Modelling of hardfacing layers deposition parameters using robust machine learning algorithms

Modelling of hardfacing layers deposition parameters using robust machine learning algorithms

K Zając[1], K Płatek[2], P Biskup[2] and L Łatka[2]

[1] Faculty of Microsystem Electronics and Photonics, Wrocław University of Science and Technology, 11 Janiszewskiego Street, 50-372 Wrocław, Poland
[2] Faculty of Mechanical Engineering, Wrocław University of Science and Technology, 5 Łukasiewicza Street, 50-371 Wrocław, Poland

piotr.biskup@pwr.edu.pl

Abstract. The study presents a data-driven framework for modelling parameters of hardfacing deposits by GMAW using neural models to estimate the influence of process parameters without the need of creating experimental samples of the material and detailed measurements. The process of GAS Metal Arc Welding (GMAW) hardfacing does sometimes create non-homogenous structures in the material not only in deposited material, but also in the heat-affected zone (HAZ) and base material. Those structures are not fully deterministic, so the modelling method should account for this unpredictable component and only learn the generic structure of the hardness of the resulting material. Artificial neural networks (ANN) were used to create a model of the process using only measured samples without any knowledge of equations governing the process. Robust learning was used to decrease the influence of outliers and noise in the measured data on the neural model performance. The proposed method relies on modification of the loss function and several of them are compared and evaluated as an attempt to construct general framework for analysing the hardness as a function of electric current and arc velocity. The proposed method can create robust models of the hardfacing layers deposition or other welding processes and predict the properties of resulting materials even for unseen parameters based on experimental data. This modelling framework is not typically used for metallurgy, and it requires further case studies to verify its generalisability.

1. Introduction

Increasing the tribological and mechanical properties of the manufactured materials is one of the main interests of material science research. One of the methods widely used for improving them is hardfacing. Hardfacing is a method, where the layer of material with specified properties is deposited on the base material to improve its surface properties such as hardness or grindability [1]. Due to the given various possibilities containing reconstruction, protective coatings or even composite materials, the allotment of hardfacing technologies increases [2, 3]. The economic advantages of its applications cannot be dissembled either since hardfacing is still one of the most financially rewarding methods for surface properties improvement.

In the hardfacing process, welding is used for depositing the alloy with specified properties on the base material (which is typically low carbon steel). The variety of alloys that can be used in the hardfacing process and methods itself gives numerous possible applications [1, 4]. Across many different methods used for hardfacing, the GMAW method, also known as gas-shielded metal arc welding, is one of the most used due to its numerous advantages like high deposition rate or no slag

formation. Similarly, to MMA welding, the heat in the GMAW process is produced by an arc formed between the workpiece surface and consumable electrode in form of a small diameter wire which melts to create the bead [5]. There are two groups of metal gas arc methods used dependently on the composition of shielding gas mixture: in Metal Inert Gas (MIG) welding argon and helium are typically used, while for Metal Active Gas (MAG) the shielding gas is usually a mixture of argon, carbon dioxide and oxygen.

Material science interests include not only improving the properties of the materials themselves, but also potentially useful methods and tools. Simulations are a useful tool for analysing many engineering processes [6–8]. They require a model of the process of interest, which can be derived using knowledge about the process itself or experimental measurements. Those are distinct cases of modelling, which are called white-box models and black-box models. White box models rely on knowledge of underlying rules governing the system, for example, laws of physics. Black box models on the other hand require only input and output measurements of the system. Machine learning is one of the largest active areas of research related to black-box modelling of various systems and it has been successfully applied to metallurgy, material science and simulation of different phenomena [9-11]. One particularly useful black-box modelling method is using artificial neural networks (ANN). Sigmoidal neural networks are universal approximators as shown in [12, 13], hence they are a perfect tool for creating a data-driven black-box model of hardfacing layers deposition parameters.

Steel demonstrates isotropic behaviour, although its microscopic structure is non-homogeneous. Alloying elements such as molybdenum, manganese, silicon or chromium even added in small amounts, highly affects steel properties and microstructure as well. However, material's structure depends not only on its chemical formula, but also on the processes parameters, which impact physical and chemical phenomena, which cannot be ignored. These phenomena are leading to creating non-homogeneous structures containing elements segregation or carbide precipitation while properties (like hardness or grindability) of these elements vary from matrix properties. The problem is that character of these phenomena is partly randomized and hard to predict. A strong impact on the material properties of these elements may lead to problems when trying to model those properties. In that case, it is important to treat highly diverged measurements as outliers. This procedure additionally can be used for the elimination of human error since measurements are usually taken manually.

The concept of robustness often comes up in statistics and in general relates to creating models which are less influenced by outliers than by inliers. There are numerous statistical techniques of achieving robustness in simple regression models, such as Huber regression [14], LASSO [15, 16], Ridge regression [17] and other [18–21]. In this study outlier robust regression problem is considered, it is explained in detail in [22]. In machine learning, particularly using neural networks, a common method of achieving desired robustness concerning data with errors or some degree of randomness is usually done by modifications of loss function during training. This method applied to hardfacing layers deposition parameters is a novel technique, which can improve existing results of modelling in terms of robustness and generalizability. All regression models were created, trained and evaluated using Python scientific environment, including libraries such as numpy, pandas, scikit-learn, tensorflow and matplotlib. Experimental measurements were used as input data required by the models.

