

Principal component analysis of low-frequency impedance spectra for hardness estimation in ferromagnetic materials

Krzysztof Fryczowski^{1*}, Maciej Roskosz², Grzegorz Psuj³,
Paweł Lonkwić⁴, Michał Maciusowicz³

¹ Department of Power Engineering and Turbomachinery, Faculty of Energy and Environmental Engineering, Silesian University of Technology, ul. Akademicka 2A, 44-100 Gliwice, Poland

² Department of Machinery Engineering and Transport, Faculty of Mechanical Engineering and Robotics, AGH University of Krakow, Al. A. Mickiewicza 30, 30-059 Krakow, Poland

³ Center of Electromagnetic Fields Engineering and High-Frequency Techniques, Faculty of Electrical Engineering, West Pomeranian University of Technology: Szczecin, ul. Sikorskiego 37, 70-313 Szczecin, Poland

⁴ The Institute of Technical Sciences and Aviation, The University College of Applied Sciences in Chełm, ul. Pocztowa 54, 22-100 Chełm, Poland

* Corresponding author's e-mail: kfryczowski@polsl.pl

ABSTRACT

This study presents a novel approach for the non-destructive evaluation of hardness in S235JR ferromagnetic steel using low-frequency impedance spectroscopy combined with advanced multivariate statistical analysis. The primary objective was to develop a correlation model enabling estimation of Vickers hardness (HV1) directly from complete electromagnetic response spectra without the need for equivalent circuit parameter identification. The experimental procedure included measurements of impedance components, namely resistance (R) and inductance (L), over a frequency range from 20 Hz to 200 kHz for specimens exhibiting various levels of strain hardening induced by plastic deformation. To reduce the dimensionality of the measurement dataset, consisting of 66 diagnostic parameters for each measurement point, principal component analysis (PCA) was employed. The results demonstrated that the first four principal components (PC1 – PC4) accounted for more than 95% of the total variance of the electromagnetic response. While individual principal components exhibited limited predictive capability, a significant synergistic effect was observed when PC1 and PC2 were combined within a multiple linear regression (MLR) model. The resulting model achieved a coefficient of determination of $R^2 = 0.855$. The reliability of the proposed methodology was verified using an independent blind-test validation dataset, confirming the robustness and repeatability of the developed approach. The obtained results indicate that low-frequency impedance spectroscopy combined with PCA constitutes a promising tool for the non-destructive assessment of mechanical properties and structural condition monitoring of ferromagnetic construction steels.

Keywords: non-destructive testing, impedance spectroscopy, principal component analysis, ferromagnetic materials, Vickers hardness estimation, S235JR steel.

INTRODUCTION

The non-destructive evaluation of mechanical properties of ferromagnetic materials remains one of the most important challenges in modern materials diagnostics [1]. Among

these properties, hardness is of particular interest because it reflects the microstructural condition of a material and is closely related to heat treatment processes, plastic deformation, and service-induced degradation. Conventional hardness measurement methods, such as Brinell,

Vickers, and Rockwell tests, provide high accuracy but are considered partially destructive and are often unsuitable for in-service structures or high-value components.

Consequently, extensive research has been devoted to the development of indirect non-destructive methods for evaluating mechanical properties [2, 3]. In recent years, electromagnetic non-destructive testing techniques have attracted significant attention owing to the strong relationship between the mechanical and electromagnetic properties of ferromagnetic materials [3–8].

Microstructural alterations induced by manufacturing processes and in-service loading simultaneously affect the mechanical, electrical, and magnetic properties of ferromagnetic materials [9, 10]. In particular, changes in dislocation density and internal stress state modify the magnetic domain structure and the dynamics of domain-wall motion, resulting in measurable variations in magnetic permeability, magnetic susceptibility, coercivity, magnetostriction, and electrical conductivity [5, 11].

Modern magneto-mechanical models indicate that the electromagnetic response of a ferromagnetic material is strongly influenced by stress and plastic deformation, thereby providing a physical basis for indirect determination of mechanical properties [12–14].

Although these techniques have proven effective in the evaluation of microstructural evolution and residual stress, modern data-acquisition systems now enable the collection of large multidimensional datasets. Consequently, increasing attention is being directed toward data-driven approaches that exploit the full information content of electromagnetic measurements rather than relying on a limited set of predefined diagnostic parameters.

The existing literature on non-destructive hardness assessment is dominated by studies based on micromagnetic techniques, including Barkhausen noise analysis, magnetic permeability measurements, and eddy-current testing [5, 6, 15].

These methods have demonstrated high effectiveness in the evaluation of microstructural changes and residual stresses using both single-parameter [16] and multi-parameter approaches [12–14]. However, rapid advances in data acquisition and computational techniques have increasingly shifted research interest toward the analysis of complete measurement datasets. Such an approach enables the extraction of substantially more information from the electromagnetic response of a

material than can be obtained from a limited set of predefined diagnostic parameters [17].

Impedance spectroscopy of ferromagnetic materials is a particularly attractive diagnostic technique because it provides access to rich frequency-dependent information about the magnetic condition of a material. The method is based on analysing variations in complex impedance or inductance as functions of excitation frequency and magnetic field. Since complex magnetic permeability is directly related to complex inductance, the resulting spectra contain information about the magnetization processes occurring over different frequency ranges [18, 19]. In particular, the low-frequency response is strongly influenced by domain-wall motion, whereas the contribution of skin effects and magnetization rotation becomes increasingly significant with increasing frequency [18].

Research on magnetoimpedance phenomena has shown that impedance characteristics are highly sensitive to microstructural condition, magnetic anisotropy, mechanical stress, and material geometry [19–22]. In recent years, several advanced approaches have been proposed, including the reconstruction of ferromagnetic hysteresis loops from excitation-coil impedance measurements [20], magnetoimpedance tomography of magnetic microwires [21], and investigations of magnetoinductive effects in thin ferromagnetic films [22]. These studies collectively demonstrate that frequency-dependent impedance responses constitute a rich source of information about the structural condition and mechanical properties of ferromagnetic materials.

Traditionally, impedance spectra have been interpreted using equivalent circuit models, in which the measured response is represented by a set of electrical elements chosen to approximate the underlying physical processes [23, 24]. In ferromagnetic materials, however, the selection of an appropriate equivalent circuit is often ambiguous because multiple overlapping physical mechanisms contribute simultaneously to the measured response. Consequently, increasing attention has been paid to data-driven approaches that enable direct analysis of complete experimental datasets, eliminating the need to identify equivalent circuit parameters prior to feature extraction [24].

