Journal of ELECTRICAL ENGINEERING, VOL. 53, NO. 1-2, 2002, 3–9

USING OF NEURAL NETWORKS FOR CLASSIFICATION AND DATA ANALYSIS IN EDDY–CURRENT TESTING

Ján Grman — Rudolf Ravas — Lívia Syrová *

At present a very prospective solution of indications classification in defectoscopy is neural network application. One of the fields is classification of indications into classes that are characterized by the signal shape or by the signatures related to the signal shape. Nondestructive defectoscopy of steam generator tubes of nuclear power plants by multifrequency eddy current method is the field in which the use of classifiers based on neural network is very prospective. The contribution concentrates on the choice of suitable neural network structures and of a suitable representation of indications. Selected structures are compared using real records of steam generator tubes and also using artificial defects and imitations of construction elements.

K e y w o r d s: defectoscopy, eddy-currents, neural network, steam generator, cluster analysis

1 INTRODUCTION

Eddy current testing is one of the methods of nondestructive testing. In our case it is testing of heat-exchanger tubing using a differential probe [1]. Tubes are produced from a nonmagnetic material. The shape of the output signal from the probe reflects the properties of the tested material [2]. The fundamental problem is to determine, according to the signal shape, whether there is some defect, structural element, roughness or impurity in the tube (Fig. 1). Potential locations of the defect in the signal are called indications.

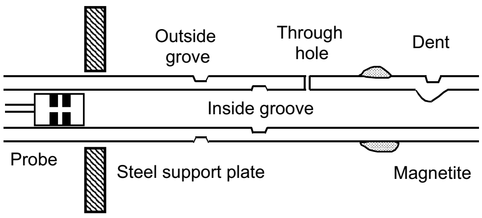

Fig. 1. Defect types

The measurement is generally made at four frequencies. Low frequencies are more suitable for evaluation of identifiable anomalies at the outside wall of the tube, *eg* of the presence of construction elements of the steam generator. At high frequencies there are dominating such signal components that correspond to changes of the internal tube wall, *eg* changes of the tube profile. Middle frequencies are chosen to obtain phase separation of 2D curves (Fig. 2) corresponding to defects that are important for quantification of percentual material drop.

The process of assessment can be separated into several phases, as: signal unification, signal filtration, classification of indication and quantification of defects. The contribution treats the topic of classification of indications into defined classes. Classification is based on the identification of the shape of 2D curves by neural network. From this point of view the choice of suitable descriptors of 2D curves plays an important role. Techniques based on Fourier transformation and Fourier descriptors belong to frequently used methods. It is necessary to transform the obtained parameters in order to obtain descriptors that are invariant to the shift, rotation, scale and to the choice of boundary values at numerical calculation of parameters.

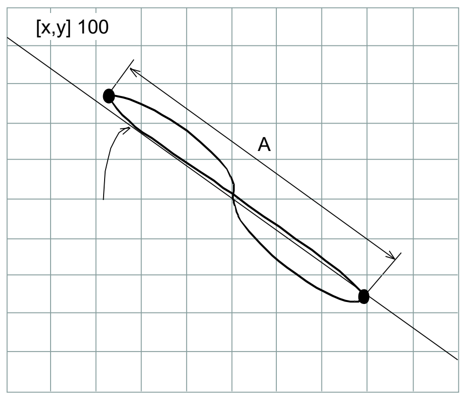

Fig. 2. Parameters of curve corresponding to 100 % defect

2 DATA REPRESENTATIONS

The choice of a suitable representation of the indication data is very important. Every representation has some advantages and disadvantages. The contribution concentrates on a comparison of data representations to be used as input of neural classifier with respect to the

* Department of Measurements, Faculty of Electrical Engineering and Information Technology STU, Ilkovičova 3, 812 19 Bratislava 1, Slovakia e-mail: Rudolf.Ravas@elf.stuba.sk

ISSN 1335-3632

measurement frequency of representation original data. Representation convenience is compared in terms of data input space separability into disjunctive clusters.

