JOURNAL OF APPLIED
COMPUTER SCIENCE
Vol. 25 No. 2 (2017), pp. 43-51

Outlier Phenomenon in Data Interpretation for One Waves Scattering Problem

Volodymyr F. Emets, Jan Rogowski

Institute of Information Technology
Lodz University of Technology
Wolczanska Str. 215, 90-924, Lodz, Poland
volodymyr.yemyets@p.lodz.pl, jan.rogowski@p.lodz.pl

Abstract. *An outlier is an observation (or measurement) that is different with respect to the other values contained in a given data set. Outliers can occur due to several causes. The measurement can be incorrectly observed, recorded or processed or otherwise is correctly measured but represents a rare event. In this paper it is shown that observed data can contain values that differ from expected ones and can be interpreted as an outlier, but in fact are caused by a specific physical phenomenon.*
Keywords: *Outlier phenomenon, Acoustical waves scattering.*

1. Introduction

Outlying observations may be errors, or they could have been recorded under exceptional circumstances, or belong to another population [1, 2, 3]. Consequently, they do not fit the model well. It is very important to be able to detect these outliers [4, 5, 6, 7, 8, 9]. Outlier detection is related to, but distinct from noise removal and noise accommodation, that both have to deal with unwanted noise in the data. Noise can be defined as a phenomenon in data which is not of interest to the analyst, but acts as an obstacle to data analysis. Noise removal is dictated by the need to remove the unwanted objects before any data analysis is performed on the data.

Removing such errors can be important in data mining and data analysis tasks. However, removing noise can lead to the removing of the important data.

Scattering of acoustic or electromagnetic waves by several objects has important applications in remote sensing, non-invasive diagnostics in medicine and non-destructive testing. The received signal can be used to determine some of the geometrical and physical properties of the scatterer. Solutions for recognised problems including half-plane, cylinder or sphere are apparently essential regarding the diffraction theory. The strip is considered to be one of the most important familiar structures due to its geometry, strips are usually accustomed to investigate the multiple diffraction phenomenon.

In this paper we consider two acoustic waves scattering problems: wave scattering by a hard strip and scattering by a hard partially debonded strip. It is shown that the observed total cross-section (TCS) data for both problems are similar. The corresponding TCS data deviation is proportional to the Gauss error function. This leads to the situation when observed TCS data for a hard partially debonded inclusion can be interpreted as TCS data with noise for a hard inclusion, despite the fact that we are dealing with two different physical problems.

2. Problem Formulation

Let us consider a thin hard plane inclusion which occupies a domain

$$S = \{|x_1| < a,\ -h < x_3 \le 0,\ |x_2| < \infty\},$$

that is located in an acoustic medium. Here h is the inclusion thickness.

A plane, incident wave of the form

$$u^i(\mathbf{x}) = \exp[\mathrm{ik}(\mathbf{l}, \mathbf{x})],\ \mathbf{x} = (\mathrm{x}_1, \mathrm{x}_3) \tag{1}$$

impinges on the inclusion (the time factor of the form $e^{-i\omega t}$ is omitted throughout the analysis, where ω is the circular frequency). Here $\mathbf{l} = (\sin\theta_0, -\cos\theta_0)$ is the direction of sounding, k is the wave number and typical wavelength kh satisfies the condition $kh << 1$ (see Fig. 1).

The scattering problem of time harmonic waves is described by the wave equation

$$(\Delta + k^2)u(\mathbf{x}) = 0,\ \mathbf{x} \in \mathrm{R}^2\backslash\mathrm{S}$$

and the following condition along the boundary S of the inclusion:

$$u(\mathbf{x}) = 0,\ \mathbf{x} \in \mathrm{S}. \tag{2}$$

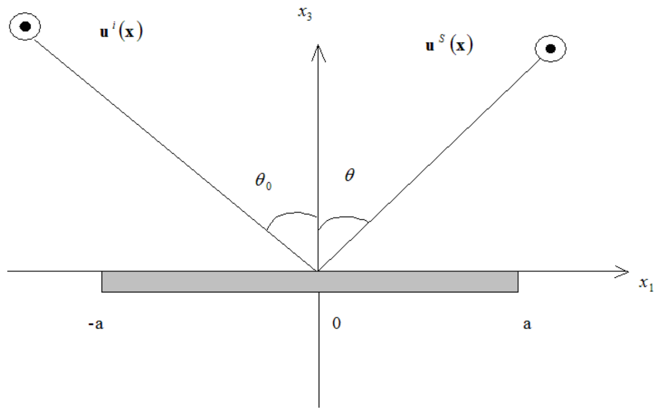

Figure 1: Geometry of the wave scattering by a hard strip problem.

