1. Krzysztof KRÓL[1], 2. Tomasz RYMARCZYK[1,2], 3. Edward KOZŁOWSKI[3]

Research and Development Center, Netrix S.A (1), University of Economics and Innovation, Lublin (2), Lublin University of Technology (3)
ORCID. 1. 0000-0002-0114-2794, 2. 0000-0002-3524-9151, 3. 0000-0002-7147-4903

doi:10.15199/48.2022.02.45

Application of logit regression to the analysis of the state of the flood embankment

***Abstract**. Non-destructive methods also include electrical impedance tomography, in which electrical measurements are made. This method, thanks to the measuring device used and the implemented algorithms, allows for a non-invasive spatial determination of the degree of moisture. The article presents the problem of identifying flood protection by means of image reconstruction in electrical impedance tomography (EIT). Reconstruction in EIT concerns the performance of a series of measurements using multiple sensors as well as image reconstruction. The reconstruction of the image allows the presentation of various inclusions in the examined object. Logit regression was used to determine the inclusions in the analyzed area. Additionally, the elasticnet method was used to select predictors in logit regression. The results of research on the development of an effective and non-invasive method of flood embankment detection were prepared.*

***Streszczenie.** Do metod nieniszczących zalicza się również elektryczną tomografię impedancyjną, w której wykonuje się pomiary elektryczne. Metoda ta, dzięki zastosowanemu urządzeniu pomiarowemu oraz zaimplementowanym algorytmom, pozwala na bezinwazyjne przestrzenne określenie stopnia zawilgocenia. W artykule został przedstawiony problem identyfikacji przesiąkania wałów przeciwpowodziowych za pomocą rekonstrukcji obrazu w elektrycznej tomografii impedancyjnej (EIT). Rekonstrukcja w EIT dotyczy wykonania szeregu pomiarów przy użyciu wielu czujników jak i rekonstrukcji obrazu. Z rekonstrukcji obrazu pozwala na przedstawienie różnych wtrąceń w badanym obiekcie. Do określenia wtrąceń w analizowanym obszarze zastosowano regresję logit. Dodatkowo do wyboru predyktorów w regresji logit zastosowano metodę elasticnet. Opracowano wyniki badań nad opracowaniem skutecznej i nieinwazyjnej metody detekcji wałów przeciwpowodziowych (**Zastosowanie regresji logitowej do analizy stanu wału przeciwpowodziowego**).*

Keywords: logit regression,tomography,seepage of embarkment,
Słowa kluczowe: regresja logit, tomografia, przesiąkanie nasypów,

Introduction

In order to be able to identify seepage in embankments, the measurement method must first be selected. Only measurement methods that will not damage the structure can be used for the study of flood embankments. Therefore, non-destructive measurements should be used in this case. In addition, areas with seepage should be determined visually by measuring the dike area. Taking these conditions into consideration, the Electrical Impedance Tomography method was chosen for this purpose. There are many methods of solving optimization problems that are part of a specific system [1-11]. In tomography, deterministic methods and machine learning are used to solve the inverse problem [12-18].

Electrical impedance tomography is a non-destructive method for creating image reconstructions in various application areas (e.g. [19-22]. The primary purpose of using tomography is image reconstruction. Image reconstruction is directly related to solving the inverse problem. Moreover, here, there is a problem that some of the signals coming from the electrodes are correlated, which introduces the problem of collinearity [23]. In this case, the Gauss-Markov theorem cannot be applied to determine unknown linear parameters does not give satisfactory results. To overcome the above problem, singular value decomposition, LARS [24,25], regularisation methods (such as Tikhonov regularisation, total variation, least angle regression, regularisation sparse, see, e.g. [26-28] are usually used. [29,25]) or a neural network [30].