2. Materials and Method

2.1. Deposition of the hardfacing layers

For the following case, the structural steel S235 had been used as a base material in the form of a plate with dimensions 150 mm x 100 mm and thickness equal to 10 mm. The chemical composition of base material according to EN 10025 is shown in Table 1 [23]. Both the filler material and shielding gas were invariable. As shielding gas, a mixture of argon with 2 vol % of carbon dioxide was applied. The gas flow rate was c.a. 18 slpm. The filler material was a wire with a diameter equal to 1.2 mm. The chemical composition of the applied wire is shown in Table 2 [24].

The hardfacing process was carried out in KGHM Polish Copper, Copper Smelter Legnica using the GMAW method. The current source was Lincoln Electric POWERTEC 505S. Before surfacing, the

surfaces of the base material were prepared using subtractive processing - grinding. The process was carried out on a magnetic plane grinder. Just before hardfacing, preheating up to 150°C was applied. On the other hand, after the process, tempering also was applied.

Table 1. Chemical composition of S235 steel according to EN 10025 standard, [wt.%].

C	Si	Mn	P	S	Fe
0.22	0.05	1.6	0.05	0.05	Balance

Table 2. The chemical composition of filler material according to EN 14700 standard, [wt.%].

C	Si	Mn	Cr	Fe
0.5	3.0	0.5	9.5	Balance

The primary aim of the study was to create a model of the impact of different input parameters on basic material properties. From all input parameters, the electric current and welding speed were chosen as the variable parameters. The electric current used was direct with positive polarity (DC+) as it is preferred MAG welding polarity [25]. Sample codes and variable parameters were collected in Table 3. Another variable factor was beads configuration, which is presented in Figure 1.

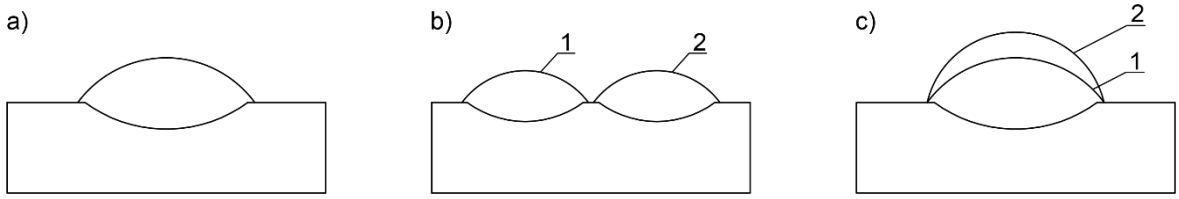

Figure 1. Diagram of bead placement: a) single bead; b) second bead (2) next to the first one (1); c) second bead (2) on top of the first one (1).

Table 3. Variable parameters and sample codes.

	1-1	**1-2**	**1-3**	**2-1**	**2-2**	**2-3**
Electric Current [*A*]	136	170	204	170	170	170
Welding Speed [*mm/s*]	10	10	10	8	10	12
	3-1	**3-2**	**3-3**	**4-1**	**4-2**	**4-3**
Electric Current [*A*]	136	170	204	170	170	170
Welding Speed [*mm/s*]	10	10	10	8	10	12
	5-1	**5-2**	**5-3**	**6-1**	**6-2**	**6-3**
Electric Current [*A*]	136	170	204	170	170	170
Welding Speed [*mm/s*]	10	10	10	8	10	12

Single bead configuration was applied for 1- 2- samples (Fig. 1a). For the 3- and 4- series second bead was placed next to the first one (Fig. 1b), while for 5- 6- samples the second bead was deposited on top of the first one (Fig. 1c).

2.2. Measurement Process

The examination process included visual testing (VT), structure examination and hardness measurements. Visual tests were taken following ISO 6520-1 [26]. Materials structure examination performed using optical microscope consider both macroscopic and microscopic examination in search of weld imperfections (open porosity and discontinuities). Each sample had to be examined and evaluated with a view to carbides (mostly chromium) what was required due to proper interpretation of hardness measurements.

The hardness of the coatings was tested by the Vickers hardness test method (HV1) following ISO 6507 standard [27]. For modelling the characteristic, it was desirable to create multiple data points. The measurement was taken under 1 kgf load (test force 9.81 N) and hardness was conditioned not only on welding speed and electric current but also on the placement of measurement on the sample. In another case, it would not be possible to accurately model the dependency using neural networks. According to this requirement, the series of hardness measurements were taken vertically in the middle of the surface bead. The number of measurements taken for each sample was 19, with the first measure taken 0.5 mm from the weld face and the distance between each measurement had a constant value of 0.5 mm (Figure ***2***).