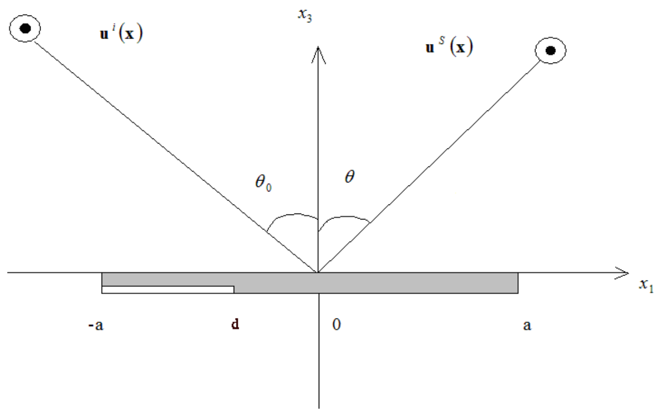

Figure 2: Geometry of the wave scattering by a hard partially debonded strip problem.

The total wave $u = u^i + u^s$ is decomposed into the given incident wave u^i and the unknown scattered wave u^s, which is required to satisfy the Somerfield radiation condition at infinity, from which it follows that

$$u^s(\mathbf{x}) = \frac{e^{ik|\mathbf{x}|+i\pi/4}}{\sqrt{8\pi k\,|\mathbf{x}|}} f(k; \mathbf{l}, \nu), \ |\mathbf{x}| \to \infty, \tag{3}$$

where $f(k;\mathbf{l},\boldsymbol{\nu})$ is the complex amplitude or far-field pattern of the scattering wave, $\boldsymbol{\nu} = \mathbf{x}/|\mathbf{x}| = (\sin\theta, \cos\theta)$ is the direction of observation and TCS is determined as $\sigma(\theta_0) = k^{-1}\mathrm{Im}f(k;\mathbf{l},\mathbf{l})$.

Using Green's theorem, the integral representation of the scattering field can be obtained as

$$u^s(\mathbf{x}) = k\int_{-a}^{a}\left[g(\mathbf{x},\mathbf{y})\,\Phi_1(y_1) - k^{-1}\,\Phi_3(y_1)\,\frac{\partial g(\mathbf{x},\mathbf{y})}{\partial y_3}\right]_{y_3=0} dy_1, \tag{4}$$

$$g(\mathbf{x},\mathbf{y}) = -\frac{i}{4}H_0^{(1)}(k|\mathbf{x}-\mathbf{y}|),\ \mathbf{y} = (y_1, y_3),$$

$$\Phi_3(x_1) = u^+(x_1) - u^-(x_1),\ k\Phi_1(x_1) = \left.\frac{\partial u^+}{\partial x_3} - \frac{\partial u^-}{\partial x_3}\right|_{x_3=0},\ u^\pm(x_1) = u(x_1,\pm 0).$$

Here $H_0^{(1)}$ is the Hankel function of the first kind.

From Eqs. (3) and (4) for the scattering amplitude we have

$$f(k;\boldsymbol{l},\boldsymbol{\nu}) = -k\int_{-a}^{a}\{\Phi_1(y_1) + i\nu_3\Phi_3(y_1)\}\,e^{-ik\nu_1 y_1}dy_1. \tag{5}$$