In order to recognise infiltration in the embankment, the imaging domain was modelled as a mesh that consists of a set of finite elements. For each finite element, a model was created that shows the relationship between the signal obtained from the electrodes and the conductivity. In the analysed case, conductivity other than background was related to the water concentration. The main goal is to recognise the location of the flooded embankment. The obtained results can be used as a preliminary resolution for image reconstruction. The paper proposes logit regression to solve the inverse problem in electrical impedance tomography.

Areas with a high probability of inclusion are determined based on measurements obtained from transducers placed on the border of the imaging area. By determining the cutoff level of the probability level, we can create a classifier that helps separate the background from the inclusions. Additionally, by changing the required level of probability, we can improve the accuracy of imaging. To overcome the problem of collinearity of measurements, the elastic net methods were used. These models allowed the computation of non-background conductivity probabilities for each of the finite elements. Based on this approach, the resolution in the imaging domain was determined, and thus the mesh was reconstructed. Additionally, the resolution of the imaging domain depends on a classification level that delimits the occurrence of non-background conductivity. By analysing the ROC [31,32] for each finite element, the classification level was determined (as the limit probability of filtration conductivity).

Methods

Let $D=\{(x_i,y_i): x_i\in R^m, y_i\in\{0,1\}, 1\le i\le n)$ denote the learning data set. Then the elements of the sequence $\{x_i\}1\le i\le n$ belong to two classes. Class membership is represented based on $y_i\epsilon\{0,1\}$ for $1\le i\le n$. For the identification of inclusions in the described problem, yi=1 is assumed for each finite element. If the finite element does not have an inclusion, $y_i=0$ is assumed. Observing the signal obtained from the sensors x, $x\epsilon R^m$, classify the existence of inclusion for each finite element. Logit regression was used to construct the classifier f: Rm →{0,1}).

Let $(\Omega,\mathcal{F},P)$ be a probabilistic space and Y a random variable with a discrete distribution, where Y: $\Omega\to\{0,1\}$. Then, the logit regression describes the dependent variable Y's realisation probability distribution based on the independent variables $X\epsilon R^m$.

The odds are the ratio of the probability of success to the probability of failure

(1) $\theta(X) = \frac{P(Y=1|X)}{1-P(Y=1|X)}$

The logistic regression task (see e.g. [19-21] is to estimate the probability of success P(Y=1|X) based on the realisation X, and we assume

(2) $P(Y = 1|X) = p(X)$

Since the probability of success p(X)ϵ (0,1), so the formula implies that the chance θ(X) ϵ(0,∞).

On logistic regression, we analyse the linear dependence of the logarithm of chance on the independent variables X. For this purpose, we consider the relationship defined by the equation

(3) $\ln\theta(X) = \ln\left(\frac{p(X)}{1-p(X)}\right) = X\beta + \varepsilon$

where ε is a random variable with a normal distribution N(0 σ^2) and β=(β_1, β_2, … βm) ϵ R^m. If there is a free word in the models, then the column corresponding to the free word consists of ones in the matrix X.

The formula gives the necessary condition for the existence of an extremum

(4) $\frac{\partial l}{\partial \beta}(\beta) = 0$, and

(5) $\frac{\partial l}{\partial \beta}(\beta) = X^T z(\beta)$

$$z(\beta) = \begin{bmatrix} y_1 - p\big(\beta, x_{(1)}\big) \\ y_2 - p\big(\beta, x_{(2)}\big) \\ \vdots \\ y_n - p\big(\beta, x_{(n)}\big) \end{bmatrix}$$

The matrix of second derivatives is negatively defined.

