LEAST SQUARE SUPPORT VECTOR MACHINE ALTERNATIVE TO ARTIFICIAL NEURAL NETWORK FOR PREDICTION OF SURFACE ROUGHNESS AND POROSITY OF PLASMA SPRAYED COPPER SUBSTRATES

Ajit Behera, Indian Institute of Technology - Kharagpur
LEAST SQUARE SUPPORT VECTOR MACHINE ALTERNATIVE TO ARTIFICIAL NEURAL NETWORK FOR PREDICTION OF SURFACE ROUGHNESS AND POROSITY OF PLASMA SPRAYED COPPER SUBSTRATES

1*Ajit Behera, 2Sudeep Behera, 3Rajesh Kumar Tripathy, 4Subash Chandra Mishra

1,4Department of Metallurgical and Materials Engineering Indian Institute of Technology, Kharagpur-721302, India
2Department of Electrical Engineering
3Department of Biomedical Engineering *Corresponding author: ajit.behera88@gmail.com

ABSTRACT
Plasma spraying technique has become a subject of intense research in many industrial structural/functional applications because its peculiarity surface properties. This investigation explains about plasma sprayed copper surface property. Here industrial waste and low grade ore (i.e. Fly-ash+ quartz+ illmenite), used as deposit material which is to be coated on copper substrates. In many applications, it is found that for structural modification, surface roughness & porosity parameters are very important. To decrease both surface roughness and coating porosity by optimizing other necessary properties, different soft computing methods like Artificial Neural Network (ANN) and Least Square support vector machine techniques used. The least square formulation of support vector machine (SVM) was recently proposed and rooted in the statistical learning theory. This technique potentially describes the approximation complexity of inter-relations between different parameters of atmospheric plasma spray process and helps in saving time & resources for experimental trials for which it is advantageous than all conventional methods. It is marked as a new development by learning from examples based on polynomial function, neural networks, radial basis function (RBF), splines or other function. From this above two methods (Multilayer Feed forward Neural Network & LS-SVM), we conclude that LS-SVM with RBF kernel gives better performance over ANN for prediction of surface roughness and coating porosity with minimum Mean Square Error. This methodology can provide clear understanding of various co-relationships across multiple scales of length and time which could be essential for improvement of product performance and process. The results of this methodology give a good generalization capability to optimize the coating surface roughness & surface porosity.

INTRODUCTION
Plasma sprayed technology have received much attention in recent years due to their outstanding surface properties when compared to those of their conventional processes. Plasma spray coating has the advantage of being able to process industrial waste and low-grade minerals to produce value-added products (Satapathy et al., 2010), which are widely used in aerospace industry to biomedical industry (Hetmanczyk et al., 2007; Muggli et al., 1999; Behera, 2012a; Ohtsu, 2012) Plasma spray processes utilize the energy contained in a thermally ionized gas to melt prepared powder particle partially/fully and propel this molten/semi-molten powder on to a substrate surface such that they adhere and agglomerate to produce coatings. This technique implements a wide variety of materials (metal, ceramic, alloy and its composite) and processes (Atmospheric plasma spraying, vacuum plasma spraying etc.) for improving surface properties (Singla et al., 2011; Davis, 2004; Sampath et al., 2001). Conventional plasma spraying refers to air or atmospheric plasma spraying (APS) process. Plasma generated by the help of an inert gas i.e. argon or argon+ hydrogen mixture (Zhang et al., 2011). Approximately 6000°C to 15000°C temperature can generated in the power heating region, which is significantly above the melting point of any known material (Heimann, 2008). Homogenized powder mixture of fly-ash+ quartz+ illmenite having particle size from 40 μm to 100 μm injected into plasma flame and then accelerated with a very high velocity to impact on substrate surface in the form of molten/semi-molten state (Fauchais, 2004; Behera, 2012b). The coating efficiency directly or indirectly depends on many other parameters during spraying, in which each one is inter-related with each other. Porosity depends on several coating properties such as thermal conductivity, coefficient of thermal expansion, elastic modulus and dielectric behavior (Deshpande et al., 2004). Also surface roughness depends on particle state (molten/semi-molten), powder feeding rate etc. Various methods are employed for quantitative measurement of surface roughness & porosity and form a necessary part of micro structural characterization of thermal spray coatings. A variety of prediction models have been proposed in the manuscript that include time series models, regression models, adaptive neuro-fuzzy inference
systems (ANFIS), artificial neural network (ANN) models and SVM models. Among these model ANN increasingly used in different optimization process. This is mainly because the effectiveness and smoothness of ANN modeling systems, which improves a great deal in the engineering area. For classification and non-linear function estimation, the SVM is introduced by Vapnik (Vapnik, 1995; Vapnik, 1999). SVM have remarkable generalization performance and many more advantages over other methods, so SVM has attracted attention and gained extensive application. Suykens and Vandevenne have proposed the use of the LS-SVM for simplification of traditional of SVM (Suykens et al., 2002). Not only LS-SVM has been used for classification in various areas of pattern recognition (Hanbay, 2009) but also it has handled regression problems successfully (Zuriani et al., 2011). LS-SVM has additional advantage as compared to SVM. In LS-SVM, a set of only linear equation (Linear programming) is solved which is much easier and computationally more simple. In this paper, by MFNN and LS-SVM models we are Predicting Surface Roughness and Porosity of Plasma Sprayed Copper. After comparing the result obtained by both these models, LS-SVM gives better performance to minimize the MSE that of MFNN.