2.1 Fourier transformation

One of the most popular methods used in signal processing is the well-known Fourier transformation [3]. By this transformation we are able to achieve the frequency domain of the tested indication. The indication can be then represented by a vector of the most significant Fourier coefficients. This transformation can reduce the dimension of indication data and it is possible to reconstruct the original data.

Let μ be a closed curve (in our case curve of indication) with parametrized representation

$$f(t) = x(t) + i \cdot y(t)\,, \text{where } t \in \langle \alpha, \alpha + \tau \rangle\,.$$

Function f is closed and periodical with periodicity τ. Let $f \in C^{\tau-1}_{\langle \alpha, \alpha+\tau \rangle}$ and on interval $\langle \alpha, \alpha + \tau \rangle$ be a defined system of orthonormal functions:

$$\left\{ h_k(t) = \frac{1}{\sqrt{\tau}} \cdot e^{i \frac{2\pi k t}{\tau}} \mid k \in Z \right\} \tag{1}$$

$$\text{Let } \omega = \frac{2\pi}{\tau}\,, \quad g_k(t) = f(t) \cdot e^{-i\omega t}$$

$$\text{and } a_k = \frac{1}{\sqrt{\tau}} \cdot \int_{\alpha}^{\alpha+\tau} g_k(t) \mathrm{d}t \mid k \in Z\,. \tag{2}$$

For $t \in \langle \alpha, \alpha + \tau \rangle$ we can assume that

$$f(t) = \sum_{k \in Z} a_k h_k(t)\,. \tag{3}$$

To reconstruct the curve we need only a limited amount of Fourier coefficients $(a_{-p}, \dots, a_0, \dots, a_p)$. So we can write

$$f(t) = \sum_{k=-p}^{p} a_k \frac{1}{\sqrt{\tau}} \cdot e^{ik\omega t}\,, t \in \langle \alpha, \alpha + \tau \rangle\,. \tag{4}$$

The fundamental problem is to calculate expression $\int_{\alpha}^{\alpha+\tau} g_k(t)\mathrm{d}t$. There are different ways to approximate this function (step functions, partially linear function, least square method, *etc*).

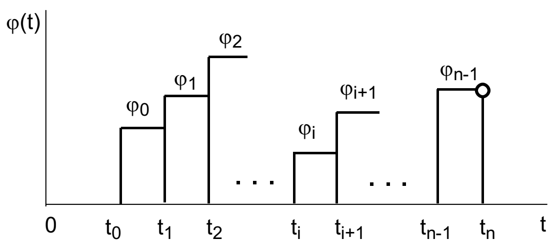

Fig. 3. Example of step function

Using a step function for integral approximation we can modify equations (2) to:

$$a_k = \frac{1}{\sqrt{\tau}} \Bigg(\int_{t=t_0}^{t_1} f_0 e^{-ik\omega t} + \int_{t=t_1}^{t_2} f_1 e^{-ik\omega t} + \dots + \int_{t=t_{n-2}}^{t_{n-1}} f_{n-2} e^{-ik\omega t}$$

$$+ \int_{t=t_{n-1}}^{t_n} f_{n-1} e^{-ik\omega t} \Bigg) = \frac{1}{\sqrt{\tau}} \bigg(\frac{1}{ik\omega} \Big(f_0 e^{-ik\omega t_0} + (f_1 - f_0) e^{-ik\omega t_1}$$

$$+ (f_2 - f_1) e^{-ik\omega t_2} + \dots + (-f_{n-1}) e^{-ik\omega t_n} \Big) \bigg) \tag{5}$$

In addition, let $\Delta_0 = f_0\,,\ \Delta_i = f_i - f_{i-1}\,,\ \Delta_n = -f_n$ so for $\omega = 0$ we can write

$$a_0 = \frac{1}{\sqrt{\tau}} \sum_{j=0}^{n} \Delta_j \tag{6}$$

and for $\omega \neq 0$ we can write

$$a_k = \frac{1}{\sqrt{\tau}} \frac{1}{ik\omega} \sum_{j=0}^{n} \Delta_j e^{-i\omega t_j}\,. \tag{7}$$

Fourier coefficients depend on the following factors: choice of starting point, curve position and rotation with respect to the absolute zero point and scale of the closed curve. For classification such data representations are suitable that are independent of the above mentioned factors. Therefore let us define invariant Fourier based (Eqs. 6 & 7) coefficients. Let

$$b_1 = \frac{|a_1| a_2}{a_1^2} \tag{8}$$

$$b_j = \frac{a_{1+j} a_{1-j}}{a_1^2}\,, \text{ for } j \geq 2\,. \tag{9}$$

Descriptor b_1 depends on the curve rotation, but other descriptors are independent of all mentioned unacceptable factors [3].

