ARTIFICIAL AND COMPUTATIONAL INTELLIGENCE

Performance of artificial neural networks in an inverse problem of laser beam diagnostics

Karol PIETRAK [ID]*, **Radosław MUSZYŃSKI, Adam MAREK** and **Piotr ŁAPKA**

Faculty of Power and Aeronautical Engineering, Warsaw University of Technology, ul. Nowowiejska 24, 00-665 Warsaw, Poland

Abstract. The presented results are for the numerical verification of a method devised to identify an unknown spatio-temporal distribution of heat flux that occurs at the surface of a thin aluminum plate, as a result of pulsed laser beam excitation. The presented identification of boundary heat flux function is a part of the newly proposed laser beam profiling method and utilizes artificial neural networks trained on temperature distributions generated with the ANSYS Fluent solver. The paper focuses on the selection of the most effective neural network hyperparameters and compares the results of neural network identification with the Levenberg–Marquardt method used earlier and discussed in previous articles. For the levels of noise measured in physical experiments (0.25–0.5 K), the accuracy of the current parameter estimation method is between 5 and 10%. Design changes that may increase its accuracy are thoroughly discussed.

Key words: inverse problems; pattern recognition; laser beams.

1. INTRODUCTION

High-power lasers are used in many areas of science and technology – for example in materials processing [1], medicine [2], defense [3], and materials characterization [4, 5]. One of their key characteristics is the spatial distribution of energy density across the beam, known as the beam profile [6]. The beam profile may be often described by the super-Gaussian function [7]. Three basic shapes of that function are recognized, i.e. Gaussian, super-Gaussian, and flattop (see Fig. 1). In laser-based devices, it is often the case that light travels through optical components before interacting with its target. The profile of the laser beam at the end of the optical track may differ from the desired one due to the heating of mirrors, lenses, and polarizers, which causes their deformation [8].

Verification of the beam characteristics and stability checks are required to assure reliable and repeatable operation of laser-based devices. It is typically achieved with the use of laser beam profiling techniques [6, 8], among which destructive and non-destructive techniques may be distinguished. In destructive measurements, laser beams interact with solid targets (e.g. thermal paper, photographic films, acrylic blocks) and leave permanent changes which allow us to identify some beam parameters [6]. This type of setup is usually unsuitable for continuous measurements, and nowadays destructive techniques have mainly historical importance. Automated non-destructive techniques, which are well-established, make use of various methods of beam attenuation before its interaction with the sensor. High power beams are particularly problematic to measure. Focused constant width lasers of power in the 1 W range and

*e-mail: karol.pietrak@pw.edu.pl
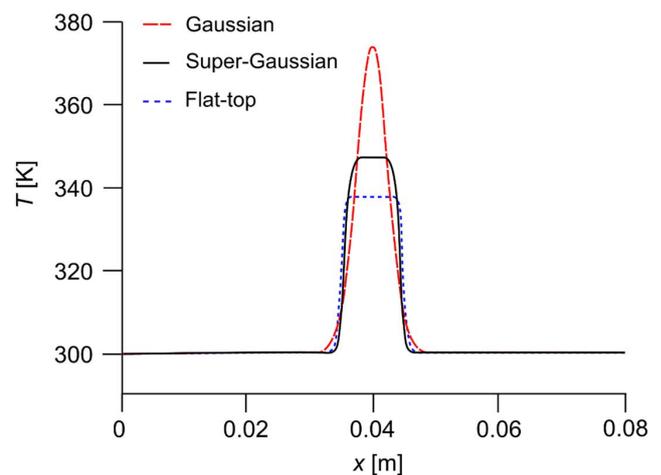
Fig. 1. Profiles of temperature at the surface of thin aluminum plate obtained by heating it with a high-power laser beam of three basic spatial energy distributions – Gaussian, super-Gaussian and flat-top (adopted from [9])

pulsed lasers in the 1 J range can easily damage scanning apertures, and beams in the kW range can damage beam samplers. One of the ways to overcome the power obstacle is the non-interceptive profiling proposed by Guttman [8]. Nevertheless, typical laser beam profiler instruments are usually very fragile and expensive. Fragility makes it difficult to use them in field conditions, and high cost limits their use in the industry. Additionally, not all of them work for pulsed beams [10].

In response to the need for less expensive and more robust testing methods applicable to pulsed laser beams, a new method has been recently proposed by Kujawińska *et al.* [9]. This method initially assumed utilization of fast infrared thermography (IRT) and digital image correlation (DIC) [11] meth-

K. Pietrak, R. Muszyński, A. Marek, and P. Łapka

ods to capture temperature maps (thermography) and displacement maps (DIC) resulting from the interaction of the short laser pulse with a thin aluminum plate. The sensing of temperature and displacement is carried out for the rear surface of the plate – opposite to the heated surface (See Fig. 2). Kujawińska and her collaborators [12] assumed that their experimental setup may be used to identify four parameters of the laser pulse, i.e. its spatial profile coefficient, beginning time, end time, and power. In their technique, the density of heat flux at the sample surface, resulting from the laser strike, was assumed in the super-Gaussian form [7] and the impulse itself was treated as rectangular with respect to time.
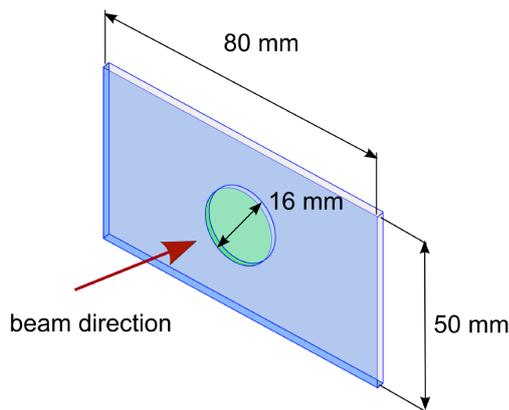
Fig. 2. The geometry of the laser-heated plate

The problem that arose may be classified as the inverse heat transfer problem (IHTP) of parameter estimation [13]. Pietrak *et al.* [12] and Łapka *et al.* [14] confirmed that useful results may be obtained with the aforementioned method even if only the thermal part of the problem is considered and the mechanical part (the measurement and simulation of displacements) is neglected. They tested the method based solely on temperature measurement using both simulated (artificial) [12] and real (physical) [14] measurement data. They obtained discrepancies between estimated and accurate parameter values of the order of 3.39% for artificial and 20–25% for physical experiments. The advantage of their technique is a simple experimental setup employing a fast-infrared camera as the only required sensor. That fact makes the method less expensive and more resistant to failure than sophisticated commercial instruments. To solve the parameter estimation problem, Pietrak *et al.* and Łapka *et al.* used the Levenberg–Marquardt inverse algorithm with sensitivity matrices obtained by numerical simulation of the forward thermal problem. That approach has one drawback, i.e. the presence of a significant delay between the physical measurement and the end of the data analysis. The delay is caused by the requirement to solve the forward problem many times using the finite volume method (FVM), which may be time-consuming. To assure fast operation of the data analyzer program in field conditions, the numerical simulation process may be moved to the time before the field measurements. This can be achieved with the artificial neural network (ANN) approach investigated in the current article.

