

# Using Statistical Algorithms for Image Reconstruction in EIT

Tomasz Rymarczyk<sup>2,1</sup> and Edward Kozłowski<sup>3</sup>

<sup>1</sup>Research and Development Center, Netrix S.A., Lublin, Poland

<sup>2</sup>University of Economics and Innovation, Lublin, Poland

<sup>3</sup>Lublin University of Technology, Department of Organization of Enterprise, Lublin, Poland

**Abstract.** The problem with image reconstruction from impedance tomography is an ill-posed inverse problem. To get quantitative information on the change in conductivity, it would be better to use a non-linear model in the differential imaging solution. Statistical methods such as PCR, PLRS, elastic net, Lars, SVR were used to reconstruct the image. The discussed techniques can be applied to the problem of electrical tomography. The algorithms used to identify unknown material coefficient.

## Introduction

This article proposes algorithms based on statistical methods to obtain more accurate and stable reconstruction results in solving the inverse problem in electrical impedance tomography (EIT) [3,8,13-16]. In the EIT, the electrical voltages are injected into the object using a set of electrodes attached to the object's surface, and the potentials are measured. The object's conductivity is reconstructed on the basis of known voltages and measured potentials. Reconstruction of electrical impedance requires accurate modelling [6,11,17,18,20]. EIT is a method of imaging in which the conductivity distribution of the tested object is estimated on the basis of measurements of electrical voltages and potentials of electrodes at the boundary. The EIT deals with the inverse problem, in which, taking into account the measured voltages on the electrodes, it estimates the distribution of conductivity by means of the image reconstruction algorithm. To reconstruct the image in the electrical impedance tomography PCR, PLRS, elastic network, Lars, SVR were used [1,4,5,7,9,10,19,21-24]. This work gives promising results as a new horizon to solve practical problems. The main objective of the tomography is to perform image reconstruction. During the measurements, we can see that the measured values from some electrodes are strongly correlated (due to the way of measurement). In this case, we have a multi-line problem [2,12,13].

## 1. Problem definition

Let the linear system be described by the state equation

$$Y = X\beta + \varepsilon, \tag{1}$$

where  $Y \in R^n$ ,  $X \in R^{n \times (k+1)}$  denote the observation matrices of response and input variables respectively,  $\beta \in R^{k+1}$  denotes the vector of unknown parameters.

<sup>2,1</sup> Corresponding author: [tomasz@rymarczyk.com](mailto:tomasz@rymarczyk.com)

When the linear model (1) contains the intercept, then the first column of matrix  $X$  is a column of ones. The object  $\varepsilon \in R^n$  in linear system (1) presents a sequence of disturbances, which is usually defined as a vector of independent identically distributed random variables with normal distribution  $N(\tilde{0}, \sigma^2 I)$ , which,  $\tilde{0} \in R^n$  is a zeros vector but  $I \in R^{n \times n}$  is an identity matrix. The classical Least Square Method depends on identification of unknown parameters  $\beta = (\beta_0, \beta_1, \dots, \beta_k)$  in (1) by solution the task

$$\min_{\beta \in R^{k+1}} \|Y - X\beta\|^2 \tag{2}$$

If  $\det(X^T X) \neq 0$ , then the best linear estimator of unknown parameters  $\beta$  is

$$\hat{\beta} = (X^T X)^{-1} X^T Y \tag{3}$$

The problem is often when  $X^T X$  is singular.

The main goal of tomography is to make an image reconstruction. During measurements we can see, that the measured values from some electrodes are strongly correlated. In this case we have a multicollinearity problem. When the independent variables (predictors) are correlated (collinear), then the matrix  $X^T X$  tends to singular matrix. Using the least square method we obtain the large absolute values of some estimators of unknown parameters. The predictions based on this model is unstable.

The most common approach depends on reduction of set of the input variables (deleting the same predictors which involve in multicollinearity). Then we have a problem of deciding which predictor variables would be

included to regression model. For example, by comparing AIC (Akaike Information Criterion) value for linear models with different set of predictors we may choose a best model.

