TECHNICAL SCIENCES
Abbrev.: Techn. Sc., No 10, Y 2007

DOI 10.2478/v10022-007-0008-2

AN ALGORITHM FOR HANDLING EXPERIMENTAL DATA OF PERIODICAL PROCESSES WITH MICROSOFT EXCEL

Jurij Dobriański, Michał Duda

Chair of Electrics and Power Engineering
University of Warmia and Mazury in Olsztyn

Key words: periodical processes, data filtering, data handling, experimental data, periodical processes, measurement noise.

Abstract

This paper describes an algorithm for handling experimental data of periodical processes with Microsoft Excel. Thanks to this method, it is possible to prepare experimental results for analysis with the use of simple, easily available tools. The algorithm filters data and removes what is called measurement noise. This makes it possible to more precisely identify the trends in the analysed data.

ALGORYTM OPRACOWANIA W EXCELU DANYCH EKSPERYMENTALNYCH PROCESÓW OKRESOWYCH

Jurij Dobriański, Michał Duda

Katedra Elektrotechniki i Energetyki
Uniwersytet Warmińsko-Mazurski w Olsztynie

Słowa kluczowe: procesy okresowe, filtrowanie danych, opracowanie danych, dane eksperymentalne, procesy okresowe, szum pomiarowy.

Streszczenie

Opisano algorytm opracowania w Excelu danych eksperymentalnych procesów okresowych. Zaprezentowana metoda pozwala, za pomocą prostych, łatwo dostępnych narzędzi, przygotować wyniki pomiarowe do analizy. Algorytm filtruje dane, usuwając tak zwany szum pomiarowy. Dzięki temu staje się możliwe bardziej precyzyjne zauważenie trendów analizowanych danych.

Introduction

The results of measurements obtained in experimental practice are frequently distorted by measurement noise. The noise is caused by external interfering factors or by excessive sensitivity of the recording equipment. Correct filtering of measurement results enables the retrieval of data, which are sometimes obtained at quite a high cost. Conventional methods of filtering measurement results are onerous and their outcome depends on subjectively adopted parameters (TALAR i in. 2003).

Analysis of such measurement data, obtained in experiments, is frequently onerous, and sometimes impossible. Properly conducted pre-analysis of such data, based on filtering algorithms, should reveal more or less generalised trends in the data being analysed. Meanwhile, the additional noise contained in the measurements should be eliminated. This means that the de-noised data enable constant differentiation of the interpolated results and their further analysis is much easier (RAUCH i in. 2004).

There are currently many techniques of data filtering. Each of the existing algorithms has advantages and disadvantages, but none of them can be considered universal, which means that it cannot be applied in the smoothing of various types of data containing measurement noise of various characteristics (RAUCH, KUSIAK 2005).

The problem of data filtering is in fact one of appropriate approximation of measurement points, which in a one-dimensional case are contained in a curve (BUADES i in.). Currently, there are many algorithms of approximation, the most popular of them are polynomial approximations, weighted average, neuronal networks, wavelet analysis and Fourier transform. Some methods are accurate but require costly calculations (polynomial approximations), others include speed of action among their characteristics (weighted average), though it is impossible to establish the appropriate condition for the algorithm to discontinue. Like trigonometrical functions in the Fourier transform, wavelet analysis is a mathematical function determined within a certain interval. Consequently, they can be hardly comprehensible for less advanced users and their application may be troublesome. Neuronal networks are even more complex, but they provide much better results. However, the crucial flaw of the method is its low degree of universality (KUSIAK i in. 2001).

With an absence of the appropriate filtering measures, a low degree of noise in experimental data, a low frequency of periodical runs, a lack of a universal and simple algorithm of data filtering, it is justified to attempt to apply the simplest logical functions, even those available in spreadsheets.

Formulating the problem

In order to calculate the characteristic features of cyclic processes, it is necessary to determine the minimum and maximum values of parameters which fluctuate periodically during the experiment.

With a considerable number of cycles (at least close to a hundred), handling data placed in measurement tables is best done with the use of computers; this ensures error-free repeated application of the same data handling technique, increases the reliability of the final results and saves the time needed to handle the data.

Obstacles and problems encountered during data handling are caused by the following properties of measurement data sets. The most important of them are:

– so-called “measurement noise”, i.e. random, short-lasting deviations of the parameter values, unrelated to the basic changes occurring in the object. It causes the appearance of small extrema which, when data are handled with the use of a computer, can be treated as the extrema of the object operation, while not being such; they should be eliminated;

– digital data transmission sometimes shows the results of rounding the transmitted values, which is typical of all digital methods and is expressed by replacing close values by identical ones. This, in turn, makes it impossible to determine the trends of a parameter change, makes it difficult to determine the extrema if only a limited number of the elements of a measurement series.

