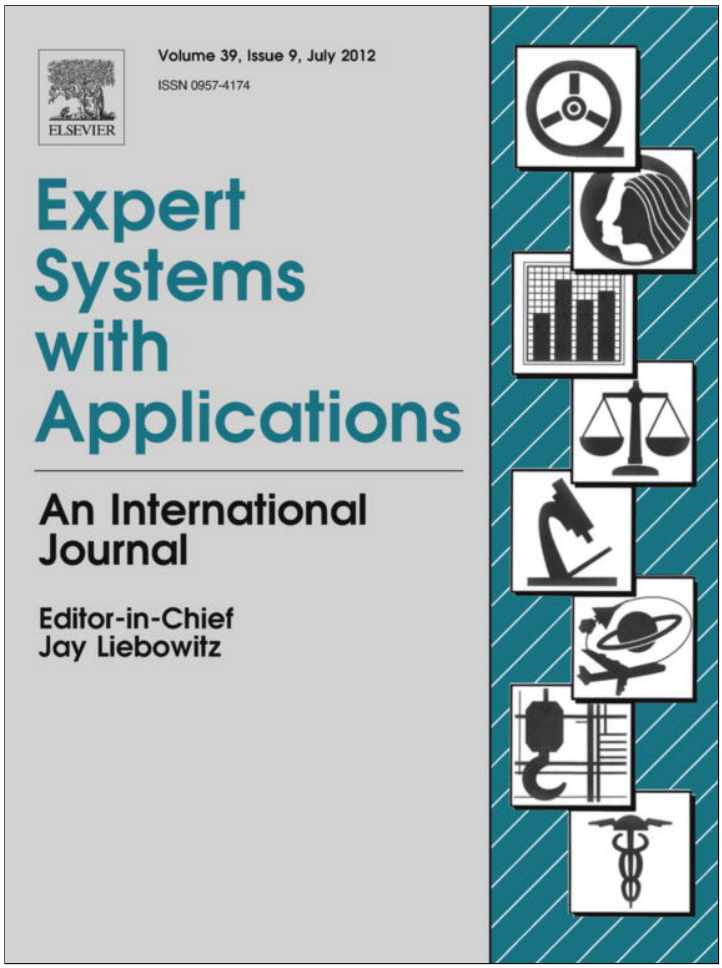

Expert Systems with Applications 39 (2012) 7776–7787

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Optimization of Radial Basis Function neural network employed for prediction of surface roughness in hard turning process using Taguchi's orthogonal arrays

Fabrício José Pontes [b], Anderson Paulo de Paiva [a], Pedro Paulo Balestrassi [a], João Roberto Ferreira [a,*], Messias Borges da Silva [b]

[a] Institute of Industrial Engineering, Federal University of Itajubá, 37500-903 Itajubá-MG, Brazil
[b] Faculty of Engineering of Guaratinguetá, Sao Paulo State University, 12516-410 Guaratinguetá-SP, Brazil

ARTICLE INFO

Keywords:
RBF neural networks
Taguchi methods
Hard turning
Surface roughness

ABSTRACT

This work presents a study on the applicability of radial base function (RBF) neural networks for prediction of Roughness Average (R_a) in the turning process of SAE 52100 hardened steel, with the use of Taguchi's orthogonal arrays as a tool to design parameters of the network. Experiments were conducted with training sets of different sizes to make possible to compare the performance of the best network obtained from each experiment. The following design factors were considered: (i) number of radial units, (ii) algorithm for selection of radial centers and (iii) algorithm for selection of the spread factor of the radial function. Artificial neural networks (ANN) models obtained proved capable to predict surface roughness in accurate, precise and affordable way. Results pointed significant factors for network design have significant influence on network performance for the task proposed. The work concludes that the design of experiments (DOE) methodology constitutes a better approach to the design of RBF networks for roughness prediction than the most common trial and error approach.

1. Introduction

Surface quality is an essential consumer requirement in machining processes because of its impact on product performance. The characteristics of machined surfaces have significant influence on the ability of the material to withstand stresses, temperature, friction and corrosion (Basheer, Dabade, Suhas, & Bhanuprasad, 2008). The need for the products with high quality surface finish keeps increasing rapidly because of new application in various fields like aerospace, automobile, die and mold manufacturing and manufacturers are required to increase productivity while maintaining and improving surface quality in order to remain competitive (Karpat & Özel, 2008; Sharma, Dhiman, Sehgal, & Sharma, 2008).

A widely used surface quality indicator is surface roughness. High surface roughness values decrease the fatigue life of machined components (Benardos & Vosniakos, 2002; Özel & Karpat, 2005). The formation of surface roughness is a complex process, affected by many factors like tool variables, workpiece material and cutting parameters. The complex relationship among the parameters involved makes it difficult to generate explicit analytical models for hard turning processes (Karpat & Özel, 2008).

In hard turning, most of process performance characteristics are predictable and, therefore, can be modeled. These models, obtained in different ways, may be used as objective functions in optimization, simulation, controlling and prediction algorithms (Tamizharasan, Sevaraj, & Haq, 2006). Al-Ahmari (2007) sustains that machinability models are important for a proper selection of process parameters in planning manufacturing operations. A better knowledge of the process could ultimately lead to the combination or elimination of one of the operations required in the process, thus reducing product cycle time and increasing productivity (Singh & Rao, 2007).

Among the strategies employed for modeling surface roughness, methods based on expert systems are very often employed by researchers (Chen, Lin, Yang, & Tsai, 2010; Zain, Haron, & Sharif, 2010). Benardos and Vosniakos (2003), in a review about surface roughness prediction in machining processes, pointed that models built by means of artificial intelligence (AI) based approaches were more realistic and accurate in the comparison to those based on theoretical approaches. AI techniques, according to the authors, "*take into consideration particularities of the equipment used and the real machining phenomena*" and are able to include them into the model under construction. Several works make use of ANNs for surface roughness prediction. It can be seen as a 'sensorless' approach for estimation of roughness (Sick, 2002), where networks

* Corresponding author. Address: Av BPS 1303, 37500-903 Itajubá/MG, Brazil. Tel.: +55 35 36291150; fax: +55 35 36291148.
E-mail addresses: fpontes@embraer.com.br (F.J. Pontes), andersonppaiva@unifei.com.br (Anderson Paulo de Paiva), pedro@unifei.edu.br (P.P. Balestrassi), jorofe@unifei.edu.br (J.R. Ferreira), messias@dequi.eel.usp.br (Messias Borges da Silva).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2012.01.058

are trained offline with historical or experimental process data and then employed to predict surface roughness. As pointed out by Coit, Jackson, and Smith (1998), neurocomputing suits modeling of complex manufacturing operations due to its universal function approximation capability, resistance to the noise or missing data, accommodation of multiple non-linear variables for unknown interactions and good generalization capability. Some works, however, report drawbacks in using ANNs for prediction (Ambro-gio, Filice, Shivpuri, & Umbrello, 2008; Bagci & Isik, 2006). An often reported problem with ANNs is the optimization of network parameters. Zhong, Khoo, and Han (2006) affirms that there is no exact solution for the definition of the number of layers and neural nodes required for particular applications.

This study proposes the application of the design of experiments (DOE) methodology for the design of neural networks of RBF (Radial Basis Function) architecture applied to the prediction of surface roughness (R_a) in the turning process of AISI 52100 hardened steel. The factors considered were the network parameters: number of radial units on the hidden layer, the algorithm employed to calculate the spread factor of radial units and the algorithm employed to calculate center location of the radial functions. This work will make use of Taguchi's orthogonal arrays to identify levels of factors that benefits network prediction skills, to assess the relative importance of each design parameter on network performance. This made it possible to evaluate the relative importance of each design factor on network performance and the accuracy attainable by RBFs as the amount of examples available for training and selection varies. Pairs of input–output data obtained from turning operations were used to generate examples for network training and for confirmation runs. Cutting speed (V), feed (f), and depth of cut (d) were employed as network inputs. The results pinpoint network configurations that presented the best results in prediction, for each size of training set. It is expected that RBF networks present good performance on the proposed task.

2. Surface roughness

Benardos and Vosniakos (2003) define surface roughness as the superimposition of deviations from a nominal surface from the third to the sixth order where the orders of deviation are defined by international standards (ISO 4287, 2005). The concept is illustrated in Fig. 1. Deviations of first and second orders are related to form. Consisting of flatness, circularity, and waviness, these deviations are due to such things as machine tool errors, deformation of the workpiece, erroneous setups and clamping, and vibration and workpiece material inhomogeneities. Deviations from third and fourth orders, which consist of periodic grooves, cracks, and dilapidations, are due to shape and condition of cutting edges, chip formation, and process kinematics. Deviations from fifth and sixth orders are linked to workpiece material structure and are related to physicochemical mechanisms acting on a grain and lattice scale such as slip, diffusion, oxidation, and residual stress (Benardos & Vosniakos, 2003).