2.2 Natural parameters

Of course, we can also use different types of 2D-curve natural parameters [4]. The output signal from the probe consists of two orthogonal parts (also called real and imaginary parts).

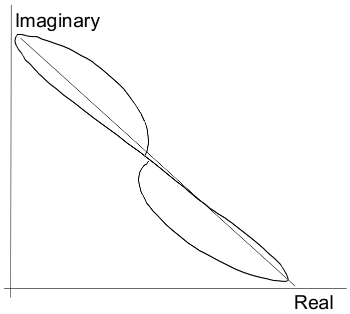

Fig. 4. Example of curve significant distances

The following list contains representations based on mentioned natural parameters:

- R1 — standard deviation of real and imaginary parts of indication signal and their covariance (or correlation) coefficient (3 numbers)

  Let us define the data of indication as

$$f(t) = x(t) + iy(t)\,, \text{ where } t \in \{1, L\} \tag{10}$$

then

$$\begin{aligned} E(x) &= \frac{1}{l}\sum_{k=1}^{L} x(k)\,; \quad E(x) = \frac{1}{l}\sum_{k=1}^{L} y(k) \\ D(x) &= E\left[\left(x - E(x)\right)^2\right] \\ D(y) &= E\left[\left(y - E(y)\right)^2\right] \\ \operatorname{cov}(x,y) &= E\left[\left(x - E(x)\right)\cdot\left(y - E(y)\right)\right] \end{aligned} \tag{11}$$

and the result vector can be defined as

$$\left[\sqrt{D(x)}, \sqrt{D(y)}, \operatorname{cov}(x,y)\right] \tag{12}$$

or

$$\left[\sqrt{D(x)}, \sqrt{D(y)}, \frac{\operatorname{cov}(x,y)}{\sqrt{D(x)}, \sqrt{D(y)}}\right] \tag{13}$$

- R2 — peak-angle-peak representation (3 numbers — peaks angle and their distances in X and Y direction [4])

  Let us define the data of indication using equation (10) and function dist as

$$\begin{aligned} \operatorname{dist}(f,m,n) &= \sqrt{\left(x(m) - x(n)\right)^2 + \left(y(m) - y(n)\right)^2)} \\ \overline{m}, \overline{n} &= \arg\max_{m,n}\left(\operatorname{dist}(f,m,n)\right) \end{aligned} \tag{14}$$

then for peaks distances we can write

$$\begin{aligned} \Delta x\left(\overline{m}, \overline{n}\right) &= \left|x\left(\overline{m}\right) - x\left(\overline{n}\right)\right| \\ \Delta y\left(\overline{m}, \overline{n}\right) &= \left|y\left(\overline{m}\right) - y\left(\overline{n}\right)\right| \end{aligned} \tag{15}$$

and for peaks angle

$$\phi\left(f(t)\right) = \arctan\frac{\Delta y\left(\overline{m}, \overline{n}\right)}{\Delta x\left(\overline{m}, \overline{n}\right)}\,. \tag{16}$$

The result vector can be defined as

$$\left[\Delta x\left(\overline{m}, \overline{n}\right), \Delta y\left(\overline{m}, \overline{n}\right), \phi\left(f(t)\right)\right]\,. \tag{17}$$

Arriving at a conclusion of this section let us define Fourier based representations used in our experiments:

- R3 — representation based on standard Fourier transformation (eqs. 6, 7) defined as a vector (15 complex numbers) of the most significant coefficients
- R4 — representation based on invariant Fourier transformation (eqs. 8, 9) defined as a vector (6 complex numbers) of the most significant coefficients

2.3 Data unification

Before using any of the mentioned representations, unification of indication data is necessary for every measurement. Reference indication (for example 100 % hole in calibration tube) is compared with the reference indication in database of the same type.