The present approach to the identification of spatio-temporal characteristics of the laser pulse is based on the experimental setup proposed by Kujawińska *et al.* [9]. Similarly to the version considered by Pietrak *et al.* [12] and Łapka *et al.* [14], it includes the thermal problem and neglects the mechanical one (the unknown characteristics of the laser pulse are estimated only based on temperatures and not on displacements). In the presented analysis it is proposed that an ANN can learn how different temperature distributions measured on the plate struck by a laser pulse translate to the characteristic parameters of the laser beam. More specifically, the aluminum plate is assumed to be heated by a short $(0.2-2\text{ ms})$ laser pulse traveling perpendicularly to the plate. The spatial distribution of the boundary heat flux resulting from the laser heating is assumed in the super-Gaussian form [7] with rectangular temporal profile. The laser beam hits the front surface of the sample and temperatures are sensed at its rear surface by employing the infrared thermography (IRT) camera. By performing many measurements with laser pulses of different parameters the ANNs can be trained to recognize beam parameters based on captured temperature fields. Nevertheless, physical experiments are costly and instead the ANNs can be trained using less expensive numerical simulations in the ANSYS Fluent software. This scenario is presented in Fig. 3, which shows the schematic of the intended method.
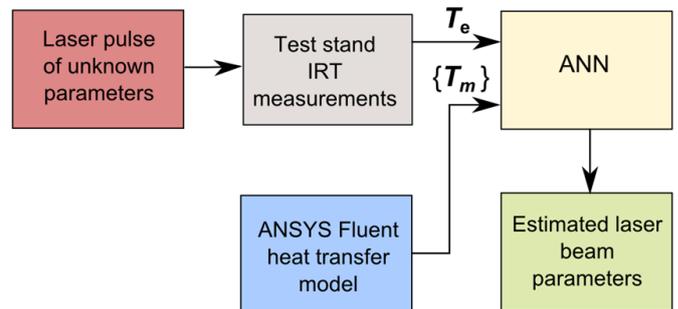
Fig. 3. The schematic of the intended method

The current work shows the details and results of the process of the optimization of ANN hyperparameters (such as topology, activation functions, and learning methods) for the intended task. In this process, the performance of the ANNs was validated using synthetic data from the Ansys Fluent simulations. In these simulations, three parameters of the laser beam were varied, i.e. the shape coefficient p, laser power Q, and pulse duration t_{end} (the start time of the pulse was considered known). Validation of the neural network using synthetic data also allowed to estimate the potential of the proposed artificial intelligence solution for further development involving real measurement data. The ANN-based data analyzer may run on portable computers with limited resources and the results of laser beam parameters estimation may be obtained with a minimal delay which makes the method advantageous.

Classical solution approaches to inverse thermal problems were briefly reviewed by Beck and Woodbury [15]. The overview of inverse methods for the determination of sur-

2

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

face heat fluxes may be found in book chapters by Taler and Taler [16, 17]. Methods applied to problems of laser-solid matter interactions were discussed by Pietrak *et al.* [12]. ANNs, which constitute an important element of the current solution method, have been previously considered an interesting alternative method to solve general IHTPs alongside genetic algorithms (GA), particle swarm optimization (PSO), and proper orthogonal decomposition (POD) [18]. Their successful application has been demonstrated in the case of inverse and optimization problems involving conduction [19–27], radiation [28–31], and convection [32–34], but also in power engineering [35–37]. In particular, Krejsa *et al.* assessed their potential and discussed various strategies regarding their application for the inverse problem of heat conduction [19]. A detailed description of the theory of neural networks, oriented towards engineering and scientific applications, can be found in Samarasinghe's book [38].

2. PROBLEM STATEMENT

2.1. Direct problem

In the direct problem considered in this paper, an aluminum (alloy AW2017A T4), rectangle-shaped sample with a circular hollow in the center is struck by a short high-power laser pulse arriving perpendicularly to the plate from the side of the hollow (see Fig. 2). The reduction of sample thickness in the heating area decreases its temperature-damping quality and improves the accuracy of the subsequent parameter estimation as temperature sensing is carried out at the backside of the plate (opposite to the illuminated one) [14]. Front-face sensing of temperatures was also considered but it was finally rejected because of the risk of damaging the IR camera by the reflected laser light. Thermophysical properties of the modeled body subjected to laser heating are grouped in Table 1.

Table 1
Assumed thermophysical properties of the heated plate

Temperature\Property	Density	Thermal conductivity	Specific heat	Surface emissivity
[K]	[kg/m^3]	[W/m/K]	[kJ/kg/K]	[–]
303		122.6	0.902	
323		125	0.911	
373		133.7	0.943	
423	2700	151.9	1.046	0.3
473		154.9	1.039	
523		130	0.832	
573		159.6	0.995	

The aluminum sample was considered homogeneous and isotropic. The presented version of the method is assumed to be non-destructive, i.e. the temperature within the heated body should not exceed the melting point of aluminum (736 – 944K depending on the alloy [39]). In such conditions the heat con-

duction process in the sample may be described by the following equation:

$$\rho c_p(T) \frac{\partial T}{\partial t} = \mathrm{div}\left[\lambda(T)\mathrm{grad}(T)\right], \tag{1}$$

where ρ denotes density, c_p – specific heat, λ – thermal conductivity, T – temperature, $\mathrm{div}[.]$ – divergence and $\mathrm{grad}(.)$ – gradient operators. The boundary heat flux at the front surface (irradiated by laser beam) was assumed in the super-Gaussian form [7] as follows:

$$q_l = \begin{cases} \varepsilon Q \dfrac{p 2^{\frac{2}{p}}}{2\pi R_0^2 \Gamma\left(\dfrac{2}{p}\right)} \exp\left[-2\left(\dfrac{R}{R_0}\right)^p\right], & 0 \le t \le t_{\mathrm{end}} \\ 0.0, & t > t_{\mathrm{end}}, \end{cases} \tag{2}$$

where ε denotes surface emissivity (assumed $\varepsilon = 0.05$ after matching the numerical model to experiments as presented in [14]), Q – laser power, p – curve shape coefficient, R_0 (assumed 5 mm) – length scale over which the profile decreases to e^{-2} of its axial value (in case of the Gaussian beam, i.e. $p = 2$), R – distance from the center of the plate and $\Gamma(.)$ – the Euler Gamma function. Let us note that by changing the parameter p in equation (2) it is possible to obtain all beam distributions shown in Fig. 1, i.e. Gaussian ($p = 2$) as well as super-Gaussian ($p \sim 2-10$) and flat-top ($p \sim 10-100$ and above).

Heat flux due to radiation and convection was included on all walls of the plate, as described by this equation:

$$q_w = H(T_\infty - T_w) + \varepsilon\sigma\left(T_\infty^4 - T_w^4\right), \tag{3}$$

where subscripts ∞ and w denote surroundings and the wall, respectively, H – heat transfer coefficient at walls ($H = 5$ W/m^2/K was assumed based on the literature [40]) and σ – Stefan–Boltzmann constant ($\sigma = 5.67 \cdot 10^{-8}$ W/m^2/K^4). Initial temperature $T_0 = 303$ K was assumed within the whole body.

The thermal problem given by equations (1)–(3) has been solved using the FVM method in the commercial software ANSYS Fluent 19.2. Firstly, a 3D model of the problem was prepared, described in detail in [12] and [14], and its thermal parameters, such as surface reflectivity were finetuned based on physical experiments with laser and actual samples (the match of experimental and modeled temperature curves can be found in [14]). The geometry of the problem has been then transformed to 2D axisymmetric to reduce the computational cost. In the reduced model, the sample was treated as a cylinder of radius 40 mm and height 1 mm with a cylindrical hollow of radius 8 mm and depth 0.5 mm. The 2D model was meshed using a regular grid containing 48450 rectangular elements with a maximal aspect ratio of 271.46 and maximal skewness equal to $7.4 \cdot 10^{-4}$. The influence of grid sizing on the result was checked and the analysis confirmed grid independence. The detailed view of the computational mesh is shown in Fig. 4. The grid density was greater in the hollow section of the plate model (where laser excitation occurred) and lower in the remaining part. Finer grid was also placed near walls where heat diffusion starts and where temperature measurement points were located.

