Plasma Spray Process Control with Neural Network

A. Jakimov, M. Hertter, H. Abdullahi,
MTU Aero Engines GmbH, Munich, Germany
S. Staudacher
Institut für Luftfahrtantriebe, Universität Stuttgart, Germany

In current aircraft engines challenging coatings or coating systems with different functions are used. Many of these coatings are applied to different components with thermal spray processes. Thermal spraying is a very sensitive and very complex process, which is influenced by numerous *controllable* variables like the powder feed rate, the gas flow rates, etc. as well as *not controllable* variables like the torch wear, varying powder properties, etc. With conventional process control based on linear algorithms it is not possible to enduringly create constant coating properties, because they cannot describe the complexity of all influencing variables. In this work, the possibility of a closed loop was investigated exemplarily for an **a**tmospheric **p**lasma **s**pray process (**APS**). During serial production a data base was collected, consisting information about torch and plume conditions as well as powder and coating properties. This data base was used to train different **n**eural **n**etworks (**NN**). With regard to the automation of the APS, the NN obtained target values of relevant coating properties and should calculate the needed control variables. The result of this work shows the difficulties in the quantification of relevant influencing variables and the feasibility of the plasma spray process control with neural network.

1 Introduction

1.1 Abradable seals in aircraft engines

To enhance the performance and the efficiency of aircraft engines, current compressor development aims at increasing the pressure ratio. Requirements like a very light and compact construction, which are achieved by reducing the number of compressor stages, additionally result in an increase of the pressure ratio between the stages. One side effect hereof is an increasing backflow from the pressure to the suction side of the compressor blades. Consequently, the sealing system reducing backflow between the rotating blades and the casing, becomes more and more important. This seal system is an important factor for the efficiency as well as for the so-called surge line and hence for the stable operation of the compressor. The surge line of a compressor is reached, when the axial flow velocity decreases too much because of an excessively high backflow rate. This causes an uncontrollable blade stall that results in a complete collapse of the compressor flow. If the sealing system is inadequate the surge line is reached much earlier. This is one of the main reasons for premature maintenance. However, the optimization of the sealing system increases the cost efficiency of the engine by reducing fuel consumption and extending maintenance intervals [1,2].

To avoid a high backflow rate it is necessary to minimize the gap between blades and casing. Because of the different strain on rotor and casing under the diverse operating conditions (acceleration, deceleration) the blade tips can touch or even run into the casing wall. Rubbing can also be caused by eccentricities of the rotor or the casing, which may be caused by flight maneuvers, especially of military aircraft. Because such an incident would lead to damage with devastating consequences for safe aircraft operation, the potential contact areas at the casing are provided with abradable coatings. The hardness of these so-called abradable seals is designed to permit rubbing of the blades without damaging them.

1.2 Required properties of abradable seals

The requirements imposed on abradable seals are diverse and often cannot be met concurrently:

- Rub-in behavior
- Small swarf size
- Erosion resistance
- Temperature, thermal cycling und fire resistance
- Thermal insulation
- Pressure- and pressure cycling resistance
- Low maintenance cost
- Good aerodynamic properties

In practice, a combination of coatings is used because one single coating could not satisfy all these partially conflicting requirements. These coatings can be optimized adequately for their tasks and their specific application (high-, intermediate- or low-pressure). These coating systems are composed, for example, of a metallic bond, a ceramic thermal barrier and titanium fire-protection and the intrinsic porous abradable coating.

When abradable seals were introduced in the 1970ies, they were made e.g. of NiC coatings. Impelled by increasing demands on aircraft engines the composition of abradable seals was getting more complex. Nowadays, so-called MCrAl(Y)-Coatings (M = Fe, Ni, Co, CoNi) with very high oxidation and corrosion resistance are used more and more. These are applied using the atmospheric plasma spray process. Because of the high complexity the diverse production lots of the powder vary marginally in their characteristics.

Although these variations are not significant and the process parameters are nearly constant, the diverse powder lots cause considerable variations of the coating properties.