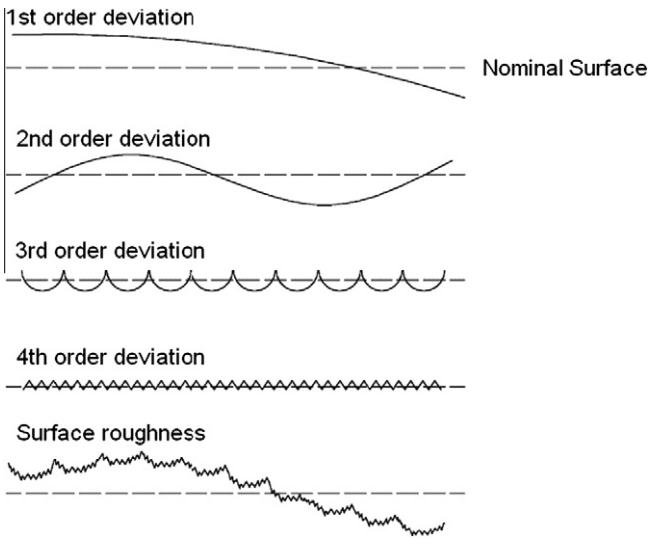

Fig. 1. Nominal surface deviations—adapted from DIN4760 (1982).

Surface roughness defines the functional behavior of a part. It plays an important role in determining the quality of a machined product. Roughness is thus an indicator of process performance and must be controlled within suitable limits for particular machining operations (Basheer et al., 2008; Karpat & Özel, 2008).

The factors leading to roughness formation are complex. Karayel (2009) declares that surface roughness depends on many factors including machine tool structural parameters, cutting tool geometry, workpiece, and cutting tool materials. The roughness is determined by the cutting parameters and by irregularities during machining operations such as tool wear, chatter, cutting tool deflections, presence of cutting fluid, and properties of the workpiece material. In traditional machining processes, Benardos and Vosniakos (2002) maintain that the most influential factors on surface roughness are: mounting errors of the cutter in its arbor and of the cutter inserts in the cutter head, periodically varying rigidity of the workpiece cutting tool machine system wear on cutting tool, and formation during machining of built-up edge and non-uniformity of cutting conditions (depth of cut, cutting speed, and feed rate). The same authors claim that statistically significant in roughness formation are the absolute values of cutting parameters such as depth of cut, feed, and components of cutting force. Still, not only the enlisted factors are influential, according to Benardos and Vosniakos (2002), but also the interaction among them can further deteriorate surface quality.

The process-dependent nature of roughness formation, as Benardos and Vosniakos (2003) explain, along with the numerous uncontrollable factors that influence the phenomena makes it difficult to predict surface roughness. The authors state that the most common practice is the selection of conservative process parameters. This route neither guarantees the desired surface finish nor attains high metal removal rates. According to Davim, Gaitonde, and Karnik (2008), operators working on lathes use their own experience and machining guidelines in order to achieve the best possible surface finish. Among the figures used to measure surface roughness, the most commonly used in the literature is roughness average (R_a). It is defined as the arithmetic mean value of the profile's departure from the mean line throughout a sample's length. Roughness average can be expressed as in Eq. (1) (ISO 4287, 2005):

$$R_a = \frac{1}{l_m}\int_0^{l_m} |y(x)|dx \tag{1}$$

where R_a stands for roughness average value, typically measured in micrometers (μm), l_m stands for the sampling length of the profile, and $|y(x)|$ stands for the absolute measured values of the peak and valley in relation to the center line average (μm). Correa, Bielza, and Pamies-Teixeira (2009) point out that being an average value and thus not strongly correlated with defects on the surface, R_a is not suitable for defect detection. Yet they also proclaim that due to its strong correlation with physical properties of machined products, the average is of significant regard in manufacturing.

Benardos and Vosniakos (2003), in a review on the subject, grouped the efforts to model surface roughness into four main groups: (1) methods based on machining theory, aimed at the development of analytical models; (2) investigations on the effect of various factors on roughness formation through the execution of

experiments; (3) design of experiment (DOE)-based approaches; and (4) methods based on artificial intelligence techniques.

Eq. (2) offers an example of a traditional theoretical model where R_a stands for roughness average (in μm), f stands for feed (in mm/rev), and r stands for tool nose radius (in mm).

$$R_a \sim 0.032x\frac{f^2}{r} \tag{2}$$

Such models, Sharma et al. (2008) tell us, take no account of imperfections in the process, such as tool vibration or chip adhesion. In some cases, according to authors like Zhong et al. (2006), Karpat and Özel (2008), results differ from predictions.

Singh and Rao (2007) describe experimental attempts to investigate the process of roughness formation. Using finish hard turning of bearing steel (AISI 52100), the authors study the effects of cutting conditions and tool geometry on surface roughness.

Empirical models are also employed for modeling surface roughness, generally as a result of experimental approaches involving multiple regression analysis or experiments planned according to DOE techniques. An example of this strategy can be found in Sharma et al. (2008). Cus and Zuperl (2006) proposed empirical models (linear and exponential) for surface roughness as a function of cutting conditions, as shown in Eq. (3):

$$R_a = C_0(V^{C_1} \times f^{C_2} \times d^{C_3}) \tag{3}$$

In Eq. (3), R_a stands for roughness average. V, f, and d stand for cutting speed (m/min), feed (mm/rev), and depth of cut (mm), respectively. C_0, C_1, C_2, and C_3 are constants that must be experimentally determined and are specific for a given combination of tool, machine, and workpiece material. Zain et al. (2010) point to the fact that in many cases, regression analysis models established using DOE techniques failed to correctly predict minimal roughness values.

3. Artificial neural networks

3.1. Radial Basis Function networks (RBF)

According to Haykin (2008), an artificial neural network (ANN) is a distributed parallel systems composed by simple processing units called nodes or neurons, which perform specific mathematic functions (generally non-linear), thus corresponding to a non-algorithmic form of computation. In its most basic format, an artificial neuron is an information processing unit composed of: a set of synapses, each one characterized by a weight value; an adder, responsible by summing the input signals properly multiplied by the weight values in the synapses; and an activation function. In an artificial neural network, the knowledge about a given problem is stored in the values of the weights of the synapses that interconnect neurons in the layers of the network. An activation function defines the output of a network node in terms of the level of activity in its inputs (Haykin, 2008).

The ability to learn by means of examples and to generalize learned information is, doubtless, the main attractive in the solution of problems using artificial neural networks, according to Braga, Carvalho, and Ludermir (2007). It is a main task of a neural network to learn a model from from its surrounding environment and to keep such a model sufficiently consistent to the real world so as to reach the goals specified for the application it is intended to perform. The use of neural networks in solving a given problem involves determining the design parameters of the network, a learning phase and a test phase, during which the performance of the network is assessed (Haykin, 2008). Fig. 2 shows, as an example, a Radial Basis Function (RBF) network.

The figure shows a typical RBF composed of three layers: an input layer composed of three radial units, a hidden layer where

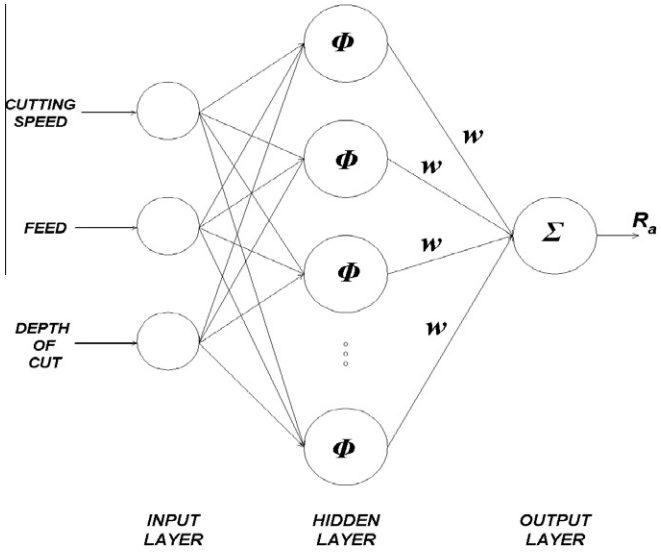

Fig. 2. Schematic diagram of a RBF network.

non-linear processing (represented by function ϕ) is carried out; and an output layer, containing a single unit. Each input unit is connected to all radial units on the hidden layer and each radial units on the hidden layer is connected by weighted synapses (represented by w) to the output layer. The synaptic weights are modified during training phase in order to teach the networks the non-linear relationship that exists between inputs and output.