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

3

K. Pietrak, R. Muszyński, A. Marek, and P. Łapka
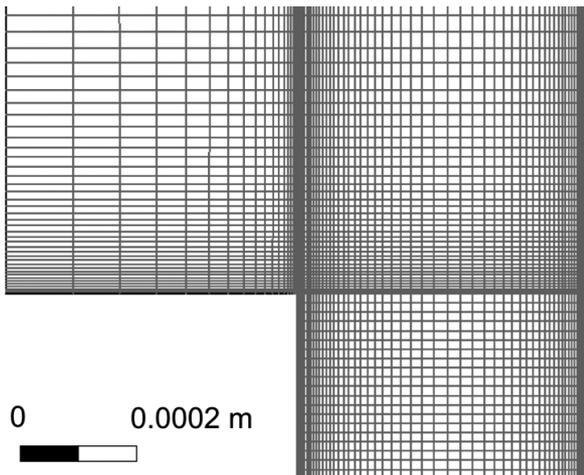
Fig. 4. Detail view of the computational grid where the plate thickness changes

Fig. 5. ANN inputs and outputs schematic

2.2. Inverse problem

For the sake of the inverse analysis, it was assumed that the results of temperature measurements carried out at the rear side of the heated plate are available. The parameters estimated by the network were:

1) laser power Q,
2) spatial profile coefficient of the beam p,
3) pulse end time t_{end}.

The start time was assumed known: $t_{\text{start}} = 0$. Before applying the ANNs it was necessary to select their inputs. Temperature measurements in real experiments are performed by fast IR cameras which leads to availability of full temperature maps of the investigated area, acquired with temporal step of the order of 0.27 ms and spatial resolution of 80×64 pixels [41]. The large array of data resulting from the IR measurement must be limited to become an input array to ANN. Appropriate selections were made based on earlier analyses of sensitivity [12]. It was assumed that the laser pulse end time t_{end} belongs to the range [0.2, 2] ms. The duration of temperature measurement was limited to 16 ms, which ensured that the decrease of temperatures due to heat loss to the surroundings is captured. The area of temperature sensing was limited to the central region of the wall placed vis-à-vis the heated area, where problem-significant temperature changes were occurring.

The input vector of measurements used to train the ANNs was constructed from two subsets:

1) temperature change history at the central point of the sample (see Fig. 5),
2) temperature values along the sample radius for a single time instant $t = 12$ ms.

The first subset of the vector reflected the temporal profile of the thermal response of the sample and the second one – its spatial profile. Such vector was sampled with 3 different resolutions ($n = 30, 60, 134$ points) as later the influence of the sampling resolution on the retrieval error was examined. Spatial and temporal responses were sampled equally, e.g. for 60 points, 30 represented the temperature profile along the radius and the other 30 – the change of temperature in the central point.
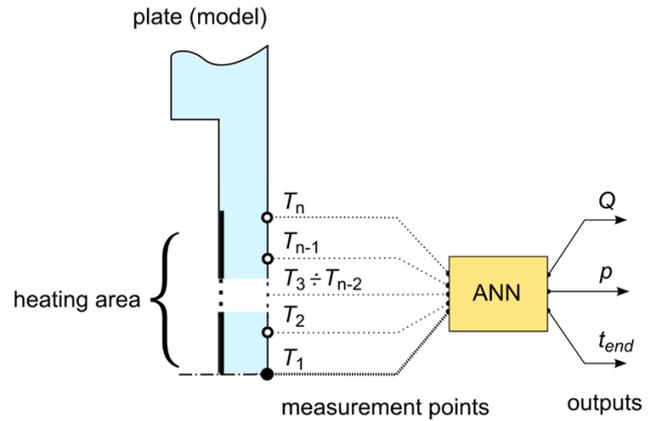
Raw data sets for ANNs were generated by simulating the axisymmetric FVM model with different values of input parameters Q, p and t_{end}. The range of pulse durations and laser power was limited to assure that melting temperature is not reached. For both scenarios, values of parameter Q (in kilowatts) were taken from the set $\{5, 6, 7, \ldots, 20\}$, the values of parameter p (dimensionless) from the set $\{2, 8, 14, \ldots, 98\}$, and values of parameter t_{end} (in milliseconds), from the set $\{0.2, 0.4, 0.6, \ldots, 2\}$. Thus, the grid of training data contained 2720 cases.

3. METHODS

3.1. Design stages and training methods

The neural networks were programmed in Python 3.7 using open-source libraries Keras and Tensorflow [42]. Initially, weights were set to random numbers. In the whole study, the maximal number of hidden layers was arbitrarily limited to two. The procedure of searching for the best ANN configuration was divided into two stages – initial and final. Firstly, the influence of network topology was tested with other hyperparameters constrained to arbitrarily chosen ones (see section 4.1). Then, two single- and double-hidden layer topologies which achieved the best performance (i.e. 60–120–3 and 60–120–240–3, respectively) were used in combination with the same arbitrarily-chosen hyperparameters treated as fixed, while different optimizers, activation functions, and data preprocessing methods were tested one at a time, successively (sections 4.2–4.4). This initial stage allowed us to draw conclusions for the second stage of tests in which allowed us to narrow down the options and select the recommended configuration of hyperparameters for the given application.

Among the training algorithms tested were all gradient descent optimization algorithms available in the Keras library, i.e. Stochastic gradient descent (SGD), Root Mean Square Propagation (RMSProp), Adagrad, Adadelta, Adam, Adamax, and Nadam. The aforementioned algorithms, among others, were thoroughly explained by Ruder [43].

Tested activation functions included Sigmoid [44, 45], Tanh [44, 45], Relu [44, 45], Elu [45, 46], Selu [45, 47], Softmax [45], Softplus [45], Softsign [45] and Hard_sigmoid [45].

4

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

To avoid overfitting, the "EarlyStopping" function [48] was used which stops training at an optimal point, when a neural network does not learn anymore. The function monitors validation loss according to the selected metric, which in this experiment was the MSE (mean squared error) [49].

The only customizable parameter was "patience" which tells how many epochs must be executed without improvement to break the process of training. For all training of neural networks mentioned in this paper, the parameter of "patience" was set to 100 epochs.

3.2. Topology study

Although the concept of machine learning is not new, there has not been found a method to calculate analytically the best number of hidden layers or number of nodes for a model [50].

Every problem is specific, so only suggestions can be found in the literature but not the final answer. Testing many possibilities is a problematic solution because of an extremely large number of combinations and the long time of training a single model. In general, the more nodes and hidden layers are used, the more complex functions may be approximated. On the other hand, large networks take longer to train and are more susceptible to overfitting [51].

The Hecht–Nielsen theorem states that a network with only one hidden layer and $2i + 1$ nodes (where i denotes the number of inputs) is enough to approximate any continuous function, provided that special activation functions are used [52].

Kurkova suggested that a second layer can be added to obtain equal capabilities with classical activation functions [53].

Later, Cybenko showed that an ANN with one hidden layer and the commonly used sigmoid activation function can approximate any continuous function of n real variables with support in the unit hypercube [54].

Given these rules, it was reasonable to start experimenting with a single hidden layer and check how adding more layers affected accuracy and generalization ability.

To gain an initial insight regarding an optimal topology for the current problem, architectures with one and two single layers were tested, starting with one hidden layer of 10 nodes, increasing gradually to 1200 nodes, and then expanding to two layers, 10 neurons each, increasing both to nearly 1000 nodes (e.g. 800 in the first and 1000 in the second). In these experiments, the optimization algorithm (Adam [43]), stopping criteria, and activation functions (sigmoid for the hidden layer, non for output layer) were the same to obtain a useful comparison. The inputs were noiseless, and their number was always 60. In later experiments selected topologies with one and two hidden layers were tried with different optimizers, activation functions, and data preparation methods.