One of the most widely used techniques for dimensionality reduction is PCA [25–27]. PCA transforms a set of correlated variables into a new orthogonal feature space composed of

principal components, each representing a specific direction of variability within the data. The first principal component is defined to capture the largest possible proportion of the total variance, while subsequent components describe progressively smaller sources of variability. In impedance spectrum analysis, PCA enables the identification of dominant patterns in the measured data, reduces the influence of measurement noise, and facilitates the extraction of diagnostic features for classification and regression tasks [28, 29]. Owing to these capabilities, PCA has increasingly been integrated with machine-learning and artificial-intelligence methods to improve the performance of modern diagnostic systems [30].

A review of the current literature reveals a distinct research gap. Existing studies on hardness estimation in ferromagnetic materials have focused primarily on micromagnetic and eddy-current-based techniques [5, 6, 15], whereas the analysis of complete low-frequency impedance spectra remains relatively underexplored. At the same time, advances in inductive and magneto-impedance spectroscopy have demonstrated that frequency-dependent electromagnetic responses contain rich information about magnetization processes and microstructural changes in ferromagnetic materials [19–22]. Despite the growing popularity of multivariate data-analysis methods, only a limited number of studies have employed PCA for the direct processing of complete low-frequency impedance spectra for hardness estimation.

In this study, PCA is applied to low-frequency impedance spectra of ferromagnetic materials for hardness estimation. The objective is to determine whether features extracted from impedance measurements can provide a reliable basis for diagnostic model development and to evaluate the applicability of this approach in electromagnetic non-destructive testing. Unlike most previously reported methods, the proposed methodology operates directly on complete low-frequency impedance spectra and does not require the determination of equivalent circuit parameters. Instead, the spectra are treated as multidimensional datasets and analysed using PCA. By exploiting the full information content of the measured electromagnetic response, the proposed approach may enable more effective extraction of diagnostic features and contribute to the development of new methods for

the non-destructive assessment of mechanical properties in ferromagnetic materials. The novelty of the present work lies in the direct application of PCA to complete low-frequency resistance-inductance spectra acquired from plastically deformed S235JR steel, without extraction of equivalent-circuit parameters or predefined diagnostic indicators.

MATERIALS AND METHODS

The study consisted of two main stages:

- measurement of impedance components at 33 excitation frequencies;
- determination of the Vickers hardness distribution.

The obtained data were subsequently used to investigate the relationship between the structural condition of plastically deformed S235JR steel specimens and their electromagnetic response measured at the specimen surface.

Test specimens

The investigations were conducted on specimens manufactured from S235JR non-alloy structural steel in accordance with EN 10025-2. This steel was selected because of its widespread industrial use in welded structures, machine elements, and civil engineering applications. Its low carbon content (not exceeding 0.17%) and good magnetic sensitivity to plastic deformation make it a suitable material for studying relationships between electromagnetic response and mechanical properties.

The main stage of specimen preparation involved introducing different levels of plastic deformation. This was achieved under static tensile loading conditions, which enabled a wide range of hardness values to be obtained. Controlled tensile loading was applied to generate a continuous hardness gradient along the specimen gauge length. As a result, a broad range of hardness values was obtained, enabling investigation of the relationship between electromagnetic response and material condition. The hardness distribution was subsequently used as the reference dataset for PCA and regression model development (presented in Figure 1).

To minimize the influence of the surface layer condition, all specimens were mechanically ground and polished prior to electromagnetic

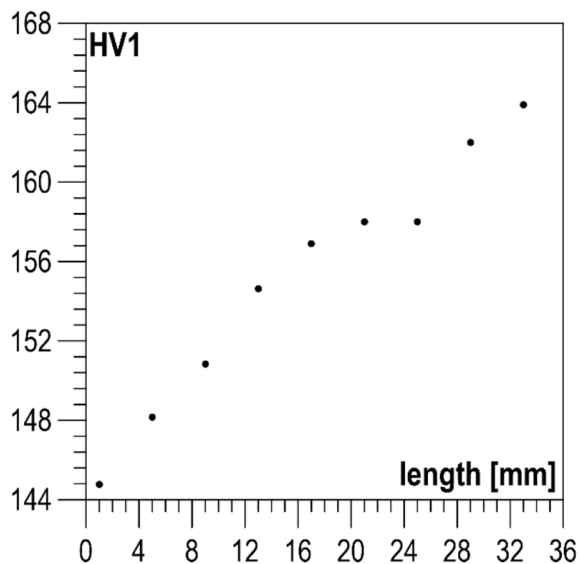


Figure 1. Distribution of Vickers hardness (HV1) along the plastically deformed specimen

measurements. This procedure ensured stable magnetic coupling between the core of the measuring transducer and the tested ferromagnetic material.

The specimen geometry was selected to minimize boundary effects during electromagnetic measurements. Both the width and thickness of the specimens were substantially larger than the effective interaction region of the sensor magnetic field. Owing to the sufficient distance from the lateral edges and the large material volume within the measurement zone, the influence of external geometry on the impedance signal was considered negligible.

This justified approximating the tested material as an infinite half-space, in which variations in the measured coil parameters, resistance R and inductance L , were attributed to local changes in physical properties, primarily magnetic permeability and electrical conductivity, related to the hardness gradient induced by plastic deformation.

The experimental dataset consisted of 5 specimens and 147 measurement points. Measurements were acquired at ~29–30 measurement locations per specimen, resulting in a total of 147 hardness-impedance pairs. The dataset was divided into a calibration dataset (118 observations) and an independent validation dataset (29 observations). Dataset assignment was performed prior to PCA decomposition and regression model development.

The measurement dataset was divided into two independent subsets: a calibration dataset, used for PCA decomposition and regression

model calibration, and a validation dataset, used exclusively for a blind test of the developed method.

The calibration dataset consisted of measurements collected from four specimens, whereas the validation dataset contained all measurements obtained from a fifth specimen that was not used during model development. Consequently, measurements originating from a given specimen were assigned entirely either to the calibration or to the validation dataset. This approach eliminated the possibility of information leakage associated with correlations between neighbouring points located along the same plastic-deformation gradient. The validation data were not included in the determination of the principal component eigenvectors.

Each reference hardness value was assigned a corresponding 66-dimensional electromagnetic feature vector, consisting of resistance (R) and inductance (L) values measured at 33 excitation frequencies.

Impedance and hardness measurements

To ensure repeatable and automated measurement conditions, a dedicated measurement system was developed in LabVIEW 2026. The system integrated Cartesian-robot-based sensor positioning with measurement data acquisition, enabling full automation of the measurement cycle. To minimize the influence of lift-off, a constant sensor-to-surface distance was maintained throughout the measurements. Communication with the HAMEG HM8118 LCR bridge was performed via a digital interface, allowing remote configuration of excitation parameters and automated acquisition of measured values at each point of the specimen.