The radial function in use is usually a Gaussian function. The output layer usually contains neurons that calculate the scalar product of its inputs. In a RBF network having k radial units in the intermediate layer and one output, this is given by Eq. (4) (Bishop, 2007):

$$y = \sum_{i=1}^{k} w_i \phi(\|x - \mu\|^2) + w_0 \tag{4}$$

where x represent an input vector, μ represents the hyper-center of radial units, ϕ represents the activation function of the radial units, as, for instance, a Gaussian function; w_i represents the weight values by which the output of a radial unit is multiplied by in the output layer and w_0 is a constant factor.

3.2. ANNs applied to surface roughness prediction

Neural network models have been widely applied to prediction tasks in hard turning processes. Networks of MLP (multi-layer perceptron) architecture are employed in most of them. Works comparing the performance of ANN models to that presented by DOE based models are not rare, with mixed results. In Erzurumlu and Oktem (2007), a response surface model RSM and an ANN are developed for prediction of surface roughness in mold surfaces. According to the authors the neural network model presented slightly better performance, though at a much higher computational cost. In Çaydas and Hasçalik (2008), an ANN and a regression model were developed to predict surface roughness in abrasive waterjet machining process. In this case, the regression model was slightly superior. Palanisamy, Rajendran, and Shanmugasundaram (2008) compared the performance of regression and ANN models for predicting tool wear in ending milling operation, with ANNs presenting better results. Karnik, Gaitonde, and Davim (2008) applied neural networks and RSM models to predict the burr size for a drilling process. The authors concluded that ANN performance was clearly superior to that obtained by the polynomial model. Bagci and Isik (2006) developed an ANN and a

response surface model to predict surface roughness on the turned part surface in turning unidirectional glass fiber reinforced composites. Both models were deemed as satisfactory. The use of neural networks in conjunction with other methods is yet another strategy adopted by some authors (Karpat & Özel, 2008).

Only a few studies make use of RBF networks for prediction in machining processes. Shie (2008) combined a trained RBF network and a sequential quadratic programming method in order to find an optimal parameter setting for an injection molding process. In Dubey (2009), they are employed in conjunction with desirability function and genetic algorithms in a hybrid approach for multi-performance optimization in electro-chemical honing process. Sonar, Dixit, and Ohja (2006) made use of RBFs for prediction of surface roughness in the turning process of mild steel with carbide tools. In that work, RBFs were outperformed by MLPs. Nevertheless, the authors emphasized that RBF definition was simple and its training fast. Cus and Zuperl (2006) performed a comparison between the performance of MLP and RBF networks applied to predict surface roughness in turning operations. Although MLP have outperformed the RBF, that work evidences that RBF is stable and converges much faster than MLPs. El-Mounayri, Kishawy, and Briceno (2005) employed RBF networks to prediction of cutting forces in CNC ball end milling operations. Results of that work reveal that RBF's achieved a high level of accuracy in the proposed task. Once more, authors stressed the easy definition and fast convergence of the network.

3.3. Network topology definition

Distinct approaches can be found in literature for the definition of the network topologies employed for roughness prediction. In a review of several publications dealing with surface roughness modeling in machining processes by means of artificial neural networks, Pontes, Ferreira, Silva, Paiva, and Balestrassi (2010) pointed to the fact that trial and error still remains as the most frequent technique for ANN topology definition, as in Erzurumlu and Oktem (2007). In some studies, heuristics are used to define the parameters (Kohli & Dixit, 2005). In other cases, a 'one-factor-at-a-time' technique is used in the search for a suitable configuration (Fredj & Amamou, 2006; Kohli & Dixit, 2005).

The use of DOE techniques for optimization is scarcely found. There are some examples as the work of Quiza, Figueira and Davim (2008), where an experimental design is employed to configure a neural network of MLP architecture intended to predict tool flank wear in hard machining of D2 AISI steel. The following factors are employed in the experimental design: learning rate, moment constant, training epochs and number of neurons in hidden layer. In regard to the use of Taguchi method as a tool for designing neural networks, Khaw, Lim, and Lim (1995) employed Taguchi's methodology to the project of MLP networks with the aim of maximize their accuracy and speed of convergence. Kim and Yum (2003) made use of Taguchi's methodology to design parameters of MLP networks in order to maximize network robustness in presence of noise signals. In Balestrassi, Popova, Paiva, and Lima (2009) the Taguchi methodology was employed for the optimization of MLP networks applied to time series prediction. The authors sustain that traditional methods of studying one-factor-at-a-time may lead to unreliable and misleading results while and error can lead to sub-optimal solutions.

In the comparison with previous papers, the present paper could be innovative in the following points:

- The use of DOE technique for the design of RBF networks in surface roughness prediction, considering a large database.
- The study of the relative importance of the design factors on network performance.
- The assessment of the attainable accuracy in surface roughness prediction for turning of AISI 52100 steel for distinct amounts of examples available for training the networks.

4. Design of experiments in Taguchi's methods

According to Montgomery (2009), the design of experiments (DOE) methodology consists in design experiments capable of generating data suitable for a statistical analysis of its results, what in turn leads to valid and objective conclusions. The DOE approach comprises execution of experiments in which factors involved in a process under analysis are varied simultaneously, with the goal of measuring its effect over the output variable (or variables) of such a process.

The strategy employed for designing experiments in Taguchi's methods is based in orthogonal arrays. They correspond to a kind of fractional factorial designs, in which not all possible combinations among factors and levels are tested. It is useful for estimation of the factor main effects over the process. The first objective of this kind of strategy is to obtain the maximum amount of information about the effect of the parameters over the process with a minimum of experimental runs (Ross, 1991).

In addition to the fact of requiring a smaller number of experiments, orthogonal array employed in Taguchi's methods allow to test factors having a different number of levels. That makes it possible to perform experiments containing some factors having two levels, and some having four levels, for example. For a given experiment designed with three factors, one of them containing four levels and the other containing two levels each, without investigating interaction among factors, the orthogonal array L8 from the Taguchi method is as shown in Table 1, where the numeral in column 'Number of experiment' specify the number of the experimental run and the numeral in columns 'Factor A', 'Factor B' and 'Factor C' specify the levels of the respective factor.

4.1. Analysis of experiments

The function loss of quality, in Taguchi's methods, varies depending on the type of problem under study. Problems may be classified as being of type "the smaller, the better", "the bigger, the better" and "nominal is better". In this work, the goal is the minimization of the output analyzed, what makes it a "the smaller, the better" type of problem. For such a problem, the signal to noise ratio to be maximized is expressed by Eq. (5) (Ross, 1991):

$$\eta = -10\log_{10}\left[\frac{1}{n}\sum_{i=1}^{n} y_i^2\right] \tag{5}$$

where η is the value of the signal to noise ratio, y_i is the value of the deviation regarding the quality attribute whose tolerance is of the type "the smaller, the better" and n stands for the number of experiments executed. Data experimentally obtained is analyzed according to their average, valor of signal to noise ratios and standard deviation

Table 1
L8 orthogonal array for an experiment involving three factor: one factor with 4 levels and two factor with 2 levels each.

Number of the experiment	Factor A	Factor B	Factor C
1	1	1	1
2	1	2	2
3	2	1	1
4	2	2	2
5	3	1	2
6	3	2	1
7	4	1	2
8	4	2	1

Fonte: Minitab®. Statistical Software Release 15.0.

of the runs. The goal of the analysis is to obtain the levels of the factors involved that lead to the minimization of the function loss of quality and to the maximization of the signal to noise ratio (Kilickap, 2010).

4.2. S. D. Ratio

The output variable chosen as measure to compare the influence of the different design factors in the performance of the network is the S. D. Ratio obtained during the testing phase. In a regression problem, S. D. Ratio is defined as the ratio between the standard deviation of the residuals by the standard deviation of data obtained experimentally. The closer to zero the value of S. D. Ratio is the better the prediction capability of the model. S. D. Ratio corresponds to one minus the variance explained by the model (Ross, 1991).