1.3 Process diagnostic system PFI

Process diagnostic systems for APS measure plasma properties, e.g. plasma geometry, electro-magnetic emission intensity, or particle properties, e.g. particle temperature, velocity, size or trajectory. In principle, they are used to determine correlations between process parameters, plasma or particle properties and coating properties. In view of the multiplicity of the diagnostic systems, in this work the **P**article **F**lux **I**maging (**PFI**) [3] was used.
The PFI system analyses the optical plasma geometry. It uses a CCD camera with a resolution of 764 x 256 pixels. The camera records monochrome pictures of the plasma jet from a position orthogonal to the torch axis (Fig. 1a). With two grey filters the pictures are divided in two parts with different luminosity. The dark filter (left section, Fig. 1b) covers the beginning of the plasma jet at the torch exit. With this filter the outer areas are blinded out so that only the high-intensity plasma plume can be seen. Whereas the brighter filter (right section Fig. 1b), which covers the rest of the plasma jet, additionally shows the outer areas, where particle emission can be seen.

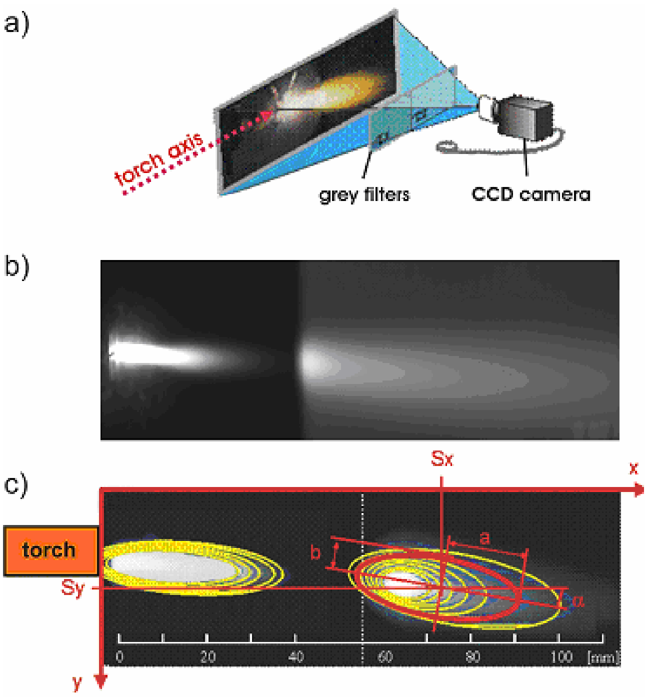

Fig. 1: Functional principle of the APS diagnostic system PFI. The CCD camera (a) takes pictures (b) of the plasma jet, that are then evaluated (c) by the PFI software.

In both sections, the picture of the plasma jet is subdivided in zones with different brightness values, which correspond to level curves with equal intensity. Using an algorithm the PFI system inscribes an ellipse in one brightness level. The ellipse characteristics, such as the long and the short axis, the coordinates of the centre point and the angle of the ellipse relative to the torch axis (Fig. 1c) are the PFI values to be used for the process diagnostics. Because of the size of the PFI camera it was not possible to mount it on the robot of the plasma torch itself. Therefore the PFI values could only be taken before and after the coating process.

2 Artificial neural networks

2.1 Functional principle of neural networks

An artificial **n**eural **n**etwork (**NN**) is a software, that copies the functional principle of the human brain.
Via nerves the human brain receives electrochemical signals, e.g. from the sense organs. Afterwards the signals are processed to send an electrochemical reaction signal to muscles, adenoids etc., if necessary. The processing of the signals occurs in the cortex. The cortex of the human brain consists of about 15 to 100 billions of processing units [4], called neurons. Every neuron is connected to 10.000 more neurons on an average, so that the collectivity builds up a network [5]. Input signals reach a neuron via the dendrites. The single output signal leaves the neuron via the axon, which can split into about 1000 thin branches at its end. The interface between one branch and the next dendrite is a synapse. The synapses biochemically amplify or inhibit the potential weights of the incoming signal before sending it to the neuron. In the neuron the incoming signals from all dendrites are summed up. Once a certain threshold value, also called bias, is reached the neuron sends a signal - via its axon - to the next synapses (Fig. 2).