Let's use the Fourier integral representation of the cylindrical wave $H_0^{(1)}$ through the plane waves:

$$H_0^{(1)}(|\mathbf{x}-\mathbf{y}|) = -\frac{i}{\pi}\int_{-\infty}^{\infty}\frac{e^{\mp(x_3-y_3)\gamma+i\alpha(x_1-y_1)}}{\gamma(\alpha)}d\alpha,\ \gamma(\alpha) = \sqrt{\alpha^2-1}, \tag{6}$$

where radical branch γ is defined by the condition $\mathrm{Im}\gamma < 0$ for $|\alpha| < 1$, sign plus in the formula (6) corresponds to the case $x_3 > y_3$, and sign minus corresponds to the case $x_3 < y_3$. This allows to deal only with symbols of corresponding pseudo-differential operators. As a result, from (1)-(6) we can obtain a singular integral equation relative to $\Phi_1(y_1)$

$$k\int_{-a}^{a}\Phi_1(p)\,K_3(k|x_1-p|)\,dp = -2\exp(ikl_1x_1),\ |x_1| < a, \tag{7}$$

$$K_3(|z|) = -\frac{1}{2\pi}\int_\Gamma \gamma^{-1}(\alpha)\,e^{\pm i\alpha z}d\alpha,$$

$$\sigma(\theta_0) = -\mathrm{Im}\int_{-\mathrm{a}}^{\mathrm{a}}\Phi_1(\mathrm{y}_1)\,\mathrm{e}^{-\mathrm{ikl_1y_1}}\mathrm{dy}_1,$$

where the contour Γ coincides with the real axis everywhere except for the branching points $\alpha = \pm 1$. The contour Γ passes these points below in the right-hand semi-plane of complex variable α and above in the left-hand one according to the limiting absorption principle. In addition, the point $\alpha = 0$ is situated below the contour Γ and for $|\alpha| < 1$ the radical $\gamma(\alpha)$ is defined by the condition $\mathrm{Im}\gamma < 0$.

Applying the Wiener-Hopf technique to a solution of the integral equation (7) (since the details on the approximation may be found elsewhere [10], only a brief summary is given here) we have

$$f_g(k; \boldsymbol{l}, \boldsymbol{\nu}) = 4i\cos\theta_0 \frac{\sin x(l_1 - \nu_1)}{l_1 - \nu_1} + O\left(x^{-3/2}\right), \ x = ka, \ x >> 1. \tag{8}$$

Let us assume now that a strip is partially debonded from the surrounding matrix (see Fig. 2). In this case we have the following boundary conditions:

$$\Phi_3(x_1) = 0, \ d < x_1 < a,$$

$$u^+(x_1) = 0, \frac{\partial u^-(x_1)}{\partial x_1} = 0, \ -a < x_1 < d,$$

where $\pm$ denote the upper and lover faces of the strip.

For this problem we can obtain a system of hypersingular integral equations for determination of $\Phi_1(x_1)$ and $\Phi_3(x_1)$ as follows [11]:

$$\Phi_1(x_1) + k\int_{-a}^{a} \Phi_3(p)\, K_1(k|x_1 - p|)\, dp = q_1 \exp(ikl_1 x_1), \ -a < x_1 < d, \tag{9}$$

$$\Phi_3(x_1) + k\int_{-a}^{a} \Phi_1(p)\, K_3(k|x_1 - p|)\, dp = q_3 \exp(ikl_1 x_1),$$

$$k\int_{-a}^{a} \Phi_1(p)\, K_3(k|x_1 - p|)\, dp = -2\exp(ikl_1 x_1), \ d < x_1 < a,$$

$$K_1(|z|) = \frac{1}{2\pi}\int_{\Gamma} \gamma(\alpha)\, e^{\pm i\alpha z} d\alpha, \ q_1 = -2i\cos\theta_0, \ q_3 = -2.$$

The scattering amplitude has a form

$$f(k; \boldsymbol{l}, \boldsymbol{\nu}) = -k\int_{-a}^{a} \Phi_1(y_1) + e^{-ik\nu_1 y_1} dy_1 - k\nu_3 \int_{-a}^{d} \Phi_3(y_1)\, e^{-ik\nu_1 y_1} dy_1. \tag{10}$$

