

# Tomographic image correction with noise reduction algorithms

Grzegorz Kłosowski<sup>1,\*</sup>, Tomasz Rymarczyk<sup>2</sup>, and Edward Kozłowski<sup>3</sup>

<sup>1</sup>Lublin University of Technology, Faculty of Management, Nadbystrzycka 38, 20-618 Lublin, Poland

<sup>2</sup>Research and Development Center, Netrix S.A., Związkowa 26, 20-148 Lublin, Poland

**Abstract.** This article presents an original approach to improve the results of tomographic reconstructions by denoising the input data, which affects output images improving. The algorithms used in the research are based on autoencoders and Elastic Net - both related to artificial intelligence or machine-learning developed controllers. Due to the reduction of unnecessary features and removal of mutually correlated input variables generated by the tomography electrodes, good quality reconstructions of tomographic images were obtained. The simulation experiments proved that the presented methods could be effective in improving the quality of reconstructed tomographic images.

## 1 Introduction

Electrical impedance tomography (EIT) and electrical capacitance tomography (ECT) are based on the processing of data generated by the electrode system [1]. In the case of reconstruction of tomographic images concerning objects with large dimensions and relatively low conductance, the data from the electrodes are often noisy. Examples of technical objects from which the tomographic data show a high level of noise are: buildings, earthworks, flood embankments, dams, landfill protections, reservoirs, reactors, as well as many industrial infrastructure facilities [2]. The noise of electrical signals is one of the main barriers hindering the development of tomographic methods for such objects [3]. Thanks to the rapid technological development including data processing techniques, access to advanced computational methods is becoming easier each year. Costs of access to computing power and storage media are falling. This promotes the development of computational techniques that use parallel computing and the processing of large data sets. The burden of compromise in which researchers are forced to choose between the time of calculation and the quality of results disappears. It also creates new possibilities in the field of application of algorithms for data smoothing and denoising [4, 5].

The term "denoising" is most often associated with the improvement of the quality of the signal generated by various types of sources. Signals can be included in the group of "time series", which can be classified according to their cyclicity, repeatability, randomness, continuity, *etc.* In the case of measurement data generated by the tomography electrodes, the time factor does not play a role, so it is omitted. During a single measurement, a system consisting of a dozen or so electrodes connected to a multiplexer generates a vector of several tens or hundreds of measured values. Many

individual measurement vectors form a set of measurement data. There are many reasons that can affect the distortion of the measurement data. Most of them result from the errors of the measurement system.

The noise associated with the data obtained from the measurements of the real object is the result of transient electrical states arising, for example, at the contacts of electrodes and the surface of the tested object. In addition, noise is generated as a result of rapid changes in the voltages applied to individual electrode pairs, which is controlled by the multiplexer. The presence of other electrodes at a short distance and the surroundings of the tested object, which cannot be isolated from it, also causes interference in the measurements and affects the noise of the vector of measurement data.

In the case of using controllers and predictors based on machine learning methods, a prerequisite for their training is to have a large set of data [6]. The measurement data is typically obtained by means of mathematical models for the generation of large sets of training data in a relatively short time. For example, thanks to the use of IT methods, it is possible to obtain tens of thousands of measurement cases in no more than one hour. A medium-class laptop is enough for this purpose.

Measurements of real objects are much more time-consuming and generate much higher costs. When creating mathematical simulation models to generate tomographic measurement signals, artificial noise is often used. This treatment aims to reproduce the real object's measurement parameters by the model with maximum accuracy. Because of a large number of unidentified factors in case of measurements performed on a real object, the noise characteristic can be slightly different from the noise generated by simulation.

Real world data is often corrupted with unwanted noise, which must be removed before further signal processing. Existing denoising algorithms, such as the

---

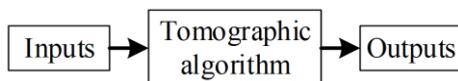
\* Corresponding author: [g.klosowski@pollub.pl](mailto:g.klosowski@pollub.pl)

least mean square (LMS) based Wiener and Kalman filtering [7], multi-scale analysis based wavelet denoising [8] and the newly developed empirical mode decomposition (EMD) method [9], are mainly designed for univariate signals. However, with the development of multichannel sensor technology, multivariate denoising is urgently needed in many applications ranging from communication system [10] to biomedicine [11] as well as in the EIT and ECT tomography.