Algorithm of unification:

Let

$$f(t) = x(t) + iy(t) \text{ and } \overline{f(t)} = \overline{x(t)} + i\overline{y(t)}$$

be reference indications where $t \in \{1, L\}$ and

$$\overline{\overline{f(t)}} = \overline{\overline{x(t)}} + i\overline{\overline{y(t)}}$$

be unification of indication $\overline{f(t)}$ similar to $f(t)$. Then

a) origin unification

$$\begin{aligned} \overline{\overline{x(t)}} &= \overline{x(t)} - E\left(\overline{x(t)}\right) + E\left(x(t)\right) \\ \overline{\overline{y(t)}} &= \overline{y(t)} - E\left(\overline{y(t)}\right) + E\left(y(t)\right) \end{aligned} \tag{18}$$

b) angle unification

Using eq. (16) we can calculate the angles for both reference indications ($\phi f(t)$) and ($\phi\overline{f(t)}$).

$$\begin{aligned} \psi &= \phi\left(f(t)\right) - \phi\left(\overline{f(t)}\right) \\ \overline{\overline{x(t)}} &= \overline{x(t)}\cos\psi - \overline{y(t)}\sin\psi \\ \overline{\overline{y(t)}} &= \overline{x(t)}\sin\psi + \overline{y(t)}\cos\psi \end{aligned} \tag{19}$$

c) scale unification

Using function dist and eq. (14) we can calculate maximal distance of peaks for both reference indications ($\mathrm{d}f(t)$) and ($\mathrm{d}\overline{f(t)}$).

$$\begin{aligned} r &= \mathrm{d}f(t)/\mathrm{d}\overline{f(t)} \\ \overline{\overline{x(t)}} &= r\cdot\overline{x(t)}\,, \quad \overline{\overline{y(t)}} = r\cdot\overline{y(t)} \end{aligned} \tag{20}$$

The result of this unification is that both reference indications then have a similar angle, scale and origin.

3 NEURAL NETWORK CLASSIFIER

Network topology depends on the choice of input data representation. The dimension of space of indications data is equal to the size of the network input layer. The size of the output network layer depends on the number of tested classes [5].

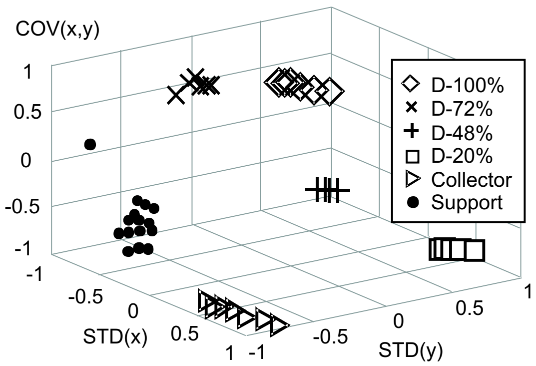

Fig. 5. Standard deviations and covariance 3D graph (R1)

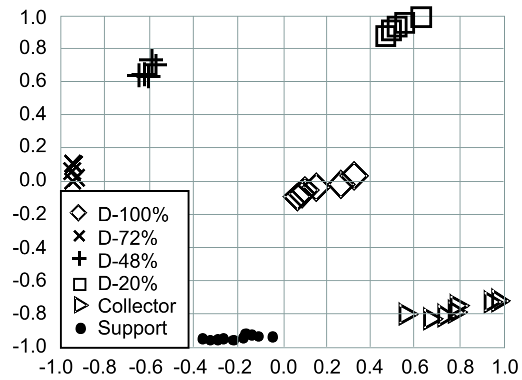

Fig. 6. Standard deviations dependency (R1)

Table 1. Groups of indications in database

group 10 – defect – 100 %
group 11 – defect – 72 %
group 12 – defect – 48 %
group 13 – outside groove – 20 %
group 15 – collector
group 16 – support plate
group 17 – defect 100 % + support edge
group 18 – support plate + defect 100 %
group 19 – defect 48 % + support edge
group 20 – support plate + defect 48 %
group 21 – outside groove 20 % + support plate
group 22 – support plate + outside groove 20 %

3.1 Multilayer perceptron

Neural networks have the ability of generalization and universal approximation as a result of the general approximation theorem. They were successfully used in this type of defectoscopy methods [4, 6, 8]. For classification usually feed-forward supervised NN are used. The most popular is the multilayer perceptron (MP). MP containing one hidden layer is adequate for approximation of any arbitrary continuous function [5]. The input space of signature vectors of indications must be separable into disjunctive subspaces (clusters). Every cluster then contains indications of the same class.