**Experimental Procedure**

Flay-ash premixed with quartz and illuminate mixture taken (weight percentage ratio of 60:20:20) and homogenized by Planetary ball mill for 3 hour. This homogenized mixture used as coating material which is to be coated on Mild Steel substrate of dimension 3 mm thickness and 1 inch diameter. Four different sizes i.e. 40 μm, 60 μm, 80 μm and 100 μm are separated out by the application of sieve. Alumina grit blasting was made at a pressure of 3 kg/cm² to create surface roughness of about 5Ra, for better bonding. By acetone substrates surface were cleaned which is followed by plasma spraying. The coating process made at the Laser & Plasma Technology Division, BARC, Mumbai. Here 40 kW DC non-transferred arc mode atmospheric plasma spray system used. Input power level varies from 11kW to 21 kW in spray gun. Coating powder material injects externally from the torch nozzle and directed towards the plasma flow.

Ar/ H2/N2 gas may be used as carrier gas. The major subsystems of the set up include the plasma spraying torch, power supply, powder feeder, and carrier gas supply, torch to substrate surface distance, control console, cooling water and spray booth. A four stage closed loop centrifugal pump water cooling used for retrieving the heat generated during the process, regulated at a pressure of 10kg/cm² supply. The specifications of spraying process parameters are given in table 1. Flow rate of plasma gas (Ar) and Secondary gas (N2) are kept constant. With increasing power level; powder feed rate, powder size and stand of distance of torch are varied. Measurement of coating surface roughness done by ‘Taylor/Hobson Surtronic 3+’ instrument and experimental value of porosity was measured by image analyzer technique. In image analyzer technique the porosity of coatings was measured by putting polished cross sections of the sample under a microscope (Neomate) equipped with a CCD camera (JVC, TK 870E). This system is used to obtain a digitized image of the object. The digitized image is transmitted to VOIS image analysis software. The total area captured by the objective of the microscope or a fraction can be accurately measured by the software. Hence the total area and the area covered by the pores are separately measured. To find out some of the better surface property, there is a software programs used by implementation of MATLAB version 10.1. Back Propagation Algorithm (BPA) is used to train the network. The sigmoid function represented by equation (3.1) is used as the activation function for all the neurons except for those in the input layer.

\[ S(x) = \frac{1}{1 + e^{-x}} \]  

(3.1)

