HAL
open science

Data Augmentation and Artificial Neural Networks for Eddy Currents Testing

Romain Cormerais, Roberto Longo, Aroune Duclos, Guillaume Wasselynck, Gerard Berthiau

To cite this version:

Romain Cormerais, Roberto Longo, Aroune Duclos, Guillaume Wasselynck, Gerard Berthiau. Data Augmentation and Artificial Neural Networks for Eddy Currents Testing. Electromagnetic Non-Destructive Evaluation (XXIII), 2020, 978-1-64368-118-4. 10.3233/SAEM200040 . hal-02982338

HAL Id: hal-02982338

https://hal.science/hal-02982338v1

Submitted on 3 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Electromagnetic Non-Destructive Evaluation (XXIII)
G.Y. Tian and B. Gao (Eds.)

© 2020 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SAEM200040

Data Augmentation and Artificial Neural Networks for Eddy Currents Testing

Romain CORMERAIS [a,b,c], Roberto LONGO [a,b], Aroune DUCLOS [b], Guillaume WASSELYNCK [c] and Gérard BERTHIAU [c]

[a] *Groupe Signal Image et Instrumentation, ESEO, Angers, France*
[b] *Laboratoire d'Acoustique de l'Université du Mans, Le Mans, France*
[c] *Institut de Recherche en Énergie Électrique de Nantes Atlantique, Saint-Nazaire, France*

Abstract. Eddy Currents (ECs) Non Destructive Testing (NDT) is widely used to determine the position and size of flaws in metal materials. Due to difficulties in estimating these parameters via inverse algorithms based on physical models, approaches focused on Artificial Neural Network (ANN) are nowadays of great interest. The main drawbacks of these techniques still reside in the complexity of the numerical models and the large number of simulated data needed to train and test the ANN, leading to a considerable amount of calculation time and resources. To overcome these limitations, this article proposes a new approach based on a data augmentation procedure via Principal Component Analysis (PCA) applied to numerical simulations.

Keywords. Eddy Currents Testing, Machine Learning, Artificial Neural Networks, Data Augmentation, Principal component analysis

1. Introduction

The goal of Non Destructive Testing (NDT) is to determine the position and size of structural flaws, in order to measure the quality and evaluate the safety of materials [1]. Various techniques, presenting different advantages and drawbacks, are used in NDT, including Ultrasonic Testing (UT), Eddy Currents (ECs), radiography, penetrating testing and thermography [2, 3]. Generally NDT make use of ECs to test metal materials [4–6], but these methods are limited to surface or sub-surface flaws because of the so called "skin effect". Other techniques based on UT are instead affected by the presence of "blind zones" in the ultrasonic beam, which could impair the detection of flaws too close to the surface of the material under test, if not specific coupling is used (i.e. wedges). Hence, this article represents a first step for a combined use of ECs and UT, with the ultimate goal to exploit the complementary between these two approaches. The present work focuses on the ECs part, while the combination with UT is still object of current study.

More in detail, ECs are generated by an electromagnetic field induced into the material under test. Variations on the coil impedance are eventually monitored to detect and characterize defects [7]. Physical models describing ECs can be quite complex, and estimations of flaw parameters like length and depth is not always possible [8]. As alternative, many approaches rely on Machine Learning (ML) methods, employing Artificial

Neural Networks (ANNs) [9–19], as they can perform any function if enough training data are provided [20–22], without requiring a physical model. Moreover, due to the flexibility of these algorithms, ANNs represent an interesting candidate for future integration with UT data. Because building an exhaustive experimental data-set can be difficult, numerical simulations are often used to obtain a training data-set. Among the different possible ways to build a simulated data-set, Finite Element Method (FEM) have been employed here. In detail, the ECs signals, expressed by the coil complex impedance, are derived following the electric vector potential and magnetic scalar potential $T-\Omega$ formulation [23]. This approach, detailed in Section 2.1, focuses on ECs propagation through Aluminum blocks with a large variety of Side Drill Holes (SDHs). Unlikely, the entire procedure needs a significant amount of calculation time and resources.

To overcome this limitation, this article presents a new approach based on data augmentation algorithms via Principal Component Analysis (PCA). The aim is to reduce the number of simulated flaws needed to build the ANN training data-set, significantly reducing computing time and resources. The PCA factorization, presented in Section 2.2, aims to concentrate useful information in a small data-set [24]. The data augmentation procedure based on 3D interpolations over the Principal Components (PCs) is also presented in Section 2.2. The collected original and augmented data-sets are employed to train an ANN whose architecture is presented in Section 2.3. Results are reported in Section 3, showing that data augmentation allows to reduce the simulated training data-set without deteriorating the ANN estimations.