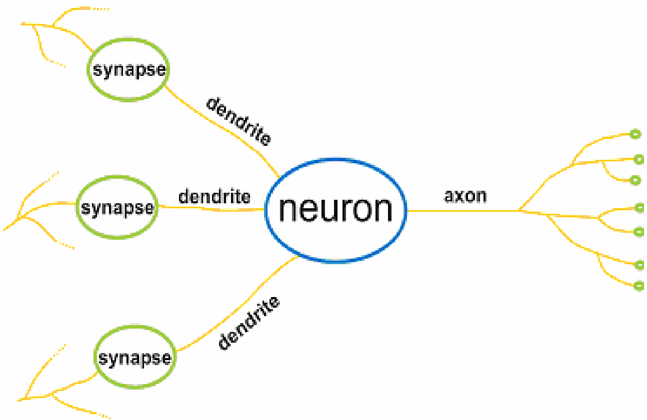

Fig. 2: General structure of neural networks. From the synapses, modified input potentials reach the neuron via dendrites, and the neuron processes the sum of the signals and sends an output signal through the axon to the adjacent synapses and neurons.

Frequently used neuron connections are trained and strengthened, whereas connections used less can even decay, which corresponds to oblivion. This means that the weights modifying effect can change in the course of time. It can be noticed, that the human brain is learning and processing information by the use of its network topology and particularly the amplifying and inhibiting effect of the synapses [5]. With a nominal magnitude of about one trillion (10^{12}) synapses, the enormous learning capability of the human brain becomes apparent.

As mentioned above, technology emulate the archetype of nature. Because the neural network of the human brain is to be imitated by software, it is necessary, first of all, to convert the biologic functional principle into a mathematical description.

2.2 Mathematical illustration of neural networks

Every input value a_j^{l-1} will be sent to every neuron in a certain layer (Fig. 3). Before it arrives, it will be amplified or inhibited in a synapse. This is mathematically implemented by multiplication of the value with a positive (amplify) or a negative (inhibit) number $w_{i,j}^l$, the so-called weight. The values thus amplified or inhibited are summed in the neuron with a bias b_i^1. Afterwards, this sum is converted with a transfer function (e.g. sigmoid, linear etc.) into the output value a_j^l of the neuron. The following equation is the proper mathematical illustration:

$$a^l{}_i = f\left(\sum_{j=1}^{n^{l-1}} w^l{}_{i,j} \cdot a_j{}^{l-1} + b^l{}_i\right)$$

All output values of one neuron layer are also the input values of the next layer. Thus, a network of neurons is built up as illustrated in Fig. 3.

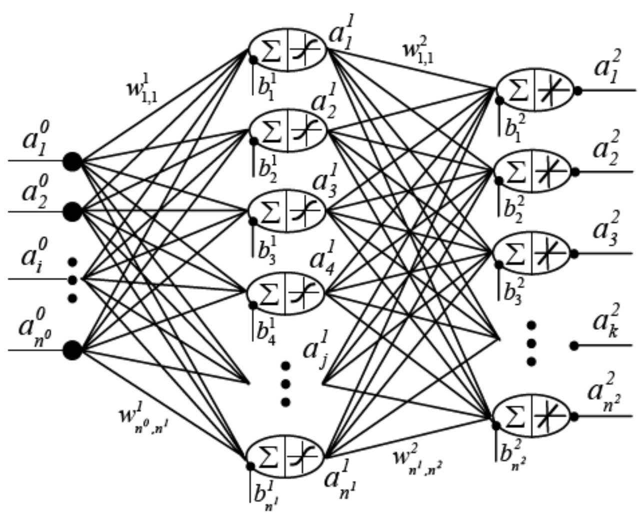

Fig. 3: Illustration of an artificial neural network with two neuron layers.

Now the question arises as to how an artificial neural network can learn to calculate the correct output from the given input values. And here, too, the biological principle is emulated, i.e. the network learns the relation between input and output values by adequate training with datasets. “Learning” in terms of mathematics means, that the neural network conveniently adjusts the weights and the bias of the neurons with a learning algorithm.