4.3. Inference about the mean of a population having known variance

The sample mean $\overline{X}$ is an unbiased estimator of the mean μ of a population, since the variance σ of the population is known. If the conditions of the central limit theorem apply, the distribution of $\overline{X}$ is approximately normal with mean μ and variance given by $\sigma/\sqrt{n}$, where n is the size of the sample. To test the null hypothesis $\mu = \mu_0$, the test statistic given by Eq. (6) can be employed (Montgomery, 2009):

$$Z_0 = \frac{\sqrt{n}(\overline{X} - \mu_0)}{\sigma} \tag{6}$$

The null hypothesis that the means are equal is rejected if the resulting P-value is inferior the level of significance adopted. This test, according to Montgomery (2009) can be used provided that the size of the sample is superior to thirty.

4.4. Levene's test

The Levene's test is employed to test the equality of variances coming from different samples. A null hypothesis that variances of the samples involved are equal is tested at a given level of significance. Such a test is recommended when there is no evidence that the samples testes follow the normal distribution. The null hypothesis is rejected provided that the situation expressed by Eq. (7) takes place (Montgomery, 2009):

$$W > F_{(\alpha,k-1,N-k)} \tag{7}$$

where F is the critical upper value of a F distribution having $k-1$ and $N-k$ degrees of freedom, at the level of significance α and W is the test statistic given by Eq. (8) (Montgomery, 2009).

$$W = \frac{(N-k) * \sum_{i-1}^{k} N_i(\overline{Z_{i.}} - \overline{Z}_{..})^2}{(k-1) * \sum_{i-1}^{k}\sum_{j=1}^{N_i}(\overline{Z_{ij}} - \overline{Z_{i.}})^2} \tag{8}$$

where N stands for the total number of reading involved in the samples under test, k is the number of samples being compared, N_i is the number of readings in each sample, $\overline{Z}_{..}$ is the overall mean of the samples, $\overline{Z}_{i.}$ is the mean of sample i and $\overline{Z}_{ij}$ is given by Eq. (9) (Montgomery, 2009).

$$Z_{ij} = |Y_{ij} - \overline{Y_{i.}}| \tag{9}$$

where Y_{ij} stands for the j-esimal reading of the i-esimal sample and $\overline{Y}_{i.}$ to the mean value of the i-esimal sample.

5. Experimental procedures

The experimental procedure consisted in the following steps:

- Cutting operations intended to build a database to train and select the ANNs.
- Generation of training and testing data sets.
- Simulation experiments, planned according to Taguchi's methods, intended to identify best network topologies.
- Confirmatory experiments intended to validate the network topologies identified during planned experiments.

5.1. Machining tests

The workpieces employed were made with dimensions of $\varnothing 49 \times 50$ mm. All of them were quenched and tempered. A total of 60 workpieces of AISI 52100 steel bars of the same lot were employed during the experiments, where chemical composition shown in Table 2. Firstly they were machined using a Romi S40 lathe. After this heat treatment, their hardness was between 53 and 55 HRC, up to a depth of 3 mm below the surface. Hardness profile was measured at six points in each workpiece and no significant differences in hardness profile were detected.

The machine tool used was a CNC lathe with power of 5.5 kW in the spindle motor, with conventional roller bearings. The mixed ceramic (Al_2O_3 + TiC) inserts used were coated with a very thin layer of titanium nitride (TiN) presenting a chamfer on the edges. The tools employed in the study were produced by Sandvik Coromant, class GC6050, CNGA 120408 S01525. The tool holder presented negative geometry with ISO code DCLNL 1616H12 and entering angle $\chi r = 95°$.

In this study, cutting speed (V), feed (f), and depth of cut (d) were employed as controlling variables. Those cutting conditions varied as follows: 200 m/min $\leqslant V \leqslant$ 240 m/min, 0.05 mm/r $\leqslant f \leqslant$ 0.10 mm/r and 0.15 mm $\leqslant d \leqslant$ 0.30 mm. The adopted values correspond to the operational limits enlisted by the toolmaker on its catalog (Sandvik Coromant, 2010). The cutting experiments used to train and test the ANN followed a RSM design. This original CCD design is formed by three distinct groups of experimental points: (i) a full factorial design with 2^3 runs, (ii) six axial points and (iii) four center points, resulting in 18 runs. Using three replicates for each run and augmenting the experimental design with 6 face centered runs, the entire design was built with 60 runs, as can be seen in Fig. 3. Then, 60 workpieces of AISI 52100 hardened steel were turned with 60 different configurations. In each of 60 workpieces, ten surface roughness measurements were done, resulting in a data set for training and testing sets for the ANN with 600 cases.

A Taylor Hobson rugosimeter, model Surtronic 3+ was employed for roughness measurements, as well as a Mitutoyo micrometer. Roughness measures were taken after the tenth machining stroke. The 10 roughness measures were collected as following: three measurements at each extremity (chuck and live centre) and four at the middle point. All measures were taken after the end of tool life. The criteria adopted for determining the end of tool life end was tool flank wear VB_{max} equal or greater than 0.3 mm.

5.2. Experimental design for selection of ANN parameters

The problem to be addressed by the designed experiment was to identify the best topology for roughness prediction. The experimental factors considered were the design parameters of the RBF networks: the algorithm for calculation of the radial spread factor (X_1), with four levels; the number of radial units present on the

Table 2
Chemical composition of the AISI 52100 steel (weight percentage).

C	Si	Mn	Cr	Mo	Ni	S	P
1.03	0.23	0.35	1.40	0.04	0.11	0.001	0.01

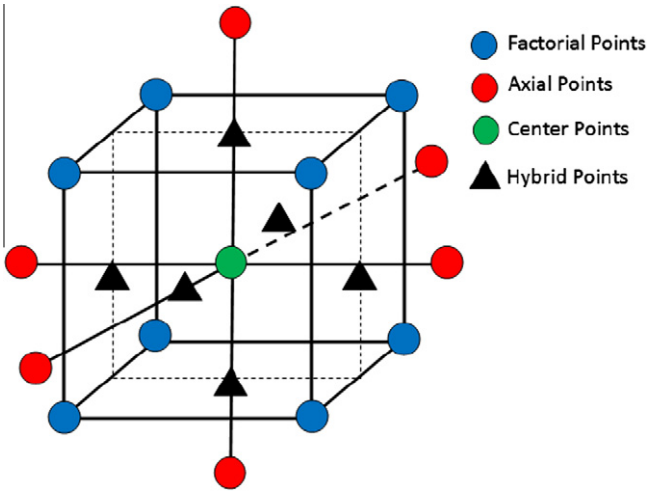

Fig. 3. Central Composite Design (CCD) augmented with hybrid points.

hidden layer of the network (X_2), with two levels, and the algorithm for calculation of the center location of radial functions (X_3), also with two levels. To achieve the established goals for the study, distinct experiments were conducted for different sizes of data sets. Eight data sets of different sizes were formed, containing 24, 30, 48, 60, 240, 300, 400 and 500 examples. The first data set contained the first 24 examples (*V, f, d,* R_a); the second training set contained the first 30 examples (*V, f, d,* R_a), and so on, up to the last training set, containing 500 examples. Two thirds of the examples contained in each data set were used as training set for networks and one third was employed as a selection set. The remaining 100 examples did not take part in any network training activity and were spared to be used as test cases during confirmation runs.

Regarding the algorithm for calculation of the radial spread factor, two distinct algorithms were tested: the isotropic and the *K*-Nearest algorithms (Haykin, 2008). For the isotropic algorithm two levels of its scaling factor were investigated, based on results of preliminary experiments. For the *K*-Nearest algorithm, the influence of its defining factor *K* was investigated. Once more, two different values of the factor were selected for testing, based on results of preliminary experiments.

Regarding the number of radial units, the levels of the factor were defined as proportions between that number of radial units and the number of training examples, as suggested by Haykin (2008). The proportions established as levels of the factor were 50% and 100% of the number of examples available for training in each experiment.

Two distinct algorithms for calculation of the center location of radial functions were tested: the Sub-Sampling algorithm and *K*-Means algorithm Haykin (2008). Each algorithm was established as a level of the experimental factor. As a consequence, the orthogonal array employed for each of the eight experiments was a L8 Taguchi array having three factors. The factors and their respective levels are detailed in Table 3.

5.3. Factor and levels adopted for experimental planning

Execution of each experimental arrangement consisted in configuring the network as specified by the experimental design and training the ANN. The Neural Networks suite of the statistical software package Statistica® release 7.1 was employed. Sixty replications were performed for each network configuration, meaning that a network configuration under test was independently initialized and trained for 60 times, in order to mitigate risks associated to the random initialization of synaptical weights (Haykin, 2008). Examples were presented in a random sequence to the network during training.