2. Method

2.1. Eddy Currents Simulations in Presence of Side Drill Holes

In this paragraph, the FEM $\vec{T}-\Omega$ formulation [23] is used to obtain signals from homogeneous Aluminum blocks containing SDH. The electrical conductivity and magnetic permeability are respectively set to $\sigma = 37.7$ MS.m^{-1} and $\mu_r = 1$. This formulation allows to compute the induced currents $\vec{J}$, the electric field $\vec{E}$, the magnetic induction $\vec{B}$ and the magnetic field $\vec{H}$. The resistance and reactance of the simulated inductor are respectively calculated through the active and reactive dissipated power, referring to Eq. (1) and Eq. (2). Eventually, the impedance variation of the probe due to the presence of the flaw is calculated. The dimensions of different cylindrical flaws vary with a radius from 0.6 to 8 mm and depth from 0.1 to 4 mm (see **Figure 1**), building a data-set of 3000 different cases. The simulated ECs probe is composed by a 2.25 mm height coil with inner and outer radius of 2.22 and 2.48 mm, driven by a sinusoidal signal at 1 kHz. Air gap is set to 0.18 mm. The probe response is recorded every millimeter along the inspection direction. Examples of simulated signals interacting with SDH from 1 to 5 mm radius at 1 mm depth are depicted in **Figure 2**. A workstation equipped with Intel I7 Processors, NVIDIA Quadro P3000 Graphic Computing Units (GPUs) and 16 Go of RAM has been employed to build the data-set, requiring a calculation time of approximately 2000 hours.

$$RI^2_{source} = \int_{D_c} \vec{E}\cdot\vec{J}dD_c = \int_{D_c} \sigma^{-1}\vec{J}\cdot\vec{J}dD_c, \tag{1}$$

$$XI^2_{source} = \int_{D} \frac{\partial\vec{B}}{\partial t}\cdot\vec{H}dD = w\int_{D} \mu^{-1}\vec{B}\cdot\vec{B}dD. \tag{2}$$

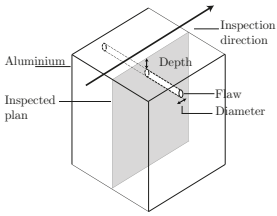

Figure 1. Example of simulated Aluminum block containing a Side Drill Hole (SDH).

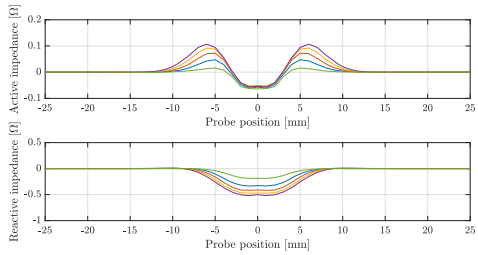

Figure 2. Examples of resistance and reactance obtained simulating five Side Drill Holes (SDHs), with radius from 1 to 5 mm at depth of 1 mm.

2.2. Data Augmentation Based on Principal Component Analysis

This section shows the data augmentation procedure used to process the previous simulated signals. As mentioned in the Introduction, this method is based on Principal Component Analysis (PCA) factorization. In detail, PCA is a statistical method that uses an orthogonal transformation to translate potentially correlated observations into linearly uncorrelated variables called PCs. This transformation is defined in a way that the first PCs present the largest possible variance, and each succeeding component in turn has the highest variance possible in the orthogonal basis [24]. As example, **Figure 3** depicts the radius (top graph) and depth (bottom graph) evolution in function of the two PCs. From this representation, a data augmentation algorithm is developed in order to create additional artificial data at given depth and radius. To start the process, a user-defined depth is selected among the simulated ones. Afterwards, the first PC for a new augmented data point at the desired radius and selected depth is computed with cubic one dimensional interpolation [25]. This interpolation is conducted over the first PCs of the simulated radius at the selected depth. Subsequently, the second PC of the new augmented data point is calculated through linear two dimensional interpolation between the two PCs of the simulated radius at the selected depth and the already interpolated first PC belonging to the new point. Following this procedure, the amount of data composing the augmented data-set can be multiplied by a user-defined factor. As example, **Figure 4** shows the new data-set augmented by a factor of five.