## 2. PCR and PLSR

From the singular value decomposition (SVD) the data matrix  $X$  can be presented as  $X = UDV^T$

where  $U$  and  $V$  are the unitary matrices,  $D$  is a diagonal matrix with non-negative real numbers on the diagonal. The matrix of scores is defined as the matrix of left singular vectors of  $X$  multiplied with diagonal matrix of singular values  $T = UD$ , but the matrix of loadings is a matrix of the right singular vectors and  $P = V$ . In PCR we use the scores to explain the linear regression (1), because the scores are orthogonal, so there is no multicollinearity problem. Analysing the Root Mean Square Error of prediction we usually approximate the data matrix  $X$  by  $a$ - number of principal components

$$X \approx (U_{(a)}D_{(a)})V_{(a)}^T = T_{(a)}P_{(a)}^T \quad (4)$$

Next, we regress the response variable  $Y$  on the scores  $T_{(a)}$  instead on  $X$ . From (3) and unitarily the unknown parameters in linear model (1) are estimated as

$$\hat{\beta} = P_{(a)}(T_{(a)}^T T_{(a)})^{-1} T_{(a)}^T Y = P_{(a)}(D_{(a)}U_{(a)}^T U_{(a)}D_{(a)})^{-1} D_{(a)}U_{(a)}^T Y = V_{(a)}D_{(a)}^{-1}U_{(a)}^T Y$$

(5)

In practice PCR is performed on the mean centered data matrix  $X$  and the intercept is calculated afterwards. The prediction based on model (1) is

$$\hat{Y} = X\hat{\beta}, \quad (6)$$

where  $\hat{\beta}$  is a vector of estimators of unknown parameters  $\beta$  is given by formula (5).

Just like PCR, the Partial Least Square Regression defines the latent variables to explain the relations in linear model (1). In PCR the principal components are chosen to explain the variance of data matrix  $X$  and come from singular value decomposition. In PLSR the components (called the latent variables) are chosen to capture as much as possible the covariance between  $X$  and  $Y$ . Like PCR, PLSR is performed on the mean centered data matrix  $X$  and the intercept is calculated afterwards. The latent variables are calculated iteratively. The PLSR algorithm will be presented below.

1. First, we put  $E_1 = X$ ,  $F_1 = Y$  and  $i = 1$ .

2. Let  $w_i$  is a first left singular vector of SVD of the cross-product matrix  $E_i^T F_i$
3. Determine the normalized score  $t_i = \frac{\tilde{t}_i}{\sqrt{\tilde{t}_i^T \tilde{t}_i}}$ , where  $\tilde{t}_i = E_i w_i$ .
4. The  $i$ -th loadings of  $X$  and  $Y$  are obtained by regressing  $E_i$  and  $F_i$  against the score  $t_i$   
 $p_i = E_i^T t_i$  and  $q_i = F_i^T t_i$ .
5. The information related to the latent variable based on score  $t_i$  is subtracted from data matrices  $E_i$  and  $F_i$   $E_{i+1} = E_i - t_i p_i^T$  and  $F_{i+1} = F_i - t_i q_i^T$ .
6. Put  $i = i + 1$  and go to step 2.