Regarding pre and post processing, data was normalized to the interval (0, 1) to be applied to network inputs and re-scaled to the original dominium at the output. Results were stored under the format of files produced by the software package, containing the prediction of the networks for test cases. The results were compiled to identify factor levels favouring network performance in prediction, to investigate relative importance of each factor.

The best network configurations for each data set were kept and subjected to confirmation runs. Those consisted in applying the networks to predict surface roughness for the 100 examples spared from training, in order to assess network generalization capability.

6. Results and discussion

All the analysis was made by using statistical software Minitab®, release 15. Table 4 displays the mean values of output S. D. Ratio for all the runs of the eight experiments conducted. Table 5 displays the standard deviation associated to the runs.

The foreseen analysis for the average output data, signal to noise ratios and for the standard deviations were performed. Those analyses consider the amount of the difference between the biggest and the smallest effect calculated for a given level of a factor.

For the experiments conducted, the analysis provides values of the main effects of each factor on each analyzed response (output average, signal to noise ratio and standard deviation), as well as a ranking of the impact of the factor on those figures. This information is summarized in Tables 6–8. In the section associated to signal to noise ratio, a value of 1 in the line *rank* indicates the most influential factor in regard to signal to noise ratio and a value of 3 in the same line, the less influential one. In the section of each table devoted to the analysis of the output average, a value of 1 in line *rank* denotes the most influential factor in reducing the average of the output and a value of 3, the less influential factor in reducing the average. In the section devoted to the analysis of the standard deviations of the output, a value of 1 in line *rank* denotes the most influential factor in minimizing the standard deviation of the output, and a value of 3, the less influential factor in reducing the standard deviation.

It is perceived that, in all the cases and for the three analyses performed, the algorithm for determination of the spread for the radial function was appointed as the most influential factor. In regard to minimization of S. D. Ratio and maximization of the signal to noise ratio, the number of radial units was appointed as the second most influential factor in all cases, but two experiments, those involving 24 and 48 training cases. In these two experiments, the second most influential factor was the algorithm for calculation of centers. Regarding minimization of standard deviation, the analysis appointed the same result for the eight experiments, being the number of radial units appointed as the second most influential one.

Figs. 4–6 show, as an example, graphs of the main effects for the experiment conducted with the training set of size 300. By those graphs one can figure out the relative importance of each effect. For each analysis performed S. D. Ratio average, signal to noise ratio and S. D. Ratio standard deviation, the bigger the difference between the main effects of a given level of a factor, the bigger the influence of that factor. In the graph one can also figure out the levels of the factors the analysis points out as those that will cause the network to perform better in the task of prediction.

In accordance to what is shown in Figs. 4–6 and analyzing the results from the Taguchi's analysis, the levels of the factors pointed as those that lead to minimum S. D. Ratio average and S. D. Ratio standard deviation, as well as maximum robustness among all

Table 3
Factors and levels involved in the experiments.

Factor	Number of levels	Level 1	Level 2	Level 3	Level 4
Algorithm selection of spread factor	4	Isotropic Deviation Scaling Factor = 1	Isotropic Deviation Scaling Factor = 10	*K*-Nearest neighbors = 5	*K*-Nearest neighbors = 10
Algorithm calculation of centers of radial function	2	Sub-sampling	*K*-Means	–	–
Number of radial units	2	Half the number of training cases	Equal to the number of training cases	–	–

Table 4
Mean values of S. D. Ratio obtained during the experiments conducted.

Number of run	Number of cases in the training set							
	24	30	48	60	240	300	400	500
1	0.628802	0.601076	0.522039	0.518149	0.525391	0.499767	0.520130	0.533464
2	6.723140	31.596046	97.639313	25.554231	8.539650	136.409844	39.964778	9.638446
3	0.020493	0.017081	0.010203	0.007873	0.000161	0.000055	0.000043	0.000027
4	0.015347	0.016409	0.001517	0.159770	0.000080	0.000059	0.000083	0.000042
5	0.300682	2.059012	0.154282	11.769162	0.302916	33.274003	28.585290	0.252391
6	0.264722	0.253606	0.132365	0.112056	0.097023	0.102912	0.090886	0.080076
7	0.209856	0.146091	0.062477	0.049657	0.188020	31.535303	0.058813	1.332922
8	0.219666	0.165157	0.075796	0.058061	0.046739	0.050059	0.040693	0.036463

Table 5
Values of standard deviation of the output obtained during the experiments conducted.

Number of run	Number of cases in the training set							
	24	30	48	60	240	300	400	500
1	0.097507	0.054823	0.056103	0.062090	0.044123	0.036352	0.038504	0.040225
2	25.610531	69.309196	284.475841	165.438803	24.592671	433.037908	116.382653	62.502353
3	0.007782	0.005116	0.001351	0.002331	0.000048	0.000032	0.000023	0.000013
4	0.001007	0.000000	0.000130	0.575948	0.000030	0.000026	0.000058	0.000035
5	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000001	0.000000
6	0.069950	0.067611	0.028554	0.022283	0.012444	0.016940	0.014007	0.009498
7	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	0.074861	0.042196	0.016699	0.013782	0.008525	0.012590	0.009017	0.007559

Table 6
Values of main effects for the experiment conducted with 30 training cases.

Level	Signal to noise ratios			Means			Standard deviations		
	Algorithm for center spread	Algorithm for center location	Number radial units	Algorithm for center spread	Algorithm for center location	Number radial units	Algorithm for center spread	Algorithm for center location	Number radial units
1	−13.6149	15.4594	19.5973	25.3903	0.3604	0,1508	13.5267	0.48849	0.15330
2	38.3478	9.2724	5.1345	0.0097	12.6813	12.8908	0.0100	6.72432	7.05952
3	5.6830			0.5951			0.7937		
4	19.0478			0.0884			0.0951		
Delta	51.9627	6.1869	14.4628	25.3806	12.3209	12.7400	13.5167	6.23583	6.90622
Rank	1	3	2	1	3	2	1	3	2

Table 7
Values of main effects for the experiment conducted with 240 training cases.

Level	Signal to noise ratios			Means			Standard deviations		
	Algorithm for center spread	Algorithm for center location	Number of radial units	Algorithm for center spread	Algorithm for center location	Number of radial units	Algorithm for center spread	Algorithm for center location	Number of radial units
1	−8.3651	29.4960	34.9376	8.42546	0.13258	0.09181	5.84575	0.17188	0.10680
2	81.4415	27.9389	22.4973	0.00008	4.16215	4.20292	0.00006	2.85952	2.92459
3	18.2929			0.10310			0.13700		
4	23.5004			0.06082			0.07999		
Delta	89.8066	1.5570	12.4402	8.42538	4.02956	4.11111	5.84570	2.68763	2.81779
Rank	1	3	2	1	3	2	1	3	2

Table 8
Values of main effects for the experiment conducted with 400 training cases.

Level	Signal to noise ratios			Means			Standard deviations		
	Algorithm for center spread	Algorithm for center location	Number of radial units	Algorithm for center spread	Algorithm for center location	Number of radial units	Algorithm for center spread	Algorithm for center location	Number of radial units
1	−23.630	25.0354	41.2886	247.207	16.761	0.086	206.280	23.694	0.106
2	89.577	23.8854	7.6323	0.000	123.541	140.216	0.000	103.055	126.643
3	−5.662			33.386			47.207		
4	37.556			0.011			0.011		
Delta	113.207	1.1499	33.6563	247.207	106.779	140.130	206.280	79.361	126.537
Rank	1	3	2	1	3	2	1	3	2

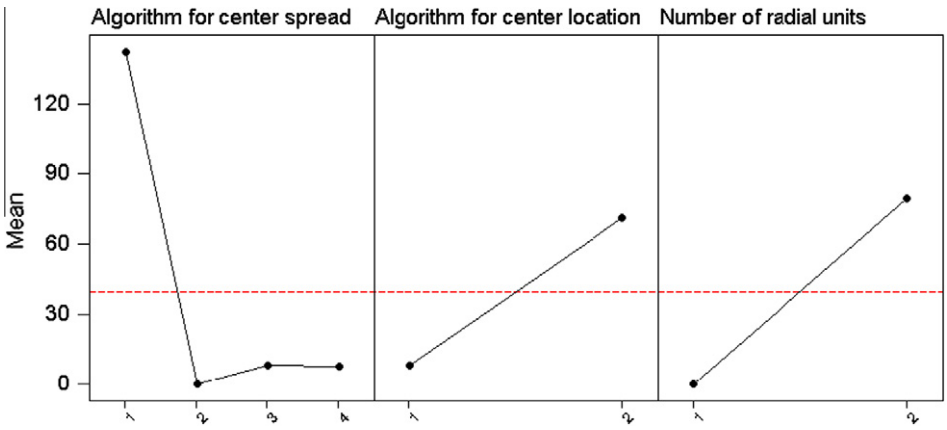

Fig. 4. Main effects on average of S. D. Ratio conducted with 300 training cases.

