the set  $\mathcal{X}$  without the points already evaluated.

### Random search

Random search (RS) is an optimization algorithm that evaluates the function to be optimized at random points until a stop criterion is reached. The searched optimum is the optimum of the evaluated points [46, 47]. The advantages are that no requirements are placed on the function to be optimized and no gradients are required [48]. In addition, RS is useful if a significant part of the inverse image  $\mathcal{X}$  maps to an optimum.

## Appendix II

**Table 4** Hyperparameters of the Matérn 5/2 kernel, its distribution, and the parameters of the distribution for the single-task approach

Hyperparameter	Distribution	Parameters of the distribution
$l$	Uniform distribution	$a = 0.1; b = 2$
$\sigma$	Logarithmic normal distribution	$\mu = 0; \sigma = 0.5$

**Table 5** Hyperparameters for the multi-task approach, its distribution, and the parameters of the distributions

Kernel	Hyperparameter	Distribution	Parameters of the distribution
$k_E$	$l$	Uniform distribution	$a = 0.1; b = 2$
	$\sigma$	Logarithmic normal distribution	$\mu = 0; \sigma = 0.5$
$k_M$	$\kappa$	Logarithmic normal distribution	$\mu = 0; \sigma = 0.5$
	$W$	Normal distribution	$\mu = 1; \sigma = 1$
$k_S$	$\kappa$	Logarithmic normal distribution	$\mu = 0; \sigma = 0.5$
	$W$	Normal distribution	$\mu = 1; \sigma = 1$
$k_I$	$\kappa$	Logarithmic normal distribution	$\mu = 0; \sigma = 0.5$
	$W$	Normal distribution	$\mu = 1; \sigma = 1$

**Table 6** Number of steps required to achieve a parameter set that was among the best 20% when using BO

Study no.	Exp. no.	Number of exp.	$\alpha_{20}$	RS mean (std. dev.)	PI		EI	
					Single-task mean (std. dev.)	Multi-task	Single-task mean (std. dev.)	Multi-task
1	1–16	16	3	4.3 (2.9)	2.6 (1.0)	1	2.7 (1.2)	1
2	17–32	16	3	4.3 (2.9)	4.1 (2.7)	1	3.1 (1.5)	1
3	33–48	16	3	4.3 (2.9)	2.6 (1.1)	1	2.8 (1.4)	1
4	49–64	16	3	4.3 (2.9)	2.6 (1.1)	1	2.6 (1.1)	1
5	65–80	16	3	4.3 (2.9)	2.8 (1.2)	1	2.8 (1.4)	1
6	81–96	16	3	4.3 (2.9)	3.6 (2.3)	1	3.0 (1.4)	1
7	97–112	16	3	4.3 (2.9)	2.9 (1.4)	11	2.8 (1.3)	11
8	113–128	16	3	4.3 (2.9)	2.9 (1.5)	1	2.8 (1.3)	1
9	129–141	13	3	3.5 (2.3)	4.4 (2.0)	5	3.2 (1.5)	5
10	142–262	121	24	4.9 (4.2)	4.5 (2.9)	1	6.6 (3.2)	1
				Ø4.3	Ø3.3	Ø2.4	Ø3.2	Ø2.4

**Table 7** Number of steps required to achieve a parameter set that was among the best 5% when using BO

Study no.	Exp. no.	Number of exp.	$\sigma_5$	RS mean (std. dev.)	PI		EI	
					Single-task mean (std. dev.)	Multi-task	Single-task mean (std. dev.)	Multi-task
1	1–16	16	1	8.5 (4.6)	2.9 (0.9)	1	3.3 (1.3)	1
2	17–32	16	1	8.5 (4.6)	13.1 (4.0)	16	9.7 (4.5)	16
3	33–48	16	1	8.5 (4.6)	3.1 (1.1)	1	3.4 (1.5)	1
4	49–64	16	1	8.5 (4.6)	4.0 (1.1)	1	6.9 (2.0)	1
5	65–80	16	1	8.5 (4.6)	3.1 (1.1)	1	8.5 (3.4)	1
6	81–96	16	1	8.5 (4.6)	4.0 (2.4)	1	3.3 (1.3)	1
7	97–112	16	1	8.5 (4.6)	3.5 (1.7)	11	3.6 (1.5)	11
8	113–128	16	1	8.5 (4.6)	5.8 (1.6)	1	7.2 (1.9)	1
9	129–141	13	1	7.0 (3.7)	5.6 (1.7)	5	5.0 (2.1)	5
10	142–262	121	7	17.4 (14.7)	13.0 (8.7)	9	16.5 (10.9)	8
				Ø9.2	Ø5.8	Ø4.7	Ø6.7	Ø4.6