The general procedure for a network training is described below:

1 At the beginning random numbers are used for all weights.
2 One dataset consisting of input and output values is taken out of the pool of training datasets. The neural network is “fed” with the input values of this dataset.
3 The network calculates its own output values with the momentary weights.
4 The output values calculated by the network are compared with the output values of the original dataset. If the values are identical, the training continues with step 2. Otherwise, the weights are modified with a proper learning rule (e.g. Hebb’s Learning or Delta Rule [5]) before continuing with step 2.

Thus the neural network is „learning“ the mathematical correlation between input and output values in numerous iterations.

3 Variation studies of neural networks

In this work, about 700 datasets recorded in the course of two years of abradable coating production were used to train neural networks. Each of these datasets contained the controllable and the not controllable process variables as well as the coating properties generated. To ensure the comparability of the networks, all network variations were trained with the same number of epochs. In one single epoch the neural network reads all allocated datasets only once, i.e. in the present case one epoch consisted of 700 training iterations. The neural networks were trained to calculate two variables, that are necessary to control the process. Different network topologies (e.g. the amount of neurons in the layers) and input parameters (e.g. PFI, powder parameters etc.) were tested.

3.1 Variation of the neural network topology

In the study of the network topology, the number of the neurons in the first layer of a two-layer network (Fig. 3) was varied. The number of neurons in the last layer is determined by the number of output values. Fig. 4 shows the mean square error of neural networks versus the quotient between the number of neurons in the first layer and the number of input parameters. The mean square error is calculated from the differences between the actual values from the 700 datasets and the values calculated by the trained network.

As can be seen, the mean square error decreases asymptotically when the number of neurons increases.

If the number of neurons of the first layer is lower than 20% of the number input parameters, the mean square error is higher than 15%. If the number of neurons and the number of input parameters are the same, the mean square error is about 4%. With a further increase of the number of neurons, the improvement of the mean square error is much smaller.

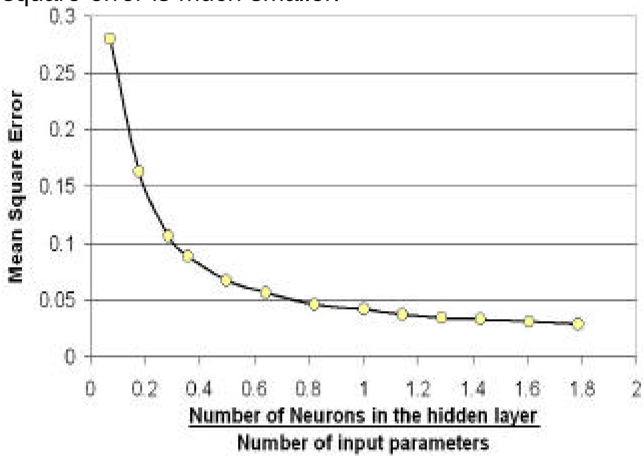

Fig. 4: Mean square error of various neural networks versus the quotient between the number of neurons in the first layer and the number of input parameters.

3.2 Variation of the input parameters

In a second study, neural networks with comparable topologies but different input configurations were generated and trained. Every network was trained with the 700 datasets recorded in the production of the coatings. The difference between the configurations was the input information. As input parameter quantified information about substrate, PFI parameters, process facilities abrasion, plasma torch and powder parameters were available.