3.3. Data preparation

Due to the characteristics of algorithms applied to update weights and biases, it is generally observed that input data preprocessing (rescaling, normalization, and standardization [55]) is beneficial for neural network performance [55–57].

Preprocessed data leads to faster training and better accuracy of estimations. In the present case, the ANN models did not work with raw data. Therefore, various preconditioning methods for the input temperature data were tested, i.e. min-max standardization to $\langle -1, 1 \rangle$ [58], dividing by the median [58], subtracting the median [59], and subtracting the midrange [60].

Furtherly, also the expected values (in our case laser pulse parameters) may be transformed to improve the prediction accuracy in machine learning [57].

It was decided to compare the accuracy without modification of the expected values (raw data) and with outputs scaled to a similar range as the inputs.

3.4. Verification methods

For each training, the data were randomly divided into 3 groups:
    2000 samples – training data,
    420 samples – validation data,
    300 samples – test data.

The first subset was used for training. The second one was monitored during training to detect overfitting. The last set was not used for training and was used to assess the network performance when it was presented with unknown data. The proportions of data division were chosen arbitrarily.

In thermal measurements, temperature data is always noisy. To check how noise affects the performance, the data sets were modified by the addition of random noise with uniform distribution. Four levels of noise amplitude were tested: 0.1 K; 0.25 K; 0.5 K and 1 K.

3.5. Error measure

The agreement between the test datasets and network predictions was quantified using mean absolute percentage error (MAPE) which can be described by equation (4), where: E_i – the value estimated by ANN, C_i – the correct value, z – the number of estimations, $|.|$ – the absolute value and $i = 1, 2, \ldots, z$ – the index of estimation. It gives information about the mean deviation of estimations from accurate parameter values expressed as percentage

$$\text{MAPE} = \frac{1}{z} \sum_{i=1}^{z} \frac{|E_i - C_i|}{|C_i|} \cdot 100\% . \qquad (4)$$

The error was measured for all 3 laser parameters separately but because there were no significant differences observed between them, for every case only the average MAPE was considered.

4. DESCRIPTION AND RESULTS OF INITIAL STUDIES

4.1. The initial topology investigation

In the first stage of topology optimization, multiple single- and double-layer topologies were verified systematically with the number of inputs constrained to 60, Adam optimizer, sigmoid activation function for hidden layers, and non-sigmoid for the output layer. Each configuration was trained only once. Figure 6 shows the behavior of the estimation error as the number of neurons in the single hidden layer case is increased.

It may be seen that the MAPE decreases steadily until the number of neurons in the hidden layer is equal to the number

K. Pietrak, R. Muszyński, A. Marek, and P. Łapka
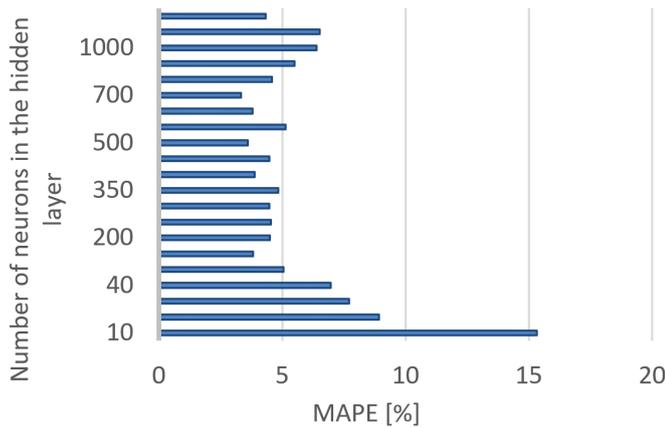
Fig. 6. Relation between estimation error and number of neurons in the hidden layer in case of one hidden layer
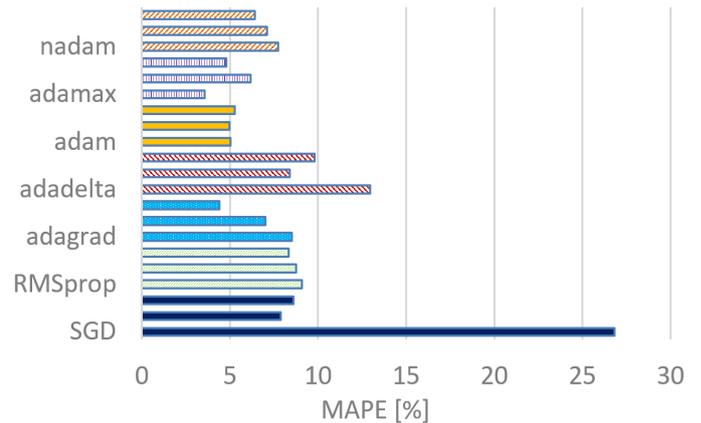
Fig. 7. MAPE for different optimizers for a single hidden layer (3 results per algorithm shown individually)

of inputs (60). The local minimum of error (MAPE = 3.82%) occurred for 120 neurons. It had been concluded that further increase of the layer size was not effective as the error fluctuated above 120 neurons with no clear improvement. The global minimum occurred for 700 neurons (MAPE = 3.33%), but it was regarded as a coincidence due to the fluctuating nature of the error value in that range.

Surprisingly, the results for two hidden layers, with the same settings, were generally worse. The sizes of the first layer in tested configurations were the same as for the single-layer cases and the sizes of the second layer were selected to be of a similar order (i.e. 60–120, 1000–800, etc.). A total number of 24 topologies with two hidden layers was examined, and no clear minima could be identified. However, the results were best if both layers contained 60 neurons and more. The overall estimation accuracy was lower than for a single hidden layer. The median and mean of error for single hidden layer experiments were 4.59 and 5.56%, respectively, whereas for the cases with two layers these were 5.78 and 6.65%.

4.2. Optimizer comparison and selection

Next, all the optimizers were tested for both single-hidden layer and double-hidden layer topologies which achieved the best results (i.e. 60–120–3 and 60–120-240–3, respectively). Three simulations were performed for each optimizer to verify the repeatability. Sigmoid activation functions were used for hidden layers and non-sigmoid for output layers.

In the case of one hidden layer, the best repeatable results (MAPE around 5%) were obtained with Adam and Adamax optimizers, as can be seen in Fig. 7. When the optimizers were tested with the double-hidden layer network (Fig. 8), even better results were obtained. In those experiments, the best accuracy was achieved with the Adagrad algorithm (MAPE around 2.5%). The typical number of epochs to complete the training for that algorithm was generally higher than for other algorithms, both in the case of a single (5000–50000 epochs for Adagrad vs 500–2000 for other optimizers) and double hidden layer (2800–5000 epochs for Adagrad vs 250–900 for other optimizers).
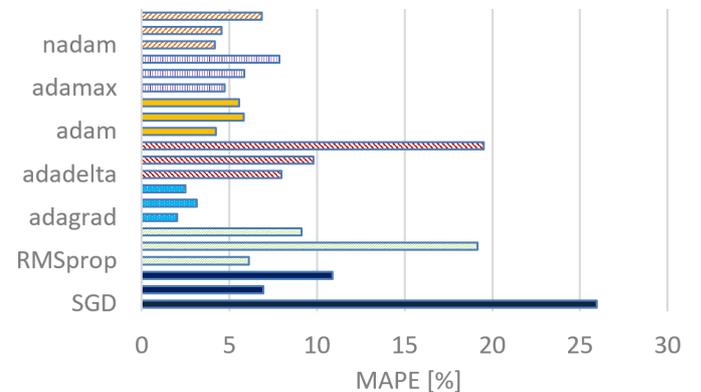
Fig. 8. MAPE error for different optimizers for two hidden layers (3 results per algorithm shown individually)