The asymptotic expansion of solutions of integral equation (9) we seek in the form $\Phi_\beta(x_1) = \Phi_\beta^+(x_1), \beta = 1, 3$ for $k(a+d) >> 1$ and $\Phi_1(x_1) = \Phi_1^-(x_1)$ for $k(a-d) >> 1$, where function $\Phi_\beta^+(\eta) = \Phi_\beta(-a+\eta k^{-1})e^{ixl_1}$ and $\Phi_1^-(\eta) = \Phi_1(a-\eta k^{-1})e^{-ixl_1}$ satisfy the convolution equations

$$\Phi_1^+(\eta) + \int_0^\infty \Phi_3^+(\zeta)K_1(|\eta-\zeta|)d\zeta = q_1 \exp(i\eta l_1),\ 0 < \eta < \infty, \tag{11}$$

$$\Phi_3^+(\eta) + \int_0^\infty \Phi_1^+(\zeta)K_3(|\eta-\zeta|)d\zeta = q_3 \exp(i\eta l_1),$$

$$\int_0^\infty \Phi_1^-(\zeta)K_1(|\eta-\zeta|)d\zeta = q_3 \exp(i\eta l_1).$$

The Fourier transform can be employed to reduce the integral equations (11) to the Wiener-Hopf equations. As follow from results obtained previously [12] the explicit expression for $\Phi_1^-(x_1)$ and $\Phi_\beta^+(x_1)$ read

$$\Phi_1^-(x_1) = \Phi_1^+(x_1) = q_1 e^{ikx_1 \sin\theta_0},$$

$$\Phi_3^+(x_1) = q_3 D \frac{e^{ik(a+x_1)}}{\sqrt{k(a+x_1)}},$$

$$D = \frac{1}{2\sqrt{2\pi}} e^{i\pi/4} e^{-ixl_1} \left[\frac{\cos\phi}{1-\nu_1}\cos\theta_0 + \sin\phi\right],\ \phi = 1/4(\pi/2+\theta_0),$$

where D is the diffraction coefficient at the left inclusion end.

Thus for the scattering amplitude $f(k; \boldsymbol{l}, \boldsymbol{\nu})$ (10) as $k(a+d) >> 1$ and $x >> 1$ we have

$$f(k; \boldsymbol{l}, \boldsymbol{\nu}) = f_g(k; \boldsymbol{l}, \boldsymbol{\nu}) + 4i\nu_3 D e^{ix\nu_1} \frac{1}{\sqrt{1-\nu_1}} \int_0^{\sqrt{x(1+\delta)(1-\nu_1)}} e^{it^2} dt,\ \delta = d/a. \tag{12}$$

3. Results

On Figs. 3 and 4 the frequency dependence of the normalised TCS $\sigma^* = \sigma(0)/2a$ for $\delta = 0$ and $\delta = -0.5$ correspondingly are plotted. The solid curve presents the numerical results obtained from Eq. (7) using the complete system of the Chebyshev polynomials of the first kind to determine the unknown function $\Phi_1(x_1)$ [10].

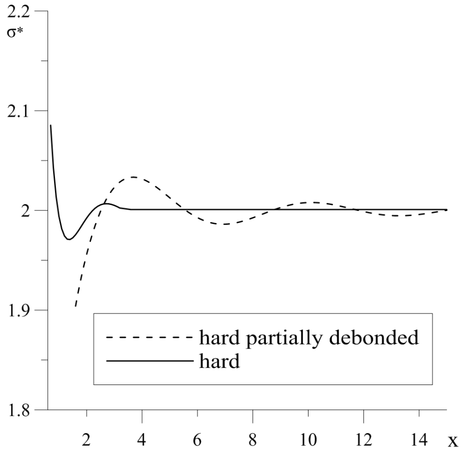

Figure 3: The normalised TCS versus x for $\theta_0 = 0$ and $\delta = 0$.

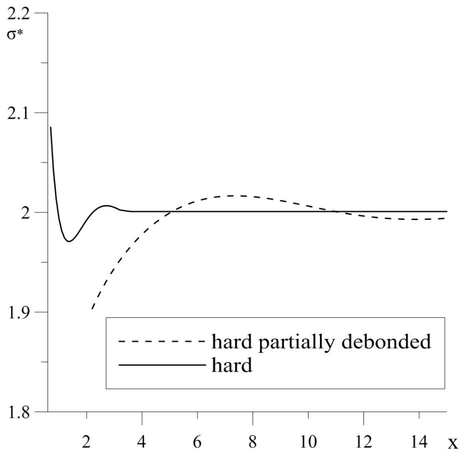

Figure 4: The normalised TCS versus x for $\theta_0 = 0$ and $\delta = -0.5$.