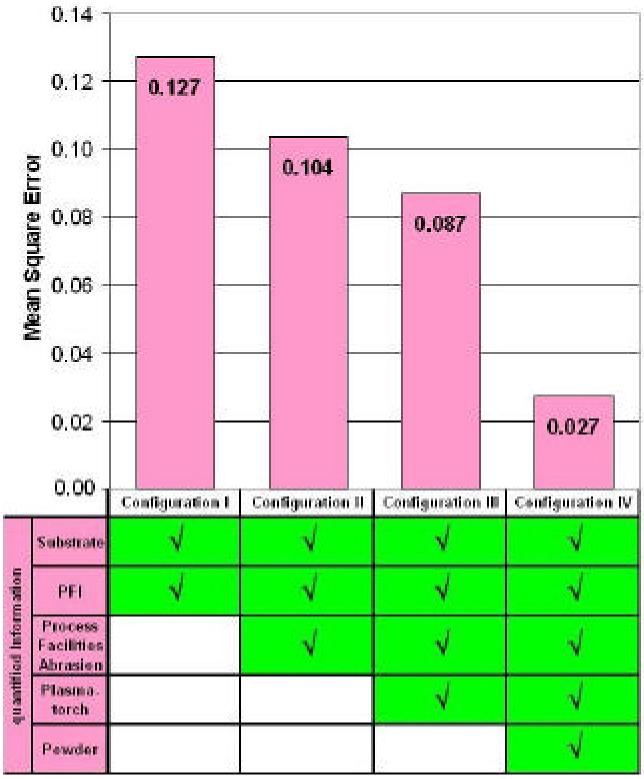

Fig. 5: Mean square error and input information of the configurations I to IV

With configuration I (Fig. 5) only the substrate and the PFI values from the datasets were used as input parameters. With every further configuration, the scope of information was increased, so that the last one (configuration IV) had all quantified parameters. In this study, too, the neural networks were trained to calculate the two important control variables of the process. Fig. 5 shows the mean square error and the input information used for each configuration. As can be seen the mean square error decreases with the increase in information. Especially, with the powder information taken into account the error decreased from 8.7% to 2.7%.

4 Training of a process controlling neural network prototype

After evaluation of the results of the topology and the input information, a neural network prototype was generated and trained to control the APS process. For the purpose, a neural network with the optimal topology and input configuration was trained with the 700 datasets from coating production and afterwards tested

using new datasets. These new datasets were not available in the training data pool. Thus, they were unknown to the trained network.
Although the training datasets in every training session were the same and training results, such as the mean square error did not vary significantly, part of the test results differed strongly.

a)

b)

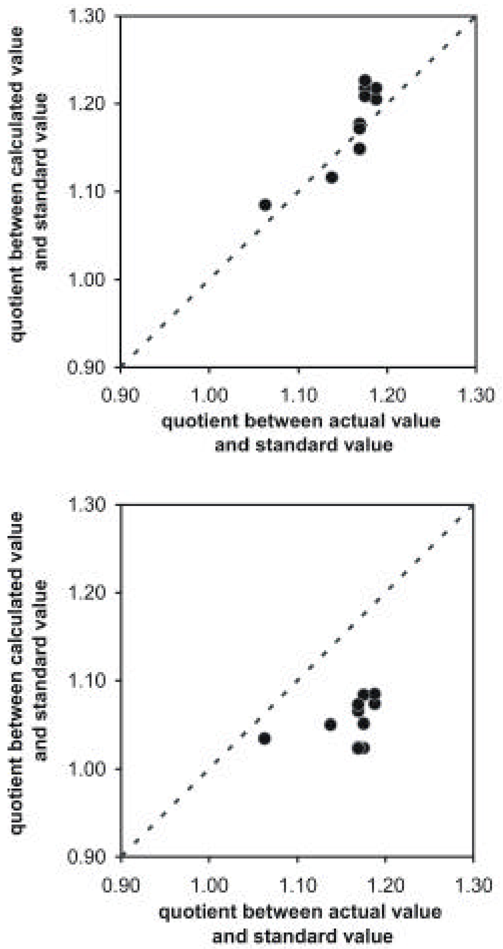

Fig. 6: Results of the tests with unknown datasets. Standardized output values calculated by the network versus standardized actual values from the datasets.

Fig. 6 shows the test results of two trainings sessions. The difference between the mean square error of the training session shown in 6a) and 6b) was 0.00046. The diagrams show the standardized output values which were calculated by the network versus the standardized actual variable values from the datasets. The ideal case would be that all points are on the diagonal starting from the origin (broken line).
The variations of the results are attributable to the different weights and bias determined in each renewed training session. The weights and bias after the sessions are not identical because their starting values are random values.
For further testing the prototype with the results shown in Fig. 6a) was used.