4.3. Comparison and selection of activation functions

The choice of activation functions may influence the ANN's approximation ability and accuracy [45, 52] and due to that, it was decided to test all activation functions available in the Keras library. The functions were compared on networks with topologies which provided the best results in the initial topology study, i.e. 60–120–3 and 60–120–240–3. The comparisons of estimation accuracy are shown in Figs. 9 and 10.
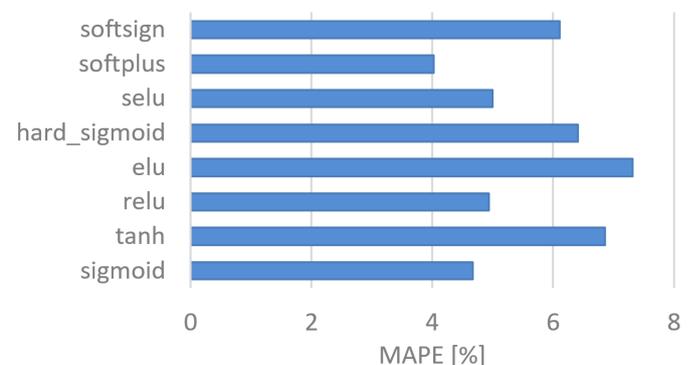
Fig. 9. Estimation errors for different activation functions (one hidden layer, averaged from 3 experiments)

Performance of artificial neural networks in an inverse problem of laser beam diagnostics
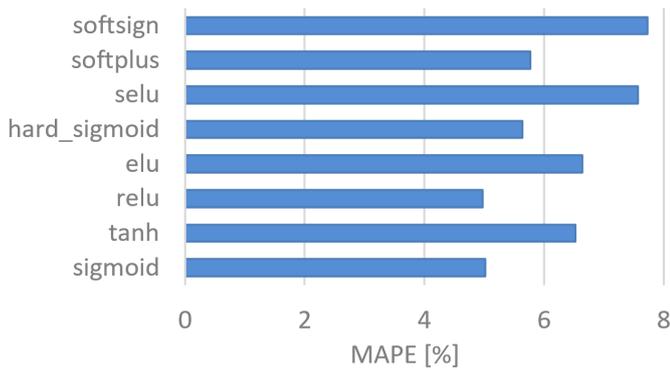
Fig. 10. Estimation errors for different activation functions
(two hidden layers, averaged from 3 experiments)

The MAPE errors shown are averaged from 3 experiments. The results for the Softmax function are not shown, because the error for this case was extremely high (90 and 136%, respectively for single and double hidden layers). The remaining activation functions gave results of comparable order, with the Softplus function being the best in the case of one hidden layer and Relu – in the case of two hidden layers.

4.4. Effect of different data modification methods

To find the best preprocessing methods, a few series of experiments were conducted. In the first series, input data transformation methods (i.e. min-max standardization, subtraction of the median, subtraction of the midrange, and division by the median) were compared. Three simulations were performed per every method to obtain averaged results shown in Fig. 11. Parameters of neural networks in the experiments were: topology 60–120–3, Activation functions -sigmoid for the hidden layer, non-sigmoid for the output layer, optimizer- adam.
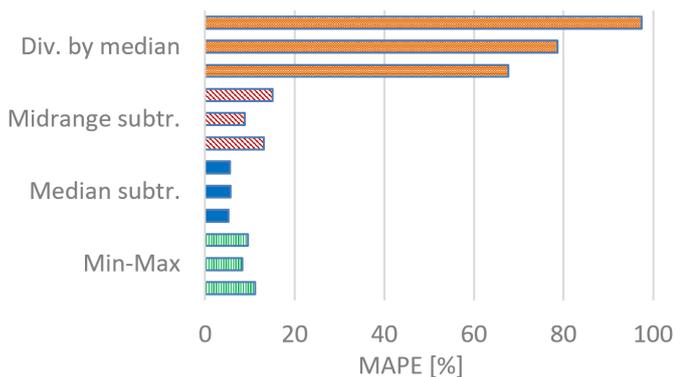
Fig. 11. Comparison of estimation errors for various methods of input
data transformation (3 results per method shown individually)

It can be seen in Fig. 11 that simply dividing the measured temperatures (in kelvins) by their median was not effective (estimation error of 68 to 98%), while other methods, which shifted the center of the data to zero, gave considerably better results. Among those methods, subtracting the median resulted in the smallest error (around 5.4%).

Secondly, the outputs had been scaled to possess a similar range to the inputs, i.e. laser power in W was divided by 10 and shape coefficient p was divided by 20, whereas the pulse duration remained unchanged. The estimation accuracy in such a case was compared with the case of unaltered outputs (raw data). The effect was checked in five experiments for each type of data (altered/unaltered). As can be seen in Fig. 12, scaling of the output data resulted in profound increase in accuracy. Figure 12 shows results for a network with two hidden layers (60–120–240–3). The results for a single-hidden layer network (60–120–3) were very much alike and hence they are not shown.
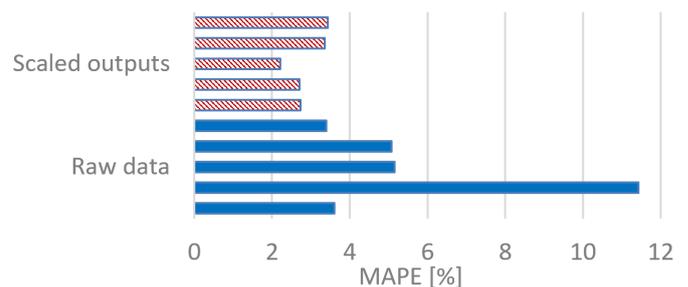
Fig. 12. Comparison of estimation errors for raw and scaled outputs
(5 results per method shown individually)

In one experiment the pulse duration parameter (in ms) was multiplied by 10 which resulted in all output parameters having similar magnitudes (1–100). However, it was still much greater than the magnitude of inputs. The study showed that this type of treatment is ineffective. The range of input and output data should be similar. Subtracting of the median from the scaled output data was also attempted to shift the center of the data to zero, but it resulted in a loss of stability (MAPE for one parameter exploded).

5. FINAL RESULTS

5.1. Selection of the final ANN topology

Although the initial topology study showed that a single hidden layer network might be the best, changing the optimizer to Adagrad caused the error to drop significantly for a network with two hidden layers (see Fig. 8). Relu and sigmoid activation functions were selected for additional tests because they demonstrated the best performance in double-hidden layer networks and good performance in single-hidden layer networks (Figs. 9 and 10).

Additional experiments have shown that the best accuracy is assured if both input and output layer operates with relu activation function (configuration relu-relu resulted in 2% error, as opposed to 2.25% for sigmoid-relu and 3.4% for sigmoid-sigmoid). It was hard to point out the optimal topology, apart from the fact that the number of neurons in the hidden layers should be greater than, or equal to, the number of inputs. Based on additional experiments with Adagrad optimizer and relu-relu activation functions, the architecture 60–1000–1000–3 was identified as best and selected for further studies. The topology

K. Pietrak, R. Muszyński, A. Marek, and P. Łapka

with a high number of neurons in the hidden layer was expected to perform well in the upcoming simulations with the number of inputs increased to 134.