It often happens that the signals are sparse in the wavelet domain, in which case, shrinking or thresholding-based nonlinear operators in wavelet denoising [12] can be used – particularly because their effectiveness was proven for real-world 1D and 2D signals. In [13] a multivariate variant of univariate wavelet denoising technique was offered. The authors presented a method of parallel denoising of multichannel signals. Combination of a principal component analysis (PCA) and univariate wavelet denoising performed on the signal after making a decorrelation of noise among channels, MWD method reaches better performance than standard channel-wise wavelet denoising. Newly, synchro-squeezing transform (SST) method was introduced to the wavelet denoising and it was confirmed that it could beat state-of-the-art methods that use wavelets for the signals with low-frequency modulation [14]. Accordingly, multivariate wavelet synchro-squeezing denoising (MWSD) algorithm, which employs the thresholding-based technique for the multivariate framework, was proposed in [15].

In the case of EIT and ECT tomographic systems, we deal with noisy measured data, which are usually set in specific measuring cases. Each of these cases has a vector of input measurements. With regard to EIT, these are voltage drops on individual electrode pairs, while for ECT they are electric capacitances. At the same time, each measuring vector is assigned a scalar or an output vector, which is generated by an algorithm that processes measurement data into output values that are the basis for generating an image of the reconstructed real object. The use of techniques belonging to the machine learning group gives great opportunities to create effective and efficient mechanisms for tomographic images reconstruction [16]. The condition for their use is to have a large set of training data, which should be free of interference, noise and redundancy of the input features.

Figure 1 presents a diagram of the computed tomographic data processing system.



**Fig. 1.** A tomographic data processing system.

If the tomographic algorithm was trained using historical data, then they had to be properly ordered. Formula (1) corresponds with Figure 1. It shows the transformation of inputs of a tomographic system into

outputs. Formulas (2) and (3) are a refinement of Formula 1.

$$\mathbf{A} \rightarrow \Psi \quad (1)$$

where:  $\mathbf{A}$  – matrix of measurement values  
 $\Psi$  – vector (or matrix) of results

$$\mathbf{A} = \begin{bmatrix} \alpha_{1,1} & \alpha_{2,1} & \dots & \alpha_{m,1} \\ \alpha_{1,2} & \alpha_{2,2} & \dots & \alpha_{m,2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1,n} & \alpha_{2,n} & \dots & \alpha_{m,n} \end{bmatrix} \quad (2)$$

where:  $\alpha_{n,m}$  – a value of the  $n$ -th measurement from the  $m$ -th pair of electrodes  
 $n$  – number of measuring cases  
 $m$  – the size of the measurement vector corresponding to the number of measured values

$$\Psi = \begin{bmatrix} \Psi_1 \\ \Psi_2 \\ \vdots \\ \Psi_n \end{bmatrix} \quad (3)$$

where:  $\Psi_m$  – the output value of the  $n$ -th measuring case

The output values  $\Psi_n$  in the formula (3) can be both scalars and vectors. For the purpose of this study, each input measurement vector is assigned a single real value at the output.

This type of measurement data is processed leading to the rejection of heavily correlated characteristics of the input vector. Consequently, the data is not only denoised but above all, provides the foundation for developing a tomographic data processing algorithm, which will be resistant to disturbances generated by the measurement system installed on the real object.

In this study, two variants of denoise data were tested by reducing correlated input values. The first experiment included the use of autoencoders. The second experiment was based on the statistical methods Lasso (Least Absolute Shrinkage and Selection Operator) and Elastic Net [17].

Autoencoders transform the input vector into the same output vector [18]. During processing, they encode the input vector for the assumed number of features, which is usually smaller than the input vector. The decoder then converts the encoded feature vector back into the input value vector, so the input and output of autoencoder should be equal. The quality of the autoencoder is demonstrated by the fact that despite coding and decoding operations, the input and output vectors are identical or similar. Autoencoders do not use output generated by the tomographic algorithm. Their mode of functioning is based on the identification of mutual dependencies occurring in the data generated by measurement cases (input vectors).

