LEAK DETECTION IN LIQUID TRANSMISSION PIPELINES USING STATISTICAL ANALYSIS

Andrzej SOBOLEWSKI*, Paweł OSTAPKOWICZ**

* Bialystok University of Technology, Faculty of Electrical Engineering
Wiejska Street 45C, 15-351 Białystok, Poland, e-mail: a.sobolewski@pb.edu.pl
** Bialystok University of Technology, Faculty of Mechanical Engineering
Wiejska Street 45C, 15-351 Białystok, Poland, e-mail: p.ostapkowicz@pb.edu.pl

Abstract

Pipeline leak detection systems based on flow and pressure data from a SCADA are commonly used for supporting maintenance and increasing the reliability and safety of pumping process. Such systems should maintain high sensitivity with minimum number of false alarms over the long term. This paper examines the application of a statistics based leak detection system in a physical model of water pipeline (380-metre-long, 34 mm in internal diameter, made of polyethylene (PEHD)) and compares it to a reference algorithm. Sophisticated statistical analysis techniques constantly evaluate all new and existing information in the database, discard invalid data, and perform leak detection basing only on high-quality information in the current database.

Key words: pipelines, leak detection, statistical analysis.

DETEKCJA PRZECIEKU PŁYNÓW W RUROCIĄGACH ZA POMOCĄ ANALIZY STATYSTYCZNEJ

Streszczenie

Systemy wykrywania przecieków w rurociągach wykorzystujące dane z pomiaru przepływów i ciśnień w układach pomiarowych typu SCADA są często stosowane do wspierania obsługi I zwiększenia niezawodności oraz bezpieczeństwa procesu pompowania. Wymaga się od tego typu systemów wysokiej czułości z jednoczesną odpornością a jednocześnie ograniczenia występowania fałszywych alarmów podczas pracy. W tym artykule przetestowano algorytm detekcji działający wykorzystujący metody analizy statystycznej w fizycznym modelu wodociągu (długość 380 metrów, średnica wewnętrzna 34 mm, materiał polietylen (PEHD)) i porównano go z algorytmem referencyjnym. Zaawansowane techniki analizy statystycznej dokonują ciągłej oceny wszystkich nowych i istniejących informacji w bazie danych, odrzucając dane nieprawidłowe. Celem takiego działania jest dokonywanie detekcji wycieków w wodociągu na podstawie tylko wysokiej jakości informacji w budowanej bazie danych.

Słowa kluczowe: rurociąg, detekcja wycieku, analiza statystyczna.

1. INTRODUCTION

The primary purpose of leak detection systems (LDS) is to assist pipeline controllers in detecting and localizing leaks. LDS alarm and provide leak related data to the pipeline controllers in order to aid the decision-making process. Pipeline leak detection systems are also beneficial because they can enhance productivity and system reliability thanks to reduced downtime and reduced inspection time. Leak detection systems should maintain high sensitivity with minimum number of false alarms over the long term. Most popular leak detection systems are developed with the use of diagnostic methods which are based on measurements of internal flow parameters (flow rate, pressure and fluid temperature) [5, 7, 8]. Elaboration of a leak detection system requires the use of at least several internal methods working concurrently [5]. Systems developed with the use of such methods are usually additional modules of SCADA systems which are used in pipelines for monitoring, regulation and control of pumping processes.

SCADA systems comprise measuring devices and transducers located along the pipeline and data transmission systems which provide parameter measured flow data (process variables). Due to large distances between measurement points and the control of a SCADA system, it is possible that the LDS will only have part of useful measurement data [6]. In addition, the measurement data are often disturbed by additive measurement noise. In many cases, this limits the use of diagnostic methods such as, for example, the method based on pressure wave detection [4, 5]. Effective application of this method requires not only accurate synchronization of measuring time and continuous reception of pressure signals from individual measurement points along the pipeline (with imposed very short sampling periods), but also a reliable and unambiguous measurement. Acquisition of data with additive measurement noise is therefore a challenge for algorithms performing continuous measurement data analysis and often limits their effectiveness. Statistical methods, on the other hand, that can be used to estimate parameters and test hypotheses about them prove to be much more effective in such conditions. The paper proposes statistical methods of leak detection that provide an alternative to the methods based on pressure wave detection.