5.2. Effects of noise and measurement resolution

The dataset of simulated temperatures from the Fluent software was characterized by the high spatial and temporal resolution of the data. This feature allowed to increase the number of ANN inputs to 134. Both spatial and temporal resolutions were increased equally. The influence of increase and decrease of the number of inputs on ANN retrieval accuracy was tested using the final ANN configuration X-1000-1000-3.

At this stage, the input data were modified by the addition of noise of different amplitudes. The influence of temperature measurement resolution and the level of noise on the accuracy of estimations are shown in Fig. 13. For non-noisy data, the MAPE in all 3 cases (30, 60, 134 input points) was close. However, the higher the noise level, the bigger the difference between them. Models with more inputs had evidently better performance for noisy datasets. To generate the graph, each configuration was trained three times and the arithmetic mean of obtained MAPEs was calculated. The scatter of results was very small (within 0.1–1% range). Figure 13 also shows the error level obtained using the LM method discussed earlier [12] for comparison. It was observed that ANN estimation results are significantly more repeatable if the training data are noisy, compared to non-noisy data.
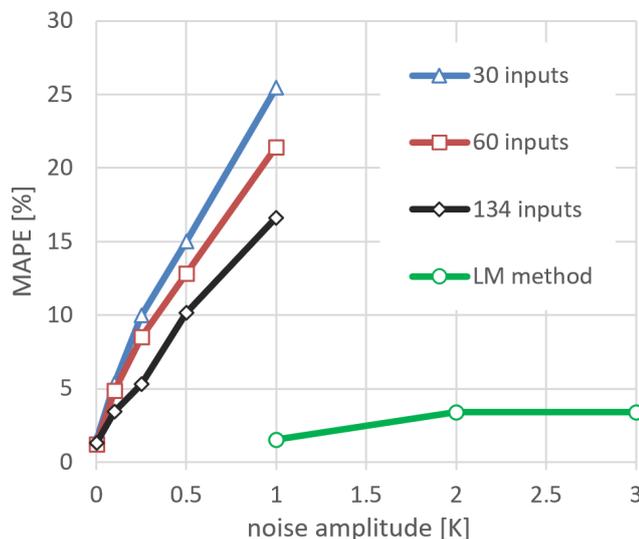
Fig. 13. Dependence of the ANN parameter estimation error on the number of inputs and the level of noise for the finally chosen topology X-1000-1000-3 (performance of the alternative LM method is also shown for comparison)

6. DISCUSSION

The lowest MAPE error achieved during all tests was 1.22% (134 inputs), but it was for non-noisy data which represents a non-realistic case. For the noise amplitude of 0.1 K, the best result was MAPE error equal to 3.44%; for 0.25 K, it rose to 5.33%; for 0.5 K, it was 10.15%, and for 1 K – 16.63%.

To judge the quality of the finally obtained result, it should be compared with similar research and analyzed from different perspectives. The same inverse problem was previously solved using the gradient-based Levenberg–Marquardt (LM) technique in which the heat transfer model in Fluent was directly simulated to obtain the sensitivity matrices [12].

With that technique retrieval errors as low as 1.55% were obtained when the level of noise was 1 K and increased to 3.39% at 2 and 3 K (lower levels of noise were not tested). For the noise level of 1 K, the current ANN method yielded a much greater error (16.63%) which makes it look rather disadvantageous (see Fig. 13). Nevertheless, there is a substantial difference – the LM method uses a direct simulation of the heat transfer model to take iterative steps and gradually minimize the difference between modeled data and presented data in the full range of possible parameters, whereas the ANN method generalizes based on a limited number of cases generated for some fixed configurations of parameters. The neural network can give a rough estimate of the unknown parameters immediately even if the details of the heat transfer model are not known. On the contrary, the LM required knowledge of the heat transfer model and long computation times. The comparison shows that the machine learning analyzer may be advantageous in situations where the immediate estimate is needed, or the model of the phenomenon is not fully known (it can rely solely on the measurement data). It can be also used to obtain the initial estimate before a more accurate estimate from the second method is available.

It must be noted that the accuracy of the neural network method can be increased by raising the signal-to-noise ratio. The signal-to-noise ratio is especially low if laser pulses of low power and low pulse duration are measured. In this case, the temperature increase sensed at the back surface of the plate is low. To increase the accuracy, thinner measuring plates may be used. A version of the current laser diagnostics method in which melting occurs has been considered in preliminary numerical tests (employing the LM algorithm) and yielded promising results where retrieval of pulse parameters was possible (see [61]). Nevertheless, the presented ANN version was intended as non-destructive, i.e. without melting. To assure a good signal-to-noise ratio while avoiding melting, it is possible to use specialized plates of different thicknesses for different ranges of laser pulse energy. Such design was not applied in the present work but should be considered in the future. Changing the material of the plates to one withstanding higher temperature and having good thermal diffusivity (e.g. tungsten) can be used to produce thin measuring plates, which should assure an even better signal-to-noise ratio. Another method to reduce the influence of noise is software noise cancellation.

In the physical experiments conducted to test the gradient-based method [14] the typical level of spatial noise in IR images was about ±0.5 K, and temporal noise ±0.25 K, which corresponds to ANN estimation error of around 8%, based on Fig. 13. Shielding of the elements of the optical track and stabilization of the electric power sources should result in a further increase of the method accuracy. The levels of noise in the data captured in the real experiments [14] were pretty high for

8

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

a modern Stirling-cooled IR camera, which suggests that a better signal-to-noise ratio is still possible.

Now moving to comparisons with a broader spectrum of works, it should be said that the ANN approach was previously successfully applied in heat transfer engineering to solve various problems of estimation and prediction. For example, Czél *et al.* [23] performed numerical simulations regarding simultaneous estimation of temperature-dependent volumetric heat capacity and thermal conductivity in solid material and obtained a 1.216% average error for noisy inputs. Their ANN structure contained 34 output and 34 or 162 input neurons. Romero-Méndez *et al.* [33] reported a 2.46% error in the problem of convective heat transfer coefficient estimation in evaporative mini-tubes based on the knowledge of saturation temperature, refrigerant mass flow rate, and applied heat flux.

A case similar to the current one, involving a radiative source and temperature measurements, was presented by Mirsepahi *et al.* [29] who used shallow neural networks trained with LM to estimate the temporal change of irradiative heat generated by a heat lamp in a batch dryer based on the temperature history measurement on one of the surfaces of the enclosure, as measured by a single thermocouple. In their first paper, the mean of absolute error (MAE) was 43.00245 W for the input heat in the range of 0–2 kW. For an average input heat of 1 kW this translates to 4.3% MAPE. Data from physical experiments were used to train the networks, and the reported error of temperature measurement was ± 1 K. In the subsequent research [30] genetic algorithms were applied to optimize the number of neurons in hidden layers, which allowed to decrease the MAE to 35.05 W (MAPE ~3.5%). In the next paper [28], the tested physical system was extended to a dryer with two chambers, each heated by a single lamp. Each chamber was monitored with one thermocouple. The presented data allow estimating the MAPE in that case around 3%.

At the same level of sensor error, the MAPE in the present solution is much greater (16%), which can be explained by the fact that the current problem suffers more from ill-conditioning. For the irradiative dryer, the heat transfer between the lamp and the thermocouple is confined within the enclosure, which makes it well shielded from ambient noise, and the sensor is separated from the source only by air. In the current solution, the heat source is obstructed by a solid body, which is necessary to prevent sensor damage. The radial spread of heat within the plate and its thermal inertia greatly weakens the effect measured at the opposite surface which results in ill-conditioning. It is still possible to overcome these problems and increase the accuracy in future designs by measuring the heat effects directly at the irradiated surface.