The data of indications are a bit specific. Training of typical feed-forward MP is quite difficult because it is difficult to obtain a big amount of real data of indications. We are forced to use the data of artificial indications.

3.2 Probabilistic neural network

Probabilistic neural networks (PNN) can be used to solve this classification problem. Their design is straightforward and does not depend on training. When an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a complete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a one for that class and a zero for the other classes.

A probabilistic neural network is able to divide the input space to disjunctive subspaces that correspond to the set of required types of indications. A PNN is guaranteed to converge to a Bayesian classifier provided it is given enough training data.

4 EXPERIMENTAL RESULTS

We have built a database of indications. These indications were then transformed by user defined algorithms to the desired representation. Table 1 shows groups of different types of indications that were used in our experiments. The database consists of indications of clear defects as well as of different combinations of indications. Representations were compared using the same subset of database indications and the following figures show selected projections of data clusters for selected representations.

4.1 Cluster analysis

Figures 5 and 6 show that representation R1 gives very compact class clusters. Representation R2 is similar, but gives worse results for some pairs of classes (*eg*, support plate and collector) and mixes of indications. The last tests show that this representation is more sensitive to signal noise (see Section 4.3).

Situation with Fourier based coefficients (R3 & R4) is more complicated. Data analysis shows the existence of class clusters, but these clusters seem to be not very compact and regular. Figure 8 shows a projection of the first coefficient for representation R3 and Figure 9 for R4 (projections for the rest of the coefficients are more confusing, especially for representation R3).

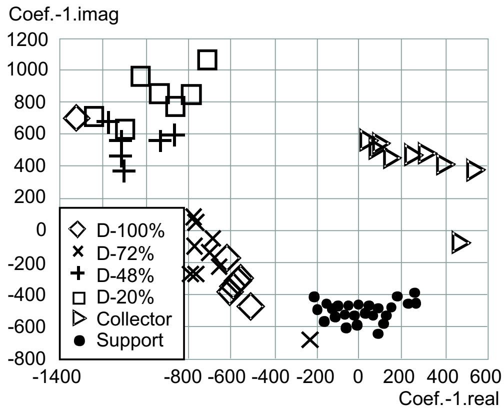

Fig. 7. Peak distances dependency (R2)

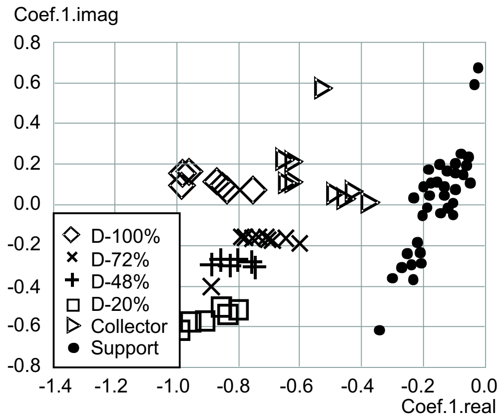

Fig. 8. Complex Fourier coefficient (R3 — $k = -1$)

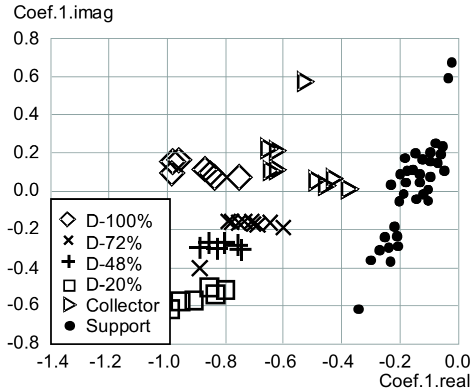

Fig. 9. First invariant Fourier based coefficient (R4)