The frequency-dependent resistance and inductance values were measured using a programmable HM8118 LCR bridge with an upper frequency limit of 200 kHz. Measurements were performed at 33 selected excitation frequencies in the range from 20 Hz to 200 kHz, with the excitation voltage amplitude set to 1.5 V. To ensure consistency of the input data used in the PCA procedure, the instrument was configured to operate in the series equivalent circuit mode. This configuration was selected due to the inductive nature of the sensing element and the need to maintain a uniform interpretation of the measured electrical parameters over the entire investigated frequency range.

An industrial Fastron 09P-332J-50 inductor with a ferrite core and a nominal inductance of 3.3 mH was employed as the measuring transducer (Figure 2). The transducer featured an open magnetic circuit with an 8.5 mm diameter ferrite core, enabling effective magnetic flux coupling with the tested ferromagnetic material.

Reference hardness measurements were performed using a Krautkrämer TIV-101 hardness tester based on the optical Vickers method, in accordance with ISO 6507-1. The measurements were carried out using an indenter load of 9.81 N, corresponding to the HV1 scale.

PCA-based development of the diagnostic correlation model

For PCA-based development of the diagnostic correlation model, each reference HV1 value was assigned a corresponding 66-dimensional electromagnetic feature vector composed of resistance and inductance L values measured at 33 excitation frequencies. The raw dataset was therefore arranged into a matrix with 66 rows, representing the frequency-dependent R and L

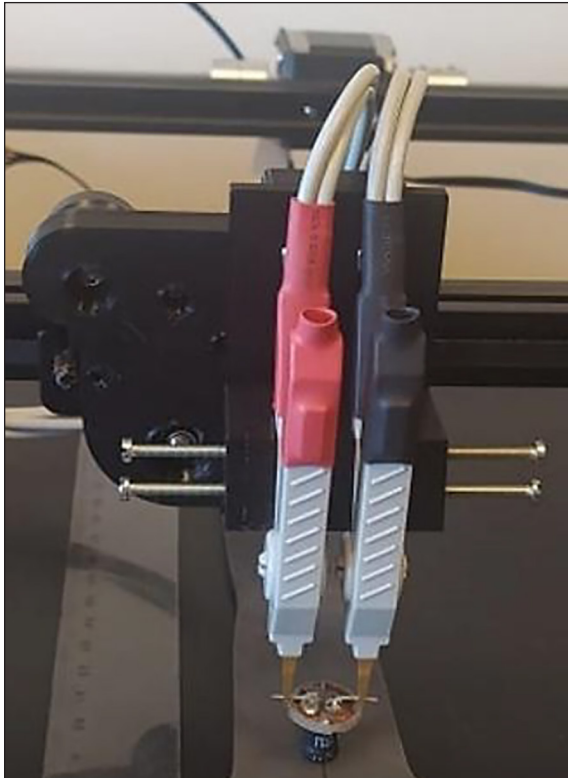


Figure 2. Experimental setup showing the inductive measuring transducer positioned above the investigated ferromagnetic specimen

parameters, and columns corresponding to individual hardness measurement points. The measurement data are presented in Figure 3a–c using a normalized impedance-plane representation, where the abscissa is defined as $R_{norm} = \frac{R-R_0}{\omega L}$, and the ordinate as $L_{norm} = \frac{\omega L}{\omega L_0}$.

The impedance characteristics presented in Figure 3 indicate that the normalized inductive component decreases with increasing excitation frequency and asymptotically approaches unity. This behavior corresponds to the coil response in air, where magnetic coupling with the tested material is negligible. In contrast, the normalized resistive component initially increases with frequency and then decreases after a characteristic frequency range is exceeded. It was also observed that the dispersion of resistance values increases with increasing material hardness.

The relatively narrow variation range of the individual normalized impedance components limits their usefulness as standalone diagnostic indicators. This justifies the use of a multivariate approach capable of extracting information from the full frequency-dependent impedance response. PCA was therefore employed to reduce the dimensionality of the dataset and identify the dominant patterns associated with hardness variation.

Principal Component Analysis transforms the original set of correlated impedance-related features into a new orthogonal feature space composed of mutually uncorrelated principal components. Although the number of components initially equals the number of original variables, dimensionality reduction is achieved by retaining only those components that explain the dominant part of the total variance. The procedure involves standardization of the measurement data, construction of the covariance matrix, and eigenvalue decomposition, which provides the eigenvectors defining the principal component directions and the corresponding eigenvalues representing the variance explained by each component.

In the first step, the raw measurement results were arranged into an input data matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,p} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,p} \end{bmatrix} \quad (1)$$

where: n is the number of diagnostic channels, comprising 33 resistance (R) values and 33 inductance (L) values; thus, $n = 66$,

p is the number of measurement points,
 x_{ij} is the value of the i -th diagnostic channel recorded at the j -th measurement point.

Because resistance (Ω) and inductance (H) values differ by several orders of magnitude, Z-score standardization was performed independently for each diagnostic channel (matrix row). This procedure placed all frequency-dependent measurement channels on a comparable scale during the determination of the principal components.

For each element x_{ij} of the input matrix, the corresponding standardized value z_{ij} was calculated as,

$$z_{i,j} = \frac{x_{i,j} - \mu_i}{\sigma_i} \quad (2)$$

where: μ_i and σ_i denote, respectively, the arithmetic mean and standard deviation of all measurements in the i -th row.

As a result of this operation, each of the 66 diagnostic channels had a mean value of zero and unit variance. The standardized matrix prepared in this way was used as the direct input to the PCA algorithm, with variables arranged in rows.

MATHEMATICAL PROCEDURE OF PCA DECOMPOSITION

After data standardization, PCA proceeds to the key stage of identifying new coordinate axes