**Table 8** Data sets used in each case to calculate suitable parameter settings for the investigations without duplicates

Study no.	Auxiliary data sets	Study no.	Auxiliary data sets
1	2, 3, 4, 5, 7, 10	6	1, 2, 3, 4, 7, 10
2	1, 3, 4, 5, 7, 10	7	1, 2, 3, 4, 5, 10
3	1, 2, 4, 5, 7, 10	8	1, 2, 3, 5, 7, 10
4	1, 2, 3, 5, 7, 10	9	1, 2, 4, 5, 7, 10
5	1, 2, 3, 4, 7, 10	10	1, 2, 3, 4, 5, 7

**Table 9** Comparison of the necessary steps until suitable parameter settings have been found with and without duplicates when using BO

Study no.	With duplicates				Without duplicates				
	PI		EI		PI		EI		
	20%	5%	20%	5%	20%	5%	20%	5%	
1	1	1	1	1	2	3	2	3	
2	1	16	1	16	1	13	1	13	
3	1	1	1	1	1	1	1	1	
4	1	1	1	1	1	3	1	3	
5	1	1	1	1	1	2	1	2	
6	1	1	1	1	1	2	1	2	
7	11	11	11	11	2	2	2	2	
8	1	1	1	1	1	2	1	2	
9	5	5	5	5	4	4	3	3	
10	1	9	1	8	1	12	1	5	
Ø2.4   Ø4.7   Ø2.4   Ø4.6   Ø1.5   Ø4.4   Ø1.4   Ø3.6									

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**References**

- Colligan KJ (2010) The friction stir welding process: an overview. In: Lohwasser D, Chen Z (eds) Friction stir welding—from basics to applications. Woodhead Publishing Limited and CRC Press LLC, Cambridge, pp 15–41 ISBN: 978-1-84569-450-0
- Zuo L, Zuo D, Zhu Y, Wang H (2018) Effect of process parameters on surface topography of friction stir welding. Int J Adv Manuf Technol 98:1807–1816. <https://doi.org/10.1007/s00170-018-2326-x>
- Rajakumar S, Muralidharan C, Balasubramanian V (2010) Optimization of the friction-stir-welding process and tool parameters to attain a maximum tensile strength of AA7075–T 6 aluminium alloy. Proc Inst Mech Eng B J Eng Manuf 224:1175–1191 8. <https://doi.org/10.1243/09544054JEM1802>
- Farzadi A, Bahmani M, Haghshenas DF (2017) Optimization of operational parameters in friction stir welding of AA7075-T6 aluminum alloy using response surface method. Arab J Sci Eng 42: 4905–4916 11. <https://doi.org/10.1007/s13369-017-2741-6>
- Montgomery DC (2017) Design and analysis of experiments. John Wiley & Sons Inc., Hoboken ISBN: 9781119113478
- Box GEP, Wilson KB (1981) On the experimental attainment of optimum conditions. J R Stat Soc 13(1–45):1