**Choice of Hidden Neurons**

The choice of optimal number of hidden neurons, \( N_h \) is the most interesting and challenging aspect in designing the MFNN. Simon Haykin (Simon, 1999) has specified that \( N_h \) should lie between 2 and \( \infty \) (Hecht-Nielsen, 1990). Hecht-Nielsen uses ANN interpretation of Kolmogorov’s theorem to arrive at the upper bound on the \( N_h \) for a single hidden layer network as 2(\( N_i + 1 \)), where \( N_i \) is the number of input neurons. However, this value should be decided very judiciously depending on the requirement of a problem. A large value of \( N_h \) may reduce the training error associated with the MFNN, but at the cost of increasing the computational complexity and time. For example, if one gets a tolerably low value of training error with certain value of \( N_h \), there is no point in further increasing the value of \( N_h \) to enhance the performance of the MFNN. The input and the output data are normalized before being processed in the network as follows:

In this scheme of normalization, the maximum values of the input and output vector components:

\[ n_{i,\text{max}} = \max\left( n_i(p) \right) \]

(3.2)

Where \( N_p \) is the number of patterns in the training set and \( N_i \) is the number of neurons in the input layer. Again,

\[ O_{k,\text{max}} = \max\left( o_k(p) \right) \]

(3.3)

Where, \( N_k \) is the number of neurons in the output layer.

Normalized by these maximum values, the input and output variables are given as follows.

\[ n_{i,\text{nor}}(p) = \frac{n_i(p)}{n_{i,\text{max}}} \]

(3.4)

and

\[ o_{k,\text{nor}}(p) = \frac{o_k(p)}{o_{k,\text{max}}} \]

(3.5)

After normalization, the input and output variable lay [13] in the range of 0 to 1.

**Choice of ANN Parameters**

The learning rate, \( \eta \) and the momentum factor, \( \alpha \) have a very significant effect on the learning speed of the BPA. The BPA provides an approximation to the trajectory in the weight space computed by the method of steepest descent method (Ghosh et al., 1999). If the value of \( \eta \) is considered very small, this results in slow rate of learning, while if the value of \( \eta \) is too large in order to speed up the rate of learning, the MFNN may become unstable (oscillatory). A simple method of increasing the rate of learning without making the MFNN unstable is by...
adding the momentum factor $\alpha$ (Rumelhart et al., 1986). The values of $\eta$ and $\alpha$ should lie between 0 and 1 (Simon, 1999).

### Weight Update Equations

The weights between the hidden layer and the output layer are updated based on the equation as follows:

$$
wb(j, k, m+1) = wb(j, k, m) + \eta \delta_k(m) \times Sb(j) + \alpha \times (wb(j, k, m) - wb(j, k, m-1))
$$

Where $m$ is the number of iterations, $j$ varies from 1 to $N_h$ and $k$ varies from 1 to $N_o$. $\delta_k(m)$ is the error for the $k$th output at the $m$th iteration. $Sb(j)$ is the output from the hidden layer (Mohanty et al., 2010). Similarly, the weights between the hidden layer and the input layer are updated as follows:

$$
wa(i, j, m+1) = wa(i, j, m) + \eta \delta_i(m) \times Sa(j) + \alpha \times (wa(i, j, m) - wa(i, j, m-1))
$$

Where $i$ varies from 1 to $N_i$ as there are $N_i$ inputs to the network, $\delta_i(m)$ is the error for the $i$th output after the $m$th iteration and $Sa(i)$ is the output from the first layer. The $\delta_i(m)$ in equation (3.6) and $\delta_k(m)$ in equation (3.7) are related as

$$
\delta_j(m) = \sum_{k=1}^{K} \delta_k(m) \times w_b(j, k, m)
$$

### Evaluation Criteria

The Mean Square Error Etr for the training patterns after the $m$th iteration is defined as

$$
E_{tr} = \left( \sum_{p=1}^{P} (V_{b1p} - V_{b2p}(m))^2 \right) \times \left( \frac{1}{P} \right)
$$

Where $V_{b1p}$ is the experimental value of (Surface roughness in first case and Porosity of Plasma Sprayed Copper in second case), $P$ is the number of training patterns and $V_{b2p}(m)$ is the estimated value of (Surface roughness in first case and Porosity of Plasma Sprayed Copper in second case) after $m$th iteration. The training is stopped when the least value of Etr has been obtained and this value does not change much with the number of iterations.