Examples of such properties are shown in Figure 1.

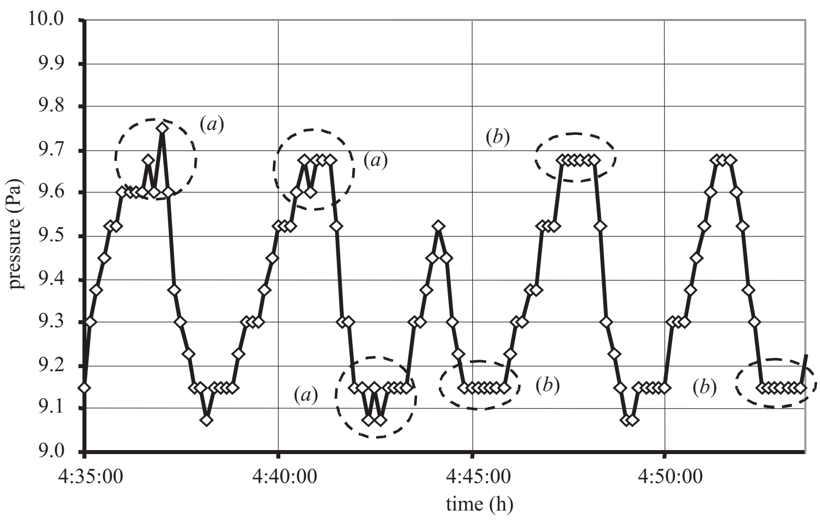

Fig. 1. Examples of additional, “noise-related” maxima and minima (a) and series of identical values (b) in measurement data transmitted with the use of digital techniques

Selection of a data handling method

Selection of data handling method depends not only on the capabilities offered by the software, but also on how advanced the user is in operating it.

Spreadsheets, especially Microsoft Excel, provide high level of calculation capabilities and convenient tools for graphic data presentation. In addition, they are commonly used by many users. Therefore, Microsoft Excel can be a good and easily available tool for experimental data handling. Excel is not intended as a tool for creating complex algorithms; this is related to significant limitations of certain of its functions and should be taken into account when selecting a method of data handling.

One of the significant limitations of Excel is that its logical function IF has only two variants of response: TRUE or FALSE. This in turn makes it necessary to apply additional levels of embedding a function within a function, which is limited to 7 levels of embedding. This affects the overall length of algorithm in a cell. When standard functions are used, it is indeed limited. Nevertheless, Excel's calculation capabilities can ensure the required level of data handling without using specialised programming languages. An example of such an application is presented in this paper.

Algorithm of elimination of minimum and maximum caused by measurement noise

Measurement noise causes sharp increase in the value of the measurement parameter, followed immediately by a return to the values close to the average. In practical terms, it shows as a pair of maximum-minimum which are situated next to or close to each other, as in Figure 1a. Therefore, an algorithm intended for elimination of such interferences should not take into account many consecutive measurement values – it could be two-three values before and after the analysed item. A block diagram of the algorithm, based on the logical function IF, is shown in Figure 2.

The rhombus elements of the block diagram contain the indexes of comparable measurement elements. The index of an analysed cell is marked by i. Over the lines connecting the diagram's elements are symbols (lines and dots). A dot in a symbol means the current position, while lines – possible meanings of other measurement values in relation to the current value, which have been analysed at this place in the algorithm. The following text of the Excel function describes the algorithm:

```
=IF(C9=B10;B10;IF(C9>B10;IF(C8>=C9;IF(C7>=C8;B10;IF(B10>=B11;B10;B9));IF(B10=B11;IF(B11>=B12;B10;B9);IF(B10>B11;B10;B9)));IF(C8<=C9;IF(C7<=C8;B10;IF(B10<=B11;B10;B9));IF(B10=B11;IF(B11<=B12;B10;B9);IF(B10>B11;B9;B10)))))
```

According to the algorithm text, the processed measurement value will be written into the C10 cell. Data for comparison are taken from previous