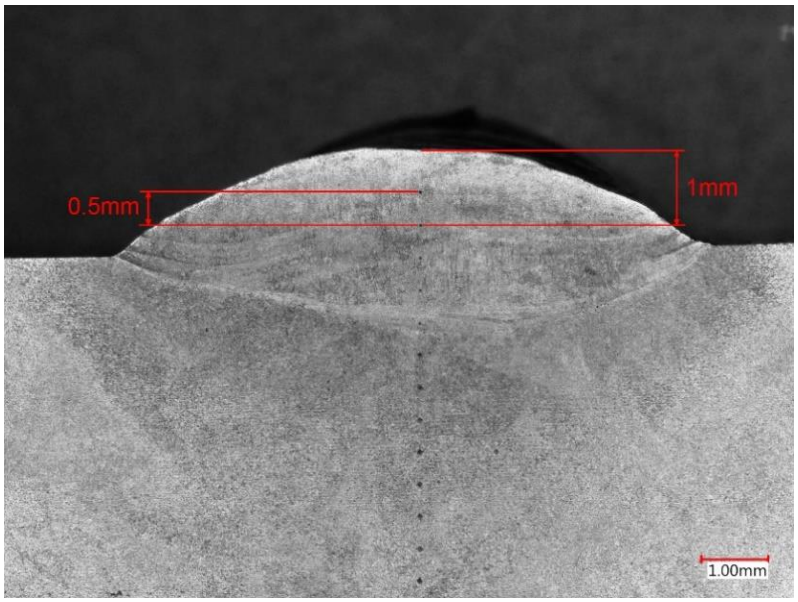

Figure 2. Scheme of HV1 measurements.

2.3. Robust Machine Learning

In this paper robust learning in terms of outlier robust regression using feedforward neural networks is discussed. The robustness is achieved by modifications of the loss function. There are other techniques of creating robust machine learning algorithms, for example, least trimmed squares algorithm for regression problems [28], robust support vector regression [29] and many others. This study aims at comparing various robust loss functions to improve the performance of neural modelling applied to modelling hardfacing layers deposition parameters. It also shows that for this problem tuneable loss function can be used, which can result in better performance and flexible robustness.

The most common loss function used for regression problems is mean squared error (abbreviated MSE), it gives the name to the least squares' algorithm [30], which is fundamental for many applications used across machine learning. When used as a neural network loss function MSE is computed as:

$$L_M(y, \hat{y}) = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad (1)$$

In the paper y is used to denote target variable observed values, also called true values and $\hat{y}$ denotes predicted values, while n denotes the number of predictions. Symbol L with subscripts is used to denote different loss functions. MSE meets the requirements for the regression loss function of being non-negative and smooth. Its disadvantage is a heavy penalty for outliers. This can be seen in Figure 3.

One simple robust loss function is the Log-Cosh function. Close to a minimum, it behaves similarly to MSE loss, but for values far from the origin, it is much smoother. It was introduced in [31] and successfully applied [32]. The Log-Cosh function is expressed as:

$$L_C(y,\hat{y}) = \sum_{i=1}^{n} log\left(cosh(y-\hat{y})\right) \tag{2}$$

This loss function is differentiable only two times, but this is sufficient to be used as a neural network loss function for regression since no currently used optimisation algorithm uses higher derivatives than second. Typically, loss functions are required to belong to a class $\mathbb{C}^{\infty}$ (infinite number of continuous derivatives), but for practical purposes $\mathbb{C}^2$ (two continuous derivatives) is enough.

Another commonly used robust loss function is Huber loss. It was introduced in [14] and has a parameter, which can control its shape, increasing or decreasing its sensitivity to outliers. The influence of the δ parameter can be seen in Figure 3. This loss function is formulated as:

$$L_H(y,\hat{y}) = \begin{cases} \frac{1}{2}(y-\hat{y})^2 & for\ |y-\hat{y}| \le \delta \\ \delta\left(|y-\hat{y}| - \frac{1}{2}\delta\right) & otherwise \end{cases} \tag{3}$$

Huber loss can be intuitively thought of as a combination of L_1 for extreme cases and L_2 for values close to the minimum.

One particularly useful robust loss function is presented in [33]. It is more generic than any other previously mentioned function since it introduces robustness as a continuous parameter of the function itself and a parameter controlling the size of a quadratic region near the origin. This function is therefore tuneable and can be used to find the optimal trade-off between sensitivity to errors and robustness. The generic form of loss is given by:

$$L_A(y,\hat{y},\alpha,c) = \frac{|\alpha-2|}{\alpha}\left(\left(1+\frac{\left(\frac{y-\hat{y}}{c}\right)^2}{|\alpha-2|}\right)^{\alpha/2} - 1\right) \tag{4}$$

$\alpha \in \mathbb{R}$ is a shape parameter that controls the robustness of the loss function, meaning it decreases the influence of outliers when it is set to small values. For large values, the function becomes more similar to MSE. $c > 0$ is a scale parameter that controls the size of the quadratic region near the origin. The influence of the α parameter for $c = 1$ is shown in Figure 4. For the function to be well-defined for all possible values of α special cases are introduced, for $\alpha = 2, \alpha = 0$ and $\alpha = -\infty, \alpha = \infty$. When $\alpha = 2$ the function resembles the MSE loss function, and it is derived by taking the limit (function is not defined for $\alpha = 2$):

$$\lim_{\alpha \to 2} L_A(y,\hat{y},\alpha,c) = \frac{1}{2}\left(\frac{y-\hat{y}}{c}\right)^2 \tag{5}$$