7. CONCLUSIONS

- The ANN method is attractive as it can work purely on measurement data, without the heat transfer model (whose parameters are often hard to establish). Once trained, estimations are given immediately.
- The best results were obtained when the number of neurons in the hidden layers was more than or equal to the number of inputs.

- Changing the optimizer from Adam to Adagrad introduced a significant diminishing of error (from around 5 to 2.5% for undistorted data) and unlocked the potential of the second hidden layer.
- During the testing of different activation functions, it was found that the Softmax function leads to disproportionally large errors (90,136%). Errors while applying other functions were far smaller.
- Both subtracting the median and scaling the inputs to a similar range to outputs resulted in a profound increase in estimation accuracy.
- Increasing the number of available inputs is especially effective in lowering the estimation errors for noisy data.
- For the levels of noise measured in physical experiments (0.25-0.5 K), the accuracy of the ANN parameter estimation method is between 5 and 10%.
- The error level of the ANN retrieval method is much greater than the LM method with the same levels of added noise (see Fig. 13), but the latter one utilizes a more costly 3D heat transfer model and is much slower. The trained neural network gives an immediate estimate of the unknown parameters and therefore it can be used to generate the starting point for the LM method or to quickly obtain a rough estimate.
- The ANN method accuracy may still be increased in future designs, either through the utilization of plates of different thicknesses for different ranges of pulse energy or by changing the temperature measurement location to the irradiated front surface. Among these two solutions, the second one appears as one with greater potential.

CREDIT AUTHORSHIP STATEMENT

Karol Pietrak: Conceptualization, methodology, writing – original draft, writing – review & editing, formal analysis; **Radosław Muszyński**: Writing – original draft, methodology, software, investigation; **Adam Marek**: Software, investigation; **Piotr Łapka**: Project administration, software, resources.

ACKNOWLEDGEMENTS

REFERENCES

[1] A.K. Dubey and V. Yadava, "Laser beam machining – A review", *Int. J. Mach. Tools Manuf.*, vol. 48, no. 6, pp. 609–628, 2008.

[2] Q. Peng *et al.*, "Lasers in medicine", *Reports Prog. Phys.*, vol. 71, no. 5, p. 056701, May 2008.

[3] G.P. Perram, M.A. Marciniak, and M. Goda, "High energy laser weapons: technology overview", *Proc. SPIE*, vol. 5414, pp. 1–25, 2004.

[4] K. Pietrak and T.S. Wiśniewski, "Methods for experimental determination of solid-solid interfacial thermal resistance with application to composite materials", *J. Power Technol.*, vol. 94, no. 4, pp. 270–285, 2014.

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

9

K. Pietrak, R. Muszyński, A. Marek, and P. Łapka

[5] L. Bruno, "Mechanical characterization of composite materials by optical techniques: A review", *Opt. Lasers Eng.*, vol. 104, pp. 192–203, May 2018.

[6] C.B. Roundy, "Current technologies of laser beam profile measurements", in *Laser Beam Shaping*, F.M. Dickey and S.C. Holswade, Eds. New York: CRC Press, 2000.

[7] J.A. Hoffnagle and C.M. Jefferson, "Design and performance of a refractive optical system that converts a Gaussian to a flattop beam", *Appl. Opt.*, vol. 39, no. 30, pp. 5488–99, Oct. 2000.

[8] J.L. Guttman, "Noninterceptive Beam Profiling of High-Power Industrial Lasers", *Laser Tech. J.*, vol. 12, no. 5, pp. 20–23, Nov. 2015.

[9] M. Kujawińska, P. Łapka, M. Malesa, K. Malowany, M. Prasek, and J. Marczak, "Investigations of high power laser beam interaction with material by means of hybrid FVM-FEM and digital image correlation methods," in *Proc. SPIE 10159*, 2016, p. 1015916.

[10] O. Aharon, "High power beam analysis", *Ind. Laser Solut. Manuf.*, vol. 28, no. 5, pp. 17–20, 2013.

[11] M.A. Sutton, J.J. Orteu, and H. Schreier, *Image Correlation for Shape, Motion and Deformation Measurements*. Springer US, 2009.

[12] K. Pietrak, P. Łapka, and M. Kujawińska, "Inverse analysis for the identification of temporal and spatial characteristics of a short super-Gaussian laser pulse interacting with a solid plate", *Int. J. Therm. Sci.*, vol. 134C, pp. 585–593, 2018.

[13] M.N. Ozisik and H.R.B. Orlande, *Inverse heat transfer*. New York: Taylor&Francis, 2000.

[14] P. Łapka, K. Pietrak, M. Kujawińska, and M. Malesa, "Development and validation of an inverse method for identification of thermal characteristics of a short laser pulse", *Int. J. Therm. Sci.*, vol. 150, Apr. 2020.

[15] J.V. Beck and K.A. Woodbury, "Inverse problems and parameter estmation: integration of measurements and analysis", *Meas. Sci. Technol.*, vol. 9, no. 6, pp. 839–847, Jun. 1998.

[16] J. Taler and D. Taler, "Measurement of Heat Flux and Heat Transfer Coefficient", in *Heat Flux, processes, measurement techniques and applications*, G. Cirimele and M. D'Elia, Eds. New York: Nova Science Publishers, Inc, 2012, pp. 1–104.

[17] J. Taler and D. Taler, "Surface-heat transfer measurements using transient techniques", in *Encyclopedia of Thermal Stresses*, R. B. Hetnarski, Ed. Dordrecht Heidelberg New York London: Springer, 2014, pp. 4774–4784.

[18] M.S. Gadala and S. Vakili, "Assessment of Various Methods in Solving Inverse Heat Conduction Problems", in *Heat Conduction – Basic Research*, V. Vikhrenko, Ed. InTech, 2011.

[19] J. Krejsa, K.A. Woodbury, J.D. Ratliff, and M. Raudensky, "Assessment of strategies and potential for neural networks in the inverse heat conduction problem", *Inverse Probl. Eng.*, vol. 7, no. 3, pp. 197–213, Jun. 1999.

[20] X. Wang and M. Yao, "Neural networks for solving the inverse heat transfer problem of continuous casting mould", in *2011 Seventh International Conference on Natural Computation*, 2011, vol. 2, pp. 791–794.

[21] M. Grabarczyk and P. Furmański, "Predicting the effective thermal conductivity of dry granular media using artificial neural networks", *J. Power Technol.*, vol. 93, no. 2, pp. 59–66, May 2011.

[22] K. Goudarzi, A. Moosaei, and M. Gharaati, "Applying artificial neural networks (ANN) to the estimation of thermal contact conductance in the exhaust valve of internal combustion engine", *Appl. Therm. Eng.*, vol. 87, pp. 688–697, Aug. 2015.

[23] B. Czél, K.A. Woodbury, and G. Gróf, "Simultaneous estimation of temperature-dependent volumetric heat capacity and thermal conductivity functions via neural networks", *Int. J. Heat Mass Transf.*, vol. 68, pp. 1–13, Jan. 2014.

[24] S. Chudzik, S. Grys, and W. Minkina, "The application of the artificial neural network and hot probe method in thermal parameters determination of heat insulation materials Part 1-thermal model consideration", in *Proceedings of the IEEE International Conference on Industrial Technology*, 2009.

[25] S. Chudzik, "Measuring system with a dual needle probe for testing the parameters of heat-insulating materials", *Meas. Sci. Technol.*, vol. 22, no. 7, p. 075703, Jun. 2011.

[26] S. Chudzik, "Measurement of thermal diffusivity of insulating material using an artificial neural network", *Meas. Sci. Technol.*, vol. 23, no. 6, 2012.

[27] S. Chudzik, W. Minkina, and S. Grys, "The application of the artificial neural network and hot probe method in thermal parameters determination of heat insulation materials. Part 2 – application of the neural network", 2009, pp. 1–5.