(6)$\frac{\partial^2}{\partial\beta\partial\beta^T}(\beta) = -\sum_{i=1}^{n} x_i^T x_i h_i(\beta) = -X^T Z(\beta) X$

(7) $\frac{\partial l(\beta)}{\partial \beta_j} = \sum(y_j - p(\beta, x_{(i)}))x_{ij} = 0$

and

(8) $\frac{\partial^2 l(\beta)}{\partial\beta_j\partial\beta_k} =$
$\sum_{i=1}^{n} x_{ij}x_{ik}p(\beta, x_{(i)})(1 - p(\beta, x_{(i)}))$

We estimate the values of the β parameters iteratively. To determine the estimators of the unknown parameters we use the Newton-Raphson algorithm, where in step k+1 the values of the estimators of the parameters β are calculated using the formula

(9) $\beta_{k+1} = \beta_k - \left(\frac{\partial^2 l}{\partial\beta\partial\beta^T}(\beta_k)\right)^{-1} \frac{\partial l}{\partial \beta}(\beta_k)$

In the results, highly correlated data are often obtained. For this reason, the elastic net method [22] was used, which is one of the regularisation methods (a combination of LASSO and ridge regression called Tikhonov regularisation. By imposing a penalty on large parameter values, the estimators are pulled towards zero. Consequently, we obtain a reduction in the variance of the estimators and an improvement in prediction quality. Linear regression parameters are determined by solving the task

(10) $\max_{\beta}\{\sum_{i=1}^{N}\big(y_i x_{(i)}\beta - \ln\big(1 +$
$e^{x_{(i)}\beta}\big)\big) - \lambda P_\alpha(\beta)\}$

where λ>0, 0≤α≤1 and P_α is the penalty given by the formula

(17) $P_\alpha(\beta) = \alpha\|\beta\|_{L_1} + \frac{1-\alpha}{2}\|\beta\|_{L_2} =$
$\sum_{j=1}^{p}\left(\alpha|\beta_j| + \frac{1-\alpha}{2}\beta_j^2\right)$

The penalty Pα(β) is the linear combination of the norms of the vector of β estimators in L_1 and L_2 spaces. Thus, for α =0 we have ridge regression, while for ¯α =1 we have LASSO.

Results

A shaft model with 32 point electrodes was built. This model consists of 9805 triangular finite elements, 2818 nodes. After obtaining the signal from the electrodes, the main objective is to reproduce the water seepage in the embankment. Logit models were created for each finite element. Figures 1-14 contain model descriptions, measurement results and image reconstruction.

