

Zofia Magdalena Łabęda-Grudziak
Politechnika Warszawska, Instytut Automatyki I Robotyki
02-525 Warszawa, ul. św. Andrzeja Boboli 8
email: z.labeda@mchtr.pw.edu.pl

ADDYTYWNY MODEL PROCESU ZMIAN CIŚNIEŃ W SIECI GAZOCIĄGOWEJ W UKŁADZIE DETEKcji NIESZCZELNOŚCI¹

Streszczenie – W artykule przedstawiono próbę oszacowania modelu procesu zmian ciśnień w sieci gazociągowej za pomocą addytywnego modelu regresji oraz technik eksploracji danych. Do wyznaczenia parametrów modeli addytywnych zastosowano algorytm dopasowania wstecznego i nieparametryczne techniki wygładzające. Celem badań jest wykorzystanie otrzymanych modeli do konstrukcji układu detekcji nieszczelności sieci gazociągowej. Badania przeprowadzono dla wybranych fragmentów rzeczywistej sieci przesyłowej gazu.

Słowa kluczowe: model addytywny, sieci przesyłowe gazu, przewidywanie ciśnienia, detekcja nieszczelności, eksploracja danych.

1 Wstęp

Systemy detekcji nieszczelności rurociągów odgrywają kluczową rolę w zadaniu zminimalizowania występowania wycieków oraz ograniczenia ich skutków. Badania nad metodami wykrywania nieszczelności są prowadzone na świecie od ponad 20 lat, jednak nie opracowano do tej pory całkowicie skutecznego systemu działającego dla każdej konfiguracji rurociągu. Ze względu na wysokie koszty budowy rurociągów gazu, które są prowadzone przeważnie w ziemi, okres eksploatacji rurociągów powinien być długi (do ok. 50 lat) i to przy minimalnym ryzyku wystąpienia awarii. Trudne warunki eksploatacji stawiają coraz większe wymagania dotyczące długotrwałości i wysokiego stopnia bezawaryjności systemów sterowania. Uszkodzenia rurociągów mogą być spowodowane m.in. korozją lub erozją ścianek wewnętrznych i zewnętrznych, pęknięciami spawów, wgnieceniami ścianek, wadami materiałowymi, degradacją materiału wskutek zmian ciśnienia, a także działaniem ludzi. Ze względu na łatwopalność i szkodliwość ekologiczną gazu, awarie powo-

¹ Praca została częściowo zrealizowana w ramach projektu rozwojowego KBN nr O R00 0013 06 oraz grantu promotorskiego KBN nr N N514 238337.

dujące rozszczelnienie instalacji i uchodzenie gazu stwarzają silne zagrożenie wystąpienia wybuchu i skażenia środowiska. Zagrożenia te wyeliminować można poprzez bieżącą detekcję [7,8], czyli wykrywanie nieprawidłowych stanów procesu oraz uszkodzeń urządzeń technologicznych, wykonawczych i pomiarowych, pozwalającą przewidzieć ewentualną konieczność wyłączenia tłoczenia lub odcięcia nieszczelnego odcinka rurociągu.

Metody detekcji nieszczelności można podzielić na dwie ogólne kategorie [1,9]: bezpośrednie (zewnętrzne), kiedy dokonuje się detekcji od zewnątrz rury poprzez zastosowanie wyspecjalizowanych urządzeń lub obserwacji wizualnej oraz pośrednie (wewnętrzne), kiedy detekcja jest oparta na pomiarach i analizie parametrów procesu przepływu (ciśnienia, strumienia, temperatury). Metody pośrednie dzielą się na metody oparte na detekcji fal akustycznych spowodowanych uchodzeniem, metody oparte na bilansowaniu medium wpływającego do i wyptywającego z rurociągu oraz metody analityczne, oparte na modelu matematycznym i danych pomiarowych z obiektu otrzymanych z systemu telemetrii.

Praca niniejsza podejmuje zagadnienie analitycznego modelowania procesu zmian ciśnień w sieci gazociągowej za pomocą addytywnego modelu regresji dla potrzeb detekcji uszkodzeń gazociągów. Ponieważ na obecnym etapie badań nie były dostępne dane opisujące warunki pracy podczas rzeczywistych uchodzeń gazu z sieci oraz żadne uszkodzenia nie zostały do tej pory zasymulowane, stąd skuteczność skonstruowanych modeli została zweryfikowana na podstawie danych testowych pochodzących ze stanu normalnej eksploatacji gazociągu. W przyszłości kontynuacją tych badań będzie wykorzystanie uzyskanych modeli dla potrzeb detekcji nieszczelności instalacji gazociągowych. Przedstawione badania zostały częściowo przeprowadzone w ramach projektu rozwojowego pt. "Badania nad systemami detekcji i lokalizacji nieszczelności rurociągów" we współpracy z Instytutem Metrologii i Inżynierii Biomedycznej Politechniki Warszawskiej i firmą Biatel S.A.

2 Układ detekcji nieszczelności

Metoda detekcji uszkodzeń oparta na addytywnym modelu regresji jest nowym podejściem w diagnostyce procesów przemysłowych i została przedstawiona w pracach, opublikowanych w pozycjach [11-16] spisu literatury. Istotę metody omówiono poniżej, w odniesieniu do sieci gazociągowej.