Table 2. Results of PNN testing (6 classes)

representation training set	all 100/200 kHz	centres 100/200 kHz
R1	100/100%	100/100%
R2	96/99%	98/98%
R3	100/100%	91/89%
R4	100/100%	90/91%

Table 3. Results of PNN testing (all classes)

representation training set	all 100/200 kHz	centres 100/200 kHz
R1	100/98.6%	100/95.6%
R2	89.7/92.1%	94/86%
R3	100/100%	94/93%
R4	100/100%	88/81%

Against all expectations experiments described in the next sections show that the neural classifier (especially PNN) is able to recognize clusters in input data for all representations.

Table 4. Selected results of MP testing

networks\ training set	all 100/200 kHz	centres 100/200 kHz
PNN (R1)	100/100%	100/100%
MP (R1)	100/100%	94.8/97.1%
PNN (R2)	96/99%	98/98%
MP (R2)	91.3/96%	88.4/89.3%

4.2 Classification experiment

The contribution of this section concentrates on a comparison of classification success of MP and PNN on defined data representations. Indications of two frequencies were used in this experiment (100 kHz and 200 kHz). Both networks were trained using data of all indications and then using only calculated centres of data clusters that were found by cluster analysis. Table 2 shows results for PNN testing using indication groups 10, 11, 12, 13, 15 and 16 (see Tab. 1).

Table 3 shows results for PNN testing using all indication groups in database.

These results show that network success depends on representation. Representation R2 is not able to describe some types of indication mixes (for example mix of support plate with small defect). Both tables show that representation R1 gives good results in spite of that the network was trained only by centers of data clusters (small amount of training data). Fourier based representation R3 and R4 give good results only when all indications were used. A PNN is guaranteed to converge to good classification results provided it is given enough training data.

The same tests as for PNN were made using MP. Selected results are in Tab. 4. Results for MP are the same or worse (about 10 %) than results of PNN.

Table 5 shows that the measurement frequency is very important. Table includes information about classification mistakes. Entry 16/15 means that indication of support plate (group 16 — see Tab. 1) was classified as collector.