### Least Square Support Vector Machine

The formulation of LS-SVM is introduced as follows. Consider a given training set $\{x_i, y_i\}, i = 1, 2, ..., N$, with input data $x_i \in R$ and output data $y_i = R$. The following regression model can be constructed by using non-linear mapping function $\phi(x)$ (Zegnini et al., 2011).

$$
y = w^T \phi(x) + b
$$

Where $w$ is weight vector and $b$ is the bias term. As in SVM, it is necessary to minimize a cost function $C$ containing a penalized regression error, as follows:

$$
\min C(w, e) = \frac{1}{2} w^T w + \frac{1}{2} \gamma \sum_{i=1}^{N} e_i^2
$$

Subject to equality constraints

The first passssrt of this cost function is a weight decay which is used to regularize weight sizes and penalize large weights. Due to this regularization, the weights converge to similar value. Large weights deteriorate the generalization ability of the LS-SVM because they can cause excessive variance. The second part of eq. (4.2) is the regression error for all training data. The parameter $\gamma$, which has to be optimized by the user, gives the relative weight of this part as compared to the first part. The restriction supplied by eq. (4.3) gives the definition of the regression error. To solve this optimization problem, Lagrange function is constructed as

$$
L(w, b, e, \alpha) = \frac{1}{2} w^T w + \gamma \sum_{i=1}^{N} e_i^2 - \sum_{i=1}^{N} \alpha_i \left( w^T \phi(x_i) + b + e_i - y_i \right)
$$

Where $\alpha_i$ are the Lagrange multipliers. The solution of eq. (4.4) can be obtained by partially differentiating with respect to $w$, $b$, $e_i$ and $\alpha_i$. Then

$$
w = \sum_{i=1}^{N} \alpha_i \phi(x_i) = \sum_{i=1}^{N} \gamma e_i \phi(x_i)
$$

Where a positive definite Kernel is used as follows:

$$
K(x_i, x_j) = \phi(x_i)^T \phi(x_j)
$$

An important result of this approach is that the weights ($w$) can be written as linear combination of the Lagrange multipliers with the corresponding data training ($x_i$). Putting the result of eq. (4.5) into eq. (4.1), the following result is obtained as

$$
y = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x) + b
$$

For a point $y_i$ to evaluate it is:

$$
y_i = \sum_{i=1}^{N} \alpha_i \phi(x_i)^T \phi(x_j) + b
$$

The vector follows from solving a set of linear equations:

$$
A \begin{bmatrix} \alpha \\ b \end{bmatrix} = \begin{bmatrix} y \\ 0 \end{bmatrix}
$$

Where $A$ is a square matrix given by

$$
A = \begin{bmatrix} K + I \\ \gamma N \end{bmatrix}
$$

Where $K$ denotes the kernel matrix with $i^j$th element in eq. (4.5) and $I$ denotes the identity matrix $N \times N$, $1N = [1 \ 1 \ 1 \ ... \ 1]^T$. Hence the solution is given by:
\[
\begin{bmatrix}
\alpha \\
b
\end{bmatrix} = A^{-1} \begin{bmatrix} y \\
0
\end{bmatrix}
\] (4.11)

All Lagrange multipliers (the support vectors) are non-zero, which means that all training objects contribute to the solution and it can be seen from eq. (4.10) to eq. (4.11). In contrast with standard SVM the LS-SVM solution is usually not sparse. However, by pruning and reduction techniques a sparse solution can easily achieved. Depending on the number of training data set, either an iterative solver such as conjugate gradients methods (for large data set) can be used or direct solvers, in both cases with numerically reliable methods.

In application involving nonlinear regression it is not enough to change the inner product of \( \phi(x_i), \phi(x_j) \) eq. (4.7) by a kernel function and the \( j^{th} \) element of matrix \( K \) equals to eq. (4.5).