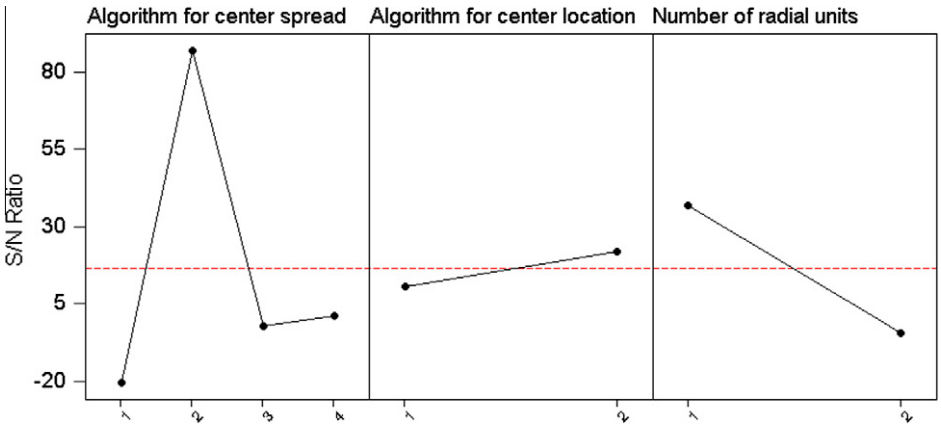

Fig. 5. Main effects on signal to noise ratio conducted with 300 training cases.

the conditions tested are summarized in Table 9. It was found that, except for one case, the configurations of factors are the same. In regard to signal to noise ratio, in six out of eight experiments, the configurations appointed as the best for robustness are the same. Regarding standard deviation, the configuration of factor pointed out as the best for reducing variance was the same in seven out of eight experiments.

For each experiment, nonetheless, tests of hypothesis for inference about the mean of a population using Z statistic (Montgomery, 2009), were applied. Tests using statistic t of Student and analysis of variance were not applied because preliminary statistical tests did not present statistical evidence of equal variance among samples, nor that they follow a normal distribution, both of which would be required for using those techniques.

In each experiment, the mean value of a given configuration was compared (by using the Z test) to the value of the mean of each other configuration, at the level of significance of 0.05. The null hypothesis assumed was the mean of the tested configuration to be equal to the other mean. The objective of these tests was to establish statistically which configuration presented the smallest S. D. Ratio among the configurations tested.

By comparing results obtained from Taguchi's analysis and those obtained from the Z tests, some discrepancies were found. In the experiments with training sets containing 24, 48 and 240 cases, the analysis of Taguchi's pointed to a configuration that was not the one possessing the smallest average of S. D. Ratio among the runs of those experiments. In addition, for the experiment with a training set containing 60 cases, the analysis pointed

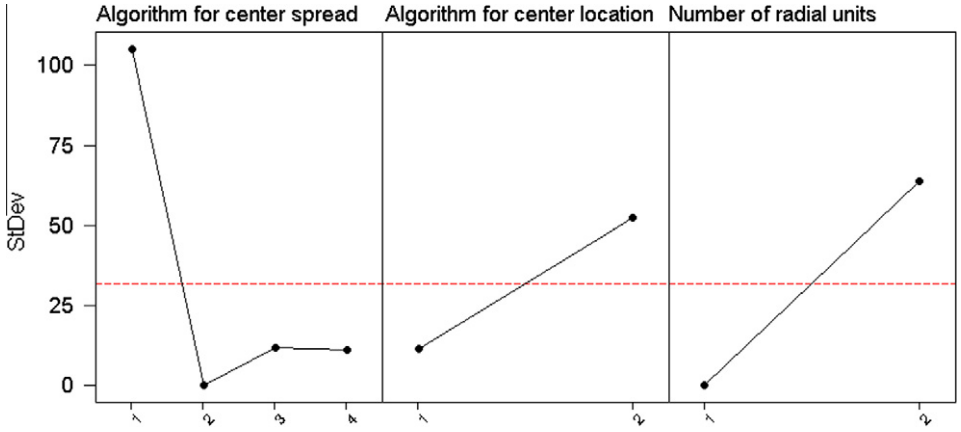

Fig. 6. Main effects on standard deviation of S. D. Ratio conducted with 300 training cases.

Table 9
Configurations pointed out by the analysis as best ones in each criterion, for each experiment.

Number of test cases	Best configuration for minimizing mean of S. D. Ratio	Best configuration for maximizing signal to noise ratio	Best configuration for reducing standard deviation
24	2:1:1	2:1:1	2:1:1
30	2:1:1	2:1:1	2:1:1
48	2:1:1	2:1:1	2:1:1
60	4:1:1	4:1:1	4:1:1
240	2:1:1	2:1:1	2:1:1
300	2:1:1	2:2:1	2:1:1
400	2:1:1	2:1:1	2:1:1
500	2:1:1	2:1:1	2:1:1

Table 10
Comparison between configurations pointed by Taguchi's analysis with smallest S. D. Ratio (discrepant cases).

Number of test cases	Mean of the configuration pointed as the best	S. D. configuration pointed by analysis	Smallest mean observed during test	S. D. to smallest observed mean	*P*-value resulting from second *Z* test
24	0.020493	0.007782	0.015347	0.001007	0.000
48	0.010203	0.001351	0.001517	0.000130	0.000
60	0.076326	0.028080	0.007873	0.002331	0.000
240	0.000161	0.000048	0.000080	0.000030	0.000

to a configuration that was not part of the orthogonal array employed (i.e. that was not part of the experiments).

For the experiments with 24, 48 and 240 cases (those where discrepancies were found), a new statistical comparison was performed with use of the *Z* test. Samples of the configurations pointed by Taguchi's analysis were compared to mean values pointed by the first *Z* test as the lowest. For the experiment with 60 cases, the configuration pointed by the Taguchi's analysis was built up and an extra experimental run was conducted, using the appropriate training set and the same number of repetitions of all the other runs. After that, a new *Z* test was performed to compare the mean values for this new run to the smallest of the original experiment. Once more, the null hypothesis was that the average of the S. D. Ratio of the sample pointed by Taguchi's analysis was equal to the average of better performance among experimental runs. Results for these tests are displayed in Table 10.

Regarding Table 10, it is noted that configurations pointed by Taguchi's analysis possess bigger means (and so a poorer performance) than those of the experimental runs, as can be verified by the *P*-values equal to zero. This can indicate the existence of interactions among the factors involved in the experiment, what could not be detected due to the simple orthogonal array employed.

The best overall configurations obtained for prediction of roughness average (R_a), average of S. D. Ratio, as well as values of standard deviations obtained for those configurations are shown in Table 11.

Data displayed in Table 11 does not allow to conclude for the equality or inequality of the outputs obtained. In order to compare statistically the best configurations obtained for each experiment, tests were applied to compare variances and mean values for each of the configurations to which data in Table 11. To test variances among configurations, Levene's tests were applied for a null hypothesis of equal variance among each possible pair of configurations at a level of significance of 0.05. This test was chosen to compare pairs of variance because preliminary tests did not present evidence that samples follow normal distribution. The results obtained from Levene's tests can be observed in Table 12, where a *P*-value superior to the significance level adopted means that there is no statistical evidence of difference among pairs of variance and a *P*-value inferior to 0.05 evidences difference between the variances of two configurations under test.

Analysis of Table 12 indicates that there is no evidence of difference in the variance for the best network obtained for the training set containing 24 cases and those networks obtained for training sets containing 30, 48, 240, 300, 400 and 500 training cases. It also evidences that variance of the best network obtained for the training set containing 30 cases is inferior to that of networks obtained for training sets containing more cases.

The results of the tests provide evidence that, at the level of significance adopted, the variance of the best network obtained for the training set containing 48 cases is inferior to that of the best network obtained from the training set containing 60 cases. The results show that there is no statistical evidence, at the level of significance adopted, of difference between the variance of the best network obtained for the 48 cases training set and that of the best networks obtained for 240 and 300 training cases. On the other hand, the results show that the best network obtained for the 48 cases training set present a variance that is superior to that observed for the best network obtained using 400 and 500 training cases. For its turn, the network obtained using 60 training cases presents strong evidence of variance superior to that of the best networks obtained for all training set containing a bigger number of cases.