Statistical regression methods such as Lasso and Elastic Net are currently undergoing a renaissance. Lasso was first described in 1986 and then popularised by Tibshirani in 1996. However, only the development of information technology has enabled the

implementation of this method on a large scale. In 2005 Zou and Hastie introduced Elastic Net as one of Lasso generalisations.

The main goal of the presented research was to develop an algorithm that uses auto-coders and the elastic net method in the reconstruction of tomographic images. Due to the hybrid (stack) combination of autoencoders + neural networks (ANN) and Elastic Net + ANN, the presented solutions are resistant to noise in the input data, hence the tomographic images fairly accurately reflect the reference objects. The presented hybrid approach is original and innovative in the field of electrical tomography.

## 2 Methods and models

This chapter presents two methods applicable to the denoising of tomographic data - autoencoders and Elastic Net. In addition to graphic diagrams and mathematical models, the charts of module training and quality indicators of trained models based on Mean Squared Error (MSE) are also presented.

### 2.1 Autoencoders

The introduction of autoencoders was a significant innovation in unsupervised learning, in which the key features of the function in terms of backpropagation were discovered [19]. The learning structure is borrowed from the neurological process underlying global learning and the behaviour of intelligent beings [3].

Denoising Autoencoders were introduced in 2009 by Bengio and Vincent. They constituted an alternative to deep belief networks by integrating stochastic corruption during the training process. The assumption of using this kind of denoise function is that good representation is immune to interference and interference of input data.

Figure 2 presents the model of the convolutional autoencoder stacked together with multilayer perceptron (MLP) in the last layer of the stack [20]. The stack consists of three encoders and one regressive artificial neural network MLP (multilayer perceptron) type. The stack works in the following way: primarily the first auto-encoder is trained, which converts 96 inputs to 96 outputs (Figure 3). The transformation causes that the system gains new characteristics contributing to its improved resistance to noisy data.

The second autoencoder converts 96 features generated by the first autoencoder into 20 output values. The third autoencoder distinguishes 8 characteristics with high significance from 20 features. At this point, a three-stage (3-stage stack) reduction of input values from 96 to 8 is made.

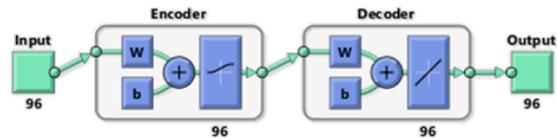


Fig. 3. Model of the first autoencoder in the stack.

As shown in Figure 2, the final fourth element of the stack is the MLP network, which converts 8 input features, generated by the preceding encoders, into one.

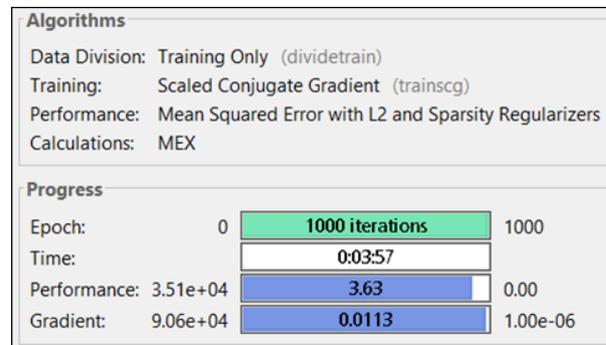


Fig. 4. First autoencoder training summary real number that reflects the colour of a single pixel of the reconstructed image.

Figure 4 shows the first autoencoder training summary. The training set quantity is 10,000. All cases were designated to the training process autoencoder.

Scaled Conjugate Gradient algorithm was used as the training method. The indicator of the quality of autoencoder training was mean squared error (MSE) with L2 norm (see equation 6) and scarcity regularises. Learning stopped after 1000 epochs (iterations). The condition for stopping the training process was in this case reaching the maximum number of epochs. The "Performance" parameter shows the average MSE error from the current epoch. The gradient reflects the rate of MSE drop. If the gradient approached 10<sup>-6</sup>, then the training process would be finished before the 1,000 epoch.