2. DESCRIPTION OF THE LEAK PHENOMENON

A pipeline usually has flowmeters installed on the ends which measure pump forced flow at the inlet and the flow at the outlet. In case of a leak, juxtaposing flows at the inlet and outlet allows for a diagnostic decision to be made. Moreover, pipelines are often equipped with additional pressure sensors installed at regular intervals along the pipeline. They are installed in order to increase the accuracy of leak detection and thus reduce the number of false alarms.

The occurrence of leakage leads to changes of pressure and flow in the pipeline. Leakage is accompanied by the formation and propagation of pressure waves. These waves are produced by a sudden pressure drop at the place of the leakage and propagate from there in both directions of the pipeline with the speed of sound. In the case of sudden leaks waves have clearly visible fronts, and in the case of leakages increasing slowly, due to a milder character of pressure changes, waves have a smoother shape. Behind the wave front, the longer is the distance from the leak point, the smaller is the pressure drop in the pipeline. Example signals of pressure and flow rate in the pipeline without and with leakage are shown in Figure 1. The signals are measured at the inlet and outlet, and in the case of pressure, additionally at several points along the pipeline.

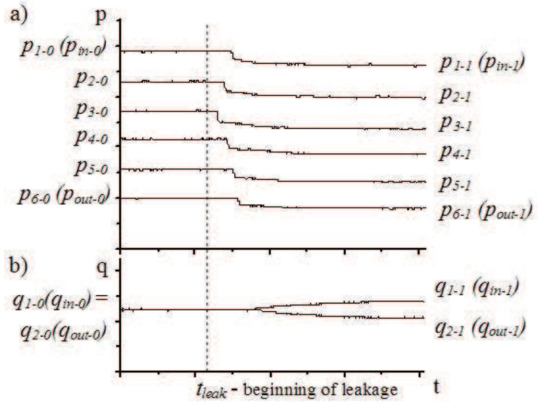

Fig. 1. Pressure profiles a) and flow profiles b) from transducers in the pipeline before and after the occurrence of leakage

For small intensity leaks, when additionally the measurement data are strongly disturbed, the detection of wave front propagation causes a problem. In such situations, it is also difficult to balance the flows. This, however, can be solved using statistical methods which verify hypotheses about changes in the average values of the measured signals basing on the analysis of data collected in n-samples time windows. Fig. 2. shows data disturbed by measurement noise in a time window for one of the signals in the state of a leak at 20th second of data recording.

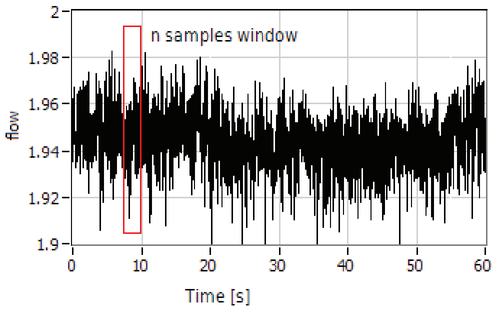

Fig. 2. Data with noise and n-samples window for statistical analysis. Leakage occurs at 20sec

3. SCOPE OF RESEARCH

An important element of this research was to compare the diagnosis of simulated leaks obtained by means of two solutions:

- the reference method. It operates in a continuous cycle and generates diagnostic results each time it receives the signals with a sampling period T_p. It involves the use of pressure and flow signals from two measuring transducers located at the inlet and outlet or, additionally, the use of signals from all pressure transducers located along the pipeline;
- the elaborated statistical-based method. The system receives flow and pressure data from a SCADA (Supervisory Control and Data Acquisition) system and calculates the probability of a leak at regular intervals e.g. every 0.3 seconds. The hypothesis test procedure is used in order to determine signal mean change probability.

2.1. Description of the reference method

A detection algorithm presented in [2] was used. The algorithm is using measurement signals from two devices (sensors) located at the beginning and end of the pipeline (Fig. 3). The signals which were used were pressure signals p_{in} and p_{out} and, for comparison, measured signals of the flow q_{in} and q_{out}. The algorithm includes:

- recursive filtering of signals using a recursive filter with a low-pass characteristics –

recursive averaging with fading memory (exponential smoothing) – the results of which are reference values for the calculation of deviations Δx_{in} and Δx_{out};

- calculation of the maximum of the correlation function of the deviations Δp_{in} and Δp_{out} for the pressure signals, and the minimum of the correlation function of deviations Δq_{in} and Δq_{out} for the flow signals for the time shift $\tau_{\max}{}^{A}$. The shift value $\tau_{\max}{}^{A}$ is determined using the velocity of the pressure wave propagation and taking into account the distance between the pressure measurment points;
- recursive filtering of the calculated extremes of the correlation functions with the resulting values Fx (Fp and Fq respectively) which, when they exceed the alarm threshold Px (Pp and Pq respectively) inform about the occurrence of leakage.