Regarding the best network obtained using 240 training cases, the tests provide no statistical evidence of difference of its variance

Table 11
Configuration having best performance obtained for the experiments and respective value of mean S. D. Ratio for roughness average (R_a).

Size of training set	Algorithm for selection of spread of radial function	Algorithm for calculation of centers of radial function	Number of radial units	Mean S. D. Ratio for configuration	Associated standard deviation
24	Isotropic Deviation Scaling Factor = 10	*K*-Means	24	0.015347	0.001007
30	Isotropic Deviation Scaling Factor = 10	*K*-Means	30	0.016409	0.000000
48	Isotropic Deviation Scaling Factor = 10	*K*-Means	48	0.001517	0.000130
60	Isotropic Deviation Scaling Factor = 10	Sub-Sampling	30	0.007873	0.002331
240	Isotropic Deviation Scaling Factor = 10	*K*-Means	240	0.000080	0.000030
300	Isotropic Deviation Scaling Factor = 10	Sub-Sampling	150	0.000055	0.000032
400	Isotropic Deviation Scaling Factor = 10	Sub-Sampling	200	0.000043	0.000023
500	Isotropic Deviation Scaling Factor = 10	Sub-Sampling	250	0.000027	0.000013

Table 12
P-values resulting from Levene's tests applied to pairs of best networks obtained from each experiment.

Size of training set		24	30	48	60	240	300	400
	Observed variances	1.013E−06	7.08E−19	1.683E−08	5.433E−06	8.959E−10	1.007E−09	5.304E−10
P-values obtained from Levene's test								
30	7.08E−19	0.153	–	–	–	–	–	–
48	1.683E−08	0.282	0.005	–	–	–	–	–
60	5.433E−06	0.000	0.000	0.000	–	–	–	–
240	8.959E−10	0.203	0.000	0.118	0.000	–	–	–
300	1.007E−09	0.203	0.000	0.125	0.000	0.938	–	–
400	5.304E−10	0.189	0.000	0.061	0.000	0.124	0.150	–
500	1.725E−10	0.174	0.000	0.024	0.000	0.000	0.001	0.010

Table 13
P-values resulting from *Z*-test for the best network configurations obtained in each experiment.

Number of training cases	Mean values of S. D. Ratio for best network obtained for training set	Number of training cases in the training set of size immediately inferior	Mean of S. D. Ratio for best network obtained for training set of size immediately inferior	*P*-value
30	0.016409	24	0.015347	0.000
48	0.001517	30	0.016409	0.000
60	0.007873	48	0.001517	0.000
240	0.000080	60	0.007873	0.000
300	0.000055	240	0.000080	0.000
400	0.000043	300	0.000055	0.000
500	0.000027	400	0.000043	0.000

and that of the best networks obtained using 300 and 400 training cases. Conversely, the test provide evidence that the variance of the best network obtained using 240 training cases is superior to that observed in the best network obtained using 500 cases.

The result of the Levene's tests for the best network obtained for the training set containing 300 cases, at the level of significance of 0.05, provides no evidence of difference between its variance and that of the best network obtained for 400 cases. Conversely, the results indicate that the variance of the best network obtained for 300 cases is superior to that obtained for 500 cases. Regarding the training set containing 400 cases, the test did not provide evidence of difference between the variance of the best network obtained for that training set and the variance of the best network obtained for 500 training cases, what can be noted by the *P*-value equal to 0.01.

In order to compare means of the S. D. Ratio output from the best network configuration obtained in each experiment, a third set of *Z*-tests (test for inference about the mean of a population) was applied. The tests compared, at a level of significance of 0.05, the mean of S. D. Ratio of the best configuration obtained in one experiment to the mean of S. D. Ratio of the best configuration obtained in the experiment with training set of size immediately inferior. The results of these tests can be observed in Table 13.

Analysis of Table 13 reveals that there is sufficient evidence, at the level of significance adopted, that mean value of S. D. Ratio of the best network obtained using 30 is superior to that obtained using 24 cases. There is also evidence that the mean value of S. D. Ratio of the best network obtained using 48 training cases is inferior to that of the best network obtained using 30 cases.

Regarding the best network obtained using 60 training cases, there is evidence that its mean value of S. D. Ratio is superior to that of the best network obtained using 48 cases. From this point on it can be noted, and there is statistical evidence from the tests, that the mean values of S. D. Ratio of the best network obtained using a given training set is inferior to the mean values of S. D Ration obtained for the training set of size immediately inferior.

Exception made to the experiments involving training sets containing 30 cases and 60 cases, there is a trend towards reduction of mean of S. D. Ratio and variance as the number of training cases increases. This fact suggests a better performance of networks in prediction of roughness average (R_a) as more training cases are made available. A box-plot graph for the best network configurations obtained in each experiment is shown in Fig. 7. In that graph one can observe that mean and dispersion are bigger in the experiments involving 24 or 30 training cases, and that mean and dispersion have a tendency to be reduced in experiments involving more training cases. The exception observed in the experiment involving 60 training cases, in which even the best network present a mean value of S. D. Ratio and dispersion superior to values obtained using 48 cases, can be clearly observed in Fig. 7.

It can be observed that, even in situations involving a small number of training cases (as the one involving 24 or 30 cases), RBF networks presented a good performance in the task of prediction of roughness average (R_a) as can be seen in Table 11. This suggests that RBF's can constitute a valid e economically viable alternative to the task in turning process of SAE 52100 – 55 HRC steel with mixed ceramic tools.

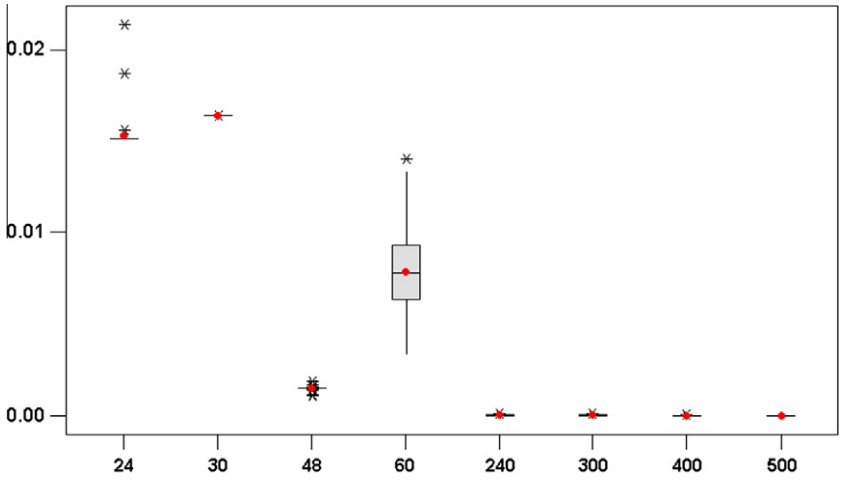

Fig. 7. Boxplot of the output S. D. Ratio for the best network configuration.

7. Conclusions

The use of DOE methodology proved to be an efficient tool for the design of neural networks of RBF architecture for surface roughness prediction in the turning of AISI 52100 hardened steel. The methodology made it possible to identify network configurations presenting high degree of accuracy and reduced variability in the proposed task. Results obtained show that RBF ANNs trained with only 30 examples can present mean value S. D. Ratio equal to 0.016409, for the worst case corresponding to a training set. This fact suggests that RBF networks designed with the use of DOE methodology can be an effective, efficient and affordable alternative for surface roughness prediction in hard turning.

The algorithm for calculation of the radial spread factor was the most influential among the three factors under investigation. The isotropic algorithm with scale factor equal to 10 yielded the best results. The influence of that factor becomes dominant as the number of examples available for training increases. The second most influential factor was the number of radial units on the network hidden layer. The less influential factor was found to be the algorithm for calculation of center locations. The influence of this factor was negligible in almost all experiments. The conclusion is that the option for any of the algorithms tested implied in no pronounced difference in network performances. Such finding can further simplify the design of the ANN to be used for roughness prediction in AISI 52100 hard turning.