Figure 5 shows the training process of the first autoencoder in the stack. Best training performance is 3.632 at epoch 1000. This indicator corresponds with the performance indicator in Figure 4. Initially, MSE error rate was very dynamic, and after a dozen or so epochs it became somewhat stabilised. The course of the learning curve resembles hyperbole, which indicates that the autoencoder was not overtrained.

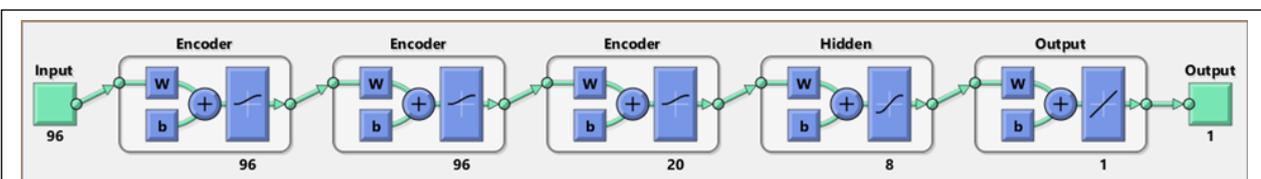


Fig. 2. Model of three convoluted autoencoders stacked with MLP with single regression output.

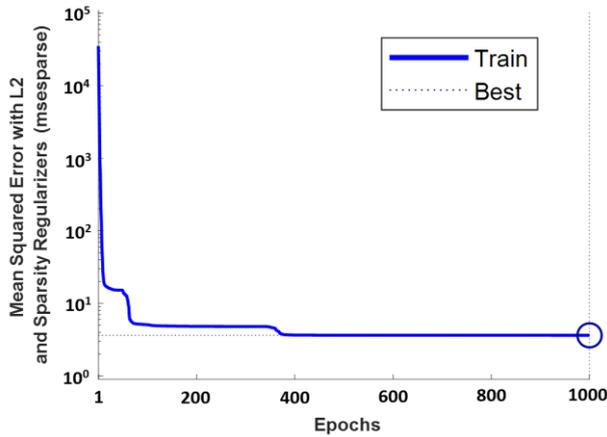


Fig. 5. First autoencoder training performance.

## 2.2 Lasso and Elastic Net

Let the linear system can be described by the state equation (4)

$$Y = X\beta + \varepsilon \quad (4)$$

where  $Y \in R^n, X \in R^{n \times (k+1)}$  denote the measurement matrices of response and input variables separately. Coefficient  $\beta \in R^{k+1}$  denotes the vector of unknown parameters. If the linear task (5) contains the intercept, then the first column of matrix  $X$  is a column of ones. The object  $\varepsilon \in R^n$  in the linear model (4) reflects a sequence of disturbances [17].

One of the possible ways to determine the linear regression when the input variables are collinear depends is the solution of the task (5)

$$\min_{(\beta_0, \beta') \in R^{k+1}} \frac{1}{2n} \sum_{i=1}^n (y_i - \beta_0 - x_i \beta')^2 + \lambda P_\alpha(\beta') \quad (5)$$

where:  $x_i = (x_{i1}, \dots, x_{ik}), \beta' = (\beta_1, \dots, \beta_k)$   
 for  $1 \leq i \leq n$

$P_\alpha$  – Elastic Net penalty value given by (6).

$$P_\alpha(\beta') = (1 - \alpha) \frac{1}{2} \|\beta'\|_{L_2} + \alpha \|\beta'\|_{L_1} = \sum_{j=1}^k \left( \frac{1 - \alpha}{2} \beta_j^2 + \alpha |\beta_j| \right) \quad (6)$$

We can notice that the penalty, in this case, is a linear combination of norms  $L_1$  and  $L_2$  of unknown parameters  $\beta'$ . The introduction of the parameter-dependent penalty function to the objective function reduces the estimators of unknown parameters.