5 Testing of the neural network prototype in experiments

The neural network prototype determined in the variation studies (see 3.2) was then tested in special experiments. In these experiments, the output values were determined directly at the spray booth so that process samples could be sprayed with these parameters.
In these **p**rototype **e**xperiments **PE**1 - **PE**8 different powder lots were used for the coating. Five lots (lot A: PE1 and 8; lot D: PE 4; lot E: PE5; lot F: PE6; lot G: PE 7) were known to the prototype, i.e. the training pool for the prototype contained datasets with these powder lots. The training pool did not contain information on the lots B and C used in experiments PE2 and PE3, therefore, they were unknown to the prototype. In each experiment, the current process parameters were entered in the prototype, and the control variables were to be calculated by the prototype. In the coatings of experiments PE1 to PE7 a newly overhauled torch (torch 1) was used. PE 8 was processed with another torch (torch 2), that was already in use in production. The results of the prototype experiments are shown in Fig. 7. Additionally, the target value of the coating property and the permitted tolerance range are depicted.

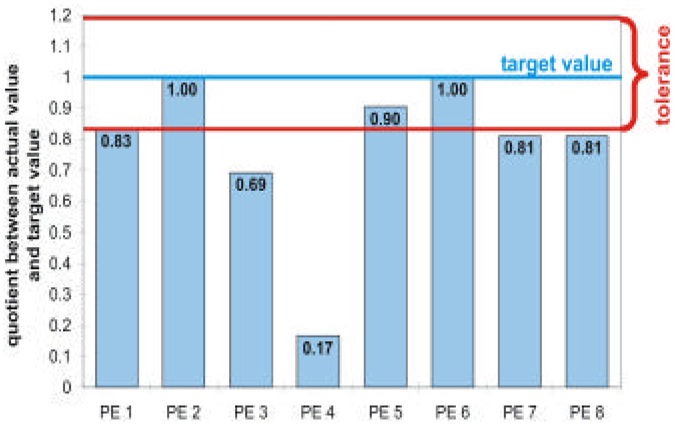

Fig. 7: Results of the neural network prototype experiments.

The coating properties of experiments PE 2 and PE 6 are in conformity with the target values. The value of PE 5 is within the range of the measurement inaccuracy. PE 1 reaches the lower limit of the tolerance range, while PE 7 and PE 8 remain just below it. The values of PE 3 and PE 4 are clearly out of tolerance.

6 Discussion of the results

6.1 Variation studies and training of neural networks

The results of the variation of the number of neurons in the first layer (Fig. 4) show an asymptotic improvement with increasing layer width. Increased network size

means an increase in calculating time. With today's computers, however, this remains within acceptable limits.
As can be seen from Fig. 6, the results of the tests with unknown datasets can differ completely after every training session. The reason for the partially great differences between the test results is the fact that the neural network determines several marginally different mathematical correlations between input and output values. This is also the explanation why the results of the training (with known datasets) vary only insignificantly – the difference between the mean square error of 6a) and 6b) was only 0.00046 – while the associated test results differ strongly (Fig. 6). Thus, an important goal is to make all process-affecting parameters available to the neural network to reduce the mathematical correlations to the ones that actually describe the physics.
The comparison of the training results of the different configurations (Fig. 5) shows that the results become better with increasing extent of input information. Particularly, the addition of the powder information led to a clearly reduced training square error. Therefore, it has to be emphasized that as much meaningful quantified process information as possible should be used in the network. Because the processor performance of the standard Pentium 4 used was sufficient, the increased neural network layer did not result in unacceptable calculating times. The neural networks that have a larger topology because of the increased input parameters, can only be trained meaningfully with large data pools, whose datasets contain all desired quantified information.
At the beginning of the input variation studies, “outliers” in the training results were noted. The values of these network-calculated parameters differed from the actual ones up to 23%. A check of the associated datasets, revealed writing errors for some value. After correction of these errors the network calculation did no longer produce any outliers in the training results.
Obviously, the neural network could not correlate random datasets, that did not consist of faultless values. Since the neural network could differentiate between actual and faulty datasets, it can be assumed that there is a mathematical correlation between the input and output values used, that is determined by the network. A mathematical correlation between the input and output values indicates a correlation between the physical values represented by the input and output.