For $\alpha = 1$ the adaptive loss function closely resembles Huber loss. It can be intuitively thought of as a smoothed version of L_1 loss. In such case, it is given by:

$$L_A(y,\hat{y},1,c) = \left(1+\frac{y-\hat{y}}{c}\right)^{\frac{1}{2}} - 1 \tag{6}$$

For $\alpha = 0$ the loss is undefined, but taking the limit yields a special case, which is significantly less sensitive to outliers than two previous particular cases. The function is derived as:

$$\lim_{\alpha \to 0} L_A(y, \hat{y}, \alpha, c) = log\left(\frac{1}{2}\left(\frac{y-\hat{y}}{c}\right)^2 + 1\right) \tag{7}$$

The final special case is the least sensitive form of the loss, which is given by taking $\alpha = -\infty$. For this special case, the function is expressed using:

$$\lim_{\alpha \to -\infty} L_A(y, \hat{y}, \alpha, c) = 1 - exp\left(-\frac{1}{2}\left(\frac{y-\hat{y}}{c}\right)^2\right) \tag{8}$$

The loss function can be tuned to be more robust by decreasing α or to be more sensitive by increasing α. Functions for $\alpha > 2$ are overly sensitive and are not used frequently, since they weight outliers heavier than the MSE loss, which is already not robust. However, in some cases, when an extremely sensitive loss function is required, those forms can also be useful. The paper also introduces a probabilistic interpretation of the function, which allows it to be self-adaptive.

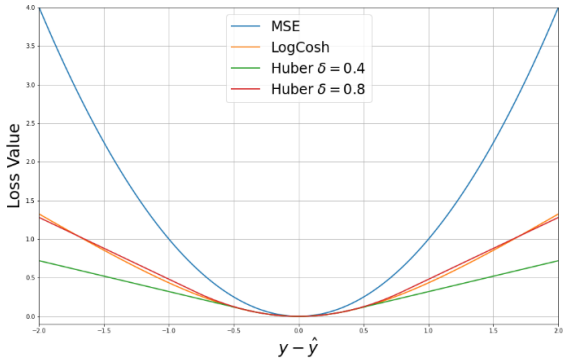

Figure 3. Generic and robust loss functions.

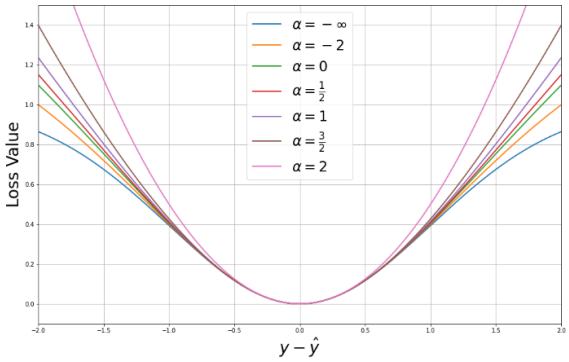

Figure 4. Adaptive loss function with different values of α

2.4. Model Development

The black-box model of the process further referred to simply as a model was developed using measurement data from experiments with different samples of the material. Stand-off distance, welding speed, electric current, number of beads and their type were used as input features and measured hardness was used as outputs of the model. This is shown on the neural model diagram in Figure ***5***. Stand-off distance, welding speed, electric current and number of beads were numerical features fed to the neural model after simple pre-processing, consisting of mean substitution and standardization. For categorical features, which had three possible values ordinal encoding was used [34]. Outputs were pre-processed using the same procedure as numerical features. Data were divided into three sets, firstly three randomly chosen material samples were left out entirely, those will be referred to as test data. Then remaining samples were organised into input-output pairs, which created 19 data points per sample since there were 19 measurements for each of them. Those pairs were randomly split into two sets, 70% was used for training the model, further referred to as training data set and 30% was used to evaluate the model on partly seen data, further referred to as validation data set. Those are commonly used ratios of train and test data used for model development [35]. Validation data and test data are both used to evaluate the models. During training one cross-validation batch was used, which consisted of 10% of randomly sampled training examples. At each epoch, loss function was computed for this batch, and if for k epochs the loss did not decrease the training was stopped. Values of k were derived experimentally [36], but kept constant for different loss functions to compare them as accurately as possible.