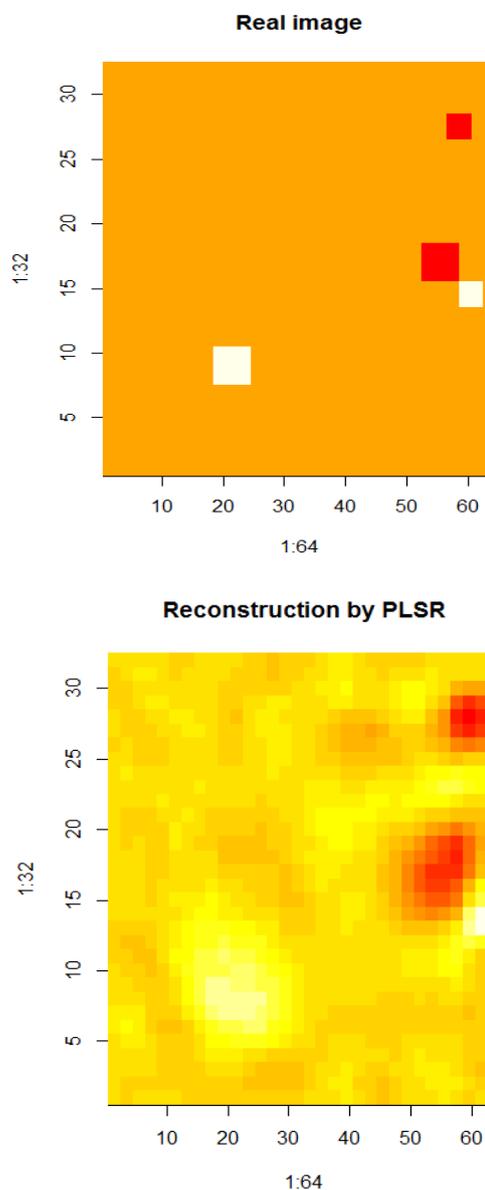


Fig. 1. PLSR image reconstruction

After every iteration the vectors  $w_i, t_i, p_i, q_i$  are saved as a  $i$ -th columns in matrices  $W, T, P$  and  $Q$  respectively. The number of latent variables is determined by analysing the root mean square error of prediction. The matrix

$$R = W(P^T W)^{-1} \quad (7)$$

is generalized inverse of  $P^T$  and the matrix of  $X$  scores satisfies the equation  $T = XR$ . From above we have  $X = TP^T$  and  $Y = TQ^T$ . Thus, we have this same situation as in PCR. Next, we regress the response variable  $Y$  on the scores  $T$  instead on  $X$  and consider the equation

$$Y = X\beta + \varepsilon = T(P^T \beta) + \varepsilon.$$

From (3) we have

$$P^T \hat{\beta} = (T^T T)^{-1} T^T Y.$$

Finally, from (7) we obtain the estimators of unknown parameters

$$\hat{\beta} = R(T^T T)^{-1} T^T Y = R(T^T T)^{-1} T^T T Q^T = R Q^T \quad (8)$$

The image reconstruction is shown in Fig.1.

### 3. Elastic net

Another way to determine the linear regression when the input variables are collinear depends on solution the task

$$\min_{(\beta_0, \beta') \in \mathbb{R}^{k+1}} \frac{1}{2n} \sum_{i=1}^n (y_i - \beta_0 - x_i \beta')^2 + \lambda P_\alpha(\beta'), \quad (9)$$

where  $x_i = (x_{i1}, \dots, x_{ik})$ ,  $\beta' = (\beta_1, \dots, \beta_k)$  for  $1 \leq i \leq n$  and  $P_\alpha$  is an elastic net penalty given by

$$P_\alpha(\beta') = (1 - \alpha) \frac{1}{2} \|\beta'\|_{L_2}^2 + \alpha \|\beta'\|_{L_1} = \sum_{j=1}^k \left( \frac{1 - \alpha}{2} \beta_j^2 + \alpha |\beta_j| \right) \quad (10)$$

We see that the penalty is a linear combination of norms  $L_1$  and  $L_2$  of unknown parameters  $\beta'$ . The introduction the penalty function dependent from parameters to the objective function allows to shrink the estimators of unknown parameters.

The parameter  $\lambda$  in task (9) denotes the coefficient of penalty, but the parameter  $0 \leq \alpha \leq 1$  creates the compromise between LASSO (Least Absolute Shrinkage and Selection Operator) and ridge regression. The ridge regression ( $\alpha = 0$ ) is called Tikhonov regularization and is one of the most commonly used for regularization of linear models. LASSO ( $\alpha = 1$ ) was introduced by Robert Tibshirani. This method performs the variable selection and regularization in linear statistical models.