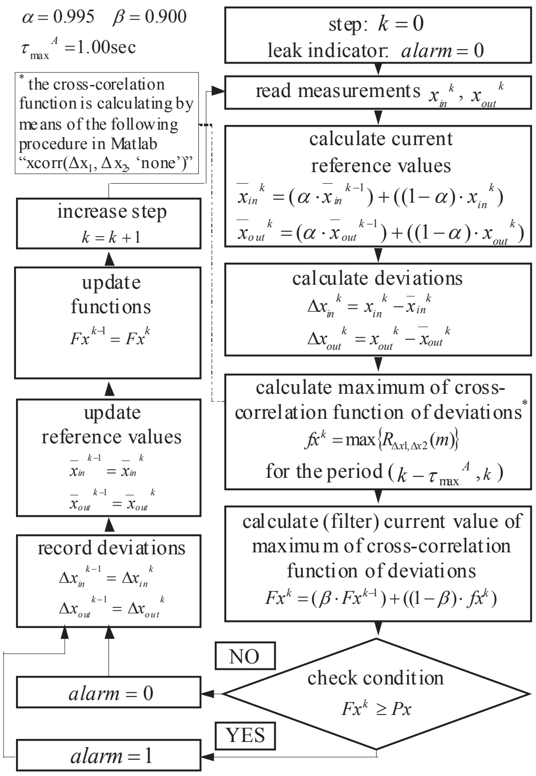

Fig. 3. The reference algorithm

The algorithm can also operate with the use of a larger number of signals. In this case, the pressure signals measured at the inlet and outlet and also in the other points along the pipeline were used. The resulting function here is the function $\overline{Fp}$. This function is the medium value of a set of $n-1$ functions Fp_n based on the calculation of the cross-correlation of the variables Δp_n. Particular functions Fp_n are the result of filtration of the functions fp_n based, in turn, on a recursive filter, referred to as recursive averaging with fading memory (exponential smoothing). The individual functions fp_n are the maximum values of the cross-correlation function which corresponds to the correlation of the variable Δp_n. The variable represents a given measuring point z_n with the variable Δp_{n+1} which, in turn, represents the neighbouring measuring point z_{n+1} located along the pipeline. The cross-correlation functions are calculated with the time shift $\tau_{\max}{}^{B}$. The shift value $\tau_{\max}{}^{B}$ is determined using the velocity of the pressure wave propagation and taking into account the distance between the pressure measuring points. If the function $\overline{Fp}$ exceeds its alarm threshold $\overline{Pp}$, the alarm is generated.

2.2. Description of the elaborated method

The proposed method is based on statistical analysis of pressure/flow signals at several points in the pipeline in order to detect a leak. A leak changes the hydraulics of the pipeline and therefore changes the pressure or flow readings after some time. Local monitoring of flow at the beginning and the end of the pipeline can therefore provide a simple way of detecting leaks. However, in order to minimize false alarms and properly identify changes in the pressure and flow signals when a leak occurs, it is necessary to read data from several points. This gives the opportunity of optimizing leak detection provided that some statistical assumptions hold. The uncertainty of the mean value of the measured data requires the estimation of its real value by calculating confidence intervals (CI) and its limits. A confidence interval for a mean specifies a range of values within which the unknown population parameter, in this case the mean, may lie. CI for means are calculated as follows:

$$P\left(\left|m - \mu_0\right| < u_{\alpha/2} \frac{s}{\sqrt{n}}\right) = 1 - \alpha \tag{1}$$

where:

P	probability value
1-a	confidence level.
s	estimated standard deviation
m	estimated mean
μ_0	unknown population mean
n	number of samples
$u_{\alpha/2}$	normal distribution = 1.96 for confidence level 0.95

Confidence intervals (CI) for means are calculated for each of the measured data intervals. Limits of the interval will be used as a reference for the detection of transducer signal mean change. The (two sided) confidence interval for a mean contains all the values of μ_0 (the true population mean) which would not be rejected in the two-sided hypothesis test of:

H0: m = μ_0 (no significant mean change)
against
H1: m not equal to μ_0 (a significant mean change)

To minimize false alarms, P-value, the strength of evidence in support of a null hypothesis, is averaged over all data sets. It allows setting a custom threshold for leak detection. A leak is detected if the average P-value exceeds the custom threshold (50% in the work). Otherwise, the leakage is not detected.