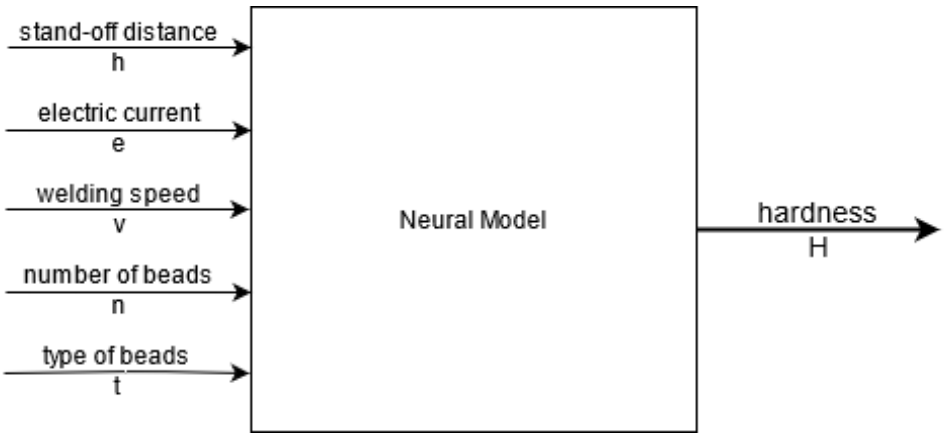

Figure 5. Block diagram of the neural model.

2.4.1. Model Selection. Base architecture of the ANN model was selected using only training and validation data and comparing between different feedforward neural networks. The variables considered were activation functions, number of hidden units and layers. Loss function and optimisation algorithms were fixed during this step, the chosen loss was a mean squared error, which is traditional for regression problems and chosen optimisation algorithm was RMS Prop [37, 38]. Results were compared to benchmark models such as linear regression and support vector machines [39, 40]. Chosen based network architecture was then fixed and trained using different loss functions, comparing robust and traditional regression losses. Models were evaluated using both test and validation data. Results are covered in Section 3.

2.4.2. Evaluation. Models were evaluated using regression metrics to compare their performance. The same set of metrics was used both for the base and robust models. Used metrics were mean square error, abbreviated MSE, mean absolute error (MAE), median absolute error (MDE) and R^2 coefficient. Additionally, for each model maximal error and its ratio with maximal true value are included, denoted as MAX. All metrics were computed for absolute and normalised values, normalised metrics were computed as:

$$e = \frac{\|y - \hat{y}\|_n}{\|y\|_n} \tag{9}$$

Where $\|\cdot\|_n$ refers to N-norm (used norms were L_1or L_2). Metrics are computed per each pair of true and predicted values and aggregated using mean or median. Normalised metrics can be greater than 1 since prediction could be off by more than 100%, but this is a very undesirable result. Median is especially useful since it does not highlight the influence of outliers, which are present in all training, validation, and test data. Additionally, for test samples, trajectories for the dependency of hardness conditioned on stand-off distance were created and empirically compared.

3. Results

3.1. Materials

In the following research, testing of hardfacing layers was focused mainly on its hardness, however, the material's structure and imperfections were also considered. In some cases, porosity in the weld bead could be observed (Figure 6). Microscopic observations revealed that in the structure of the hardfacing layers silicon oxides and chromium carbides could be found. Moreover, the amount of these carbides was explicitly dependent on the electric current value (Figure 7).

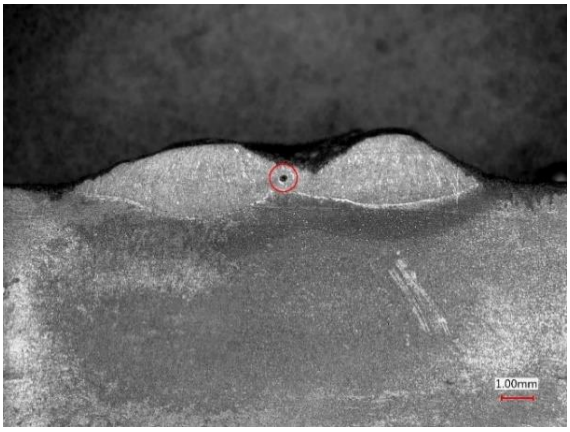

Figure 6. Microstructure of samples 4-3 with marked porosity. Light microscopy, etched with 5% Nital.

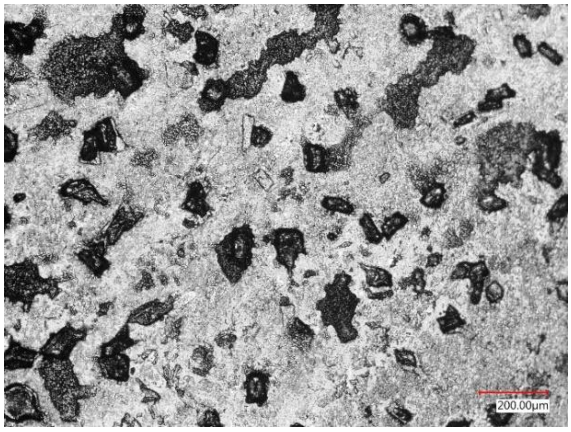

Figure 7. Microstructure of 6-3 samples. Light microscopy, etched with 5% Nital.

The Vickers hardness measured 1mm from the surface ranged from 377 for 1-3 up to 557 HV1 for samples 5-3, while the average base material hardness was c.a. 130 HV1. The average hardness for samples made by depositing one layer on top of another is visibly higher. For single bead samples, the correlation between welding speed and hardness is transparent and in accordance with the literature [41–43].