This show to the following nonlinear regression function:

\[
y = \sum_{i=1}^{N} \alpha_i K(x_i, x) + b
\] (4.12)

For a point \( x_j \) to be evaluated, it is:

\[
y_j = \sum_{i=1}^{N} \alpha_i K(x_i, x_j) + b
\] (4.13)

For LS-SVM, there are lot of kernel function (Radial basis function, Linear polynomial), sigmoid, bspline, spline, etc. However, the kernel function more used is a simple Gaussian function, polynomial function and RBF. They are defined by:

\[
K(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{\sigma_{sv}^2} \right)
\] (4.14)

\[
K(x_i, x_j) = (x_i^T x_j + 1)^d
\] (4.15)

Where \( d \) is the polynomial degree and \( \sigma_{sv}^2 \) is the squared variance of the Gaussian function, to obtain support vector it should be optimized by user. For \( \alpha \) of the RBF kernel and \( d \) of the polynomial kernel, in order to achieve a good generalization model it is very important to do a careful model selection of the tuning parameters, in combination with the regularization constant \( \gamma \).

**Evaluation Criterion**

The Mean Square Error \( E_m \) for the training patterns after the \( m^{th} \) iteration is defined as

\[
E_m = \left( \sum_{p=1}^{P} \left( V_{p1p} - V_{k2p} (m) \right)^2 \right) \times \left( \frac{1}{P} \right)
\] (4.16)

Where \( V_{p1p} \) is the experimental value of (Surface roughness in first case and Porosity of Plasma Sprayed Copper in second case), \( P \) is the number of training patterns and \( V_{k2p} (m) \) is the estimated value of (Surface roughness in first case and Porosity of Plasma Sprayed Copper in second case) after \( m^{th} \) iteration. The training is stopped when the least value of \( E_m \) has been obtained and this value does not change much with the number of iterations.

**RESULT AND DISCUSSION**

With the help of 56 sets of experimental input-output patterns, the proposed modeling are carried out; 44 sets of input-output patterns used for training both networks and for testing purpose the remaining 12 sets are used. The software programs developed are used for implementation using MATLAB version 10.1.

**Table 1. Experimental plasma sprayed operating parameters.**

<table>
<thead>
<tr>
<th>Operating parameters</th>
<th>Parametric variations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma arc current (Amp)</td>
<td>270, 300, 400 &amp; 420</td>
</tr>
<tr>
<td>Arc voltage (volt)</td>
<td>40,45 &amp; 50</td>
</tr>
<tr>
<td>Torch input power (kW)</td>
<td>11,15,18 &amp; 21</td>
</tr>
<tr>
<td>Plasma gas (argon) flow rate (IPM)</td>
<td>28</td>
</tr>
<tr>
<td>Secondary gas (N2) flow rate (IPM)</td>
<td>3</td>
</tr>
<tr>
<td>Carrier gas (Ar) flow rate (IPM)</td>
<td>12</td>
</tr>
<tr>
<td>Powder feed rate (gm/min)</td>
<td>12, 15 &amp; 18</td>
</tr>
<tr>
<td>Powder Size (µm)</td>
<td>40, 60, 80 &amp; 100</td>
</tr>
<tr>
<td>Torch to base distance (nm)</td>
<td>100</td>
</tr>
</tbody>
</table>

**Surface Morphology of plasma spray coating**

Figure 2 shows SEM image of coating surface. Here the fly-ash+ quartz+ ilmenite composite powder is deposited on copper at 21 kW, 12 gm/min feed rate, 100nm to base distance with varying the power level. It is observed that surface contains some amount of pore with little surface roughness. In spraying the exact power level cannot be specified, because thermally sprayed coatings are very complex and incorporate process dependent defects such as splat gaps/interlamellar pores, globular pores, cracks (for ceramics), etc. It is clear that there is a close agreement of porosity measurement by the MFNN, LS-SVM and the experimental study, which indicates that the both the models can be used for predicting the amount coating porosity. Porosity formation by inter-splat position also depends on the mode of heat transfer through the substrate or prior deposited splats. This also similar to splat quenching by the Duwez gun technique. Inadequate amount of heat flow which results molten/semi-molten particles gather in a way to form more pore.