For the ridge regression the penalty is calculated in norm  $L_1$  but for LASSO in  $L_2$ . Difference between ridge regression and LASSO is symbolic, only the norms are changed. The ridge regression shrinks coefficients for correlated predictors towards each other. When the correlated predictors depend on any latent factor, then ridge regression allows to uniformly distribute the strength of latent factor on these predictors. Whereas LASSO is indifferent to correlated predictors. This method allows to determine the preferred predictor and to ignore the rest. By applying LASSO method, we obtain a model, where the many coefficients to be close to zero, and as a result we receive a sparse model. The elastic net is a connection of ridge regression and LASSO. Choosing the appropriate  $\alpha$  we may create the compromise between ridge regression and LASSO.

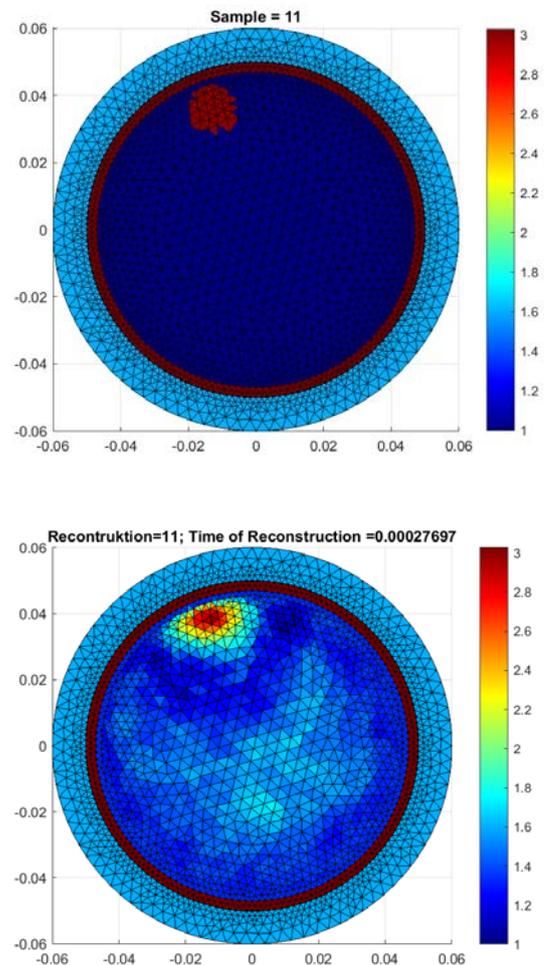


Fig. 2. Image reconstruction by the elastic net.

By solution the task (9) for fixed  $\lambda$  and  $\alpha$  we estimate the unknown parameters of linear system (1), where predictors are correlated. Then the prediction based on model (1) is given by formula (6), where the vector of estimators of unknown parameters  $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$  is estimated by solution the task (9).

Fig. 2 presents the image reconstruction by the elastic net.

#### 4. Lars

Next possible way of reduction of multicollinearity problem between predictors depends on applying Least Angle Regression algorithm. This algorithm includes to linear model only causal variables should be included (from a set of predictors such input variables should be chosen which influence directly to response variable). In this case the linear model is built by employing the forward stepwise regression, where at each step the best variable is added to model (Fig. 3).

