Neural network modeling of the laser material-removal process

Yousef, Basem F.; Knopf, George K.; Bordatchev, Evgueni V.; Nikumb, Suwas K.

Publisher's version / Version de l'éditeur:
https://doi.org/10.1117/12.452660
Sensors and Controls for Intelligent Manufacturing II, 2001-12-27

Neural network modeling of the laser material-removal process

Basem F. Yousef [a], George K. Knopf [a], Evgueni V. Bordatchev[b], Suwas K.Nikumb[b]
[a] Mechanical and Materials Engineering, The University of Western Ontario, London, Canada

[b] Integrated Manufacturing Technologies Institute, National Research Council of Canada, London, Canada

ABSTRACT

Industrial lasers are used extensively in modern manufacturing for a variety of applications because these tools provide a highly focused energy source that can be easily transmitted and manipulated. For micro-machining, the quantity of material removed and the roughness of the finished surface are a function of the crater geometry formed by a laser pulse with specific energy (power). Laser micro-machining is, however, a complex nonlinear process with numerous stochastic parameters related to the laser apparatus and the material specimen. Consequently, the operator must manually set the process control parameters by "trial-and-error". This paper describes how an artificial neural network can be used to create a nonlinear model of the laser material-removal process in order to automate micro-machining tasks. The multi-layered neural network predicts the pulse energy needed to create a crater of specific depth and average diameter. Laser pulses of different energy levels are impinged on the surface of the test material in order to investigate the effect of pulse energy on the resulting crater geometry and volume of material removed. Experimentally acquired data from several sample materials are used to train and test the network's performance. The key system inputs for the modeler are the mean depth and diameter of the crater, and the system outputs are pulse energy and variance of depth and diameter. The preliminary study using the experimentally acquired data demonstrates that the proposed network can simulate the behavior of the physical process to a high degree of accuracy. Future work involves investigating the effect of different input parameters on the output behavior of the process in hopes that the process performance, and the final product quality, can be improved.

Keywords: Laser micro-machining, pulse energy, crater geometry, artificial neural network.

1. INTRODUCTION

Laser beams are used extensively as a cutting tool in a variety of industrial applications because they provide highly concentrated energy sources that can be easily transmitted and manipulated. Micro-mechanical structures are becoming pervasive, and their demand is accompanied with demand for new micro-fabrication tools. Conventional mechanical machining can produce work pieces and assemblies with feature sizes larger than a few hundred microns (thousands of inch), nevertheless, the steadily increasing demand for smaller sizes requires new tools and processes. As feature sizes fall below 25 μm (one thousandth of an inch), mechanical approaches to cutting, drilling and shaping these materials may be replaced with photon or particle beam techniques such as lasers. Laser processing is among the most promising tools for micro-machining because it can process features down to the size of the laser wavelength (smaller than micrometer) with extremely high precision, superior repeatability, and cost-effectiveness. Further, the process is non-contact, dry and clean, and can be easily automated.

Laser micro-machining is a complex nonlinear process as shown in Figure 1[1]. The stochastic nature and large number of process parameters needed to control precisely the laser material removal process, makes it very difficult to improve. Moreover, all these factors make the modeling of such a process by conventional analytical and numerical methods a difficult task. In addition, the laser machining system operator has to make numerous decisions in order to set appropriate process control parameters. This "trial-and-error" method is time consuming and costly especially for small batch production or prototyping, and does not ensure near *optimality* with a given set of process conditions and manufacturing objectives.

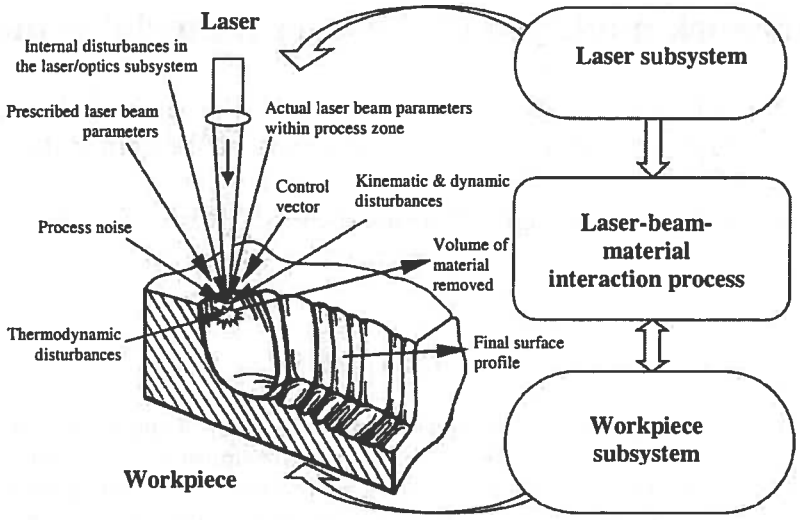

Fig. 1: Laser machining system and controlling parameters [1].