2.3. Description of the algorithm

The existence of noisy data makes it necessary to choose a time window length, e.g. 0.3 sec, and calculate the confidence interval and its limits for all considered signals. The limits may be used as a start up parameters. The next step is to move the time window and perform a hypothesis test for all signal means, taking as the start up references the CI limits, and then calculate *P*-values as a hypothesis test result. Each signal has its *P*-value. If the average of *P*-values is smaller than the threshold, the current measurements should be added to the previous ones and new confidence limits should be calculated. This operation extends the database so that mean estimation can be performed more precisely in further calculations (Fig. 4). If the average of *P*-values exceeds the threshold, the leak is detected (Fig. 5).

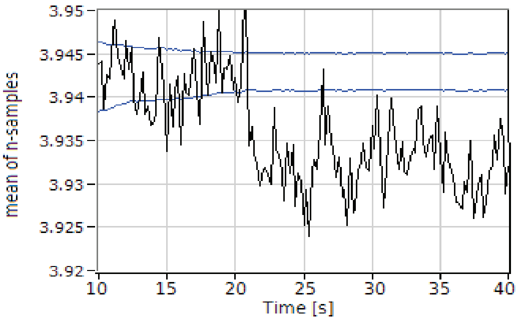

Fig. 4. The n-samples window means and confidential interval limits for a chosen signal. Leakage occurs at 20sec

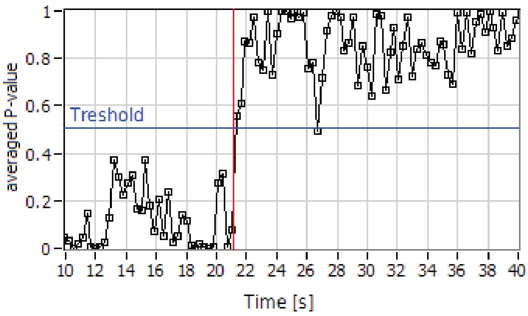

Fig. 5. *P*-value averages and the chosen threshold. Leakage occurs at 20sec

3. VERIFICATION EXPERIMENTS ON A PHYSICAL MODEL OF A WATER PIPELINE

The elaborated technique has been tested on a test stand with a physical model of a water pipeline. The scheme of the test stand is shown in Figure 6, Figures 7 and 8 show the general scheme and the view of the pipeline model.

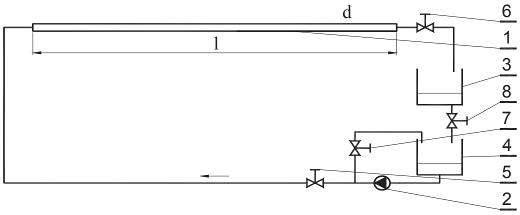

Fig. 6. Laboratory plant scheme1 – pipeline; 2 – pump; 3, 4 – tanks; 5, 6 – stations of control valves; 7 – control valve; 8 – bleeding valve