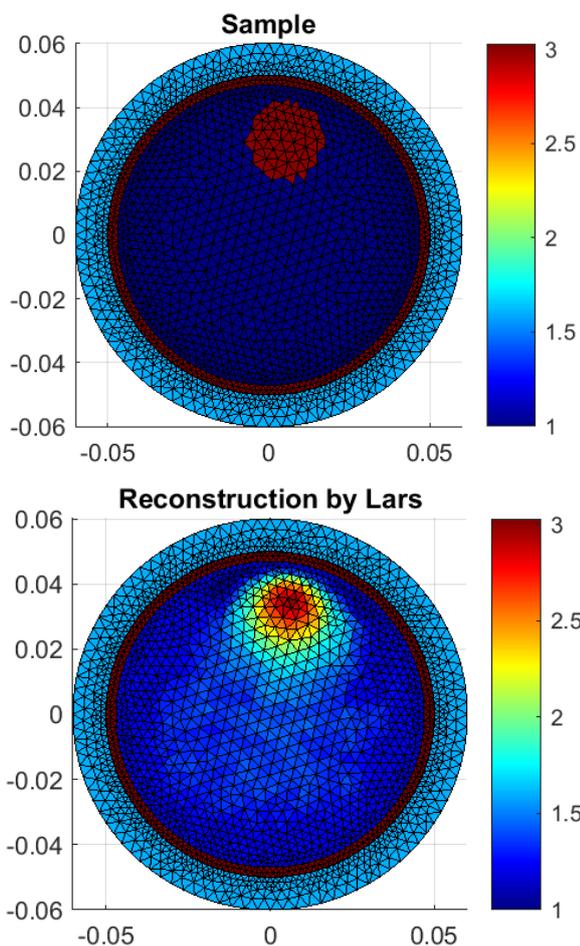


Fig. 3. Image reconstruction by LARS.

Below we present a short version of algorithm of Least Angle Regression.

1. The predictors should be standardized. The intercept  $\beta_0$  in expression (1) is equal a mean of response variable and we put  $\beta_1 = \beta_2 = \dots = \beta_k = 0$ . Active set  $A$  (set of predictors) is empty.
2. Calculate the residuals  $r = Y - \beta_0 - X_{(A)}\beta_{(A)}$  for the linear model with all predictors from active set  $A$ . Determine the predictor  $X_j$  (which is not in

active set) most correlated with residuals  $r$  and attach to the active set  $A$ .

3. Move coefficient  $\beta_j$  from 0 towards its least-squares coefficient  $\langle X_j, r \rangle$  until some other competitor  $X_k$  has a much correlation with the current residuals as does  $X_j$ .
4. Move  $\beta_j$  and  $\beta_s$  in the direction defined by their joint least square coefficient of the current residual on  $\langle X_j, X_s \rangle$  until some other competitor  $X_l$  has a much correlation with the current residual.
5. Go to step 2 and continue in this way until all  $k$  predictors have been entered.

#### 5. SVR

Let  $D = \{(x_i, y_i) : x_i \in R^m, y_i \in R, 1 \leq i \leq n\}$  means a set of training data. The regression spread on support vectors (Support Vector Machine for Regression) consists in determining the regression coefficients defined by the formula (10). Let  $\|y - f_h(x)\|_\varepsilon = \max\{0, \|y - f_h(x)\| - \varepsilon\}$

means the loss intensity function above  $\varepsilon > 0$ , i.e. we omit losses below  $\varepsilon$ . To determine the regression, we solve the task

$$\min_{\beta, \beta_0, \varepsilon_1, \dots, \varepsilon_n} \frac{\|\beta\|^2}{2} + \frac{C}{n} \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (11)$$

with limitations

$$\begin{cases} f_h(x_i) - y_i \leq \varepsilon - \xi_i & \text{dla } 1 \leq i \leq n, \\ y_i - f_h(x_i) \leq \varepsilon - \xi_i^* & \text{dla } 1 \leq i \leq n, \\ \xi_i, \xi_i^* > 0 & \text{dla } 1 \leq i \leq n. \end{cases} \quad (12)$$

We are looking for a solution where the coefficients are in form

$$\beta = \sum_{i=1}^n (\alpha_i^* - \alpha_i) h(x_i). \quad (13)$$

From the formula (13) we can present a hyperplane in the form

$$f_h(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) \langle h(x_i), h(x) \rangle + \beta_0 \quad (14)$$

Dual task to task (11) with restrictions (12)

$$\max_{\alpha_1, \dots, \alpha_n} \sum_{i=1}^n (\alpha_i^* - \alpha_i) y_i - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) k(x_i, x_j) - \varepsilon \sum_{i=1}^n (\alpha_i^* + \alpha_i) h(x_i) \quad (15)$$

with limitations

$$\begin{cases} 0 \leq \alpha_i \leq \frac{C}{n} & \text{dla } 1 \leq i \leq n \\ \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0. \end{cases} \quad (16)$$

Fig. 4. Shows the image reconstruction by SVR.