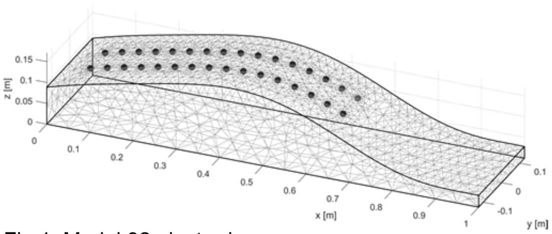

Fig.1. Model 32 electrodes

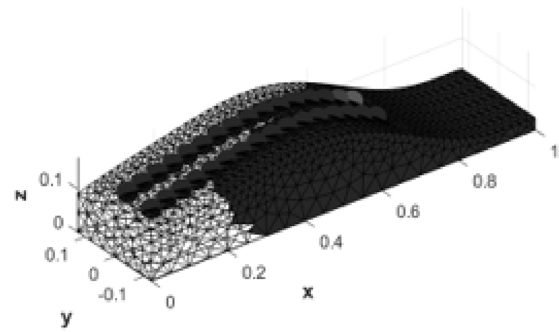

Fig.2. Model 1 pattern

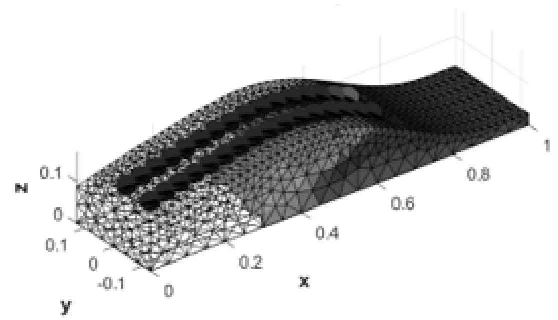

Fig.3. Model 1 reconstruction with Elasticnet

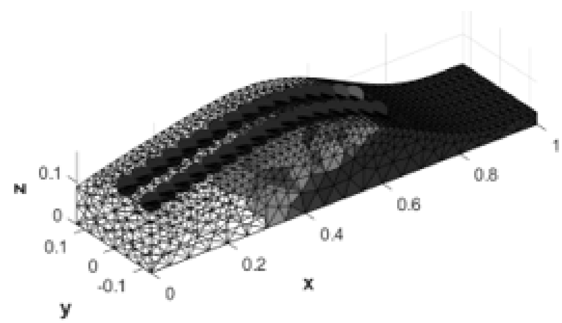

Fig.4. Model 1 reconstruction by logit regression

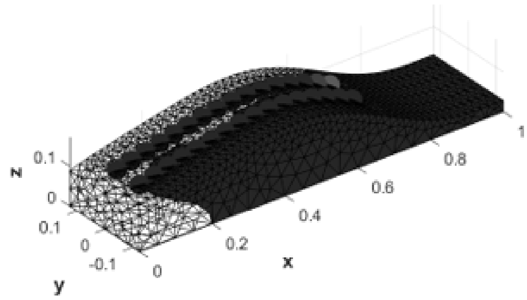

Fig.5. Model 2 pattern

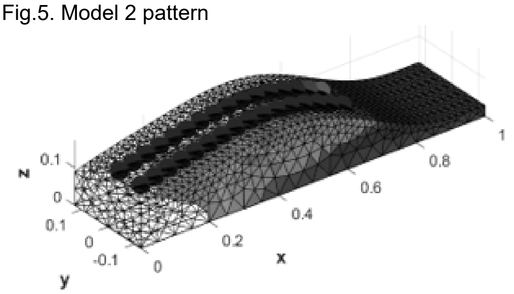

Fig.6. Model 2 reconstruction with Elasticnet

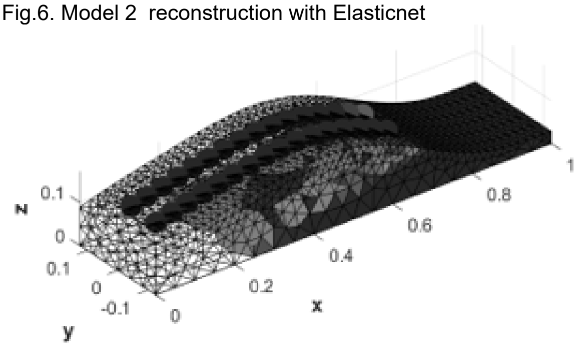

Fig.7. Model 2 reconstruction by logit regression

A second shaft model with 16 point electrodes was built. This model consists of 7657 triangular finite elements, 2159 nodes. After obtaining the signal from the electrodes, the main objective is to reproduce the water seepage in the embankment. Logit models were created for each finite element.