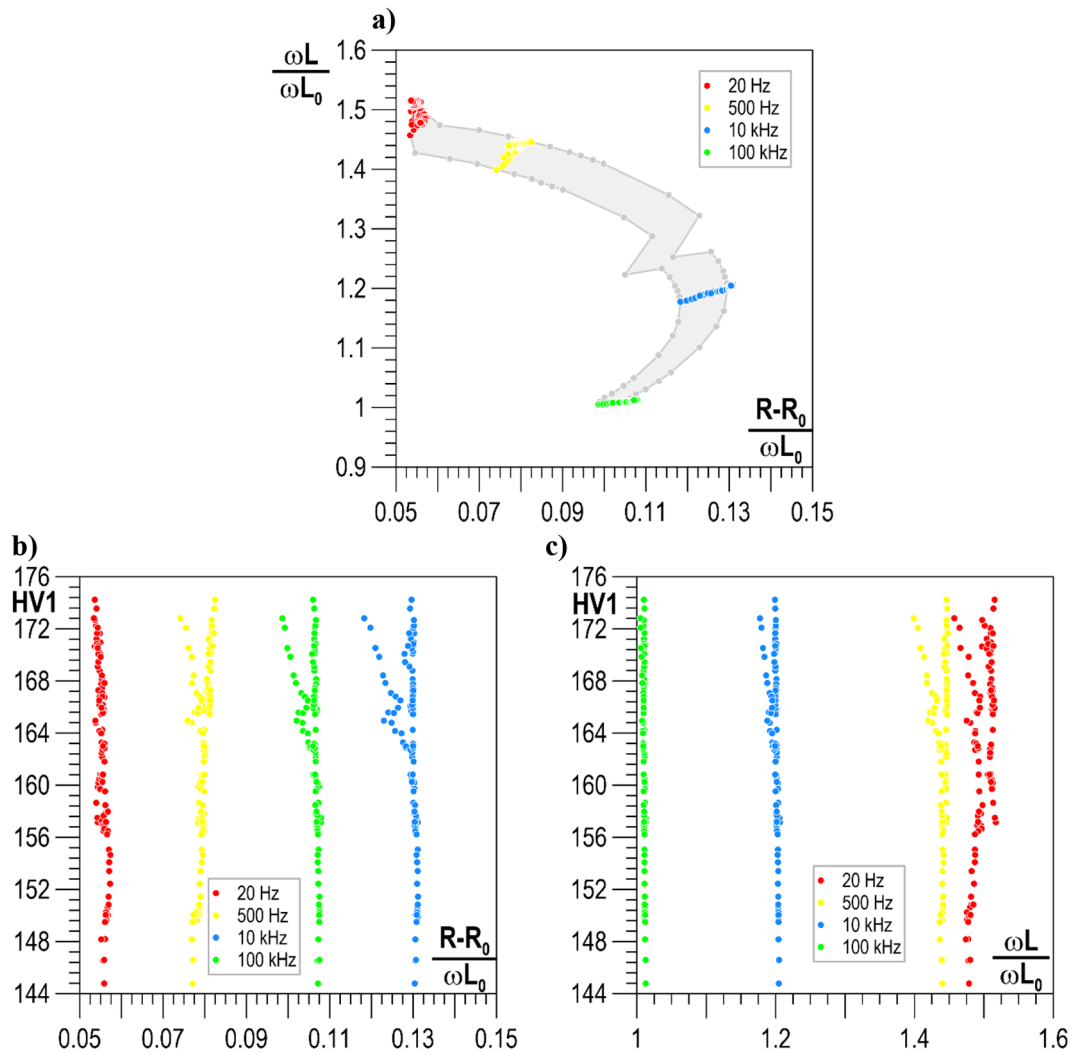


Figure 3. Frequency-dependent impedance response and its relationship with reference HV1:
 a) normalized impedance plane showing the family of measurement curves across the investigated frequency range, b) Relationship between reference HV1 and normalized resistance R_{norm} for selected excitation frequencies: 20 Hz, 500 Hz, 10 kHz, and 100 kHz, c) Relationship between reference HV1 and normalized inductance L_{norm} for selected excitation frequencies: 20 Hz, 500 Hz, 10 kHz, and 100 kHz

that best describe the variability of the electromagnetic response. This procedure consists of the following steps:

Calculation of the covariance matrix $\mathbf{C}$

First, the covariance matrix $\mathbf{C} \in \mathbb{R}^{n \times n}$ was calculated, where n corresponds to the number of diagnostic channels, i.e., resistance R and inductance L values measured at the selected excitation frequencies. The covariance matrix describes the mutual relationships between all measurement channels. Each element of this matrix indicates how variations in the impedance response at one frequency co-vary with variations observed at the remaining frequencies.

Determination of eigenvalues

To determine the eigenvalues λ_i , the characteristic equation of the covariance matrix was solved. This yielded a set of n eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$, which can be represented as a vector of dimensions $n \times 1$. Each eigenvalue λ_i represents the amount of variance explained by the corresponding principal component. The largest eigenvalues indicate the components that contain the dominant information within the dataset. In the present analysis, the first four components, PC1–PC4, were associated with the highest eigenvalues and therefore retained most of the variability of the electromagnetic response.

Determination of eigenvectors

For each eigenvalue λ_i , the corresponding eigenvector $\mathbf{v}_i$ was calculated. For n diagnostic channels, each eigenvector $\mathbf{v}_i$ consists of n coefficients that define the contribution of individual measurement channels, for example resistance R measured at 100 Hz, to the corresponding principal component. These eigenvectors define the directions of the principal components in the original feature space and are subsequently used to project measurement data into the PCA space.

Calculation of principal component scores

The principal component scores were obtained by projecting the standardized data matrix onto the eigenvectors determined in the previous step. If the eigenvectors are arranged as columns of the matrix $\mathbf{V}$, the score matrix $\mathbf{T}$ is calculated as:

$$\mathbf{T} = \mathbf{V}^T \mathbf{Z} \quad (3)$$

where: $\mathbf{Z}$ is the standardized data matrix, $\mathbf{V}$ is the matrix of eigenvectors, and $\mathbf{T}$ contains the principal component scores for individual measurement points.

For the j -th measurement point, the score of the k -th principal component ($k = 1, 2, \dots, n$) was obtained as:

$$PC_{k,j} = \sum_{i=1}^n z_{i,j} w_{k,i} \quad (4)$$

where: $z_{i,j}$ denotes the standardized value of the i -th diagnostic channel at the j -th measurement point, and $w_{k,i}$ is the loading coefficient, i.e., the i -th element of the eigenvector defining the k -th principal component.

The level of dimensionality reduction was determined by ranking the eigenvalues in descending order and analysing the cumulative explained variance. The number of retained components was selected so as to preserve a predefined proportion of the total variance. In this study, the first four principal components, PC1–PC4, were sufficient to explain 95% of the cumulative variance as presented in Figure 4.

To facilitate interpretation of the PCA results, the loading distributions of the first two principal components were analysed. Figure 5 presents the loading coefficients obtained for resistance (R) and inductance (L) channels measured at individual excitation frequencies.