[28] A. Mirsepahi, L. Chen, and B. O'Neill, "An Artificial Intelligence Solution for Heat Flux Estimation Using Temperature History; A Two-Input/Two-Output Problem", *Chem. Eng. Commun.*, vol. 204, no. 3, pp. 289–294, Mar. 2017.

[29] A. Mirsephai, M. Mohammadzaheri, L. Chen, and B. O'Neill, "An artificial intelligence approach to inverse heat transfer modeling of an irradiative dryer", *Int. Commun. Heat Mass Transf.*, vol. 39, no. 1, pp. 40–45, Jan. 2012.

[30] A. Mirsepahi, L. Chen, and B. O'Neill, "A comparative artificial intelligence approach to inverse heat transfer modeling of an irradiative dryer", *Int. Commun. Heat Mass Transf.*, vol. 41, pp. 19–27, Feb. 2013.

[31] A. Mirsepahi, L. Chen, and B. O'Neill, "A comparative approach of inverse modelling applied to an irradiative batch dryer employing several artificial neural networks", *Int. Commun. Heat Mass Transf.*, vol. 53, pp. 164–173, Apr. 2014.

[32] R.-M. Ricardo, H.-L. Juan Manuel, D.-G. Héctor Martín, and P.-V. Arturo, "Use of Artificial Neural Networks for Prediction of Convective Heat Transfer in Evaporative Units", *Ing. Investig. y Tecnol.*, vol. 15, no. 1, pp. 93–101, Jan. 2014.

[33] R. Romero-Méndez, P. Lara-Vázquez, F. Oviedo-Tolentino, H.M. Durán-García, F.G. Pérez-Gutiérrez, and A. Pacheco-Vega, "Use of Artificial Neural Networks for Prediction of the Convective Heat Transfer Coefficient in Evaporative Mini-Tubes", *Ing. Investig. y Tecnol.*, vol. 17, no. 1, pp. 23–34, Jan. 2016.

[34] M. Mohanraj, S. Jayaraj, and C. Muraleedharan, "Applications of artificial neural networks for thermal analysis of heat exchangers – A review", *Int. J. Therm. Sci.*, vol. 90, pp. 150–172, 2015.

[35] O. Cortés, G. Urquiza, and J.A. Hernández, "Optimization of operating conditions for compressor performance by means of neural network inverse", *Appl. Energy*, vol. 86, no. 11, pp. 2487–2493, Nov. 2009.

[36] J.A. Hernández, D. Colorado, O. Cortés-Aburto, Y. El Hamzaoui, V. Velazquez, and B. Alonso, "Inverse neural network for optimal performance in polygeneration systems", *Appl. Therm. Eng.*, vol. 50, no. 2, pp. 1399–1406, Feb. 2013.

[37] Y. El Hamzaoui, B. Ali, J. A. Hernandez, O. Cortés-Aburto, and O. Oubram, "Search for Optimum Operating Conditions for a Water Purification Process Integrated to a Heat Transformer with Energy Recycling using Artificial Neural Network Inverse Solved by Genetic and Particle Swarm Algorithms", *Chem. Prod. Process Model.*, vol. 7, no. 1, Jan. 2012.

[38] S. Samarasinghe, *Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition.* Boca Raton: Auerbach, 2007.

[39] Engineering ToolBox, "Metals – Melting Temperatures," 2005. [Online]. Available: https://www.engineeringtoolbox.com/melting-temperature-metals-d_860.html. [Accessed: 28-May-2018].

[40] Y.A. Cengel and A.J. Ghajar, *Heat and Mass Transfer, Fundamentals & Applications*, 5th-th ed. New York: McGraw-Hill Education, 2015.

[41] "FLIR SC7500 instrument specifications", *Flir Systems Inc.*, 2014. [Online]. Available: https://issuu.com/netservis/docs/flir-sc7500.

[42] A. Gulli and S. Pal, *Deep Learning with Keras.* Packt Publishing Ltd, 2017.

[43] S. Ruder, "An overview of gradient descent optimization algorithms," Sep. 2016.

[44] B. Ding, H. Qian, and J. Zhou, "Activation functions and their characteristics in deep neural networks", in *Proceedings of the 30th Chinese Control and Decision Conference, CCDC 2018*, 2018, pp. 1836–1841.

[45] C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation Functions: Comparison of trends in Practice and Research for Deep Learning," Nov. 2018.

[46] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)", *4th Int. Conf. Learn. Represent. ICLR 2016 – Conf. Track Proc.*, Nov. 2015.

[47] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-Normalizing Neural Networks", *Adv. Neural Inf. Process. Syst.*, vol. 2017-December, pp. 972–981, Jun. 2017.

[48] Keras Documentation, "Earlystopping", 2020. [Online]. Available: https://keras.io/api/callbacks/early_stopping/. [Accessed: 29-Jul-2020].

[49] R.J. Hyndman and A.B. Koehler, "Another look at measures of forecast accuracy", *Int. J. Forecast.*, vol. 22, no. 4, pp. 679–688, Oct. 2006.

[50] J. Brownlee, "How to ConFigure the Number of Layers and Nodes in a Neural Network," 2018. [Online]. Available: https://machinelearningmastery.com/how-to-conFigure-the-number-of-layers-and-nodes-in-a-neural-network/. [Accessed: 30-Jul-2020].

[51] N. Srivastava, G. Hinton, A. Krizhevsky, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," 2014.

[52] D. Stathakis, "How many hidden layers and nodes?", *Int. J. Remote Sens.*, vol. 30, no. 8, pp. 2133–2147, 2009.

[53] V. Kůrková, "Kolmogorov's theorem and multilayer neural networks", *Neural Networks*, vol. 5, no. 3, pp. 501–506, Jan. 1992.

[54] G. Cybenko, "Approximation by superpositions of a sigmoidal function", *Math. Control. Signals, Syst.*, vol. 2, no. 4, pp. 303–314, Dec. 1989.

[55] W.S. Sarle, "Should I normalize/standardize/rescale the data?", 2002. [Online]. Available: http://www.faqs.org/faqs/ai-faq/neural-nets/part2/section-16.html. [Accessed: 24-Aug-2020].

[56] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift", in *32nd International Conference on Machine Learning, ICML 2015*, 2015, vol. 1, pp. 448–456.

[57] J. Brownlee, "How to use Data Scaling Improve Deep Learning Model Stability and Performance", 2019. [Online]. Available: https://machinelearningmastery.com/how-to-improve-neu ral-network-stability-and-modeling-performance-with-data-sca ling/. [Accessed: 29-Jul-2020].

[58] S. Nayak, B.B. Misra, and H.S. Behera, "Impact of Data Normalization on Stock Index Forecasting", *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, vol. 2014, pp. 257–269, 2014.

[59] Tibo Spotfire Documentation, "Normalization by subtracting the median," 2014. [Online]. Available: https://docs.tibco.com/pub/spotfire/6.5.0/doc/html/norm/norm_subtract_the_median.htm. [Accessed: 25-Aug-2020].

[60] J. Jones, "Stats: Measures of Central Tendency," 2020. [Online]. Available: https://people.richland.edu/james/lecture/m170/ch03-ave.html. [Accessed: 25-Aug-2020].

[61] K. Pietrak, P. Łapka, and M. Kujawińska, "Numerical tests of an inverse method for the characterization of pulsed and continuous laser beams", in Contemporary Issues of Heat and Mass Transfer, 2019, vol. 1, pp. 481–502.

Bull. Pol. Acad. Sci. Tech. Sci., vol. 70, no. 1, p. e140100, 2022

11