2.3. Artificial Neural Network Architecture and Performance Evaluation

Artificial Neural Networks (ANNs) are proposed in this article to estimate radius and depths from ECs signals, offering an alternative to inverse problems procedures based on physical models. Generally speaking, an ANN is composed by basic interconnected computing units trained to perform a given function. An example is given in **Figure 5**, which shows an ANN composed by N_E inputs, an hidden layer of N neurons and N_S outputs. Each neuron returns the inputs weighted sum $S_v(k)$ or $S_w(k)$, according to the predefined activation functions f_v or f_w, as indicated in Eq. (3). Weights v_{ij} and w_{ij}, firstly randomly initialized, are calculated by minizing an error function defined as the difference between the network outputs S_W and the desired ones T. To do so, a strategy based on error back-propagation has been implemented through the Levenberg-Marquardt algorithm [20, 21, 26]. Eventually, the ANN performance is calculated in function of the

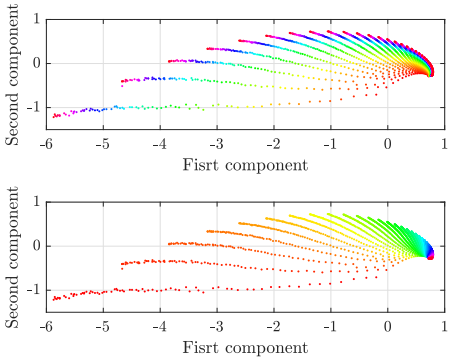

Figure 3. Simulated flaws radius (top graph) and depths (bottom graph) represented in function of the two Principal Components (PCs).

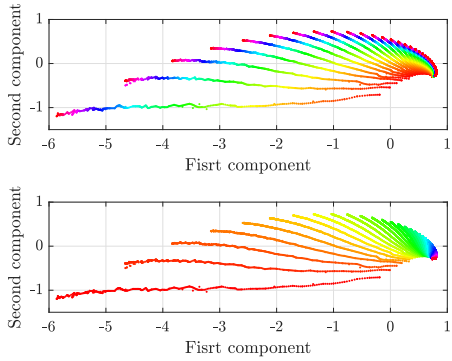

Figure 4. Augmented data for flaws radius (top graph) and depths (bottom graph) represented in function of the two Principal Components (PCs).

Mean Squared Error (MSE) between estimated and real SDHs parameters, following a 10-fold cross validation approach. For that purpose, the simulated data-set has been split into ten groups evenly distributed through the SDHs parameters space, with validation samples evenly distributed inside each training data-set.

In the present study, the ANN is composed by two inputs corresponding to the two first PCs extracted from the simulated signals, as the cumulative sum of eigenvalues shows that they cover 99% of the PCA space information. This choice also reduce the number of weights composing the ANN, and consequently the number of training samples necessary for the optimization process. The ANN inputs are connected to an hidden layer of 20 neurons. This value has been set empirically. The final layer is composed by two outputs neurons leading to radius and depth estimations. The ANN inner and output activation functions, respectively f_v and f_w, are reported in Eq. (4), in order to form a universal function estimator [20, 22].

Eventually, each training data-set of the 10-fold cross validation is progressively decimated until a factor of five in order to evaluate the proposed data augmentation procedure. For each decimation factor, a first ANN is trained with the decimated training data-set and a second one with augmented data. For both, the corresponding test data did not change. As last step, the MSE is monitored in function of the decimation factor. The main results are reported in the next section.

$$\mathbf{S_v}(k) = f_v\left(\sum_{n=1}^{N_E} v_{nk} \times \mathbf{E}(n)\right) \quad \text{and} \quad \mathbf{S_w}(k) = f_w\left(\sum_{m=1}^{N} w_{mk} \times \mathbf{S_v}(m)\right), \tag{3}$$

$$f_v(x) = \frac{2}{1+e^{-2x}} - 1 \quad \text{and} \quad f_w(x) = ax + b. \tag{4}$$

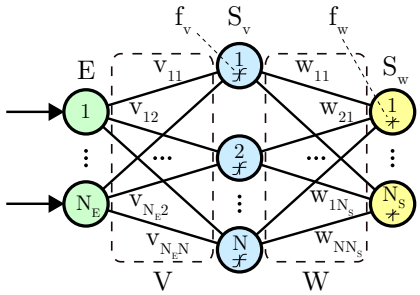

Figure 5. Scheme of an ANN composed by N_E inputs, an hidden layer of N neurons and N_S outputs.