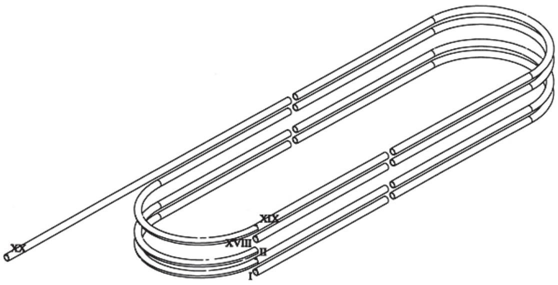

Fig. 7. Scheme of the pipeline
I – the inlet, II – the outlet

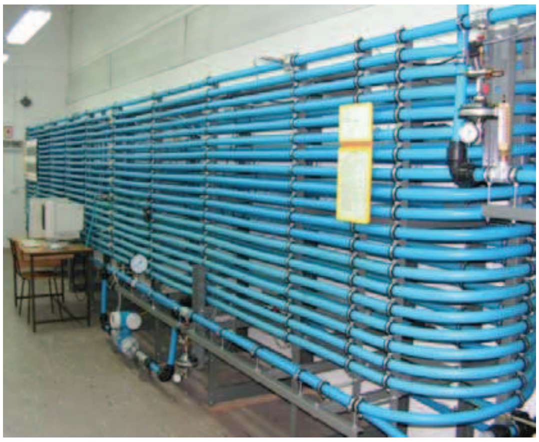

Fig. 8. View of the pipeline

The pipeline is 380 meters long and is made of polyethylene (PEHD) pipes which are 34 mm in internal diameter and 40 mm in external diameter. It consists of three sections each of which is over one hundred meter long. The sections: 0÷140 m, 140÷280 m and 280÷380 m, are joined with the use of special connectors which have the same diameter as the pipeline. The model pipeline is equipped with standard measuring devices: two electromagnetic flow meters, six pressure transducers and two thermometers. They are connected to a PC equipped with a 16 bit A/D converter. Table 1 contains information about the location and meteorological characteristics of the sensors and the measurement system. Proportionally controlled solenoid valves were used to simulate leakage. The valves were installed in several selected points in the pipeline.

Table. 1. Characteristics of the elements of the measuring system

devices	pressure transducers	flow rate transducers
location [m]	$P_{1(in)}=1$; $P_3=61$; $P_4=141$; $P_5=201$; $P_6=281$; $P_{2(out)}=341$	$Q_{1(in)}=-6$; $Q_{2(out)}=382{,}2$;
range	0÷10 [bar]	0÷200 [l/min]
accuracy	0.1 % of range	0.2 % of range
uncertainty of measurement*	±0.012 [bar]	±0.44 [l/min]
* uncertainty of measurement = transducer + 16-bit converter		

3.1. Research conditions

Before each simulation of a leakage the pipeline operated under steady-state conditions. The tests included slowly simulated leakages. The following settings of the operating point of the pipeline were used: inlet pressure $p_{in-0} \approx 6.0$ bar, outlet pressure $p_{out-0} \approx 1.0$ bar, nominal flow rate $q_{in-0} \approx 120$ l/min, temperature of the pumped medium $T_{water} \approx 20\,°\mathrm{C}$. In order to simulate leakages, solenoid valves were installed in several points along the pipeline. Measured signals were sampled with the frequency $f_p = 100$ Hz. The tests included simulated leaks of a size of $0.02 \div 4.00\,\%$ of the nominal flow rate q_{in-0} at 155th and 235th meter of the pipeline. The leaks were slowly increasing, which means that they reached the nominal value some time after the moment of their occurrence. Two tests were performed for a given leakage size.
From the diagnostic point of view the chosen way of simulating leaks was not too favourable, since the resulting phenomenon of pressure wave propagation is characterized by a smooth shape of wave fronts. Commonly used diagnostic method based on the detection of pressure waves can be in this case not very effective.

4. RESULTS OF TEST WITH SIMULATED LEAKS

An important element of the research was the comparison of the detection times (i.e. response times) of simulated leaks obtained by means of the reference method with the results obtained using the elaborated statistical method. The results are shown in Table 2. All times are given in seconds. They were measured from the moment of the opening of the valves. The results are the mean values of the two tests for a given size of the leak. The sign "–" in the table means that the algorithm did not detect the given leakage.

4.1. Application of the reference method

Three variants were considered for the method: variant "A"– involving the use of two pressure signals p_{in} and p_{out}, variant "B" – involving the use of two stream signals q_{in} and q_{out} (the signals in both variants were measured at the inlet and outlet of the pipeline), and additional variant "C" – involving the use of a package of two pressure signals measured at the inlet and outlet and additional four pressure signals measured along the pipeline.

The choice of alarm thresholds was based on the established minimum threshold values that ensured no alarm in states without leakage (in the whole series of tests). In the case of the algorithms based on two signals the values of alarm thresholds were $Pp = 0.0115$ and $Pq = -0.0100$, respectively, whereas in the case of the algorithm with six pressure signals the alarm threshold was $\overline{Pp} = 0.00490$. For individual algorithms, the following parameter values were used: $\alpha = 0.995$, $\beta = 0.900$, $\tau_{\max}{}^{A} = 1$ sec, and $\tau_{\max}{}^{B} = 0{,}25$ sec.