For samples 3 -1, 3-2, 3-3, 4-1, 4-2 and 4-3, which beads are placed next to each other, it is noticeable that the hardness of beads vary from each other and for each sample it is possible to indicate one with higher average hardness. The results of hardness measurements taken for each sample are shown in Table 4. For samples with two beads next to each other, the higher score was taken into consideration.

Table 4. Selected hardness measurements.

	1-1	**1-2**	**1-3**	**2-1**	**2-2**	**2-3**
Hardness [*HV1*]	442	405	377	469	526	506
	3-1	**3-2**	**3-3**	**4-1**	**4-2**	**4-3**
Hardness [*HV1*]	507	529	471	543	515	488
	5-1	**5-2**	**5-3**	**6-1**	**6-2**	**6-3**
Hardness [*HV1*]	528	538	557	538	492	515

3.2. Model Architecture

The chosen architecture was an MLP with 5 hidden layers with 10 units in each hidden layer and sigmoid activation functions. The first and last layers used linear activation. This model did not achieve the best results, but was marginally worse and is a good entry point for evaluating the influence of robust loss functions, which is the goal of the study. The model was compared against a simple linear regression model, SVR with parameters chosen optimally for the problem, and a neural model with a single hidden layer (also with 10 neurons), referred to as simple MLP (Table 5). Computed metrics are rounded to whole numbers since for hardness measurements fractions do not have any meaningful physical interpretation.

Table 5. Results for base hardness models

	Absolute				Normalised				
	MSE	MAE	MDE	MAX	MSE	MAE	MDE	MAX	R^2
Linear Regression	10941	88	89	235	1.198	0.856	0.678	0.377	0.563
SVR	6766	58	33	231	0.127	0.255	0.160	0.369	0.730
Simple MLP	6323	57	40	**222**	0.144	0.269	0.152	**0.355**	0.747
MLP	**5387**	**48**	**30**	266	**0.103**	**0.212**	**0.125**	0.425	**0.785**

3.3. Robust Model

For chosen architecture, the model was trained using robust loss functions. Results for three modified functions and one benchmark (MSE) are compared. All used loss functions are described in detail in Section 2.3For Huber loss function, the shape parameter was chosen to be $\delta = 0.4$ and for adaptive loss function best shape parameter found was $\alpha = 3/2$. Values were chosen using the manual trial and error method, tested values were in range $(0.1, 10)$.

Results for validation (Table 6) and test data sets (Table 7) both show the properties of robust and traditional loss functions. MSE loss shows the best performance in terms of maximal errors from the whole data set, which is the expected behaviour. It is since it heavily penalises points, which are far from target values. All robust losses perform better in all other metrics, the most interesting one is a median normalised error (MDE), which says by what fraction of true values models were typically off. It clearly shows models with robust losses are usually closer to true values. R^2 coefficient is not the most informative metric for such problems, since it quantifies the proportion of variance explained by the model. It is not suitable for problems with outliers, since the goal of the model is not to explain all variance in measurements, but to learn the underlying characteristics. This is also contained in the results, since there is only a very minor difference between R^2 for MSE, Log-Cosh and Adaptive loss.

These results are not fully reproducible, since used neural networks are sensitive to initialization. They should however not vary by much when trained with the same data and parameters. Differences between initialization outcomes are a minor factor, but they should be considered when conducting experiments.

Table 6. Results for robust hardness models for validation data.

	Absolute				Normalised				
	MSE	MAE	MDE	MAX	MSE	MAE	MDE	MAX	R^2
MSE	5358	50	28	**239**	0.104	0.220	0.133	**0.381**	0.786
Log-Cosh	5276	41	15	292	0.085	0.169	0.075	0.464	0.789
Huber	6440	41	**13**	351	0.075	0.154	0.082	0.559	0.743
Adaptive	**4559**	**40**	15	275	**0.054**	**0.143**	**0.072**	0.538	**0.790**

Test data was completely unseen by the models, and it can be used to generate the dependency between stand-off distance and predicted hardness. Those results were created using fitted models and a linear spacing of stand-off distances, generated between first and last measured samples. There were 19 measurements for each sample, but model trajectories were generated with 100 equally spaced points to analyse the behaviour of the model, especially in the region of the fusion line. Learned trajectories are shown in Figure 8 and Figure 9. Model trajectories are smoother than measurements, especially for samples 6-3. It is possible to see that models learn underlying characteristics and do not exhibit partly

random behaviour of the measurements. Differences between loss functions are difficult to see empirically, but are quite clearly distinguishable in computed metrics.

Table 7. Results for robust hardness models for test data.