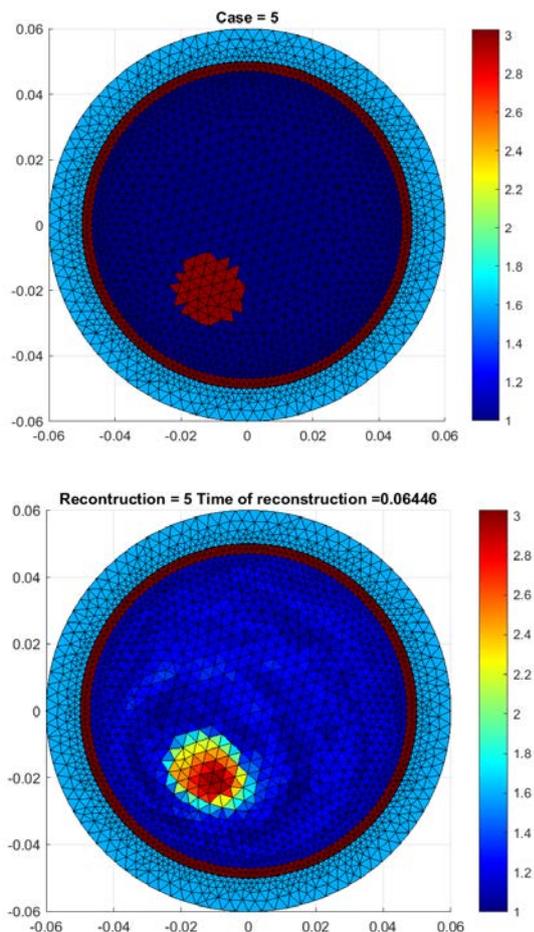


Fig. 4. Image reconstruction by SVR.

## Conclusion

This article proposes algorithms based on statistical methods to obtain more accurate and stable reconstruction results in solving the problem of inverse EIT. We can see that the algorithm allows us to estimate the parameters adequate for linear models that describe linear relationships between conductivity and voltage measurements on electrodes. The method presented in the article allows us to accurately predict conductivity for image reconstruction. An effective algorithm for solving inverse problems would also improve many numerical efficiencies. The reconstruction process is sufficient because the region borders are almost at the edges of the objects being searched. The presented method has been successfully applied in many areas of scientific modelling in the EIT.

## References

1. B. Efron, T. Hastie, I. Johnstone, R. Tibshirani, *Least Angle Regression*, *The Annals of Statistics*, **32** (2), 407-409, (2004).
2. P. Fiderek, T. Jaworski, R. Banasiak, J. Nowakowski, J. Kucharski, R. Wajman, *Intelligent system for the two-phase flows diagnosis and control on the basis of raw 3D 3D ECT data*, *Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Środowiska*, **7** (1), 17-23, (2017).