The goal of this research is to investigate, analyze and model how the geometry of the final surface profile forms and its dependence on the laser pulse energy. An artificial neural network model of the laser material removal process has been developed in order to find more satisfactory conditions and parameters for a particular setup. The significance of this work lies with the fact that the developed "adaptive" neural network model will provide a powerful platform from which further expansion of the knowledge base in the field can be conveniently launched. The improved knowledge will lead to the development of better predictive capabilities when novel applications or materials are considered. The results and experience gained in this project will provide valuable insights into future development of quantitative prediction capabilities in other areas of the field, which is also fast advancing. Practically, the knowledge to be obtained in this project will enable better process design, reduce the need for trial-and-error and eliminate the need for frequent human oversight. This will also reduce the need for post-machining operation(s) as a result of improved quality. The expanded technology base, along with improved quality and thus economic viability, will greatly encourage more and new laser micro-machining industrial applications in diverse fields and other micro device manufacturing.

2. NEURAL NETWORK MODEL OF THE PROCESS

The only information available from the laser system is a set of labeled input-output data generated by the system. Because of the lack of mathematical formulas that describe the laser-material interaction, there aren't many options other than artificial intelligence solutions, which have capabilities to formulate relationships between the various process parameters without the need for information about the internal details of the process. One of those solutions that perform fast non-linear function mapping capabilities between inputs and outputs is the artificial neural network (ANN). The strength of neural structures lies in the speed of parallel computing as well as their learning and adaptive abilities in a changing environment. And since any laser machining process requires a laser beam, light guiding optics and optical devices for detection, which may deteriorate with age and use, the adaptability of the modeling approach can be a very attractive aspect to meet this challenge.

The input signal x_i (i=1,2,...,p) to each individual neuron is modified by a weight w_{ji} where the subscript means the strength of the interconnection between the i^{th} input signal and neuron j. The weighted inputs are then summed with a bias weight w_{jo} and the resultant is passed through nonlinearity, called an activation function, modeled by a sigmoid or hyperbolic tangent. The operation can be mathematically expressed as

$$y_j = f(W^T X) \tag{1}$$

where y_j is the output of neuron j, W is the interconnection weights vector, X is the input signals vector, and f is the activation function given by

$$f(\cdot) = \frac{1}{1+e^{-a(\cdot)}} \tag{2}$$

Individual neurons are able to perform only simple mathematical or logic operations [2]. The modeling of complex input-output behavior is achieved by connecting numerous simple neurons in multi-layered structures. Often the neurons are fully connected between layers as shown in Figure 2.

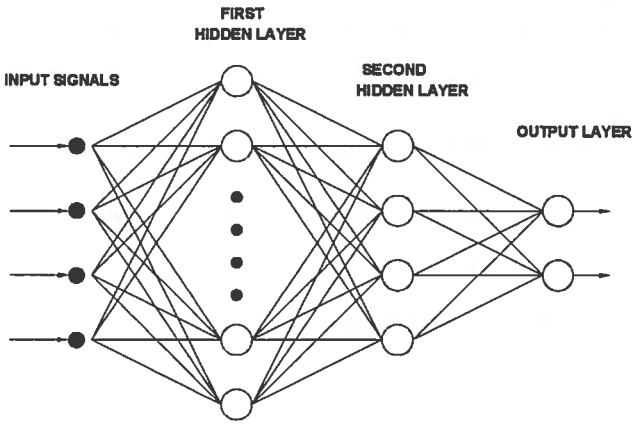

Fig. 2: Typical multi-layer perceptron.

To model a manufacturing process such as laser-material removal, the neural network must undergo both learning and operation phases. During the learning phase, the network first layer's neurons are fed with a randomly selected input $x_i(n)$. Each neuron j in layer l sends its calculated output $y_j^l(n)$, based on the interconnection weights $w_{ji}^l(n)$, to the next layer's neurons as an input, and so on, until this signal is forwarded up through the network. At the output layer, 0, the desired output d_k of neuron k is compared with the actual (simulated) output of the neuron $y_k^0(n)$. Since the outputs, in general will not be the same as the desired values, an error signal $e_j(n)$ initiates at those output neurons and back propagates through the layers of the network updating the interconnection weights by a step change $\Delta w_{ji}^l(n)$. After the network completes the feedforward pass of the computed output signal by each neuron, and the backpropagation pass of the error signal that updates the weights, another input/output training pair is implemented to the network and the process is repeated. This learning is continued until obtaining a set of weights and biases that satisfies the all input output pairs.

The backpropagation algorithm used by the neural network during the learning phase is summarized as follows:

Step 0. Set the learning rate parameter α and the momentum constant μ to small values from 0.1 to 1.
Determine the number of hidden layers as well as the number of neurons per layer.
Determine the maximum number of epochs.
Set the minimum system error E_{av} to be maintained.

Step 1. Initialize the weights w_{ji} and biases w_{jo} for all layers to small random values between -0.5 to +0.5

Step 2. While the stopping conditions are false, do Steps 3 - 10.

Step 3. For each training pair $(X(n), d(n))$ do Steps 4 - 10.

***Feedforward*:**

Step 4. Determine the response $Y^1(n)$ of the neurons in the first layer using

$$u_j^1(n) = W^1(n)^T X(n) \tag{3}$$

$$y_j^1(n) = f\ (u_j^1(n)) \tag{4}$$

where f is the activation function – given by equation (2) - which is used by neurons of the first layer. Notice that the superscript index indicates the layer's number.

Step 5. Determine the response of the neurons in each of the next hidden layers and the output layer using

$$u_j^l(n) = W^l(n)^T\ Y^{l-1}(n) \tag{5}$$

$$y_j^l(n) = f\ (u_j^l(n)) \tag{6}$$

where f is the activation function in layer l and $Y^{l-1}(n)$ are the responses of neurons of the layer preceding layer l.