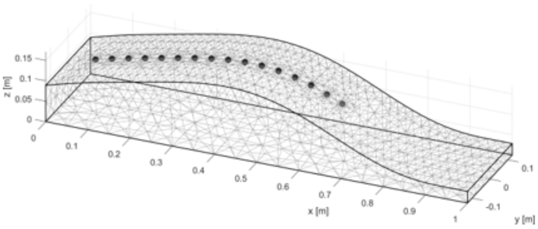

Fig.8. Model 16 electrodes

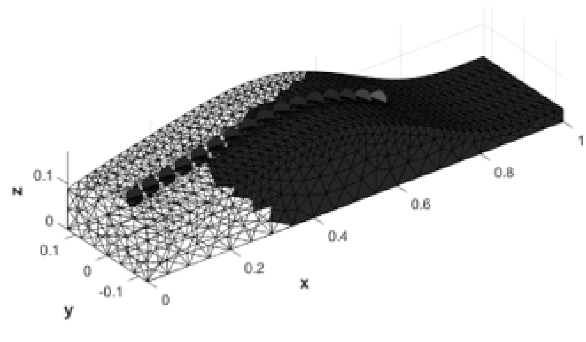

Fig.9. Model 3 pattern

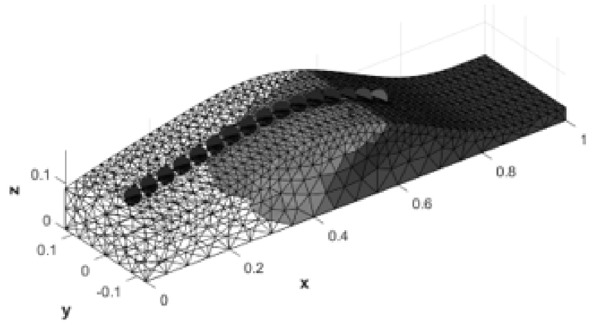

Fig.10. Model 3 reconstruction with Elasticnet

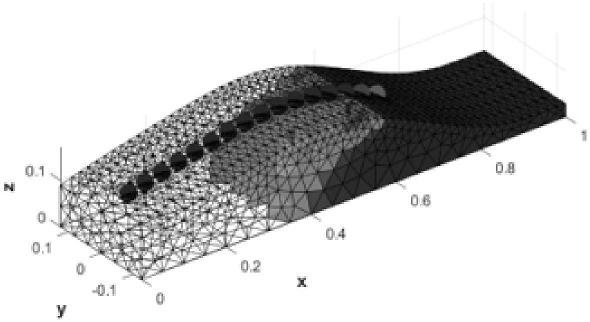

Fig.11. Model 3 reconstruction by logit regression

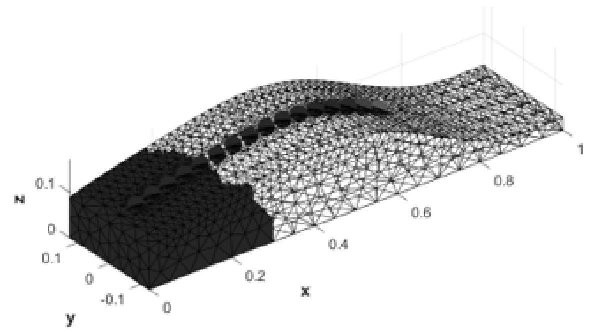

Fig.12. Model 4 pattern

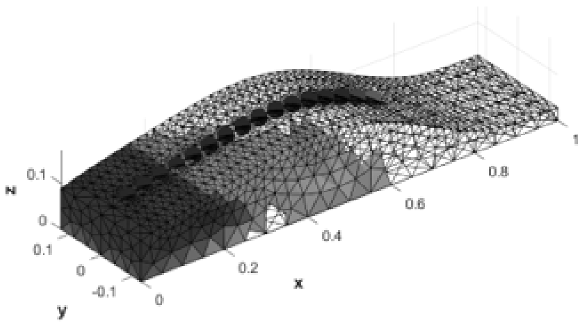

Fig.13. Model 4 reconstruction with Elasticnet

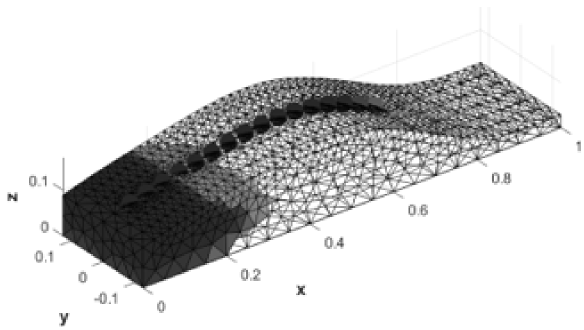

Fig.14. Model 4 reconstruction by logit regression

Conclusion

The article presents examples of the use of logit regression in spatial imaging. The image was obtained on the basis of a reconstruction with the use of electrical impedance tomography. A novelty in the presented concept is the training of many subsystems operating simultaneously, thanks to which each of them generates the binary value of a single pixel of the reconstructed image. With this approach, when each element based on several or many hundreds of input fields handles one output, multiple predictors can be correlated with each other.