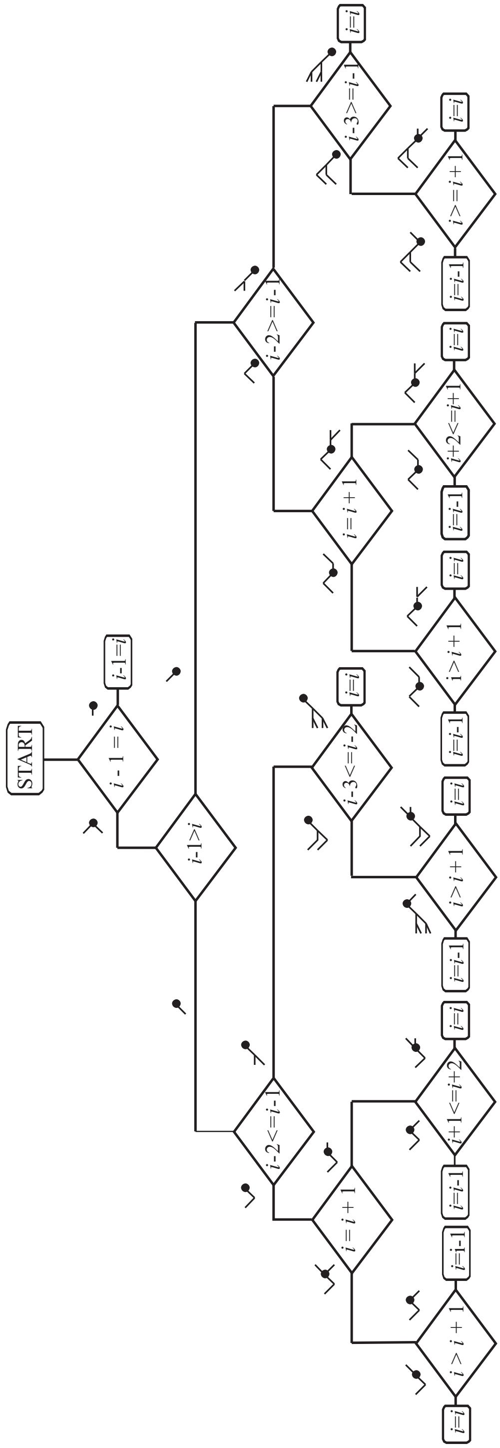

Fig. 2. A block diagram of an algorithm used to eliminate neighbouring minimum and maximum, caused by measurement noise

cells of column C and from several cells of column B, whose values have been determined in previous steps of calculation.

The diagram of the positions of cells from which data are taken is shown

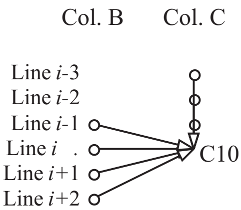

Fig. 3. The diagram of the positions of cells from which data are taken for eliminating the neighbouring minimum and maximum

Table 1

An example of elimination of measurement noise with the use of an algorithm for the 'IF' function in Microsoft Excel

Time, h	The value in the pressure cell before application of the algorithm (Pa)	The value in the pressure cell after application of the algorithm (Pa)
4:35:30	9.450	9.450
4:35:40	9.525	9.525
4:36:00	9.600	9.600
4:36:40	9.675	9.675
4:36:50	9.600	9.675
4:37:00	9.750	9.750
4:37:10	9.600	9.600
4:37:20	9.375	9.375
4:37:30	9.300	9.300
4:37:40	9.225	9.225
4:37:50	9.150	9.150
4:38:10	9.075	9.075
4:38:20	9.150	9.150
4:39:00	9.225	9.225
4:39:10	9.300	9.300
4:39:40	9.375	9.375
4:39:50	9.450	9.450
4:40:00	9.525	9.525
4:40:30	9.600	9.600
4:40:40	9.675	9.675
4:40:50	9.600	9.675
4:41:00	9.675	9.675
4:41:30	9.525	9.525
4:41:40	9.300	9.300
4:42:10	9.150	9.150
4:42:20	9.075	9.075
4:42:30	9.150	9.075