Also some additional information may be useful for determining the possible level of leak detection. Such a piece of information is the difference between the mean values for individual flow and pressure signals, before and after the leak, calculated basing on the profiles and taking into account the fact of exceeding the field of the measurement uncertainty. The results of this analysis are shown in Table 2 which, for a given signal, provides the comparison of its mean value (calculated for a 5-second interval which started 5 seconds after the start of the leak) with the mean value before the leak (calculated for three consecutive 5-second intervals, where the end of the last interval ended just before the start of the leak). If, for the two tests, for a given simulated leak size, the difference in mean values exceeded the field of measurement uncertainty, then such a case was marked by a "+", and the reverse case was marked by a "–". If, however, for individual tests, for a given size of the leak, it was ambiguous, such a case was marked by "–/+".

4.2. Application of elaborated method for simulated leak detection

The following paragraph presents examples of two averaged P-values in the case of leakage occurring at 20s. Fig. 8. shows *P*-value averages for a 0,37% leak. The leak was detected at 24.1s. In contrast, the 50% probability threshold is exceeded at 21.30s in the case of 1.19 % q_0 leakage (Fig. 9). Table 2 presents the comparison of leak detection time for leakages at 155m and 235m respectively in relation to the size of the leakage.

Table. 2. Times of detection of simulated leakages

leakages			changes of mean values				reference method			elaborated method		
							A	B	C	D	E	F
[m]	[l/min]	[% q_0]	p_{in}	p_{out}	q_{in}	q_{out}	p_{in}, p_{out}	q_{in}, q_{out}	six pressure signals	p_{in}, p_{out}	q_{in}, q_{out}	six pressure signals
155	0.09	0.07	–	–	–	–	–	–	–	-0,80	7,30	11.70
	0.17	0.14	–	–	–	–	–	–	–	-5,30	7,75	9.10
	0.28	0.23	–	–	–	–	–	–	–	-4,70	8,50	6.10
	0.44	0.36	–	–	–	–	–	–	–	-5,60	6,70	1.90
	0.77	0.63	+	–	–	–/+	1.62	2.89	1.12	1,20	2,65	1.45
	1.19	0.97	+	–	–	+	1.69	3.38	1.16	-6,80	2,65	1.30
	1.55	1.28	+	+	–/+	+	1.55	2.87	1.12	0,60	3,10	1.00
	1.84	1.51	+	+	+	+	1.43	3.04	1.07	1,30	2,50	1.30
	2.20	1.81	+	+	+	+	1.75	2.86	1.23	-4,40	2,80	1.30
	2.55	2.10	+	+	+	+	1.59	2.81	1.08	1,15	2,35	1.45
	2.89	2.37	+	+	+	+	1.56	2.75	1.15	-0,80	2,85	1.30
	3.19	2.62	+	+	+	+	1.58	2.95	1.10	-0,50	2,80	1.45
235	0.04	0.03	–	–	–	–	–	–	–	19,90	12,10	20.35
	0.11	0.09	–	–	–	–	–	–	–	7,55	12,10	27.25
	0.37	0.31	–	–	–	–	–	–	–	1,90	7,30	3.40
	0.78	0.65	+	–	–	+	–	4.44	1.37	1,60	3,10	1.60
	1.18	0.97	+	+	–	+	2.06	3.58	1.38	1,90	2,95	1.45
	1.55	1.28	+	+	–	+	2.00	2.85	1.36	1,90	2,50	1.60
	1.77	1.46	+	+	–	+	1.95	3.19	1.34	1,75	3,10	1.60
	2.06	1.70	+	+	–/+	+	1.97	3.25	1.34	1,85	3,10	1.60
	2.37	1.96	+	+	–/+	+	1.93	3.38	1.35	-0,80	7,30	11.70

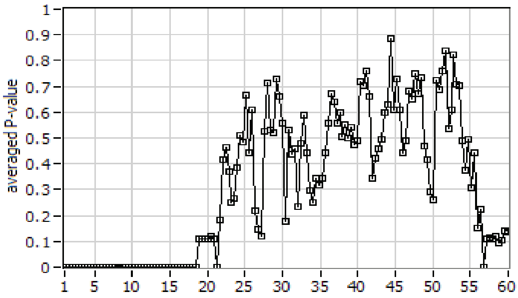

Fig. 8. *P*-value averages. 0,37% q_0 leak at 20s. Average P-value exceeds 50% at 24.10s