7. Taguchi G (1986) Introduction to quality engineering. Asian Productivity Organization, Tokyo ISBN: 9283310845
8. Unal R, Dean EB (1991) Taguchi approach to design optimization for quality and cost: an overview. In: International Society of Parametric Analysts (ed) Proceedings of the 13th Annual Conference of the International Society of Parametric Analysts. International Society of Parametric Analysts, Vienna, pp 1–20
9. Lakshminarayanan AK, Balasubramanian V (2008) Process parameters optimization for friction stir welding of RDE-40 aluminium alloy using Taguchi technique. *Trans Nonferrous Metals Soc China* 18(548–554):3. [https://doi.org/10.1016/S1003-6326\(08\)60096-5](https://doi.org/10.1016/S1003-6326(08)60096-5)
10. Ugender S, Kumar A, Somi Reddy A (2015) Effect of friction stir welding process parameters on the mechanical properties of AA 6061 aluminium alloy using Taguchi orthogonal technique. *Appl Mech Mater* 813–814:431–437. <https://doi.org/10.4028/www.scientific.net/AMM.813-814.431>
11. Ganapathy T, Lenin K, Pannerselvam K (2017) Process parameters optimization of friction stir welding in aluminium alloy 6063-T6 by Taguchi method. *Appl Mech Mater* 867:97–104. <https://doi.org/10.4028/www.scientific.net/AMM.867.97>
12. Abbas AA, Abdulkadhun HH (2019) Optimization of friction stir welding process parameters to joint 7075-T6 aluminium alloy by utilizing Taguchi technique. *J Eng* 25(1–15):5–15. <https://doi.org/10.31026/j.eng.2019.05.01>
13. Ma Z, Li Q, Ma L, Hu W, Xu B (2019) Process parameters optimization of friction stir welding of 6005A-T6 aluminum alloy using Taguchi technique. *Trans Indian Inst Metals* 72(1721–1731):7–1731. <https://doi.org/10.1007/s12666-019-01639-7>
14. Vijayan S, Raju R, Rao SRK (2010) Multiobjective optimization of friction stir welding process parameters on aluminum alloy AA 5083 using Taguchi-based Grey relation analysis. *Mater Manuf Process* 25(1206–1212):11–1212. <https://doi.org/10.1080/10426910903536782>
15. Deng J (1989) Introduction to grey system theory. *J Grey Syst* 1(1–24):1
16. Cochran WG, Cox GM (1957) Experimental designs. John Wiley & Sons, Inc., Hoboken ISBN: 0-471-16204-3
17. Mehri Khansari N, Berto F, Karimi N, Ghoreishi SMN, Fakoor M, Mokari M (2018) Development of an optimal process for friction stir welding based on GA-RSM hybrid algorithm. *Frattura ed Integrità Strutturale* 12(106–122):44–122. <https://doi.org/10.3221/IGF-ESIS.44.09>
18. Sivanandam SN, Deepa SN (2008) Introduction to genetic algorithms. Springer, Berlin ISBN: 978-3-540-73189-4
19. Tansel IN, Demetgul M, Okuyucu H, Yapici A (2010) Optimizations of friction stir welding of aluminium alloy by using genetically optimized neural network. *Int J Adv Manuf Technol* 48:95–101. <https://doi.org/10.1007/s00170-009-2266-6>
20. Tansel IN, Yang SY, Shu C, Bao WY, Mahendrakar N (1999) Introduction to genetically optimized neural network systems (GONNS). In: Dagli CA (ed) Smart engineering systems: neural networks, fuzzy logic, evolutionary programming, data mining, and rough sets. ASME Press, New York, pp 331–336 ISBN: 0791800989
21. Trueba L, Torres MA, Johannes LB, Rybicki D (2018) Process optimization in the self-reacting friction stir welding of aluminum 6061-T6. *Int J Mater Form* 11(559–570):4–570. <https://doi.org/10.1007/s12289-017-1365-4>
22. Shigematsu I, Kwon Y-J, Saito N (2009) Dissimilar friction stir welding for tailor-welded blanks of aluminum and magnesium alloys. *Mater Trans* 50(197–203):1–203. <https://doi.org/10.2320/matertrans.MER2008326>
23. Hartl R, Bachmann A, Liebl S, Zens A, Zaeh MF (2019) Automated surface inspection of friction stir welds by means of structured light projection. *IOP Conf Ser Mater Sci Eng* 480:12035. <https://doi.org/10.1088/1757-899X/480/1/012035>