Step 6. Determine the mean squared error (or system error) $E(n)$ associated with each pattern n using

$$E(n) = \frac{1}{2}\sum_{j\in O}(d_j(n) - y_j^0(n))^2 \tag{7}$$

where the set O includes all the neurons in the output layer and $d_j(n)$ is the desired output at neuron j .

Step 7. Determine the averaged (normalized) system error E_{av} using

$$E_{av} = \frac{1}{N}\sum_{n=1}^{N} E(n) \tag{8}$$

where N is the total number of training patterns.

Back-propagation of error:

Step 8. Compute the error information terms, $\delta_j^0(n)$, and calculate the weight correction terms for the output layer neurons using

$$\delta_j^0(n) = (d_j(n) - y_j^0(n))\ f'(u_j^0(n)) \tag{9}$$

$$\Delta w_{ji}^0(n) = \mu\Delta w_{ji}^0(n-1) + \alpha\delta_j^0(n)\ y_i^{0-1}(n) \tag{10}$$

$$\Delta w_{j0}^0(n) = \mu\Delta w_{jo}^0(n-1) + \alpha\delta_j^0(n) \tag{11}$$

where f' is the first derivative of the transfer function f .

Step 9. Compute the error information terms and the weight correction terms for all neurons in the previous (hidden) layers starting with layer l that precedes the output layer and propagating backwards up to the 1st hidden layer,

$$\delta_j^l(n) = (\sum_{k=1}^{m}\delta_k^{l+1}(n)\ w_{kj}^{l+1}(n))\ f'(u_j^l(n)) \tag{12}$$

$$\Delta w_{ji}^l(n) = \mu\Delta w_{ji}^l(n-1) + \alpha\delta_j^l(n)\ y_i^{l-1}(n) \tag{13}$$

$$\Delta w_{jo}^l(n) = \mu\Delta w_{jo}^l(n-1) + \alpha\delta_j^l(n) \tag{14}$$

where m is the number of neurons in layer l+1.

Update weights:

Step 10: Update weights for all layers using:

$$w_{ji}^l(n)_{(new)} = w_{ji}^l(n)_{(old)} + \Delta w_{ji}^l(n) \tag{15}$$

Step11: Test for stopping condition:
If the chosen maximum number of epochs is reached or if the normalized system error calculated in step 7 is smaller than pre-set value in step 0, then STOP; otherwise continue.

3. EXPERIMENTAL SET-UP

Experiments were conducted to provide the neural network with the necessary information needed to train the system. The experimental data are used by the neural network as a reference and restored during the operation phase at which the network is performing the simulation of the actual process for specific test conditions.

3.1 Training data

A brass foil was impinged with laser pulses of different energies to investigate the effect of pulse energy on the resulting crater geometry and the volume of material removed. Using a diode pumped Nd:YAG laser 50 craters were produced with each pulse energy. The simple geometry of Figure 3 was used to resemble the crater geometry and simplify further necessary calculations such as the volume of the crater, which was approximated using:

$$V=\frac{\pi}{2}abh \tag{16}$$

where, h is the depth and a and b are the diameters along two directions.

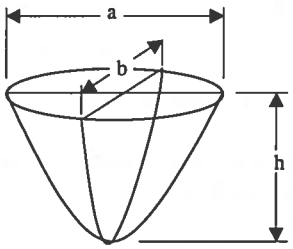

Fig. 3: Crater parameters.

Crater depths and diameters along two profiles (a and b) were measured and average diameter *d* was calculated, and when and when the crater depth was plotted against the pulse energy, the behavior of the process showed that craters with different depths were produced with the same pulse energy as shown in Figure 4. This behavior, which is the variation in depth, will be referred to as "haze around the mean". Figure 4 illustrates that the "haze around the mean" increases as the pulse energy increases. Same behaviors were found in Figures 5 and 6 that show crater average diameter and the volume of material removed versus pulse energy, respectively.

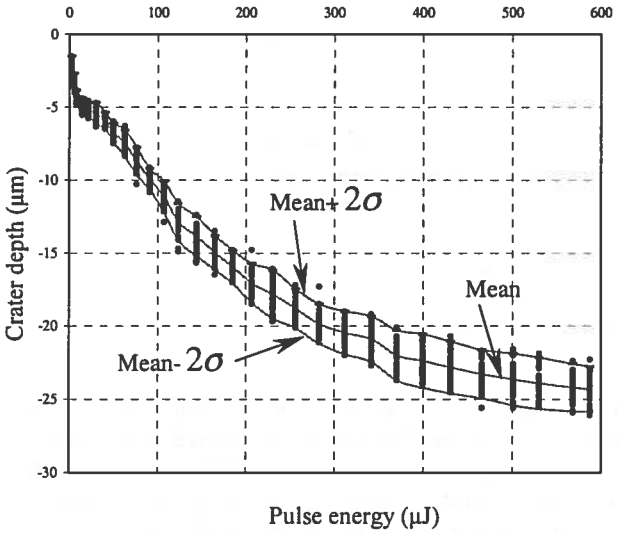

Fig. 4: Crater depth vs. pulse energy (Brass).