Authors: *Tomasz Rymarczyk, D.Sc, Ph.D. Eng., University of Economics and Innovation, Projektowa 4, Lublin,. E-mail: tomasz@rymarczyk.com; Krzysztof Król, Research&Development Centre Netrix S.A., Email: krzysztof.krol@netrix.com.pl; Edward Kozłowski Ph.D.Eng.,Lublin University of Technology, Nadbystrzycka 38, Lublin, E-mail: e.kozlowski@pollub.pl;*

REFERENCES

[1] Miłak, M., Leszczyńska, A., Grudzień, K., Romanowski, A., & Sankowski, D. (2019). Slug flow velocity estimation during pneumatic conveying of bulk solid materials based on image processing techniques. Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska, 9(1), 11-14.
[2] Kryszyn, J., Wanta, D., Smolik, W. T. (2019). Evaluation of the electrical capacitance tomography system for measurement using 3d sensor. Informatyka, Automatyka, Pomiary W Gospodarce I Ochronie Środowiska, 9(4), 52-59.
[3] Kłosowski G., Rymarczyk T., Wójcik D., Skowron S., Adamkiewicz P., The Use of Time-Frequency Moments as Inputs of LSTM Network for ECG Signal Classification, Electronics, 9(9), 1452, 2020.
[4] Korzeniewska E., Krawczyk A., Stando J., Torsion field - an example of pseudo-scientific concept in physics, Przeglad Elektrotechniczny, Volume97, Issue1, Page196-199, DOI10.15199/48.2021.01.41, 2021.
[5] Korzeniewska, E; Szczesny, A; Lipinski, P; Drozdz, T; Kielbasa, P; Miernik, A, Prototype of a Textronic Sensor Created with a Physical Vacuum Deposition Process for Staphylococcus aureus Detection, SENSORS Volume: 21 Issue: 1 Article Number: 183, 2021.
[6] Wajman, R; Banasiak, R; Babout, L, On the Use of a Rotatable ECT Sensor to Investigate Dense Phase Flow: A Feasibility Study, SENSORS Volume: 20 Issue: 17 Article Number: 4854, 2020.
[7] Banasiak, R.; Wajman, R.; Jaworski, T.; Fiderek, P.; Fidos, H.; Nowakowski, J.; Sankowski, D. Study on two-phase flow regime visualization and identification using 3D electrical capacitance tomography and fuzzy-logic classification. Int. J. Multiph. Flow 2014, 58, 1–14.
[8] Jan Dusek, Jan Mikulka, Measurement-Based Domain Parameter Optimization in Electrical Impedance Tomography Imaging, Sensors 2021, 21(7), 2507
[9] Daniewski K., Kosicka E., Mazurkiewicz D., Analysis of the correctness of determination of the effectiveness of maintenance service actions. Management and Production Engineering Review 9 (2018); No. 2, 20-25.
[10] Romanowski, A. Contextual Processing of Electrical Capacitance Tomography Measurement Data for Temporal Modeling of Pneumatic Conveying Process. In Proceedings of the 2018 Federated Conference on Computer Science and Information Systems (FedCSIS), Poznan, Poland, 9–12 September 2018; 283–286.
[11] Chen, B.; Abascal, J.; Soleimani, M. Extended Joint Sparsity Reconstruction for Spatial and Temporal ERT Imaging. Sensors 2018, 18, 4014.
[12] Rymarczyk T., Kłosowski G., Hoła A., Sikora J., Wołowiec T., Tchórzewski P., Skowron S., Comparison of Machine Learning Methods in Electrical Tomography for Detecting Moisture in Building Walls, Energies, 14(10), 2777; 2021.
[13] Rymarczyk T., Kozłowski E., Kłosowski G., Electrical impedance tomography in 3D flood embankments testing – elastic net approach, Transactions of the Institute of Measurement and Control, 42(4), 680-690, 2020.