The parameter  $\lambda$  in the task (5) represents the coefficient of the penalty. The parameter  $0 \leq \alpha \leq 1$  makes the compromise between Lasso (Least Absolute Shrinkage and Selection Operator) and ridge regression (RR). The ridge regression ( $\alpha = 0$ ) is also known as Tikhonov regularisation [21]. RR is one of the most popular methods of linear models regularisation. Lasso ( $\alpha = 1$ ) was introduced by Robert Tibshirani [22], [23]. This statistical method performs the selection of variables and regularisation in statistical linear models [24], [25]. In the case of the ridge regression, the penalty coefficient is calculated in the norm  $L_1$  but for Lasso in  $L_2$ . The difference between

Lasso and ridge regression is subtle because it concerns differences in relation to both norms. The ridge regression reduces coefficients of correlated predictors relative to each other. The correlated predictors depend on any latent factor. The ridge regression allows for the uniform distribution of the strength of latent factor on the predictors.

The Lasso method is indifferent to correlated predictors. This method allows specifying the preferred predictor and ignoring the rest. Using the Lasso method, we get a model in which many coefficients are close to zero, resulting in a sparse model. The elastic net combines the ridge regression and Lasso [26], [27]. Setting the appropriate value of  $\alpha$  the compromise between ridge regression and Lasso is formed.

Solving the formula (5) with fixed  $\lambda$  and  $\alpha$  factors we estimate the unknown parameters of the linear system in which some predictors are mutually correlated. In such a case, the prediction based on formula (4) is given by the equation  $\hat{Y} = X\hat{\beta}$ , where the vector of unknown estimators  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$  is estimated by solving the task (5).

Figure 6 presents the course of the Elastic Net learning process based on the Lambda ( $\lambda$ ) and MSE characteristics.

The green circle with the broken line indicates the Lambda with the minimum cross-validation error. The blue circle and the broken line indicate the point with the minimum cross-validation error plus one standard deviation. The parameter Lambda represents the coefficient of penalty. It can be noticed that the minimal MSE was reached at the beginning of the learning process.

During the tests, 100 reconstructions of various reference objects were made. To train and test the neural network, learning data generated using the appropriate script was used. Both the script for generating data and the algorithms for tomographic reconstructions were developed using MatLab. Eidors (MatLab toolbox) was used to develop the mesh using the finite element method.

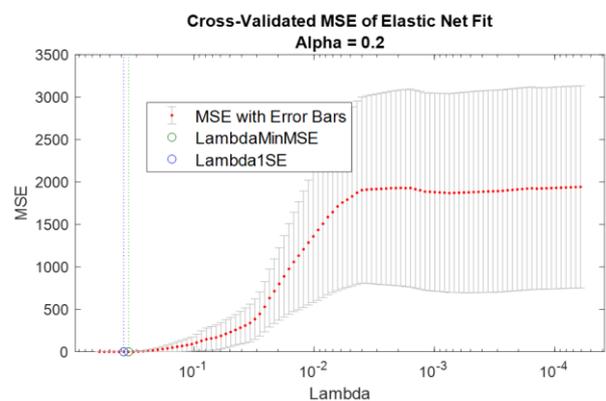
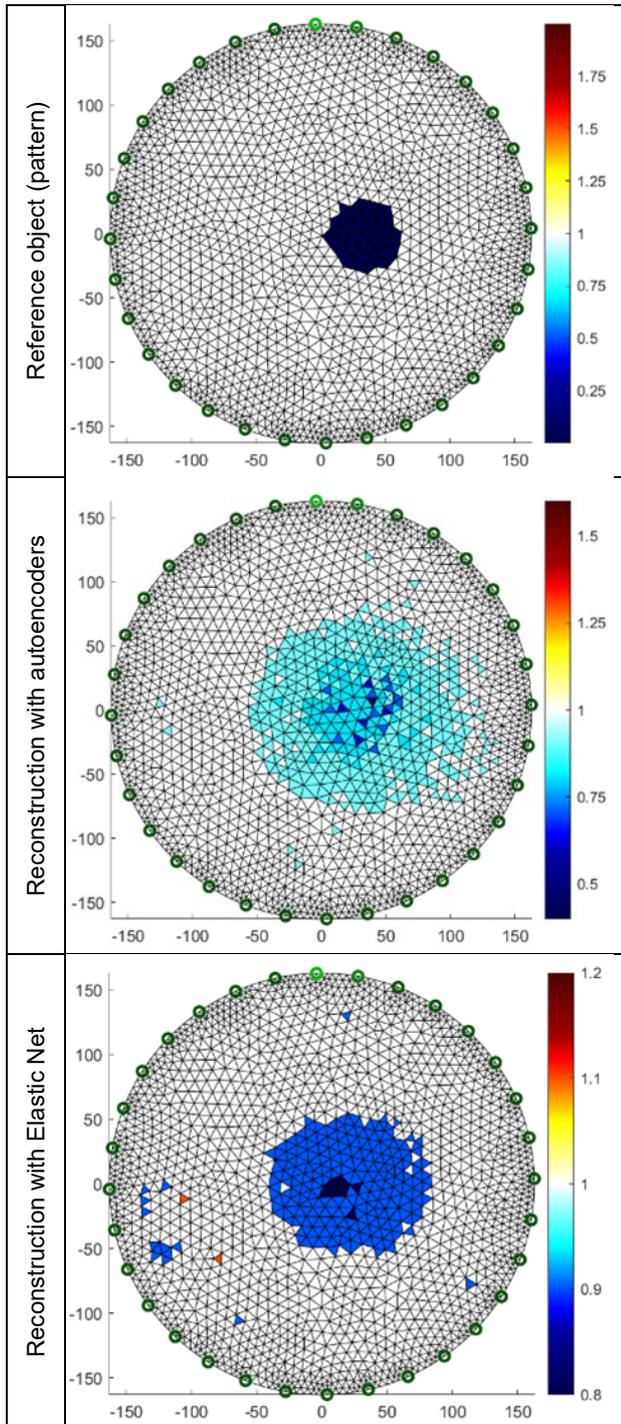