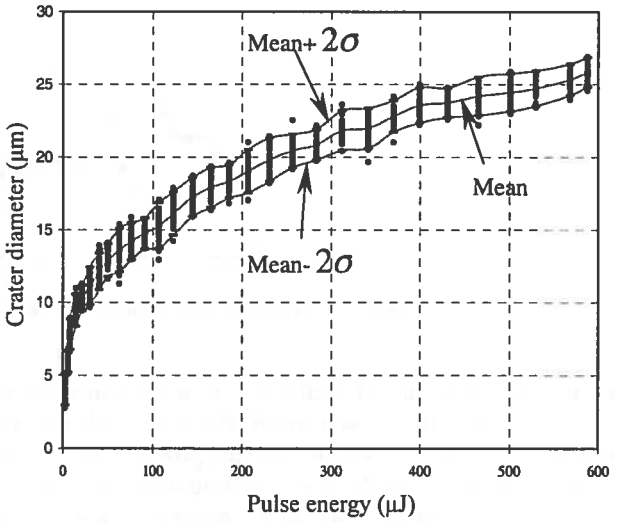

Fig. 5: Diameter vs. pulse energy (Brass).

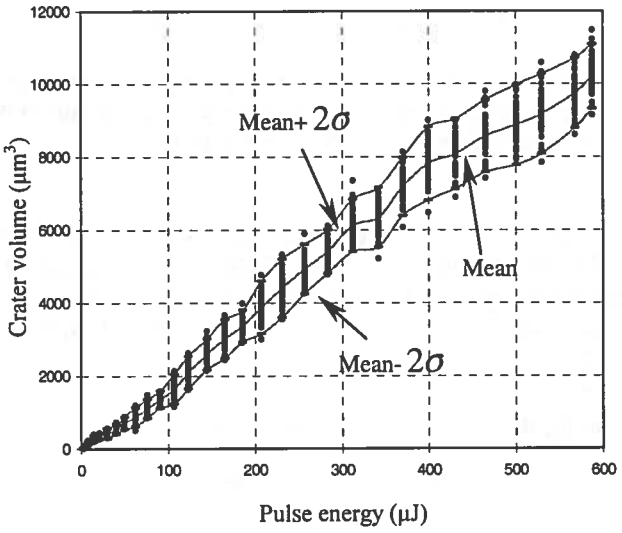

Fig. 6: Volume of material removed vs. pulse energy.

The curves behaviors are attributed to some *process parameters*. Since cutting speed (or feed rate), frequency and duration of pulse are not involved, there are two main factors causing this behavior:

(1) Variation within the pulse energy.
(2) The dynamics of the laser-material interaction.

The first factor gives the explanation for the haze around the mean. Theoretically, the pulse energy is assumed to be constant if the same Q-switch is used. But in fact, this is not a valid assumption, since the actual pulse energy changes from one pulse to another although the same Q-switch is used. Figure 7 shows the variation in depth of material removed by pulses having pulse energy of 40.4 μJ. The same curvilinear behavior is obtained if the diameter or the volume is plotted against pulse number.

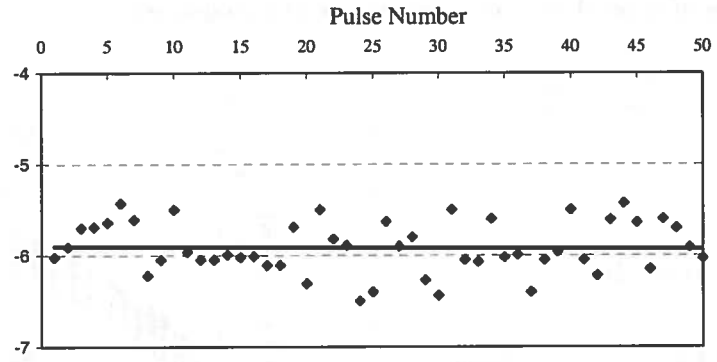

Fig. 7: Variation of depth for craters produced by pulses with a pulse energy of 40.4 μJ.

The effect of the second factor could be explained by referring to Figure 8. The layer of volume removed from the material depends on the wavefront of the focused laser beam profile, beam mode characteristics, surface reflectivity of the work piece material and its original profile which include the surface irregularities at each point of that local area that had been impinged by the wavefront photons. Subsequent impingement of photons within the single pulse defines the conformed shape of the crater as shown in Figure 8. This adds up to the chaos in material removal inside the crater once the crater starts to deepen. When experiments were carried out using other materials such as stainless steel, and copper, similar curve behaviors were obtained.

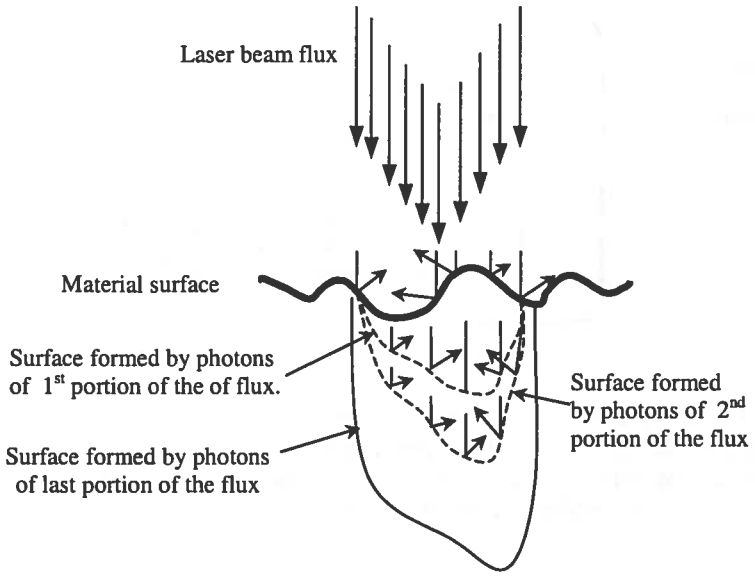

Fig. 8: Mechanism of material removal by a laser pulse.