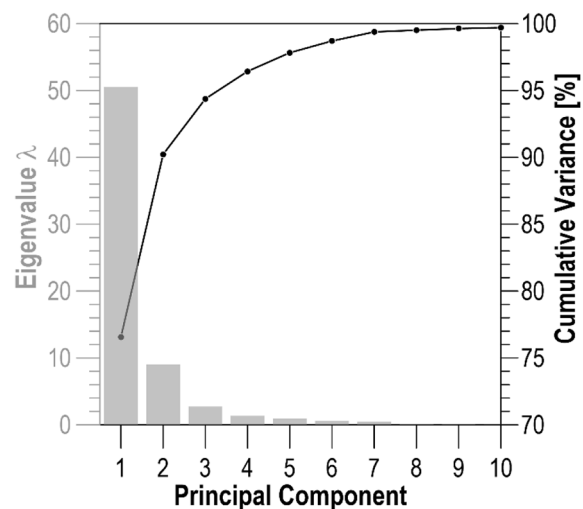


Figure 4. Pareto chart of PCA eigenvalues and cumulative explained variance

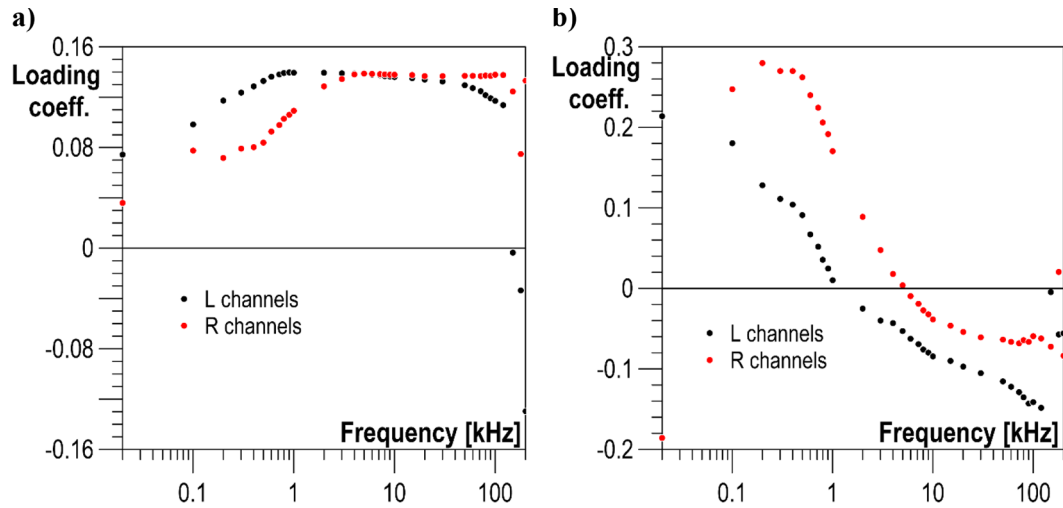


Figure 5. PCA loading distributions for the first two principal components obtained from the impedance-spectroscopy dataset: (a) PC1 loadings and (b) PC2 loadings calculated for resistance (R) and inductance (L) measurement channels over the investigated frequency range

The first principal component (PC1) is characterised by predominantly positive loading coefficients distributed over a broad frequency range for both resistance and inductance channels. This indicates that PC1 captures variations affecting a large portion of the measured impedance spectrum simultaneously.

In contrast, the second principal component (PC2) exhibits a much stronger dependence on excitation frequency, including substantial variations of loading magnitude and sign. This behaviour suggests that PC2 captures frequency-dependent spectral variations that are complementary to those represented by PC1.

Development of the diagnostic model

The diagnostic correlation model was developed using multiple linear regression (MLR). In this approach, the reference HV1 values were used as the dependent variable, while selected principal component scores served as predictors. The predictor matrix was constructed from the principal component scores obtained during the PCA stage.

For p measurement points, the general matrix form of the model can be written as:

$$HV = X_{PC}A + E \quad (5)$$

where: HV is the vector of reference hardness values, X_{PC} is the predictor matrix composed of selected principal component scores, A is the vector of regression coefficients, and E is the vector of residual errors.

For a model based on q selected principal components, the full matrix notation is given by:

$$\begin{bmatrix} HV_1 \\ HV_2 \\ \vdots \\ HV_p \end{bmatrix} = \begin{bmatrix} 1 & PC_{1,1} & PC_{2,1} & \cdots & PC_{q,1} \\ 1 & PC_{1,2} & PC_{2,2} & \cdots & PC_{q,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & PC_{1,p} & PC_{2,p} & \cdots & PC_{q,p} \end{bmatrix} \quad (6)$$

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_q \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_p \end{bmatrix}$$

In this notation, $X_{PC} \in \mathbb{R}^{p \times (q+1)}$, where the first column consists of ones corresponding to the intercept term, and the remaining q columns contain the scores of the selected principal components.

Solving this system using the least-squares method enabled the determination of the optimal regression coefficients $a_0, a_1, \dots, a_q$. These coefficients define the contribution of the selected principal components to the final HV1 estimate.

For the i -th measurement point, the regression model can therefore be expressed as:

$$HV_i = a_0 + \sum_{k=1}^q a_k PC_{k,i} + e_i \quad (7)$$

where: HV_i is the reference Vickers hardness measured at the i -th point, $PC_{k,i}$, for $k = 1, 2, \dots, q$, denotes the score of the k -th principal component calculated for the i -th measurement point, q is the number

of principal components included in the regression model, a_0 is the intercept of the regression model, a_k is the regression coefficient assigned to the k -th principal component, e_i is the residual error for the i -th measurement point.

Different combinations of principal components were subsequently tested to identify the predictor set providing the best agreement between estimated and reference HV1 values.

To systematize the mathematical notation, Table 1 summarizes the evolution of data dimensions and the symbols assigned to each processing stage of the PCA algorithm.

After completion of the calibration stage, i.e., after developing the diagnostic model, the hardness of an unknown measurement point can be estimated from the measured n -channel impedance response. For each measurement channel,

the newly acquired value is first standardized using the mean μ_i and standard deviation σ_i determined during calibration. The standardized values are then projected onto the retained principal component directions to obtain the corresponding principal component scores.

For the selected principal component scores, the estimated hardness value is calculated from the regression model:

$$\widehat{HV} = a_0 + \sum_{k=1}^q a_k PC_k \quad (8)$$

where: q denotes the number of principal components included in the regression model.