3. Results and Discussion

This section presents the different radius and depth estimations obtained using the proposed data augmentation procedure presented in Section 2.2 applied to the ANNs. In detail, the simulated training data-set is decimated by a factor varying from two to five. Radius and depth estimations accuracy is monitored over the test data-sets through 10-fold cross validation while each training data-set is progressively decimated, as shown in **Figure 6**. The ANN error evolution using the augmented data-set, where decimated data are replaced by augmented ones, is reported in **Figure 7**. Using the decimated data-set, for a decimation factor of five an average error of 50% is recorded over the fifty training. This high error value was expected as the ANN has been trained with a reduced data-set and tested on the original test data. The best performance was reached at the error of 7.5% (see **Figure 6**). Always referring to a decimation factor of five, using the augmented data-set an average error of 3% and minimum error of 2% are recorded (see **Figure 7**). Moreover, the estimations error variance is drastically reduced using the augmented data-set. Eventually, more complex ANN architectures have been tested, without registering significant improvements in the estimations accuracy. To complete this section, it is important to mention that the creation of the augmented data-set employed only 400 hours of calculation time, instead of the 2000 hours needed to build the original one.

4. Conclusion

Artificial Neural Networks (ANNs) are efficient tools largely used for Eddy Currents (ECs) Non Destructive Testing (NDT). One of their drawbacks still resides in the amount of training data needed, often leading to significant calculation time and resources. To overcome this limitation, a data augmentation method based on Principal Component Analysis (PCA) has been presented in this article. The proposed procedure allows to reduce the time needed to build the training data-set by replacing simulated flaws by artificial ones, without deteriorating ANNs performances. Promising results for Side Drill Hole (SDH) radius and depth estimations are compared in terms of Mean Squared Error (MSE). The proposed method also reduces the ANN estimations error variance, facilitating the training procedure. As future goal, the authors count on validating the ANN performances with experimental data. The ultimate goal will be to couple ECs and ultrasonic pulse-echo signals, exploiting the advantages given by the ANNs.

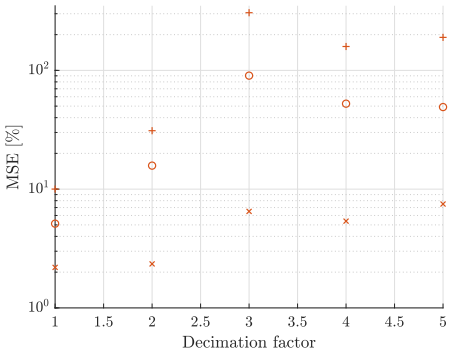

Figure 6. Maximum (+), mean (o) and minimum (x) MSE **obtained decimating the training data-set**. For a decimation factor of five, the radius and depth estimations present an average error of 50%, within a minimum value of 7.5% and a maximum of 190%.

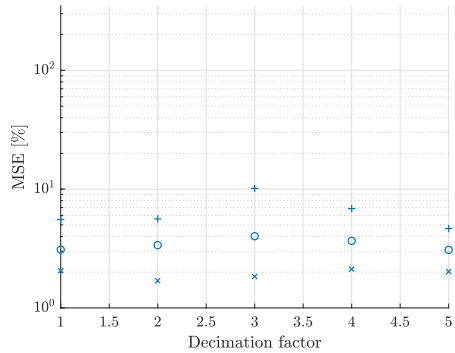

Figure 7. Maximum (+), mean (o) and minimum (x) MSE **obtained using data augmentation**. For a decimation factor of five, the radius and depth estimations present an average error of 3%, within a minimum value of 2% and a maximum of 4.6%.

5. Acknowledgment

The authors would like to thank Matthieu FEUILLOY, Fabien CHHEL and Nicolas GUTOWSKI for their help and constructive comments. This research is funded by RFI-WISE, Pays de Loire, France.

References

[1] Chuck Hellier. *Handbook of Non Destructive Evaluation*. McGraw-Hill Professional, 2001.

[2] Jacques Dumont-Fillon. Contrôle non destructif. *Techniques de l'Ingénieur*, 1(0):1–42, 1996.

[3] R Heideklang and P Shokouhi. Application of data fusion in nondestructive testing (NDT). In *Proceedings of the 16th International Conference on Information Fusion, FUSION 2013*, pages 835–841. IEEE, 2013.

[4] Yunze He, Bin Gao, Ali Sophian, and Ruizhen Yang. *Transient electromagnetic-thermal nondestructive testing*. Elsevier, 1st editio edition, 2017.