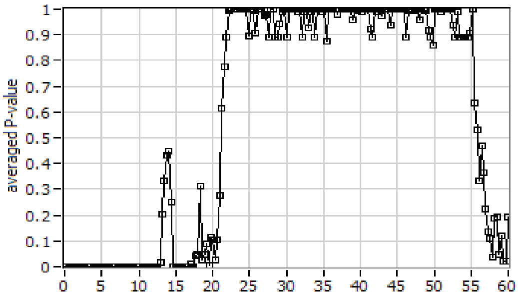

Fig. 9. P-value averages. 1,19% q_0 leak at 20s. Average P-value exceeds 50% at 21.30s

The results in columns "C" and "F" (including the results for all available pressure signals) clearly show that for all leakage sizes the statistically oriented method is more sensitive and faster in some cases than the reference method. The use of the statistical approach ensures detection even for leakages smaller than 0.5%, whereas the reference method is able to detect leakages of approximately 0.77% q_0. Even though the use of the statistical method allows for detecting very small leakages, the time of detection is about ten times longer than for leakages over 1% q_0. However, the column "D" includes results only for two pressure signals p_{in}, p_{out}. Some of the time values are negative, which should be interpreted as false alarms. When the method uses only few signals, the results are less certain.

4. SUMMARY

The use of algorithms operating on a continuous basis may be exposed to the risk of significant errors and disturbances occurring in individual moments of measuring pressure and flow signals or even the lack of data. The measurement data for a LDS should be provided periodically in a continuous manner and without significant delays. However, even the most advanced measuring systems installed on transmission pipelines are not completely resistant to interference, including errors associated with the transmission of data on such long distances.

In such situations, statistical methods prove to be a good solution. They are characterized by a short time of leak detection, high sensitivity, and accuracy. They are also very easy in application and do not involve advanced analysis methods. Statistically oriented approach gives more accurate results, because the detection is more sensitive and fast as well in comparison with the reference method.

REFERENCES

[1] Billman L., Isermann R., *Leak detection methods for pipelines*, *Automatica,* 1987, 23 pp. 381-385.

[2] Ostapkowicz P., *Improving the efficiency of diagnosing of leaks from liquid transmission pipelines by using the new diagnostic information - the signals of weak interactions between objects*, *Solid State Phenomena*, 2009, 147-149 pp. 492-497.

[3] Siebert H., Isermann R., *Leckerkennung und Lokalisierung bei Pipelines durch on-line Korrelation mit einen Prozesrechnes*, Regelungstechnik, 1977, 25 pp. 69-74.

[4] Sobczak R., *The pipeline leakage location method by following the pressure wave front*, Chemical Industry, 2004, 83 pp. 296-299.

[5] Turkowski M., Bratek A., Słowikowski M., *Methods and systems of leak detection in long range pipelines*, J. of Automation, Mobile Robotics & Intelligent Systems, 2007, 1 pp. 39-46.

[6] Zhang J., *Statistical pipeline leak detection for all operating conditions*, Pipeline & Gas Journal, 2001, 2 pp. 42-45.

Reference to a chapter in an edited book:

[7] Kowalczuk Z., Gunawickrama K., Detecting and locating leaks in transmission pipelines, in.: Korbicz K.J., Koscielny J.M., Kowalczuk Z., Cholewa W. (Eds.), *Fault Diagnosis: Models, Artificial Intelligence, Applications, Springer-Verlag*, Berlin, 2004, pp. 822-864.

[8] ADEC - Alaska Department of Environmental Conservation, Technical review of leak detection technologies for crude oil transmission pipelines, *Proceedings of the Pipeline Reliability Conference*, 1996, Houston, USA.

The research work financed with the means of the Ministry of Science and Higher Education (Poland) in the years 2010-2013 as the research project Nr N N504 494439.

Andrzej SOBOLEWSKI, Ph.D, Employee of the Faculty of Electrical Engineering since 1998 Bialystok University of Technology. Doctoral thesis in the field of technical diagnostics of electrical machines defended in 2008 Currently employed as a lecturer in the Department of Automation and Electronics. The main directions of research are diagnostics and fault detection and modern control methods.

Paweł OSTAPKOWICZ, Ph.D. Graduate and employee of the Faculty of Mechanical Engineering at Bialystok University of Technology since 1998. Doctoral thesis in the field of technical diagnostics of machines in 2007. Currently employed as a lecturer in the Department of Automatics and Robotics. The main directions of research are diagnostics and fault detection methods, design and operation of machines, vibroacoustics.