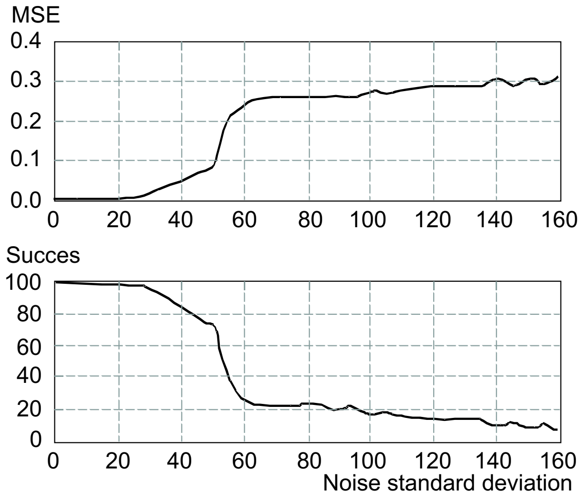

Fig. 10. PNN Error for representation R1

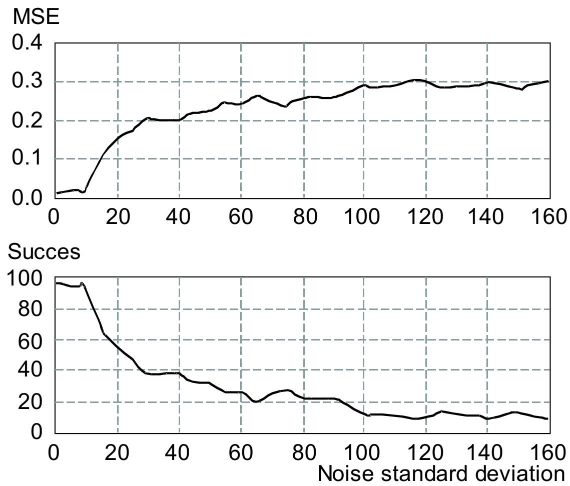

Fig. 11. PNN Error for representation R2

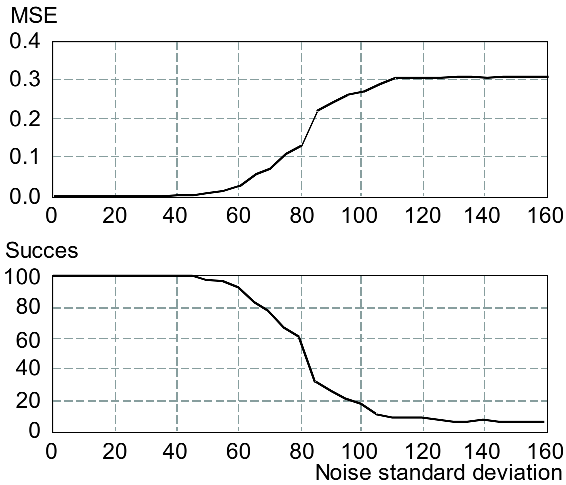

Fig. 12. PNN Error for representation R3 (15 coefficients – 30 real numbers)

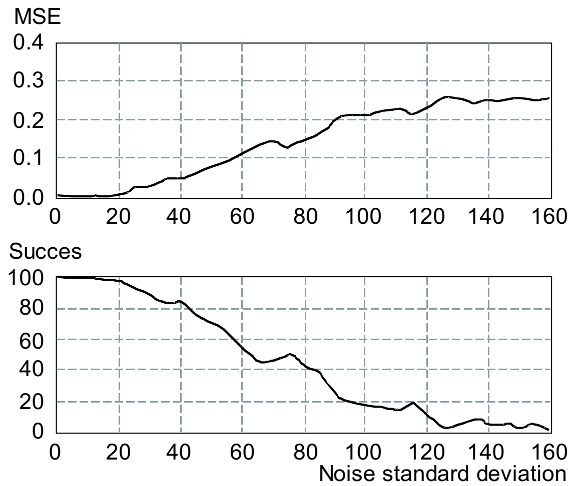

Fig. 13. PNN Error for representation R4 (6 coefficients – 12 real numbers)

Table 5. Results of PNN testing for representation R1 using different frequencies

Freq	all	centres
25 kHz	100 % –	97.3 % 12 / 13; 13 / 21
100 kHz	100 % –	100 % –
200 kHz	98.6 % 17 / 10	95.6 % 10 / 17; 13 / 21
700 kHz	89 % 16 / 15; 12 / 19; 13 / 21	88.7 % 16 / 15 ; 12 / 19; 13 / 21

Experiments show that classification errors depend on the measurement frequency of indications. Low frequencies give good results for support plate and collector classes and high frequencies especially for internal defects. There is a chance that using the results of the classification of all signal frequencies will increase classification success and stability. One of the possible solutions is to define a modular neural network where each module is implemented by a neural classifier. The result of calculation of the major net is a combination of the original input and outputs from individual modules. At the input of modules is the same indication but at different frequency (representation may be different too).

4.3 Representation stability against additive noise

The contribution of this section concentrates on a comparison of defined data representations to be used as an input of the neural classifier implemented by PNN. Stability of representations against noise was compared using the success ratio of PNN classification.

PNN was created by artificial indications without noise and then tested by the same set of indications with added Gaussian white noise in different levels. All representations were tested using indications with added noise whose standard deviation was from the interval (0, 160).

The presented results make us believe that representation R1 and R3 are the best of all tested representations. Representation R1 gives very compact class clusters and gives also good classification results for noisy data of indications. Stability of representation R2 was very influenced by noise. The best results in this test were obtained for representation R3. Successfulness of the final result for Fourier based coefficients depends on the number of coefficients used.

It is interesting that PNN using representation R1 is able to give very good results (good answer in more than 80 % cases) even if noise with standard deviation less than 45 was added. Remember that standard deviation of real [imaginary] parts of indications data without noise are from the interval $(28, 137)$ $[(22, 413)]$.