[5] Cyril Ravat. *Conception de multicapteurs à courants de Foucault et inversion des signaux associés pour le contrôle non destructif*. phdthesis, Université Paris Sud-Paris XI, 2008.

[6] Harold A Sabbagh, R. Kim Murphy, Elias H. Sabbagh, John C. Aldrin, and Jeremy S Knopp. *Computational Electromagnetics and Model-Based Inversion*. Scientific Computation. Springer New York, 2013.

[7] Jain Nikhil. The rebirth of eddy current nondestructive testing: advanced techniques such as ACFM and Eddy Current Array are making rapid strides. *Quality*, 53(8):S19+, feb 2014.

[8] R L Stoll. *The Analysis of Eddy Currents*. Monographs in electrical and electronic engineering. Clarendon Press, 1974.

[9] L. Udpa and S S Udpa. Application of Neural Networks for Classification of Eddy Current NDT Data. In *Review of Progress in Quantitative Nondestructive Evaluation*, pages 673–680. Springer US, Boston, MA, 1990.

[10] Salvatore Caorsi and Paolo Gamba. Electromagnetic detection of dielectric cylinders by a neural network approach. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2 I):820–827, 1999.

[11] Andrea Bernieri, Luigi Ferrigno, Marco Laracca, and Mario Molinara. Crack shape reconstruction in Eddy current testing using machine learning systems for regression. *IEEE Transactions on Instrumentation and Measurement*, 57(9):1958–1968, 2008.

[12] Y L Bihan, J Pavo, and C Marchand. Characterization of small cracks in eddy current testing. *European Physical JournalApplied Physics*, 43:231–237, 2008.

[13] A. Babaei, A. A. Suratgar, and A. H. Salemi. Dimension estimation of rectangular cracks using impedance changes of the eddy current probe with a neural network. *Journal of Applied Research and Technology*, 11(3):397–401, 2013.

[14] Luis S. Rosado, Fernando M. Janeiro, Pedro M. Ramos, and Moises Piedade. Defect characterization with eddy current testing using nonlinear-regression feature extraction and artificial neural networks. *IEEE Transactions on Instrumentation and Measurement*, 62(5):1207–1214, 2013.

[15] B. Helifa, M. Féliachi, I. K. Lefkaier, F. Boubenider, A. Zaoui, and N. Lagraa. Characterization of surface cracks using eddy current NDT simulation by 3D-FEM and inversion by neural network. *Applied Computational Electromagnetics Society Journal*, 31(2):187–194, 2016.

[16] S. Harzallah, R. Rebhi, M. Chabaat, and A. Rabehi. Eddy current modelling using multi-layer perceptron neural networks for detecting surface cracks. *Frattura ed Integrita Strutturale*, 12(45):147–155, 2018.

[17] Lianshuang Dai, Hao Feng, Ting Wang, Wenbo Xuan, Ziqian Liang, and Xinqi Yang. Pipe crack recognition based on eddy current NDT and 2D impedance characteristics. *Applied Sciences (Switzerland)*, 9(4), 2019.

[18] Peipei Zhu, Yuhua Cheng, Portia Banerjee, Antonello Tamburrino, and Yiming Deng. A novel machine learning model for eddy current testing with uncertainty. *NDT and E International*, 101(October):104–112, 2019.

[19] Jinhyun Park, Seong Jin Han, Nauman Munir, Yun Taek Yeom, Sung Jin Song, Hak Joon Kim, and Se Gon Kwon. MRPC eddy current flaw classification in tubes using deep neural networks. *Nuclear Engineering and Technology*, 51(7):1784–1790, 2019.

[20] Yann LeCun. Connectionist Learning Models. *Intellectica, special issue apprentissage et machine*, 1987.

[21] Martin T. Hagan, Howard B. Demuth, and Mark Hudson Beale. Neural Network Design. *Boston Massachusetts PWS*, 2:734, 1995.

[22] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006.

[23] Thomas Henneron. *Contribution à la prise en compte des Grandeurs Globales dans les Problèmes d'Electromagnétisme résolus avec la Méthode des Eléments Finis*. phdthesis, Université Lille1 - Sciences et Technologies, 2004.

[24] I T Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer-Verlag, New York, 2002.

[25] Hiroshi Akima. A New Method of Interpolation and Smooth Curve Fitting Based on Local Procedures. *Journal of the ACM (JACM)*, 17(4):589–602, 1970.

[26] Martin T. Hagan and Mohammad B. Menhaj. Training Feedforward Networks with the Marquardt Algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993, 1994.