Although PC1 captures the largest proportion of total variance, its individual correlation with hardness is relatively weak. In contrast, PC2 exhibits a substantially stronger relationship

Table 1. Mathematical notation and matrix dimensions at successive stages of PCA-based data processing, assuming diagnostic channels arranged in rows

Algorithm stage	Data type	Symbol	Matrix dimensions
Raw data	Impedance matrix	X	$n \times p$
Standardization	Z-score matrix	Z	$n \times p$
Decomposition	Covariance matrix	C	$n \times n$
Principal component directions	Eigenvector matrix	V	$n \times n$
Variance	Eigenvalue vector	λ	$n \times 1$
PC scores	Principal component score matrix	T	$n \times p$

Table 2. Coefficients of determination obtained for MLR models constructed using different combinations of principal components

Principal components included in the regression model	Coefficient of determination (R^2)
PC1, PC2, PC3, PC4	0.8576
PC1, PC2, PC3	0.8569
PC1, PC2	0.8550
PC1, PC2, PC4	0.8550
PC2, PC3	0.7002
PC2, PC3, PC4	0.6959
PC2	0.6952
PC2, PC4	0.6903
PC1, PC3, PC4	0.1493
PC1, PC3	0.1482
PC1, PC4	0.1446
PC1	0.1435
PC3, PC4	0.0040
PC3	0.0026
PC4	-0.0024

with hardness. The loading analysis indicates that PC1 and PC2 represent different patterns of variability within the impedance spectra. Their combination provides complementary information, resulting in the observed improvement in prediction performance.

To optimize the diagnostic model, the agreement between the estimated hardness values and the reference hardness measurements was analysed for all tested combinations of principal components PC1 – PC4. Table 2 summarizes calibration-stage coefficients of determination obtained for different predictor combinations. The individual PC2 component exhibited a moderate correlation with hardness ($R^2 = 0.6952$), whereas PC1 alone showed low

predictive capability ($R^2 = 0.1435$). A key result was the synergistic effect observed when PC1 and PC2 were combined, which increased the coefficient of determination to ($R^2 = 0.855$). Including higher-order components, namely PC3 and PC4, resulted in only a marginal improvement in model accuracy, below 0.2%. Therefore, the two-predictor model based on PC1 and PC2 was considered the optimal solution. The value $R^2 = 0.855$ corresponds to model calibration and should not be interpreted as an independent validation result.

The graphical comparison between electromagnetically estimated hardness values ($HV1_e$) and reference hardness measurements ($HV1_m$) is presented in Figure 6a–f.

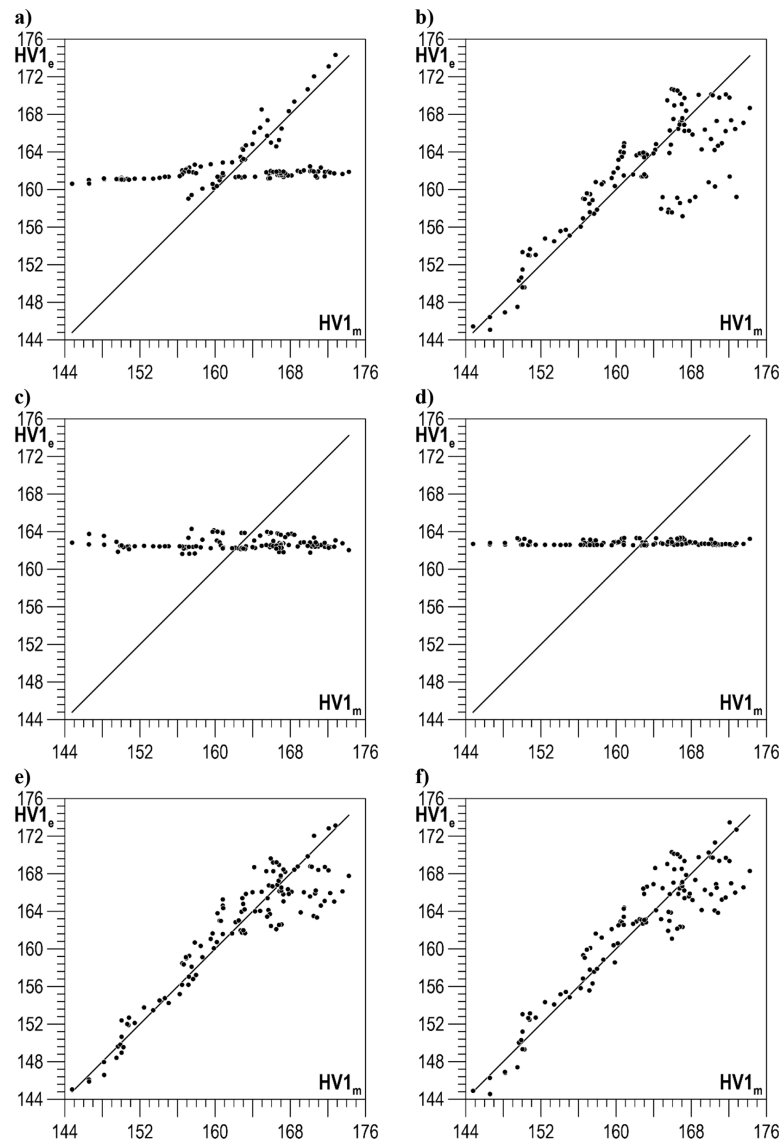


Figure 6. Parity plots comparing the electromagnetically estimated hardness ($HV1_e$) with the reference hardness values ($HV1_m$): a) Single-predictor model (PC1), b) Single-predictor model (PC2), c) Single-predictor model (PC3), d) Single-predictor model (PC4), e) Four-predictor model (PC1+PC2+PC3+PC4), f) Optimal dual-predictor model (PC1+PC2)

The slightly negative R^2 obtained for the PC4-only model indicates that this component does not contain meaningful hardness-related information and yields predictions that are marginally worse than those obtained using the mean hardness value as a predictor.

The predictive performance of the developed model was verified using an independent validation dataset (Figure 7). The obtained results show good agreement between the reference HV1 values and the values estimated from the electromagnetic impedance response. This confirms the stability and repeatability of the proposed non-destructive approach for assessing the hardness of S235JR steel along the plastic deformation gradient.

To provide a quantitative assessment of model performance, additional validation statistics were calculated for both the calibration and validation datasets. The obtained performance metrics are summarized in Table 3. To evaluate the presence of systematic prediction errors, a residual analysis was performed using the independent validation dataset (Figure 8).

The slightly higher validation performance is attributed to the favourable distribution of hardness values in the independent specimen and does not indicate model overfitting.

Residuals were calculated as the difference between electromagnetically estimated hardness (HV_{est}) and reference hardness (HV_{ref}) for the independent validation dataset. The residuals are distributed around zero without a pronounced monotonic trend as a function of hardness, which is consistent with the low mean bias error reported in Table 3. Although several individual observations exhibit larger deviations, no clear evidence of a significant hardness-dependent prediction bias was identified within the investigated hardness range.

The validation results demonstrate that the developed model retains its predictive capability

when applied to measurement data that were not used during PCA decomposition or regression model calibration.