Referring to Figures 4,5,and 6, it is obvious that the haze around the mean is less at the very beginnings of the curves due to the fact that at lower energies, there is less chance for heat loss and dissipation in the workpiece than the case when higher energies are used. And since the haze is very small at low energies, this indicates that cutting with low energies is accompanied by less damage for the workpiece and therefore, better cut quality.

3.2 Neural network modeler

After analyzing the experimental data and closely studying the process output, the next step was to simulate the process behavior. Feedforward back-propagation neural network modelers were used to simulate the process. Two stages of simulations were conducted as follows:

- An artificial neural network (ANN1) modeler simulates the energy needed to produce a crater with specific dimensions (depth and diameter) for a specified material, i.e. the modeler takes the crater depth and diameter as inputs and gives the required pulse energy as an output.
- The energy simulated in the first stage along with the depth and diameter are used as inputs to the second-stage (ANN2) modeler, and the outputs are the standard deviations σ_{depth} and $\sigma_{diameter}$, associated with that depth and diameter- respectively.

It should be noticed that the mentioned stages describe the modeler in its operation mode, which comes after the training -or learning- phase (described previously in the algorithm steps). Figure 9 shows the control box that describes the modelers' performance in its operation mode.

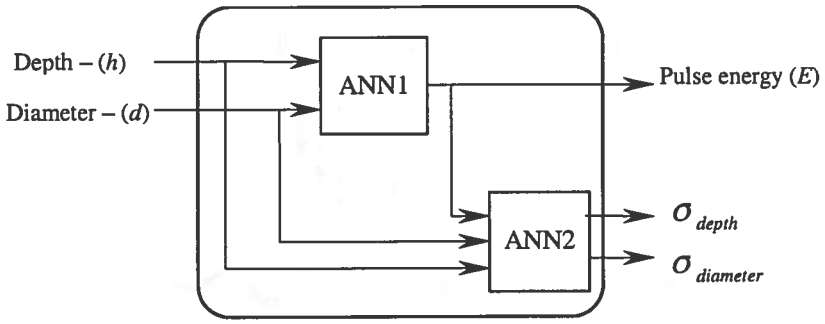

Fig. 9: ANN model in operation mode.

One can notice that ANN2 modeler works as a certainty measure for the first modeler's output. For example, the first modeler ANN1 predicts that an energy E_1 is needed to produce a crater of depth h and diameter d. The second modeler predicts the uncertainty associated with h and d by giving the standard deviations of depth and diameter, and these in tern give the expected variations in the produced crater depth and diameter.

4. SIMULATIONS AND RESULTS

Simulations were carried out for the following cases:

Case 1 The first run for ANN1 modeler was performed to simulate the energy corresponding to a specific mean depth and mean diameter. The 6X6 network simulation results are shown in Figures 10 and 11. The Figures show that ANN1 modeler was successful in modeling the process behavior to a high degree of accuracy.

Case 2 In order to simulate the variation in depth and diameter associated with each energy, ANN2 modeler was used. The simulation results of the 20X5 network are shown in Figures 12 and 13. Again, the figures show that the modeler ANN2 could simulate the anticipated output successfully.

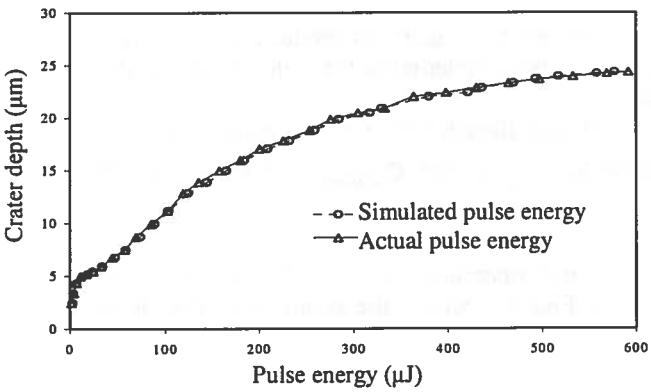

Fig. 12: Crater depth vs simulated and actual energy (brass)

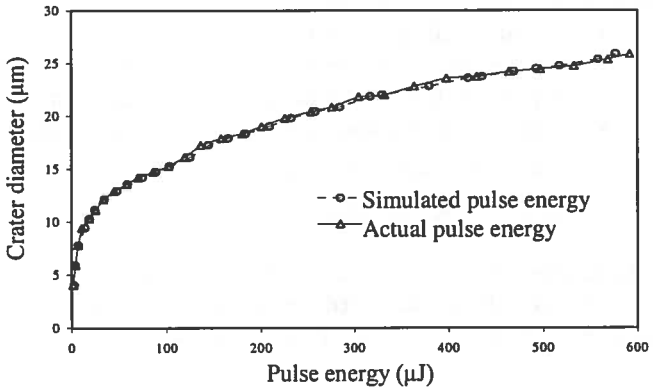

Fig. 13: Crater diameter vs simulated and actual energy (brass)