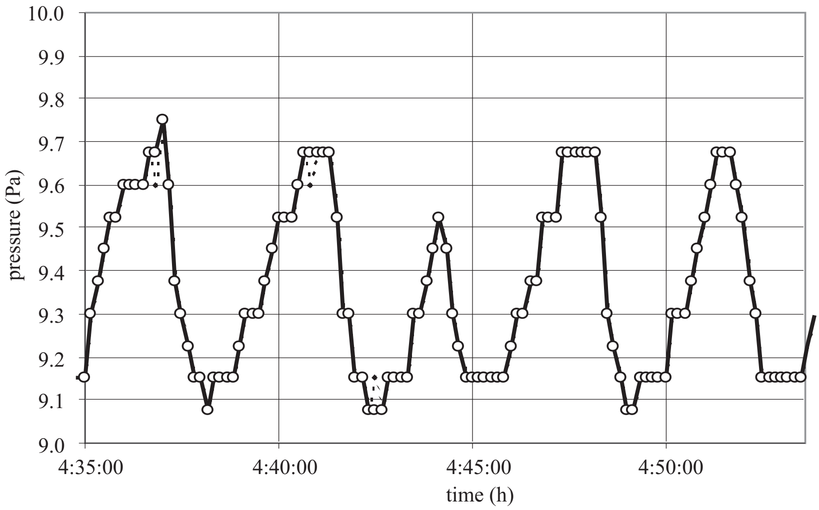

Fig. 4. The result of applying the algorithm of elimination of neighbouring minimum and maximum, caused by measurement noise: the dot line shows the original data

in Figure 3. The calculations produce a column with new, corrected data (Tab. 1). The results of processing, which eliminates the maximum and minimum situated close to each other are shown in Figure 4.

An algorithm for reducing the size of a series of identical measurement values

It is useful to reduce the size of a series of identical measurement values so as to avoid analysing considerable numbers of consecutive values of measurement data. Which in turn causes an increase in the level of embedding of the logical IF functions in one another.

The algorithm provides for checking the character of consecutive measurement data. If a series of values is found of two identical values, then the value on the limit of the series will be replaced with one that is calculated as the mean of the neighbouring cells. In other cases, the values will be assigned with no change. A block diagram of the algorithm is shown in Figure 5; its formula is as follows:

=IF(B9<>B10;IF(B10=B11;(B9+B11)/2;B10);IF(B8=B9;(B9+B11)/2;B10)).

The outcome of the double application of the algorithm is shown in Figure 7.

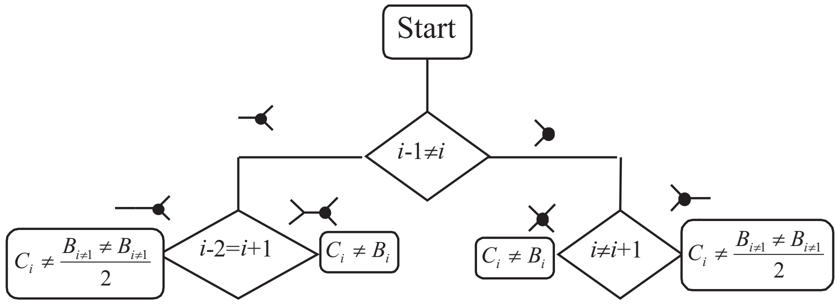

Fig. 5. A block diagram to reduce the size of a series of identical measurement values

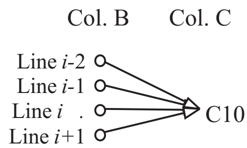

Fig. 6. A diagram of positions of the cells from which data are taken to reduce the size of a series of identical measurement values

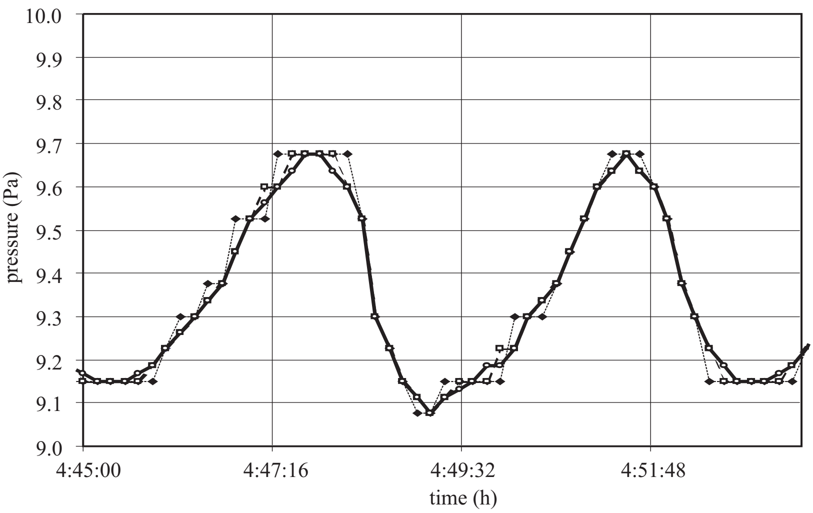

Fig. 7. The outcome of double application of the algorithm for reducing the size of a series of identical measurement values: the stroke line denotes the values before processing