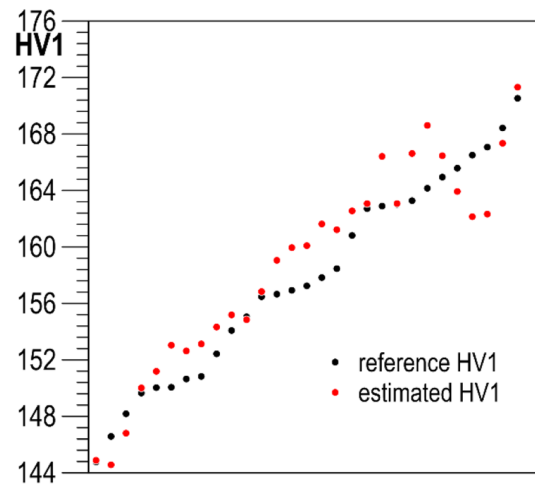


Figure 7. Validation results comparing reference HV1 values with hardness values estimated using the developed PCA-based regression model

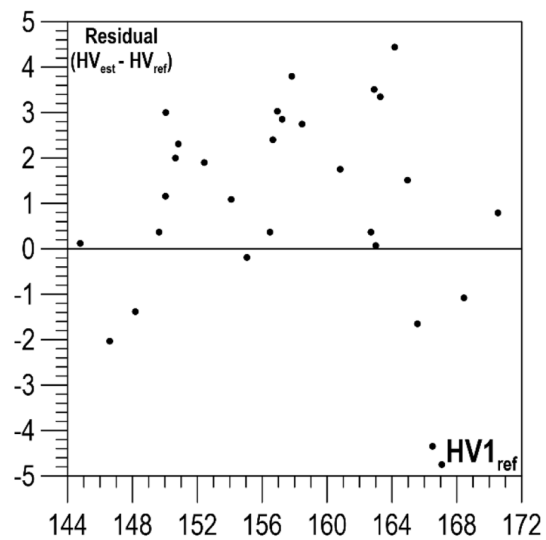


Figure 8. Residual analysis of the final PC1+PC2 regression model

Table 3. Performance of the PCA-MLR hardness estimation model

Metric	Calibration	Validation
Number of observations	118	29
R^2 (ordinary)	0.855	0.881
RMSE (HV1)	2.728	2.426
MAE (HV1)	2.090	2.013
Mean Bias Error, MBE (HV1)	-0.201	+0.949
Standard deviation of residuals (HV1)	2.732	2.272
Maximum absolute error (HV1)	7.013	4.745
95% CI of mean bias (HV1)	(-0.694; 0.292)	(0.122; 1.776)

CONCLUSIONS

The obtained results indicate that the observed electromagnetic response is consistent with deformation-induced microstructural changes reported in the literature. The study verified the applicability of low-frequency impedance spectroscopy combined with PCA for the non-destructive assessment of plastic strain hardening in S235JR ferromagnetic structural steel.

The application of PCA enabled effective concentration of diagnostic information contained in the 66-dimensional input vector, consisting of resistance R and inductance L values measured at 33 excitation frequencies. It was shown that more than 95% of the total variability of the electromagnetic response was represented by the first four principal components (PC1–PC4). The first two principal components contained most of the hardness-related information used for regression modelling. This allowed substantial reduction of the dimensionality of the regression model while retaining the dominant information contained in the measured impedance spectra.

The analysis of individual predictors showed that PC1, despite representing the largest mathematical variance, exhibited a weak correlation with material hardness ($R^2 = 0.143$). The key finding was the synergistic effect observed when PC1 and PC2 were combined. The loading analysis indicates that these principal components represent different patterns of variability within the measured impedance spectra. Consequently, they provide complementary information, resulting in a substantial improvement in prediction accuracy when both components are simultaneously included in the multiple linear regression model, reaching a calibration coefficient of determination of $R^2 = 0.855$. The inclusion of higher-order components, PC3 and PC4, did not provide a meaningful improvement in prediction accuracy.

The developed PCA-based impedance spectroscopy approach demonstrates considerable potential for non-destructive hardness assessment of plastically deformed S235JR steel. The results indicate that complete low-frequency impedance spectra contain diagnostically relevant information that can be effectively extracted using multivariate statistical analysis. The predictive capability of the proposed approach was confirmed using an independent validation dataset, which yielded $R^2 = 0.881$ and RMSE = 2.426 HV1. These results demonstrate that reliable hardness estimation can

be achieved directly from the complete impedance spectrum without the need for equivalent-circuit parameter identification. The applicability of the method may be influenced by geometrical factors, surface condition, and lift-off effects, which should be considered when interpreting impedance-based hardness estimates. Further work is required to assess the applicability of the proposed approach to other steel grades, heat-treatment conditions, microstructures, and measurement conditions encountered in industrial environments.

Acknowledgments

The presented work was supported by the Silesian University of Technology from statutory research funds.

REFERENCES

1. Hu X., Wang X., Cai H., Yang X., Pan S., Yang Y., et al. Numerical simulation of electromagnetic non-destructive testing technology for elasto-plastic deformation of ferromagnetic materials based on magneto-mechanical coupling effect. *Sensors*. 2024; 24(22): 7103. <https://doi.org/10.3390/s24227103>
2. Wang P., Zhang Y., Yao E., Mi Y., Zheng Y., Tang C. Method of measuring the mechanical properties of ferromagnetic materials based on magnetostrictive EMAT characteristic parameters. *Measurement (Lond)*. 2021; 168: 108187. <https://doi.org/10.1016/j.measurement.2020.108187>
3. Vourna P., Falara P.P., Ktena A., Hristoforou E. V., Papadopoulos N.D. Magnetic Barkhausen noise sensor: A comprehensive review of recent advances in non-destructive testing and material characterization. *Sensors*. 2026; 26: 258. <https://doi.org/10.3390/s26010258>
4. Kumar A., Arockiarajan A. Evolution of nonlinear magneto-elastic constitutive laws in ferromagnetic materials: A comprehensive review. *J. Magn. Magn. Mater.* 2022; 546: 168821. <https://doi.org/10.1016/j.jmmm.2021.168821>
5. Altinses D., Salazar Torres D.O., Holstein V., Hermes M., Schwung A. Non-destructive evaluation of material hardness using micromagnetic sensors and complex neural networks. *Applied Intelligence*. 2026; 56: 36. <https://doi.org/10.1007/s10489-026-07089-1>
6. Jedamski R., Epp J. Non-destructive micromagnetic determination of hardness and case hardening depth using linear regression analysis and artificial neural networks. *Metals (Basel)*. 2021; 11: 18. <https://doi.org/10.3390/met11010018>
7. Shokralla S., Morelli J.E., Krause T.W. Principal components analysis of multifrequency eddy current