Fig. 6. Cross-validated MSE of Elastic Net for  $\alpha = 0.2$ .

## 3 Results

Figure 7 presents the results of simulation experiments performed using noisy data generated by properly prepared computer scripts. The generated measurement

cases have the signal to noise ratio  $SNR = 14$  dB. SNR is defined in terms of power as a signal to noise ratio. This noise parameter was calculated for a given measurement system configuration and included measurement patterns and electrode positions.



**Fig. 7.** Denoised image reconstructions performed with autoencoders and Elastic Net.

The first segment of Figure 7 shows the reference image of the reconstructed object. It was a single, rounded element that clearly stands out from the background. The colours reflect the conductivity of the interior of the tested object. The background pixels have a conductivity value of 1, while the conductivity of the

test object is close to zero. The conductivity values below 1 are marked in blue. This can be deduced from the comparison of the images with the colour bars placed next to them.

In the second segment of Figure 7, reconstruction with autoencoders was presented. The pixels that distinguish the mapping of the object from the background are pale blue. Their values oscillate mostly around the level of 0.85. Several points in the centre of the sample object have lower values of conductance - around 0.7. The location of the object on the tested cross-section is correct. Its size, however, is too fuzzy.

In the last, third segment of Figure 7, the reconstruction with Elastic Net was presented. Comparing the reconstructed image of the object with the pattern, it should be stated that the location of the object, its shape and dimensions have been reproduced with high accuracy. The quality of Elastic Net reconstruction exceeded the reconstruction with the use of autoencoders. You can see that the colour bar of Elastic Net reconstruction is slightly different from the colour bar of reconstruction with autoencoders. The colours are different but the conductance values on Elastic Net reconstructions and autoencoders are similar. In both cases, they are around 0.85. There are several individual pixels on the left in the Elastic Net reconstruction image. These are the remains of unreduced noise.

#### 4 Conclusions

The analysis of the results of simulation experiments presented in Figure 7 indicates that both denoising algorithms based on autoencoders and Elastic Net are capable of performing the tasks of tomographic images reconstruction effectively. Comparing the reconstructions generated by the encoders and Elastic Net, you can see that the image created using Elastic Net is sharper.