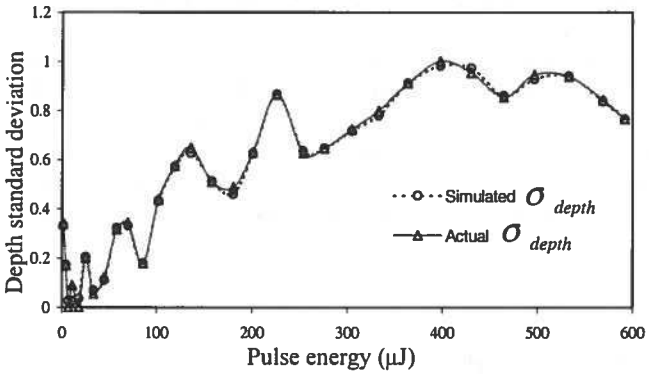

Fig. 14: Diameter standard deviation vs. pulse energy.

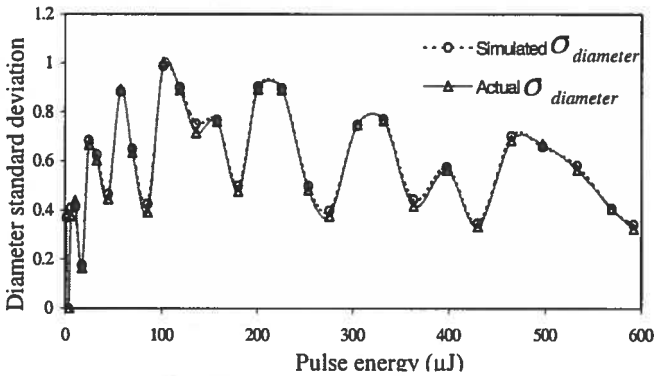

Fig. 15: Diameter standard deviation vs. pulse energy.

When comparing the network structures used in ANN1 and ANN2, the difference was in the number of neurons of the first hidden layer and the number of iterations (epochs). These two parameters have significant role in achieving the anticipated goal. While Figure 13 shows that the simulation was successful with 20 units in the first hidden layer, Figure 14 shows that simulating $\sigma_{diameter}$ using 10 units in the first hidden layer has failed.

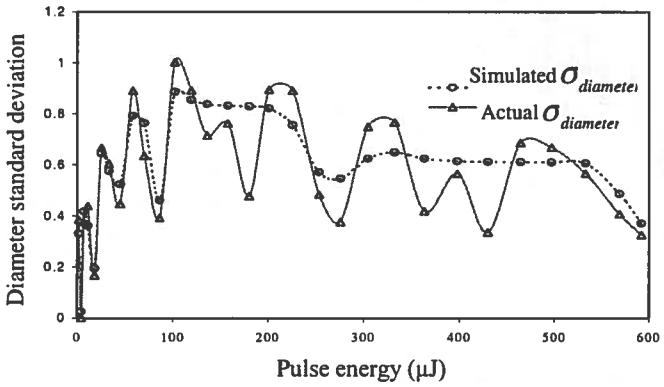

Fig. 16: Diameter standard deviation vs. pulse energy (10 neurons in first hidden layer).

The work of the number of epochs and the number of neurons in the first hidden layer could be summarized as follows:

1) The number of neurons in the first layer is responsible for defining the main features of the anticipated output curve and mapping the inputs to high order output curves.
2) The number of epochs and the number of neurons in other hidden layers are responsible for tuning the curve features and details –which are picked up by the first layer neurons- and move it closer to the solution.

Case 3 These simulations were conducted to investigate the effect of pulse energy on diameter. Simulations were carried out using ANN1 modeler. In this case, mean diameter and mean depth were fed as inputs to the network and the output was the pulse energy. Simulations were repeated by increasing and decreasing the diameter by a certain percentage. Each diameter, together with the depth that was used in the first trial of this case, was used as input to the modeler. For instance, while the inputs of the first trial were the mean depth and mean diameter, the inputs for the second trial were the same mean depth (unchanged) and 105% of the mean diameter, and so on. In this way, the change in simulated energy is only a result of the change in diameter since the depth was kept constant for all simulations.

Simulations were performed for the following diameter changes: 50%, 80%, 90%, 95%, 105% and 110% of mean diameter. The results of simulations -superimposed on the experimental data- are shown in Figure 16, which shows the effect of changing the energy on the diameter.

Bearing in mind that each point on any curve of Figure 16 represents a crater with certain diameter and depth that correspond to certain energy, the figure can be red as follows: Considering the crater with 20 μm diameter and 18.1 μm depth, according to the simulation curve labeled "Mean depth-mean diameter curve", the corresponding simulated pulse energy E =246.6 μJ. Now if the diameter is increased by 10% (i.e. diameter=22 μm) while the depth is kept constant, the simulated energy corresponding to this crater is red from the simulation curve "110% mean dia." and found to be E_1=282 μJ, and so on. In this way the simulation gives the relationship between the energy and diameter. Although E_1 corresponds to 110% mean dia and E_2 corresponds to 90% mean dia., one can observe the nonlinear relationship by comparing the unequal energy difference (E_1-E) with (E-E_2). It is noteworthy to mention that the same technique may be used to investigate the effect of the energy on depth.