It must be emphasized that conclusions obtained in this work cannot be extrapolated to other neural network architectures, other kind of machining operations, other materials or tools. The approach can, nonetheless, be recommended to different network architectures. Further investigation is required in order to evaluate the nature and the impact of the interactions among design factors on the network performance. In some experiments some discrepancies were found among configurations pointed out by the analysis as the best ones and values experimentally observed. This fact suggests the existence of interaction between factors tested, what could not be quantified due to the orthogonal arrays employed. The results show, nonetheless, that RBF networks can present good performance in the task of prediction of roughness average.

Acknowledgements

The authors would like to express their gratitude to FAPEMIG, CNPq and to the CAPES for its support to this research by means of project PE 024/2008 – Programa Pró-Engenharias.

References

Al-Ahmari, A. M. A. (2007). Predictive machinability models for a selected hard material in turning operations. *Journal of Materials Processing Technology, 190*, 305–311.

Bagci, E., & Isik, B. (2006). Investigation of surface roughness in turning unidirectional GFRP composites by using RS methodology and ANN. *International Journal of Advanced Manufacturing Technology, 31*, 10–17.

Balestrassi, P. P., Popova, E., Paiva, A. P., & Lima, J. W. M. (2009). Design of experiments on neural network's training for nonlinear time series forecasting. *Neurocomputing, 72*, 1160–1178.

Basheer, C., Dabade, U. A., Suhas, S. J., & Bhanuprasad, V. V. (2008). Modeling of surface roughness in precision machining of metal matrix composites using ANN. *Journal of Materials Processing Technology, 197*, 439–444.

Benardos, P. G., & Vosniakos, G. C. (2002). Prediction of surface roughness in CNC face milling using neural networks and Taguchi's design of experiments. *Robotics and Computer Integrated Manufacturing, 18*, 343–354.

Benardos, P. G., & Vosniakos, G. C. (2003). Prediction of surface roughness in CNC in machining: A review. *International Journal of Machine Tools and Manufacture, 43*, 833–844.

Bishop, C. M. (2007). *Pattern recognition and machine learning* (1st ed.). New York: Springer Business-Media.

Braga, A. P., Carvalho, L. F., Ludermir, T. B. (2007), Redes Neurais Artificiais: Teoria e Aplicações, Rio de Janeiro, LTC Livros Técnicos Editora S. A.

Çaydas, U., & Hasçalik, A. (2008). A study on surface roughness in abrasive waterjet machining process using artificial neural networks and regression analysis method. *Journal of Materials Processing Technology, 202*, 574–582.

Chen, H., Lin, J., Yang, Y., & Tsai, C. (2010). Optimization of wire electrical discharge machining for pure tungsten using a neural network integrated simulated annealing approach. *Expert Systems with Applications, 37*, 7147–7153.

Coit, D., Jackson, B., & Smith, A. (1998). Static neural network process models: Considerations and case studies. *International Journal of Production Research, 36*, 2953–2967.

Correa, M., Bielza, C., & Pamies-Teixeira, J. (2009). Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process. *Expert Systems with Applications, 36*(3), 7270–7279.

Cus, F., & Zuperl, U. (2006). Approach to optimization of cutting conditions by using artificial neural networks. *Journal of Materials Processing Technology, 173*, 281–290.

Davim, J. P., Gaitonde, V. N., & Karnik, S. R. (2008). Investigations into the effect of cutting conditions on surface roughness in turning of free machining steel by ANN models. *Journal of Materials Processing Technology, 205*, 16–23.

DIN4760 (1982). *Form deviations: Concepts; classification system*. Berlin: Deutches Institut Fuer Normung, Ev.

Dubey, A. K. (2009). A hybrid approach for multi-performance optimization of the electro-chemical honing process. *International Journal of Advanced Manufacturing Technology*. doi:10.1007/s00170-008-1422-8.

El-Mounayri, H., Kishawy, H., & Briceno, J. (2005). Optimization of CNC ball end milling: A neural network-based model. *Journal of Materials Processing Technology, 166*, 50–62.

Erzurumlu, T., & Oktem, H. (2007). Comparison of response surface model with neural network in determining the surface quality of moulded parts. *Materials and Design, 28*, 459–465.

Fredj, N. B., & Amamou, R. (2006). Ground surface roughness prediction based upon experimental design and neural network models. *International Journal of Advanced Manufacturing Technology, 31*, 24–36.

Haykin, S. (2008). *Neural networks and learning machines* (3rd ed.). New Jersey: Pearson Prentice Hall.

ISO 4287:1997/Cor 2:2005 (2005). Geometrical product specifications (GPS)—Surface texture: Profile method—Terms, definitions and surface texture parameters. International Organization for Standardization.

Karayel, D. (2009). Prediction and control of surface roughness in CNC lathe using artificial neural network. *Journal of Materials Processing Technology, 209*, 3125–3137.

Karnik, S. R., Gaitonde, V. N., & Davim, J. P. (2008). A comparative study of the ANN and RSM modeling approaches for predicting burr size in drilling. *International Journal of Advanced Manufacturing Technology, 38*, 868–883.

Karpat, Y., & Özel, T. (2008). Multi-objective optimization for turning processes using neural network modeling and dynamic-neighborhood particle swarm optimization. *International Journal of Advanced Manufacturing Technology, 35*, 234–247.

Khaw, J. F. C., Lim, B. S., & Lim, L. E. N. (1995). Optimal design of neural networks using the Taguchi method. *Neurocomputing, 7*, 225–245.

Kilickap, E. (2010). Optimization of cutting parameters on delamination based on Taguchi method during drilling of GFRP composite. *Expert Systems with Applications, 37*, 6116–6122.

Kim, Y. S., & Yum, B. J. (2003). Robust design of multilayer feedforward neural networks: An experimental approach. *Engineering Applications of Artificial Intelligence, 38*, 1–15.

Kohli, U., & Dixit, D. (2005). A neural-network-based methodology for the prediction of surface roughness in a turning process. *International Journal of Advanced Manufacturing Technology, 25*, 118–129.

Montgomery, D. C. (2009). *Design and analysis of experiments* (7th ed.). New York: Wiley.

Özel, T., & Karpat, Y. (2005). Predictive modeling of surface roughness and tool wear in hard turning using regression and neural networks. *International Journal of Machine Tools and Manufacture, 45*, 467–479.

Palanisamy, P., Rajendran, I., & Shanmugasundaram, S. (2008). Prediction of tool wear using regression and ANN models in end-milling operation. *International Journal of Advanced Manufacturing Technology, 37*, 29–41.

Pontes, F. J., Ferreira, J. R., Silva, M. B., Paiva, A. P., & Balestrassi, P. P. (2010). Artificial neural networks for machining process surface roughness modelling. *The International Journal of Advanced Manufacturing Technology, 49*, 879–902.

Quiza, R., Figueira, L., & Davim, J. P. (2008). Comparing statistical models and artificial neural networks on predicting the tool wear in hard machining D2 AISI steel. *International Journal of Advanced Manufacturing Technology, 37*, 641–648.

Ross, P. (1991). *Tachuchi techniques for quality engineering*. New York: McGraw Hill.

Sandvik Coromant (2010). *Cutting tools-Main catalogue*. Sandviken: Elanders.

Sharma, V. S., Dhiman, S., Sehgal, R., & Sharma, S. K. (2008). Estimation of cutting forces and surface roughness for hard turning using neural networks. *Journal of Intelligent Manufacturing, 19*, 473–483.

Shie, J. R. (2008). Optimization of injection molding process for contour distortions of polypropylene composite components by a radial basis neural network. *International Journal of Advanced Manufacturing Technology, 36*, 1091–1103.

Sick, B. (2002). On-line and indirect tool wear monitoring in turning with artificial neural networks: A review of more than a decade of research. *Mechanical Systems and Signal Processing, 16*, 487–546.

Singh, D., & Rao, P. V. (2007). A surface roughness model for hard turning process. *International Journal of Advanced Manufacturing Technology, 32*, 1115–1124.

Sonar, D. K., Dixit, U. S., & Ohja, D. K. (2006). The application of a radial basis function neural network for predicting the surface roughness in a turning process. *International Journal of Advanced Manufacture Technology, 27*, 661–666.

Tamizharasan, T., Sevaraj, T., & Haq, A. N. (2006). Analysis of tool wear and surface finish in hard turning. *International Journal of Advanced Manufacturing Technology, 28*, 671–679.

Zain, A. M., Haron, H., & Sharif, S. (2010). Prediction of surface roughness in the end milling machining using artificial neural network. *Expert Systems with Applications, 37*, 1755–1768.

Zhong, Z. W., Khoo, L. P., & Han, S. T. (2006). Prediction of surface roughness of turned surfaces using neural networks. *International Journal of Advanced Manufacturing Technology, 28*, 688–693.