The reason may be the way of training the Elastic Net unit, which during the machine learning process uses both the input vectors and output values of the consecutive measuring cases. Thanks to this, each pixel of the output image has its own unique Alpha and Beta coefficients mentioned in the formula (6).

Autoencoders are a kind of self-organizing networks (SOA). To train them all you need is a set of input data - in this case, measurements taken from the tomographic electrodes. Compared to Elastic Net the training encoder process is poorer because it does not take into account patterns that determine the real object. This fact is the reason why compared to the supervised learning methods, which are exemplified by Elastic Net, the autoencoders give worse results. This is especially important when the reconstructed image is created pixel by pixel, by separately trained modules, each of which generates a single pixel of the image.

One can presume that for images generated in full at once (not pixel by pixel) autoencoders could give better results.

## References

1. T. Rymarczyk, G. Kłosowski, *Ekspluat. Niezawodn.* **20**, 3 (2018)
2. T. Rymarczyk, P. Tchórzewski, P. Adamkiewicz, K. Duda, J. Szumowski, J. Sikora, *Sensors-Basel* **17**, 24 (2017)
3. C. Ferles, Y. Papanikolaou, and K. J. Naidoo, *Neural Networks*, **105** (2018)
4. L. Yang and Y. Hong, *Comput. Stat. Data Anal.* **108** (2017)
5. K. Hu, Q. Cheng, B. Li, and X. Gao, *Biomed. Signal Process. Control*, **39** (2018)
6. G. Kłosowski, T. Rymarczyk, *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska*, **3** (2017)
7. B. Widrow, J.R. Glover, J.M. McCool, J. Kaunitz, C.S. Williams, R. Hearn, J.R. Zeidler, J.E. Dong, R.C. Goodlin, *Proc. IEEE*, **63**, 12 (1975).
8. D.L. Donoho, I.M. Johnstone, *Biometrika*, **81**, 3 (1994)
9. M. Park, D. Kim, H.S. Oh, *IEEE Trans. Instrum. Meas.* **64**, 7 (2015).
10. P. Ching, H. So, S.Q. Wu, *IEEE Trans. Signal Process.* **47**, 10 (1999)
11. C. Park, D. Looney, A. Ahrabian, D.P. Mandic, *IEEE Trans. Neural Syst. Rehab. Eng.* **21**, 1 (2013)
12. D.L. Donoho, De-noising by soft-thresholding, *IEEE Trans. Inf. Theory* **41**, 3 (1995)
13. M. Aminghafari, N. Cheze, J.M. Poggi, *Comput. Stat. Data Anal.* **50**, 9 (2006).
14. F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, H.-T. Wu, *IEEE Signal Process. Mag.* **30**, 6 (2013)
15. A. Ahrabian, D.P. Mandic, *IEEE Trans. Signal Process.* **63**, 9 (2015)
16. T. Rymarczyk, G. Kłosowski, E. Kozłowski, *Sensors-Basel* **18**, 7 (2018)
17. R. Alhamzawi and H. T. M. Ali, *Math. Biosci.* **303** (2018)
18. H. Liu, J. Zhou, Y. Zheng, W. Jiang, and Y. Zhang, *ISA Trans.* **77** (2018)
19. A. Majumdar, *Magn. Reson. Imaging* **52** (2018)
20. X. Zeng, M. R. Leung, T. Zeev-Ben-Mordehai, and M. Xu, *J. Struct. Biol.* **202**, 2 (2018)
21. A. Lechleiter, A. Rieder, *Inverse Probl.* **24**, 6 (2008)
22. G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning with Applications in R* (Springer, New York, 2013).
23. R. Tibshirani, *J. R. Stat. Soc.* **58** (1996)
24. E. Demidenko, A. Hartov, K. Paulsen, *IEEE Transaction on Medical Imaging* **23**, 7 (2004)
25. J. Friedman, R. Tibshirani, T. Hastie, *J. Stat. Softw.* **33**, 1 (2010)
26. R. Wehrens, *Chemometrics with R. Multivariate Data Analysis in the Natural Science and Life Sciences*, (Springer, New York, 2011).
27. H. Zou, T. Hastie, *J. R. Stat. Soc.*, **67**, 2 (2005)