- data used to measure pressure tube to calandria tube gap. *IEEE Sens. J.* 2016; 16(9): 3147–54. <https://doi.org/10.1109/JSEN.2016.2529721>
8. Alamin M., Tian G.Y., Andrews A., Jackson P. Principal component analysis of pulsed eddy current response from corrosion in mild steel. *IEEE Sens. J.* 2012; 12(8): 2548–53. <https://doi.org/10.1109/JSEN.2012.2195308>
9. Wolter B., Straß B., Jacob K., Rauhut M., Stephani T., Riemer M., et al. Nondestructive material characterization and component identification in sheet metal processing with electromagnetic methods. *Sci. Rep.* 2024; 14: 6274. <https://doi.org/10.1038/s41598-024-55927-4>
10. Ghanei S., Kashefi M., Mazinani M. Comparative study of eddy current and Barkhausen noise non-destructive testing methods in microstructural examination of ferrite-martensite dual-phase steel. *J. Magn. Magn. Mater.* 2014; 356: 103–10. <https://doi.org/10.1016/j.jmmm.2014.01.001>
11. Zhang J., Qin Y., Liu X., Zhao Y., Dang W., Fan X., et al. Stress-induced magnetic anisotropy in Fe-based amorphous/nanocrystalline alloys: mechanisms, advances and challenges. *Materials*. 2025; 18(7): 1499. <https://doi.org/10.3390/ma18071499>
12. Xu C., Wang X., Dong H., Li J., Wang L., Liu X., et al. Quantitative prediction of residual stress, surface hardness, and case depth in medium carbon steel plate based on multifunctional magnetic testing techniques. *Sensors*. 2025; 25(9): 2812. <https://doi.org/10.3390/s25092812>
13. Dobmann G., Boller C., Herrmann H.G., Altpeter I. Micromagnetic and electromagnetic NDT for lifetime management by monitoring ageing of structural materials. *International Journal of Microstructure and Materials Properties*. 2014; 9(3–5): 348–59. <https://doi.org/10.1504/IJMMP.2014.066915>
14. Psuj G. Fusion of multiple parameters of signals obtained by vector magnetic flux observation for evaluation of stress loaded steel samples. *International Journal of Applied Electromagnetics and Mechanics*. 2015; 49(1): 1–7. <https://doi.org/10.3233/JAE-152003>
15. Munir N., Huang J., Wong C.N., Song S.J. Machine learning based eddy current testing: A review. *Results in Engineering*. 2025; 25: 103724. <https://doi.org/10.1016/j.rineng.2024.103724>
16. Roskosz M., Fryczowski K., Schabowicz K. Evaluation of ferromagnetic steel hardness based on an analysis of the Barkhausen noise number of events. *Materials*. 2020; 13(9): 2059. <https://doi.org/10.3390/MA13092059>
17. Taheri H., Bocanegra M.G., Taheri M. Artificial intelligence, machine learning and smart technologies for nondestructive evaluation. *Sensors*. 2022; 22(11). <https://doi.org/10.3390/s22114055>
18. Valenzuela R., Montiel H., Gutiérrez M.P., Betancourt I. Characterization of soft ferromagnetic materials by inductance spectroscopy and magnetoimpedance. *J. Magn. Magn. Mater.* 2005; 294(2): 239–44. <https://doi.org/10.1016/j.jmmm.2005.03.040>
19. Valenzuela R. Giant magnetoimpedance and inductance spectroscopy. *J. Alloys Compd.* 2004; 369(1–2): 40–2. <https://doi.org/10.1016/j.jallcom.2003.09.057>
20. Thielges S., Youssef S., Hartmann U. Reconstruction of the ferromagnetic hysteresis in the rayleigh regime by means of impedance analysis of the excitation coil. *J. Nondestr. Eval.* 2024; 43(2): 37. <https://doi.org/10.1007/s10921-024-01055-y>
21. Buznikov N.A., Kurlyanskaya G. V. Theoretical study of microwires with an inhomogeneous magnetic structure using magnetoimpedance tomography. *Sensors*. 2024; 24: 3669. <https://doi.org/10.3390/s24113669>
22. Zhang Z., Matsushima Y., Shibata Y., Hatakeyama T., Matsuzaka M., Funato T., et al. Emergent magneto-inductance effect in NiFe thin films on glass substrates at room temperature. *J. Magn. Magn. Mater.* 2024; 610: 172500. <https://doi.org/10.1016/j.jmmm.2024.172500>
23. Lukács Z., Kristóf T. A generalized model of the equivalent circuits in the electrochemical impedance spectroscopy. *Electrochim. Acta.* 2020; 363: 137199. <https://doi.org/10.1016/j.electacta.2020.137199>
24. Schaeffer J., Gasper P., Garcia-Tamayo E., Gasper R., Adachi M., Pablo Gaviria-Cardona J., et al. Machine Learning benchmarks for the classification of equivalent circuit models from electrochemical impedance spectra. *J. Electrochem. Soc.* 2023; 170(6): 060512. <https://doi.org/10.1149/1945-7111/acd8fb>
25. Zou H., Xue L. A selective overview of sparse principal component analysis. *Proceedings of the IEEE*. 2018; 106(8): 1311–20. <https://doi.org/10.1109/JPROC.2018.2846588>
26. Tonin F., Tao Q., Patrinos P., Suykens J.A.K. Deep kernel principal component analysis for multi-level feature learning. *Neural Networks*. 2024; 170: 578–95. <https://doi.org/10.1016/j.neunet.2023.11.045>
27. Jiang J., Ma J., Chen C., Wang Z., Cai Z., Wang L. SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* 2018; 56(8): 4581–93. <https://doi.org/10.1109/TGRS.2018.2828029>
28. Sophian A., Tian G.Y., Taylor D., Rudlin J. A feature extraction technique based on principal component analysis for pulsed Eddy current NDT. *NDT&E International*. 2003; 36(1): 37–41. [https://doi.org/10.1016/S0963-8695\(02\)00069-5](https://doi.org/10.1016/S0963-8695(02)00069-5)
29. BinFeng Y., FeiLu L., Dan H. Research on edge identification of a defect using pulsed eddy current based on principal component analysis. *NDT&E International*. 2007; 40(4): 294–9. <https://doi.org/10.1016/j.ndteint.2006.12.005>
30. Nafiah F., Tokhi M.O., Member S., Shirkoohi G., Duan F., Zhao Z., et al. Parameter analysis of pulsed eddy current sensor using principal component analysis. *IEEE Sens. J.* 2021; 21(5): 6897–903. <https://doi.org/10.1109/JSEN.2020.3036967>