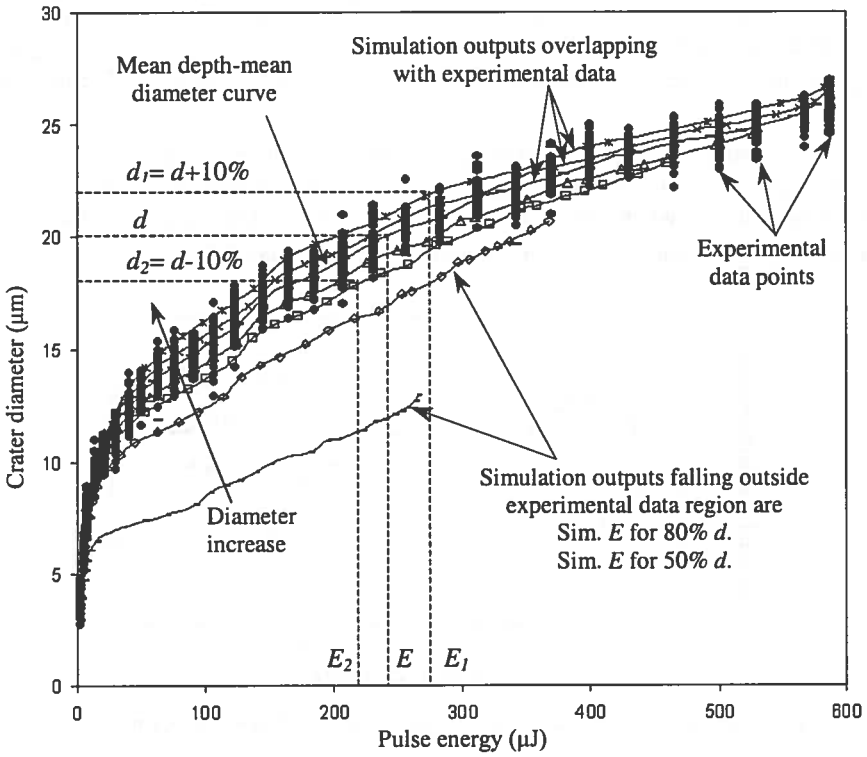

Fig. 17: Superimposing simulation outputs on experimental data.

In order to check the validity of the model results, the simulation outputs are compared with the available experimental data by superimposing the simulation results on the same graph of the experimental data. Figure 17 illustrates that the experimental data were consistent with simulation results for simulation curves that lay inside the experimental data region. One can observe that simulation curves for energy corresponding to 80% and 50% of mean diameter lay outside the experimental data region. This makes it difficult to prove the validity of the simulation outputs for these curves. In the next section, a visualization method is presented to enhance both the interpretation of the available experimental data and the validation of the simulation outputs.

5. DISCUSSION

In this section, the experimental data and simulation outputs were presented graphically using 3D modeling software to enhance simulated and experimental data visualization and comparison. When the pulse energy is plotted as the x-axis, the depth as the y-axis and the diameter as the z-axis, the data are found to be randomly spreading in planes parallel to the diameter-depth plane or (y-z plane). The regions in which the data are spreading were confined by ellipses. The axes of each of these ellipses were obtained using ($6\sigma_{depth}$) and ($6\sigma_{diameter}$), where σ_{depth} and $\sigma_{diameter}$ are the standard deviations for the depths and diameters respectively, i.e. a 2-diamensional (3σ normal-distribution) standard deviation was used to cover the experimental data area associated with each pulse energy (see Figure 18). Then using a 3D modeling software (3D Studio MAX), the ellipses were extruded to form a mesh, which represents the volume that embodies the experimental data points. In other words, points inside the mesh represent craters of specific depths and diameters that correspond to specific pulse energies (Figure 18). When superimposing the simulations results of Case 3 on the same graph of the mesh, Figure 19 shows that all simulation curves fall inside the experimental-data mesh except those belonging to 80% mean diameter and 50% mean diameter.