5 CONCLUSION

In regard to the tested representations, we suppose that representation R1 and Fourier based representations R3 and R4 give good results. The benefit of representation R4 is that it is independent of position, scale and angle (except for it's first coefficient). Representation R2 is very sensitive to noise and we believe that it can be fully substituted by representation R1.

Presented results indicate that PNN gives better results than MP and is easy to use. Training of MP has a lot of difficulties: learning parameters, learning time and especially overfitting. PNN has many advantages but it suffers from one major disadvantage. PNN is slower to operate because it uses more computation than other kinds of networks to do their function approximation or classification.

Nevertheless, the speed is not so important. Important is that PNN does not depend on training. Every measurement probe is a unique device and its output may be a bit different than the output of another probe. The probe time evolution was notified too. So, it is necessary to calibrate the probe before every measurement and because PNN does not need training, there is possibility to save a lot of time.

Finally, let us define the main topics of possible experimental steps:

- representation based on representation R1 invariant to change of measurement probe (there is dependency) = unification of representation
- classification of indication using its representations based on data in different frequencies (to increase classification stability)
- classification of indication using representation based on wavelet data analysis [7, 8].

References

[1] PALANISAMY, R.—LOYD, W.: Finite Element Simulation of Support Plate and Tube Defect Eddy Current Signals in Steam Generator NDT. Materials Evaluation, Vol. 39, June 1981.

[2] HARTANSKÝ, R.—KOVÁČ, K.—HALLON, J.: Manufacturing Imperfection Effect on Resistive Dipole Characteristics, In proc. of 7th International Scientific Conference RADIO-ELEKTRONIKA 97, Bratislava 23.–24. April 1997, pp. 287–290.

[3] YAHN, C. T.—ROSKIES, R. Z.: Fourier Descriptors for Plane Closed Curves, IEEE Transactions on Computers, March 1972.

[4] CHARLTON, C.: Investigation into the Suitability of a Neural Network Classifier for Use in an Automated Tube Inspection System, British Journal of NDT Vol. 35, No. 8, August 1993.

[5] LIMIN, FU: Neural Networks in Computer Intelligence, McGraw-Hill Companies, March 1994.

[6] RAJAGOPALAN, C.—BALDEV, R.—KALYANASUNDARAM, P.: The Role of Artificial Intelligence in Non-Destructive Testing and Evaluation, INSIGHT, Vol. 38, No. 2, February 1996.

[7] STRANG, G.—NGUYEN, T.: Wavelets and Filter Banks, Wellesley-Cambridge Press, 1996.

[8] ALVAREZ HAMELIN, J. I.—D'ATTELIS, C. E.—MENDONCA, H.—RUCH, M.: A New Method for Quasi On-Line Eddy Current Signal Analysis: Wavelet Transform, INSIGHT, Vol. 38, No. 10 (1996), pp. 715–717.

Received 28 February 2001

Ján Grman (Mgr), was born in Topoľčany, Slovakia. He graduated from the Faculty of Mathematics, Physics and Informatics of the Comenius University in Bratislava. At present he is a PhD student at the Department of Measurement of Faculty of Electrical Engineering and Information Technology in Bratislava. The main field of his research are topics of nondestructive defectoscopy by eddy-current techniques.

Rudolf Ravas (Doc, Ing, CSc), was born in Borský Jur, Czechoslovakia, on May 29, 1952. He received the Ing (MSc) degree with honors and the CSc (PhD) degree in electrical engineering from the Faculty of Electrical Engineering of the Slovak Technical University in Bratislava, Czechoslovakia, in 1975 and 1983, respectively. From 1975 till 1978 he was with the Research Institute of Civil Engineering in Bratislava. Since 1978 he has been with Department of Measurement of Faculty of Electrical Engineering and Information Technology, the Slovak Technical University. Current research interests of him are in the fields of theory of measurement and technical diagnostics.

Lívia Syrová (Doc, Ing, CSc), was born in Levice, Czechoslovakia. She graduated from the Faculty of Electrical Engineering of the Slovak Technical University in Bratislava and gained her CSc (PhD) degree in Electrical Engineering. At present she is Associate Professor at the Department of Measurement of Faculty of Electrical Engineering and Information Technology. The main field of her research are optical diagnostic measurements.