At the same time $\sigma^* = 2$ for $x >> 1$ as follows from the Eq. (8). The corresponding numerical results (dashed curve) for partially debonded inclusion are obtained using the formula (12). It is easy to notice that these results with high probability can be interpreted as one-dimensional dataset with some noise for a hard inclusion in high frequency domain. At the same time, these deviations are the contribution of the edge waves in TCS data for the debonded inclusion.

4. Conclusion

In many signal processing applications the noise is irrelevant or erroneous data. Usually noise is the result of an imperfect data collection process. Data derived from sensors may contain measurement errors. Removing the noise is an important task in data cleaning process as noise hampers most types of data analysis. However, in some applications the outlying data can be interpreted as the noise, but in reality are caused by a specific physical phenomenon.

In this paper the study of high-frequency scattering of acoustic plane waves by a hard strip and debonded hard strip is used to show that the one-dimensional datasets (TCS data) for both problems are similar in the high frequency domain and the use of different filters in signal processing for removing outliers and smoothing the input data is not always justified.

References

[1] Aggarwal, C. C., *Outlier Analysis, 2nd ed.*, IBM T. J. Watson Research Center, Yorktown Heights, New York, 2016.

[2] Angiulli, F. and Pizzuti, C., *Fast outlier detection in high dimensional spaces*, In: Principles of Data Mining and Knowledge Discovery, Lecture Notes in Artificial Intelligence, vol. 2431, edited by T. Elomaa, H. Mannila, and H. Toivonen, Springer, Berlin, Heidelberg, 2002.

[3] Cateni, S., Colla, V., and Vannucci, M., *Outlier detection methods for industrial applications*, In: Advances in robotics, automation and control, edited by J. Arámburo and A. R. Treviño, I-Tech, Vienna, 2008.

[4] Loureiro, A., Torgo, L., and Soares, C., *Outlier Detection Using Clustering Methods: a Data Cleaning Application*, In: Proceedings of the data mining for business workshop, 2004.

[5] Pahuja, D. and Yadav, R., *Outlier Detection for Different Applications: Review*, International Journal of Engineering Research and Technology, Vol. 2, No. 3, 2013.

[6] Yu, D., Sheikholeslami, G., and Zhang, A., *FindOut: Finding outliers in very large datasets*, Knowledge and Information Systems, Vol. 4, No. 4, 2002.

[7] Duraj, A. and Szczepaniak, P., *Information Outliers and Their Detection*, In: Information Studies and the Quest for Transdisciplinarity, edited by M. Burgin and W. Hofkirchner, World Scientific Publishing Company, 2017.

[8] Chomątek, L. and Duraj, A., *Multiobjective Genetic Algorithm for Outliers Detection*, In: IEEE International Conference on INnovations in Intelligent SysTems and Applications INISTA 2017, Gdynia, Poland, 3-5 July, 2017.

[9] Smoliński, M., *Resolving Classical Concurrency Problems UsingAdaptive Conflictless Scheduling*, In: IEEE International Conference on INnovations in Intelligent SysTems and Applications INISTA 2017, Gdynia, Poland, 3-5 July, 2017.

[10] Emets, V. F. and Rogowski, J., *Mathematical-numerical modeling of ultrasonic scattering data from a closed obstacles and inverse analysis*, Academic Publishing House EXIT, Warsaw, 2013.

[11] Emets, V. F. and Rogowski, J., *Scattering from a strip with PEC and partially PMC boundaries*, In: 18th International Conference on Computational Problems of Electrical Engineering, 2017.

[12] Emets, V. F. and Rogowski, J., *Diffraction by a Strip with Different Boundary Conditions on Its Surfaces*, In: 16th International Conference on Computational Problems of Electrical Engineering, 2015.