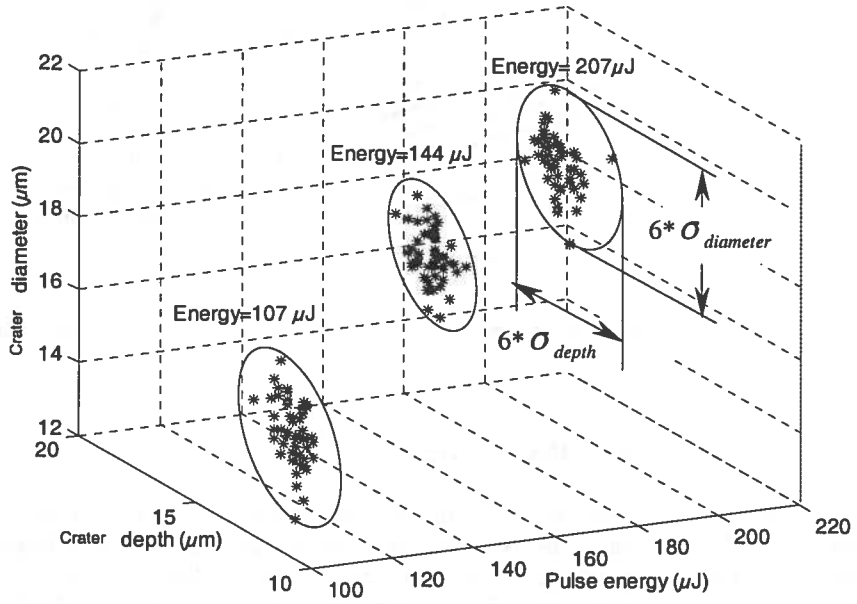

Fig. 18: Elliptical regions confining the experimental data areas associated with 3 energy levels.

Since it is difficult to experimentally check the validity of such simulation curves because of the lack of experimental data, the following technique was used. In Figure 19, simulation curves, which fall inside the mesh (i.e. those which are assisted by available experimental data), were connected, by curve "A", at points (craters) which all have the same depth of 19.85 μm. Then this curve is extended (extrapolated) in the direction of the simulation curve corresponding to 80% mean diameter. This extension should intersect the later simulation curve at the point (crater) that has a depth of 19.85 μm of all other craters on curve "A". The end point of the extension did not perfectly intersect the simulation curve at the anticipated point. Based on simulation results, the energy at the anticipated point of intersection that corresponds to a depth of 19.85 μm was supposed to be 241 μJ while the energy at the extension endpoint is found to be 236 μJ. This corresponds to an error of 2%. For verification purposes, the same technique was repeated for other depths and the similar results were obtained. This shows that the simulation results are consistent for curves inside and outside the experimental data region. For verification purposes, the same previous procedure (of extending curve "A") was repeated for the crater depths and results similar to curve "A" was obtained.

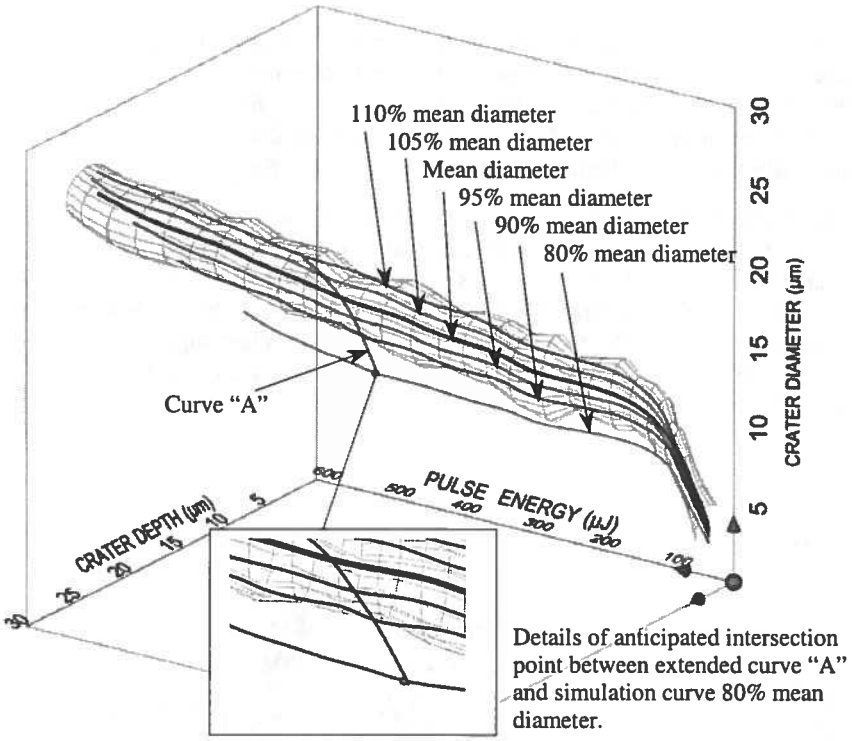

Fig. 19: All simulation curves are inside the mesh except 80% mean-diameter curve. Curve "A" corresponds to pulses having depth =19.84 μm.

6. CONCLUSIONS

In this project, a neural network model was developed to predict the pulse energy needed to create a crater with specific geometry (depth and diameter), as well as the variations associated with the depths and diameters. Experiments were carried out to provide the neural network with the data necessary for the training phase during which the network learns the process behavior. The neural networks modelers proved to be successful in simulating the pulsed laser material removal process to a high degree of accuracy. Since the process environment is likely to change and the components of laser micro-machining equipment might deteriorate with age and use, the adaptive properties of the network were found to be favorable for simulating such processes. Moreover, the same neural network model could be easily used numerous times to perform input/output function mapping without the need for any additional time for model amendment. The successful simulation results set the stage for valuable future work in the field and further process performance improvement.

REFERENCES

1. Bordatchev E.V., and Nikumb S.K., "Laser Material-Removal as a Subject of Automatic Control," *Proceedings of the ASPE 14th Annual Meeting*, Monterey, California, USA, pp.236-239, 1999.
2. Haykin S., *Neural Networks- A Comprehensive Foundation*, Ch.1-4, Prentice-Hall Inc., Upper Saddle River, New Jersey, USA, 1999.