	Absolute				Normalised				
	MSE	MAE	MDE	MAX	MSE	MAE	MDE	MAX	R^2
MSE	5631	49	23	**245**	0.052	0.162	0.120	**0.430**	0.820
Log-Cosh	5132	36	**11**	306	**0.028**	**0.097**	**0.049**	0.537	0.836
Huber	8526	43	**11**	368	0.040	0.110	0.054	0.646	0.727
Adaptive	**3682**	**33**	12	253	0.041	0.109	0.053	0.495	**0.864**

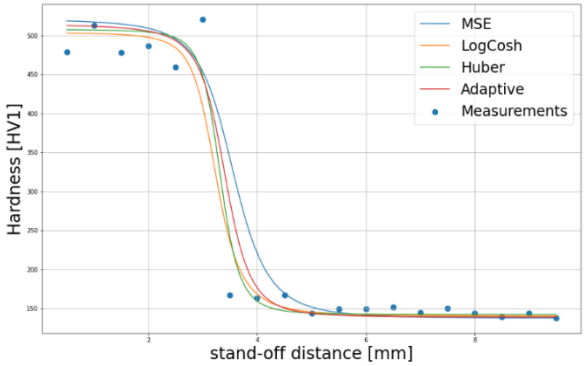

Figure 8. Predicted hardness for 3-2 samples.

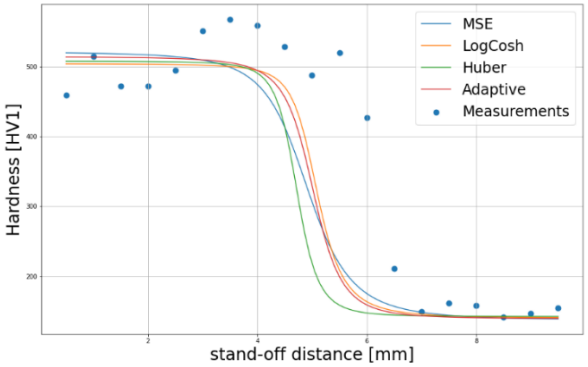

Figure 9. Predicted hardness for 6-3 samples.

4. Conclusions

In this study, neural networks fitted using robust machine learning algorithms were used to predict the hardness value of hardfacing layers manufactured by the GMAW method. The general conclusions can be drawn as follow:

1. Hardfacing layers have been successfully deposited on S235 based material. Method and parameters used to provide appropriate quality. The number of imperfections (mostly open porosity) should not have a high impact on the overall quality of the hardfacing layer.
2. The experiments taken examined the impact of three different variables on hardfacing layer properties. The influence of electric current, welding speed and beads number were taken into consideration. Noticeable correlations between variables provide a rationale for using machine learning to create black-box models of the considered welding process.
3. Neural modelling was shown to be successful for considered material and welding processes. For unseen test data models predicted material hardness after hardfacing layer was deposited with typical errors around 5% of true hardness. Those predictions were based only on the processing parameters. Moreover, applying robust learning methods for neural models enables those models to learn true underlying characteristics, without being affected by partly non-deterministic processes and human measurement errors.
4. This modelling framework is not typically used in welding technologies, so it requires further studies. However, the proposed method is promising, and can be employed for modelling welding processes with more potential improvements possible using robust learning algorithms.

Acknowledgements
The part of the work was financed by KGHM Polish Copper, Copper Smelter Legnica (316700/E1XT811).

References

[1] Garbade R R and Dhokey N B 2021 Overview of hardfacing processes, materials and applications *IOP Conference Series: Mat. Sci. and Eng*. **1017** 12-33

[2] Prabanjan S, Karthick K, Rejvin K J, Ramkumar S and Riswan A A 2020 Wear behavior and metallurgical characteristics of particle reinforced metal matrix composites produced by hardfacing *Mat. T.: Proc*. **33** 599-606

[3] Shibe J and Chawla V 2017 Depositing Fe-C-Cr based hardfacing alloys on steel substrate for enhancement in wear resistance *Int. Jour. of Eng. and Tech* **9** 105-11

[4] Łatka L and Biskup P 2020 Development in PTA Surface Modifications – A Review *Adv. In Mat. Sci.* **20** 39-53

[5] Saleem H, Gilmar F B, Chester J and Bekir Y 2014 Welding processes and technologies *Comp. Mat. Proc.* 3-48

[6] Nomura K, Fukushima K, Matsumura T and Asai S 2021 Burn-through prediction and weld depth estimation by deep learning model monitoring the molten pool in gas metal arc welding with gap fluctuation *Jour. of Man. Proc.* **61** 590-600

[7] Li J, Xie B, Fang Q, Liu B, Liu Y and Liaw P K 2021 High-throughput simulation combined machine learning search for optimum elemental composition in medium entropy *Jour. of Mat. Sci. & Tech.* **68** 70-5

[8] Schmidt J, Marques M R G, Botti S and Marques M 2019 Recent advances and applications of machine learning in solid-state materials science *Npj Comp. Mat.* **5** 1-36

[9] Szala M, Łatka L, Awtoniuk M, Winnicki M and Michalak M 2020 Neural modelling of APS thermal spray process parameters for optimising the hardness, porosity and cavitation erosion resistance of Al2O3-13 wt% TiO2 Coatings Processes **8**

[10] Gao G, Zhang Z, Cai C, Zhang J and Nie B 2019 Cavitation damage prediction of stainless steels using an artificial neural network approach *Metals* **9**