3. S.F. Filipowicz, T. Rymarczyk, *Measurement Methods and Image Reconstruction in Electrical Impedance Tomography*, *Przełąd Elektrotechniczny*, **88** (6), 247-250, (2012).
4. J. Friedman, Tibshirani R., Hastie T., *Regularization paths for generalized linear models via coordinate descent*, *Journal of Statistical Software*, **33** (1), (2010).
5. T. Fonseca, L. Goliatt, L. Campos, F. Bastos, L. Barra, R. Santos, *Machine Learning Approaches to Estimate Simulated Cardiac Ejection Fraction from Electrical Impedance Tomography*, *Advances in Artificial Intelligence - IBERAMIA 2016*, 235-246, (2016).
6. H. Garbaa, L. Jackowska-Strumiłło, K. Grudzień, A. Romanowski, *Application of electrical capacitance tomography and artificial neural networks to rapid estimation of cylindrical shape parameters of industrial flow structure works*, *Archives of Electrical Engineering* **65** (4), 657-669, (2016).
7. T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, Springer, New York, (2009).
8. J. Hoła, Z. Matkowski, K. Schabowicz, J. Sikora, K. Nita, S. Wójtowicz, *Identification of moisture content in brick walls by means of impedance tomography*, *COMPEL-The international journal for computation and mathematics in electrical and electronic engineering*, **31** (6), 1774-1792, (2012).
9. G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning with Applications in R*, Springer, New York, (2013).
10. G. Kłosowski, E. Kozłowski, A. Gola, *Integer linear programming in optimization of waste after cutting in the furniture manufacturing*, *Advances in Intelligent Systems and Computing 2018*; **637**, 260-270, (2018).
11. J. Kryszyn, P. Wróblewski, M. Stosio, D. Wanta, T. Olszewski, W. Smolik, *Architecture of EVT4 data acquisition system for electrical capacitance tomography*, *Measurement: Journal of the International Measurement Confederation*, **101**, 28-39, (2017).
12. K. Polakowski, S.F. Filipowicz, J. Sikora, T. Rymarczyk, *Quality of imaging in multipath tomography*, *Przełąd Elektrotechniczny*, **85** (12), 134-136 (2009).
13. T. Rymarczyk, P. Tchórzewski, P. Adamkiewicz, K. Duda, J. Szumowski, J. Sikora, *Practical Implementation of Electrical Tomography in a Distributed System to Examine the Condition of Objects*, *IEEE Sensors Journal*, **17** (24), 8166-8186, (2017).
14. T. Rymarczyk, *New Methods to Determine Moisture Areas by Electrical Impedance Tomography*, *International Journal of Applied Electromagnetics and Mechanics*; **52**, 79-87, (2016).
15. T. Rymarczyk, P. Adamkiewicz, K. Duda, J. Szumowski, J. Sikora, *New Electrical Tomographic Method to Determine Dampness in Historical Buildings*, *Archives of Electrical Engineering*, **65** (2), 273-283, (2016).
16. T. Rymarczyk, G. Kłosowski, Application of neural reconstruction of tomographic images in the problem of reliability of flood protection facilities. *Eksploatacja i Niezawodność -- Maintenance and Reliability*, **20** (3), 425-434 (2018).
17. T. Rymarczyk, G. Kłosowski, E. Kozłowski, *A Non-Destructive System Based on Electrical Tomography and Machine Learning to Analyze the Moisture of Buildings*.

- Sensors, **18** (7), 2285; (2018)  
<https://doi.org/10.3390/s18072285>
18. W. Smolik, J. Kryszyn, B. Radzik, M. Stosio, P. Wróblewski, D. Wanta, L. Dańko, T. Olszewski, R. Szabatin, *Single-shot high-voltage circuit for electrical capacitance tomography*, Measurement Science and Technology, **28** (2), 025902, (2017).
  19. R. Tibshirani, *Regression shrinkage and selection via the lasso*, Journal of the Royal Statistical Society, Series B, **58** (1), 267–288, (1996).
  20. R. Wajman, P. Fiderek, H. Fidos, T. Jaworski, J. Nowakowski, D. Sankowski, R. Banasiak, *Metrological evaluation of a 3D electrical capacitance tomography measurement system for two-phase flow fraction determination*, Measurement Science and Technology, **24**, 065302, (2013).
  21. R. Wehrens, B.H. Mevik, *The pls package: Principal Component and Partial Least Square Regression in R*, Journal of Statistical Software, **18** (2), (2007).
  22. R. Wehrens, *Chemometrics with R. Multivariate Data Analysis in the Natural Science and Life Sciences*, Springer, (2011).
  23. H. Zou, T. Hastie, *Regularization and variable selection via the elastic net*, Journal of the Royal Statistical Society, Series B, **67** (2), 301–320, (2005).
  24. Y. Xin, G.S. Xiao Gang, *Linear regression analysis: Theory and computing*, World Scientific, (2009).