24. Hartl R, Praehofer B, Zaeh MF (2020) Prediction of the surface quality of friction stir welds by the analysis of process data using artificial neural networks. *Proc Inst Mech Eng L J Mater Des Appl* 234:732–751 5. <https://doi.org/10.1177/1464420719899685>
25. Sutton RS, Barto A (2018) Reinforcement learning. The MIT Press, Cambridge ISBN: 9780262039246
26. Hartl R, Landgraf J, Spahl J, Bachmann A, Zaeh MF (2019) Automated visual inspection of friction stir welds: a deep learning approach. In: Stella E (ed) Multimodal sensing: technologies and applications. Society of Photo-Optical Instrumentation Engineers (SPIE). Bellingham, Washington, pp 1–24. <https://doi.org/10.1117/12.2525947>
27. Richter B (2017) Robot-based friction stir welding for E-mobility and general applications. *Biuletyn Instytutu Spawalnictwa* 2017:103–110 5. <https://doi.org/10.17729/ebis.2017.5/11>
28. DIN EN ISO (2012) DIN EN ISO 25178-2:2012 geometrical product specifications (GPS)—surface texture: areal—part 2: terms, definitions and surface texture parameters; German version. Beuth Verlag GmbH, Berlin
29. Grande R, Walsh T, How J (2014) Sample efficient reinforcement learning with Gaussian processes. In: Xing EP, Jebara T (eds) *JMLR: W&CP volume 32*. PMLR, London, pp 1332–1340
30. Mononen J, Sirén M, Hänninen H (2003) Cost comparison of FSW and MIG welded aluminium panels. *Weld World* 47(32–35):11–12. <https://doi.org/10.1007/BF03266406>
31. Rodrigues DM, Leitão C, Louro R, Gouveia H, Loureiro A (2010) High speed friction stir welding of aluminium alloys. *Sci Technol Weld Join* 15:676–681 8. <https://doi.org/10.1179/136217110X12785889550181>
32. Bischoff B (2015) Reinforcement learning for industrial applications. Dissertation, Technical University of Munich
33. Wiering M, van Otterlo M (2012) Reinforcement learning. Springer, Berlin ISBN: 9783642276453
34. Kaelbling LP, Littman ML, Moore AW (1996) Reinforcement learning: a survey. *J Artif Intell Res* 4:237–285
35. Rasmussen CE, Williams CKI (2006) Gaussian processes for machine learning. The MIT Press, Cambridge ISBN: 978262182539
36. Sheffield machine learning group (2015) GPy documentation
37. Murphy KP (2012) Machine learning. MIT Press, Cambridge ISBN: 9780262018029
38. Alvarez MA, Rosasco L, Lawrence ND (2011) Kernels for vector-valued functions: a review. *Found Trends Mach Learn* 4:195–266 3
39. Sheffield ML (2019) *GPy/coregionalize.py* at 40137cc8f7e0794bff55639ec55d4884c72e86b5. SheffieldML/GPy. GitHub. <https://github.com/SheffieldML/GPy/blob/40137cc8f7e0794bff55639ec55d4884c72e86b5/GPy/kern/src/coregionalize.py>. visited on: September 29, 2019
40. Frazier P I (2018) A tutorial on Bayesian optimization. arXiv: 1807.02811
41. Gelman A, Carlin J B, Stern H S, Dunson D B, Vehtari A and Rubin D B (2014) Bayesian data analysis. CRC Press Taylor and Francis Group, Boca Raton, Florida. ISBN: 9781439840955
42. Brochu E, Cora V M and Freitas N (2010) A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv: 1012.2599v1
43. Berger-Tal O, Nathan J, Meron E, Saltz D (2014) The exploration-exploitation dilemma: a multidisciplinary framework. *PLoS One* 9(e95693):4. <https://doi.org/10.1371/journal.pone.0095693>
44. Jones DR, Schonlau M, Welch WJ (1998) Efficient global optimization of expensive black-box functions. *J Glob Optim* 13:455–492 4. <https://doi.org/10.1023/A:1008306431147>
45. Swersky K, Snoek J, Adams RP (2014) Multi-task Bayesian optimization. In: Burges C (ed) *Advances in neural information processing systems* 26. Curran, Red Hook, pp 2004–2012 ISBN: 9781632660244

46. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. *J Mach Learn Res* 13:281–305
47. Brooks SH (1958) A discussion of random methods for seeking maxima. *Oper Res* 6:244–251 2. <https://doi.org/10.1287/opre.6.2.244>
48. Hutter F, Kotthoff L, Vanschoren J (2019) Automated machine learning. Springer International Publishing, Cham, Switzerland ISBN: 978-3-030-05317-8

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.