[14] Rymarczyk T., Nita P., Vejar A., Woś M., Stefaniak B., Adamkiewicz P.: Wearable mobile measuring device based on electrical tomography, Przegląd Elektrotechniczny, 95(4), 211-214, 2019.
[15] Rymarczyk T., Kłosowski G., Tchórzewski P., Cieplak T., Kozłowski E.: Area monitoring using the ERT method with multisensor electrodes, Przegląd Elektrotechniczny, 95(1), 153-156, 2019.
[16] Koulountzios P., Rymarczyk T., Soleimani M., A quantitative ultrasonic travel-time tomography system for investigation of liquid compounds elaborations in industrial processes, Sensors, 19(23), 5117, 2019.
[17] Kłosowski G., Rymarczyk T., Kania K., Świć A., Cieplak T., Maintenance of industrial reactors based on deep learning driven ultrasound tomography, Eksploatacja i Niezawodnosc – Maintenance and Reliability; 22(1), 138–147, 2020.
[18] Kłosowski G., Rymarczyk T., Cieplak T., Niderla K., Skowron Ł., Quality Assessment of the Neural Algorithms on the Example of EIT-UST Hybrid Tomography, Sensors, 20(11), 3324, 2020.
[19] Rymarczyk T., Tchórzewski P., Sikora J., Implementation of Electrical Impedance Tomography for Analysis of Building Moisture Conditions, Compel The international journal for computation and mathematics in electrical and electronic engineering, 37 (2018), No. 5, 1837-1861
[20] Rymarczyk T., Kłosowski G., Kozłowski E., NonDestructive System Based on Electrical Tomography and Machine Learning to Analyse Moisture of Buildings, Sensors, 18 (2018), No. 7, 2285
[21] Rymarczyk T., Koz łowski E., Kłosowski G., Electrical impedance tomography in 3D _ood embankments testing elastic net approach, Transactions of the Institute of Measurement and Control, 2019,
[22] Kłosowski G., Rymarczyk T., Gola A., Increasing the reliability of flood embankments with neural imaging method, Applied Sciences 8.9 (2018), 1457.
[23] Rymarczyk T., Kłosowski G., Kozłowski E., Tchórzewski P., Comparison of Selected Machine Learning Algorithms for Industrial Electrical Tomography, Sensors, 19 (2019), No. 7, 1521
[24] Hastie T., Tibshirani R., Friedman J., The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer, 2009.
[25] James G., Witten D., Hastie, T., Tibshirani R., An Introduction to Statistical Learning: with Applications in R, Springer, 2017.
[26] Wehrens R., Chemometrics with R. Multivariate Data Analysis in the Natural Science and Life Sciences, Springer, 2011.
[27] Tibshirani R., Regression shrinkage and selection via the lasso, Journal of the Royal Statistical Society, Series B, 58 (1996), No. 1, 267–288
[28] Zou H., Hastie T., Regularisation and variable selection via the elastic net, Journal of the Royal Statistical Society, Series B, 67 (2005), No. 2, 301–320
[29 Friedman J., Tibshirani R., Hastie T., Regularisation paths for generalised linear models via coordinate descent, Journal of Statistical Software, 33 (2010), No. 1, 1–22
[30] Rymarczyk T., Kłosowski G., Innovative methods of neural reconstruction for tomographic images in maintenance of tank industrial reactors, Eksploatac jaiNiezawodnosc – Maintenance and Reliability, 20 (2019), No. 3, 425–434
[31] T. Fawcett, An introduction to ROC analysis, Pattern Recognition Letters. 27 (2006) 20 861–874.
[32] D.J. Hand, R.J. Till, A simple generalisation of the area under the roc curve for multiple class classification problems, Machine Learning. 45 (2001) 171–186.