[11] Szala M and Awtoniuk M 2019 *IOP Conf. Ser.: Mater. Sci. Eng.* 710 012016

[12] Cybenko G 1989 Approximation by superpositions of a sigmoidal function *Math. of Cont., Sign. and Syst.*

[13] Pinkus A 1999 Approximation theory of the MLP model in neural networks *Acta Numerica* **8** 143-95

[14] Huber P J Robust estimation of a location parameter *The Ann. of Math. Stats.* **35** 73-101

[15] Tibshirani R 1996 Regression shrinkage and selection via the lasso *Jour. of Royal Stats. Society* **58** 267-88

[16] Friedman J, Hastie T and Tibshirani R 2010 Regularization path for generalized linear models by coordinate descent *J Stat Softw* **33**

[17] Hoerl A E and Kennard R W 2000 Ridge regression: biased estimation for nonorthogonal problems *Technometrics* **42** 80-6

[18] Efron B, Hastie T, Johnstone T and Tibshirani R 2003 Least angle regression *Ann. of Stat.* **32** 407-99

[19] Choi S, Kim T and Yu W 2009 Performance evaluation on RANSACK family *Proc. of the Brit. Mach. Vis. Conf.*10.5244/C.23.81

[20] Zhou W. and R Serfling 2008 Multivariate spatial u-quantiles: a bahadur–kiefer representation, a theil–sen estimator for multiple regression, and a robust dispersion estimator *Jour. of Stats. Planning and Inference* **138** 1660-78

[21] Kim S J, Koh K, Lustig M, Boyd S and Gorinevsky D 2007 An interior-point method for large-scale L_1-regularized least squares *IEEE J. of Sel. Top. in Sign. Proc.* **1**

[22] Klivans A R, Kothari P K and Meka R 2018 Efficient algorithms for outlier-robust regression *CoRR* **1803**.

[23] 2019 EN 10025-european standards for structural steel

[24] 2014 EN 14700:2014 Welding consumables - welding consumables for hard-facing

[25] Solodskiy S, Saraev Y, Malchik A and Korotkov S 2016 Technology Of MIG-MAG welds strength enhancement *IOP Conf. S.: Mat. Sci. and Eng* **12**-6

[26] 2007 ISO 6520-1 Welding and allied processes — Classification of geometric imperfections in metallic materials — Part 1: Fusion welding

[27] 2018 Metallic Materials—Vickers hardness test—part 1: test method. ISO

[28] Shen Y and Sanghavi S 2019 Iterative least trimmed squares for mixed linear regression *CoRR* abs/1902.03653

[29] Singla M, Ghosh D and Shukla K K 2020 Robust optimisation for deep regression *Int. J. of Mach. Learn. and Cyber* **11**

[30] Stigler S M 1981 Gauss and the invention of least squares *The Ann. of Stat.* **9** 465-74

[31] Neuneier R and Zimmermann H G 1998 Neural Networks: Tricks of the Trade Orr G B and Muller K L, Berlin-Heidelberg Springer Berlin Heidelberg 372-432.

[32] Zhang N, Shen S L, Zhou A and Xu Y S 2019 Investigation on performance of neural networks using quadratic relative error cost function *IEEE Access* **7**

[33] Barron J T 2017 A more general robust loss function *CoRR* abs/1701.03077

[34] Potdar K, Pardawala T and Pai C 2017 A comparative study of categorical variable encoding techniques for neural network classifiers *Int. J. of Comp. Appl.* **175** 7-9

[35] Stone M 1974 Cross-validatory choice and assessment of statistical predictions *J. of the Roy. Stat. Soc.* **36** 111-47

[36] Bishop C M 2006 *Pattern recognition and machine learning*, Springer-Verlag Berlin, Heidelberg Springer-Verlag 259-61.

[37] Tieleman T and Hinton G 2012 Lecture 6.5 - RMSProp, COURSERA: Neural Networks for Machine Learning

[38] Choi D, Shallue C J, Nado Z, Lee J, Maddison C J and Dahl G E 2019 On empirical comparisons of pptimizers for deep learning *CoRR* 1910.05446

[39] Bishop C M 2006 *Pattern recognition and machine learning* Springer-Verlag Berlin Chapter 7 Sparse Kernel Machines

[40] Awad M and Khanna R 2015 Support vector regression *Efficient Leararing Machines* Apress Berkeley, CA 67-80

[41] Hasan A, Ali O and Alsaffawi A 2018 Effect of welding current on weldments properties in MIG and TIG welding *Int. J. of Eng. and Tech. (UAE)* **7** 192-97

[42] Wichan C and Loeshpahn S 2014 Effect of welding speed on microstructures, mechanical properties and corrosion behavior of GTA-welded AISI 201 stainless steel sheets *J. of Mat. Proc. Tech.* **214** 402-8

[43] Widodo E, Iswanto I, Nugraha M and Karyanik K 2018 Electric current effect on mechanical properties of SMAW-3G on the stainless steel AISI 304 *MATEC Web of Conf.* **197**