Table 2

An example of applying an algorithm to reduce the size of a series of identical measurement values. The shaded cells indicate the values which have been changed by the algorithm

Time	Pressure	The value of the pressure cell after the application of the algorithm	The value of the pressure cell after the secondary application of the algorithm
1	2	3	4
4:45:00	9,150	9.150	9.169
4:45:10	9,150	9.150	9.150
4:45:20	9,150	9.150	9.150
4:45:30	9,150	9.150	9.150
4:45:40	9,150	9.150	9.169
4:45:50	9,150	9.188	9.188
4:46:00	9,225	9.225	9.225
4:46:10	9,300	9.263	9.263
4:46:20	9,300	9.300	9.300
4:46:30	9,375	9.338	9.338
4:46:40	9,375	9.375	9.375
4:46:50	9,525	9.450	9.450
4:47:00	9,525	9.525	9.525
4:47:10	9,525	9.600	9.563
4:47:20	9,675	9.600	9.600
4:47:30	9,675	9.675	9.638
4:47:40	9,675	9.675	9.675
4:47:50	9,675	9.675	9.675
4:48:00	9,675	9.675	9.638
4:48:10	9,675	9.600	9.600
4:48:20	9,525	9.525	9.525
4:48:30	9,300	9.300	9.300
4:48:40	9,225	9.225	9.225
4:48:50	9,150	9.150	9.150
4:49:00	9,075	9.113	9.113
4:49:10	9,075	9.075	9.075
4:49:20	9,150	9.113	9.113
4:49:30	9,150	9.150	9.131
4:49:40	9,150	9.150	9.150
4:49:50	9,150	9.150	9.188
4:50:00	9,150	9.225	9.188
4:50:10	9,300	9.225	9.225
4:50:20	9,300	9.300	9.300
4:50:30	9,300	9.338	9.338
4:50:40	9,375	9.375	9.375
4:50:50	9,450	9.450	9.450
4:51:00	9,525	9.525	9.525
4:51:10	9,600	9.600	9.600

cont. tab. 2

1	2	3	4
4:51:20	9,675	9.638	9.638
4:51:30	9,675	9.675	9.675
4:51:40	9,675	9.638	9.638
4:51:50	9,600	9.600	9.600
4:52:00	9,525	9.525	9.525
4:52:10	9,375	9.375	9.375
4:52:20	9,300	9.300	9.300
4:52:30	9,150	9.225	9.225
4:52:40	9,150	9.150	9.188
4:52:50	9,150	9.150	9.150
4:53:00	9,150	9.150	9.150

An algorithm for determining the minimum and maximum of the parameters of the process under study

When determining the minimum and maximum, it is necessary to take into account a possibility of remaining a series of identical measurement values at the extreme points of the measurement records. The aim of the function used to determine the minimum and maximum is to ascertain whether the point under examination has the maximum or minimum value as compared with the neighbouring points. It will also check whether the analysed value is situated on an upward or downward section. If the next two neighbouring measurement records have the same value as the one under examination, the analysed point will be marked as ‘unknown’.

The block diagram of the algorithm which determines the extreme values is shown in Fig. 8, and the text of the algorithm function is presented below.

=IF(B9=B10;IF(B11=B10;”0";IF(B11>B10;”g”;”d”));IF(B9>B10;IF(B11=B10;IF(B12= B11;”unknown”;IF(B12>B11;”min”;”d”));IF(B11>B10;”min”;”d”));IF(B11= B10;IF(B12=B11; ”unknown”;IF(B12>B11;”g”;”max”));IF(B11>B10;”g”;”max”))))

A diagram of the cells from which data are taken to determine the extreme values of measurement parameters is shown in Figure 9.

The data processing results in a column with markings in respective lines: min, max, g (upwards), d (downwards) and unknown. The last indicator allows for checking whether there are any series left with repeated (three or more times) data values. The table clearly shows the places where the maxima and minima are situated and enables their calculation.