**Table 2. Variation of MSE (E_m) with N_h (\eta_f=0.3, \alpha_f=0.6 and iteration = 400)**

<table>
<thead>
<tr>
<th>iteration</th>
<th>E_m</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.2090x10^7</td>
</tr>
<tr>
<td>2</td>
<td>6.0379x10^6</td>
</tr>
<tr>
<td>3</td>
<td>1.9029x10^6</td>
</tr>
<tr>
<td>4</td>
<td>9.9219x10^7</td>
</tr>
<tr>
<td>5</td>
<td>8.5546x10^7</td>
</tr>
<tr>
<td>6</td>
<td>5.8662x10^7</td>
</tr>
<tr>
<td>7</td>
<td>3.3196x10^7</td>
</tr>
<tr>
<td>8</td>
<td>1.9942x10^7</td>
</tr>
<tr>
<td>9</td>
<td>1.6227x10^7</td>
</tr>
<tr>
<td>10</td>
<td>1.4296x10^7</td>
</tr>
</tbody>
</table>

**MFNN Modeling**

From Tables 2 and Figure 3, it is quite obvious that when Nh = 10, \( \alpha_f = 0.6 \) and \( \eta_f = 0.3 \), the MSE for training data is the lowest at 1.4296 \times 10^{-7}. The sequential mode of training has been adopted. It may be noted that the range of the values of \( \eta_f \) and \( \alpha_f \) should be between 0 and 1 and value of Nh should not more than 10 as per the Hecht-Nielsen criteria. Hence we have stopped at \( N_h = 10 \) in Table1, we have stopped at \( \eta_f = 0.3 \) and \( \alpha_f = 0.6 \).
Fig. 1. Multilayer feed forward neural network

Fig. 2. Surface morphology of plasma coated copper specimen (Spraying at 21KW power level).

Fig. 3. Variation Mean Square Error \( (E_m) \) with number of iterations \( (\eta_1=0.1, \alpha = 0.6 \text{ and } N_\theta=10) \)

LS-SVM Modeling

From Figure 4, it is quite obvious that when \( \sigma =0.31 \), \( \text{Sig}^2=5780.87 \) and iteration = 400, the data is well separated from LS-SVM hyper plane and mean square error (MSE) for training data is the lowest and equals to \( 7.15 \times 10^{-7} \). It may be noted that if the value of iteration is increased MSE will decrease. Finally the test data are calculated by simply passing the input data in the LS-SVM and using \( \alpha \) and \( b \) for the kernel parameter.

RESULT OBTAINED

Table 3 depicts the comparison between the MSE for the training data and testing data obtained from two different models using different techniques. As may be seen from Table-3 that LS-SVM model is better than MFNN model. This can be confirmed from the values of MSE.

Fig. 4. Shows the change of output parameter with respect to different input for training sample.
AC- Arc current, V-Voltage, PFR- Powder Flow Rate, PD- Powder size, JSA- Jet Spray Angle, CT-Coating Thickness, ESR- Experimental Surface roughness, EPSSR- Experimental Plasma Sprayed Copper, PSRN- Predicted Surface roughness using neural network, PSRSRn- Predicted Plasma Sprayed Copper using neural network, PSRS- Predicted Surface roughness using LSSVM, PSRSSR- Predicted Plasma Sprayed Copper using LSSVM.

**Conclusions**

In this study, a LS-SVM with RBF Kernel and feed forward neural network (MFNN) using Back Propagation algorithm was designed, trained and tested for prediction of surface roughness and porosity of plasma spray coating. The selected LS-SVM model successfully exhibited a good prediction capability with a low MSE for training and testing data than that of MFNN network. This can lead to reduce the time and cost of the process and improve the quality of the product.

**REFERENCES**