6.2 Neural network prototype tests

The results of the Prototype experiments are to a large extent still not acceptable. Of eight experiments the neural network prototype could convince only in three cases. This result is attributable to the fact that torch 1 was used in seven of the eight prototype experiments. Torch 1 was completely overhauled directly before the experiments. The method of quantification of wear is obviously not yet adequate. More detailed experiments are required.
In prototype experiments 3 and 4 the largest deviations were observed (Fig. 7). The reason for the unacceptable result of PE 3 might be due to the fact that the powder lot (Lot C) used here differs very strongly from the others in a way not quantified yet and that this lot was missing in the training data pool. Powder lot D of PE 4, however, was represented with many datasets in the training pool. But this powder lot was the first used at the beginning of the production of this particular coating, i.e. the datasets of this powder lot where up to 18 months old. Outside influences not yet considered might have changed the process since then. Moreover, it has to be assumed that at the beginning of coating production a lot of errors was made in the operation sequence as well as in the data acquisition.
In the three well forecast experiments PE 2, PE 5 and PE 6 the powder lots B, E and F were used. Powder lot B was not represented, powder lot E was represented with 120 and F with only one dataset in the training pool. That means that the neural network could identify well known, hardly known and totally unknown powder lots. This permits the conclusion that at least here the method of quantification of the powder information was adequate.
The samples of experiments PE 1 and PE 8 were coated with powder lot A using two different torches (torch 1 and 2). Indeed, the results show that the calculation of the neural network was wrong in both cases, while the coating properties were in the same range. From the production it is well-known that with two torches different control variables are needed to achieve the same coating properties. The network could obviously recognize and convert the difference between the two torches with the quantified input information on the torch change. Apparently, the wrongly calculated control variables were the result of incorrect interpretation of the powder information. It can be concluded that the quantification of the powder information for the neural network has succeeded in part, but still needs to be improved.

7 Conclusion

Practice in the manufacturing shows that it is only possible with a relatively large error probability and under certain conditions to produce acceptable coating properties under the influence of changing process parameters. Above all, however, it is not possible yet to describe the plasma spraying process with generally valid conventional physical models since the number of process parameters is too large and the complexity of the process too high [6].
The results of this work show that it is possible to configure neural networks which can allow for the complexity of the plasma spraying process and control the process. Special attention must be paid here on the quality of the data records used for training. These are essential for the validity and the flexibility of a neural network. In principle, any number of input data can be fed into the network, in order to determine correlations

with desired output data. But the quantification of some input parameters often represents a problem, because the information would be very important for the network but cannot be meaningfully expressed in numbers.
A major advantage of a neural network is that it does not need additional information about possible correlations between the input and output data, to be processed by it. Rather, it determines these on its own and finds correlations, which are too complex for humans. For this reason, it can be concluded that neural networks have an enormous technological potential, not only regarding the plasma spraying process, but also regarding other fields of technology.

8 Acknowledgment

This paper is based on a research work sponsored by MTU Aero Engines GmbH. The authors would like to thank Stefan Schneiderbanger for his intensive support and all other colleagues at MTU, who were always available for discussions about this work.

9 Literature

[1] Schneiderbanger, S.; Cosack, T.: Thermisches Spritzen im Triebwerksbau, MTU Aero Engines GmbH, München

[2] Buckl, F.; Daeubler, M. A.; Bamberg, J; Schweitzer, K. K.; Hertter, M.; Eichner, J.: Fortschrittliche Dichtsysteme in Flugtriebwerken – Eine interdisziplinäre Herausforderung, MTU Aero Engines GmbH, München

[3] Zierhut, J.; Hartmann, R.; Landes, K. D.: Indirektes Plasmatron, Europäisches Patent 97810823.1-2208 (1997)

[4] Vester, F.: Denken, Lernen, Vergessen, Dtv, München, 2001

[5] Kinnebrock, W.: Neuronale Netze: Grundlagen, Anwendungen, Beispiele, Oldenbourg Verlag, München, 1994

[6] Lugscheider, E.; Seemann, K.: Prognose von Schichteigenschaften mittels neuronaler Netze, Tagungsband zum Werkstofftechnischen Kolloquium, Chemnitz, 2004, pp.50/59