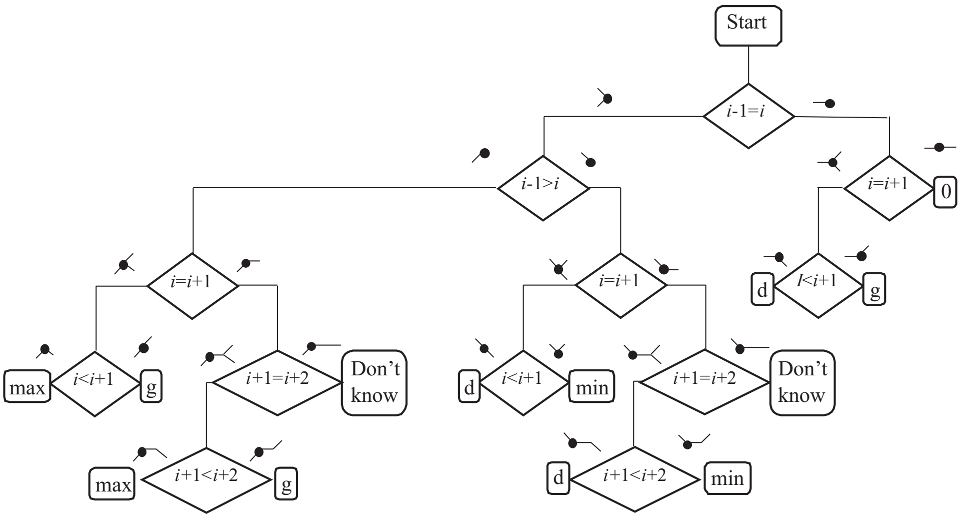

Fig. 8. A block diagram of the algorithm for determining the extreme values of measurement parameters

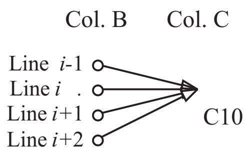

Fig. 9. A diagram of the cells from which data are taken to determine the extreme values of measurement parameters

Table 3

An example of application of an algorithm for determining the minimum and maximum of the process param

Time	Pressure, Pa	Change of Pressure	Min–max	Unknown
0:01:30	7.950	max	7.95	0
0:01:40	7.913	down	0	0
0:01:50	7.875	down	0	0
0:02:00	7.838	down	0	0
0:02:10	7.800	min.	7.8	0
0:02:20	7.875	up	0	0
0:02:30	7.950	max	7.95	0
0:02:40	7.913	down	0	0
0:02:50	7.894	down	0	0
0:03:00	7.875	min.	7.875	0

In order to calculate the periods of parameter change, i.e. the time intervals between the extrema, it is useful to select the lines containing extreme values. To this end, one should apply the 'Data' tool in MS Excel, select the 'Filter' tab, then 'Advanced filter' with options "Copy elsewhere' and 'Only unique records'.

The algorithm proposed here enables finding the minimum and maximum values of the measured parameters; it has been tested and successfully applied in handling the results of experiments to determine the thermal condition of a laboratory model and solar installation in the study conducted at the Faculty of Technical Sciences of the University of Warmia and Mazury in Olsztyn.

Summary

This study is an attempt to analyse measurement data obtained in experimental determination of the thermal condition of a laboratory model and solar installation. The tools contained in Microsoft Excel were used in filtering data. The filtration results presented in tables and figures indicate that the analysis conducted with the use of Microsoft Excel is highly effective in the filtration of data containing interfering signals in the form of measurement noise.

The results show the usefulness of the technique in filtering measurement data and is an example of using Microsoft Excel, while at the same time indicating its high capabilities in processing and analysing large sets of numerical data. Further research should be aimed at assessment of filtration errors.

Reference

TALAR J., RAUCH Ł., KUSIAK J. 2003. *Filtrowanie danych pomiarowych przy wykorzystaniu analizy falkowej i sztucznych sieci neuronowych.* Informatyka w Technologii Materiałów, 3–4: 180–188.

RAUCH Ł., TALAR J., ZAK T., KUSIAK J. 2004. *Filtering of termomagnetic data curve using artificial neural network and warelet analysis.* Proc. 7th ICAISC 2004, Conf. Zakopane, Poland, Springer-Verlag, 1093–1098.

RAUCH Ł, KUSIAK J. 2005. *Filtrowanie danych pomiarowych przy wykorzystaniu metody cząstek dynamicznych*. Przegląd mechaniczny, 7–8: 39.

BUADES A., COLL B., MOREL J.M. *On image denoising methods.* Centre de Matematiques et de Leurs Applications (http://www.cmla.ens-cachan.fr)

KUSIAK J., PIETRZYK M., GAWĄD J., ROBERTS C.M., WAJDA W. 2001. F*iltering of the measurement data by ANN approach.* Proc. Conf. "Computer Science in Metals Technology KomPlasTech'2001", eds, Grosman F., Piela A., Kusiak J., Pietrzyk M., Korbielów, ss. 65–70.

Accepted for print 12.10.2007