W instalacjach przemysłowych jakimi są sieci gazociągowe dostępne są wartości archiwizowanych zmiennych procesowych. To stwarza możliwość budowy modeli procesów na podstawie danych pomiarowych z obiektu oraz wiedzy eksperckiej o strukturze modelu. Dane te mogą być niekompletne, sprzeczne, niedokładne czy obciążone niepewnością. W

związku w tym powstało zapotrzebowanie na nowe metody i narzędzia informatyczne, które mogą wspomóc człowieka w odkrywaniu wiedzy z danych. Jest to obszarem zainteresowania nowej dziedziny nazywanej eksploracją danych (ang. *data mining*) [5,10]. Metody odkrywania wiedzy z danych polegają na analizie statystycznej wyników pomiarów parametrów istotnych z punktu widzenia funkcjonowania rurociągu (ciśnienia, przepływy, temperatury) tak aby szybko i trafnie rozpoznać zmiany stanu procesu w trakcie jego eksploatacji i nie wymagają istotnej wiedzy na temat rurociągu ani zjawisk w nim zachodzących. Należy podkreślić dużą rolę użytkownika w tym procesie. Użytkownik systemu odkrywania wiedzy powinien posiadać dobre zrozumienie dziedziny zastosowania, tak aby wybrać właściwy podzbiór danych, określić, jakie są zadania analizy, jaka powinna być reprezentacja poszukiwanej wiedzy, których algorytmów należy użyć. Dlatego system odkrywania wiedzy powinien być oprogramowaniem interaktywnym, a nie narzędziem w pełni automatycznym.

Dla potrzeb eksploracyjnej analizy danych z sieci gazociągowej w pracy wykorzystano model addytywny (ang. *additive models*) [2,6,11], który jest następującej postaci:

$$Y = \alpha + \sum_{j=1}^p \varphi_j(X_j) + \varepsilon, \quad (1)$$

gdzie Y jest sygnałem wyjściowym procesu, X_1, X_2, \dots, X_p są sygnałami wejściowymi procesu dla $p > 1$, błąd ε jest niezależny od (X_1, X_2, \dots, X_p) , $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$ oraz φ_j są jednowymiarowymi funkcjami zmiennej X_j niekoniecznie liniowymi, szacowanymi na podstawie danych za pomocą nieparametrycznych technik estymacji [6,11,16]. Podkreślmy, że nie zakładamy, że sygnały X_j są niezależne. W celu estymacji modelu addytywnego, funkcje φ_j są znajdowane za pomocą nieparametrycznych funkcji wygładzających i iteracyjnego algorytmu dopasowania wstecznego (ang. *backfitting algorithm*) [6,15].

Model addytywny jest matematycznym modelem procesu, który wiążąc wielkości fizyczne w procesie oraz czas może posłużyć do opisu zachowania się procesu i może naśladować jego działanie. Stąd w oparciu o addytywny model możemy obliczyć rozkład ciśnień w sieci na podstawie znanych, zmierzonych w rzeczywistej sieci wartości dostępnych zmiennych pomiarowych. Mając więc dwa sygnały wyjściowe możemy porównywać czy wartości ciśnienia zmierzone w rzeczywistej sieci y , znacznie odbiegają od wartości ciśnienia, estymowanych z modelu addytywnego y_m . Dla większości prawidłowych stanów pracy sieci ga-

zociągowej, przy założeniu prawdziwości modelu, różnice (tzw. residua $r = y - y_m$) powinny być bliskie bądź równe zero, zaś tam gdzie mamy do czynienia z odchyleniami od stanu prawidłowej pracy, charakteryzujących wyciek, różnice pomiędzy danymi sygnałami powinny w widoczny sposób zmieniać swoją wartość. Stąd w celu bieżącej analizy residuów uruchamiana jest następująca procedura detekcji [3,8]:

$$s(r_i) = \begin{cases} 0 & \text{gdy } K_1 \leq r_i \leq K_2 \\ 1 & \text{gdy } r_i < K_1 \vee r_i > K_2 \end{cases} \quad (2)$$

Symptom uszkodzenia zostaje wykryty, jeśli sygnał diagnostyczny $s(r_i)$ przyjmuje wartość 1, to znaczy kiedy wartość K_1 lub K_2 została przekroczona przez i -te residuum r_i . Jest to zatem typowy sposób zastosowania metody analitycznej redundancji do problemu detekcji uszkodzeń w instalacjach przemysłowych. Wartości progowe K_1 i K_2 są wyliczane z wykorzystaniem danych charakteryzujących przebieg wartości residuum w stanie normalnej eksploatacji w następujący sposób:

$$\begin{aligned} K_1 &= \min\{r_i\} - 3 \cdot \hat{\sigma}\{r_i\} \\ K_2 &= \max\{r_i\} + 3 \cdot \hat{\sigma}\{r_i\} \end{aligned} \quad (3)$$

gdzie $\hat{\sigma}\{r_i\}$ jest nieobciążonym estymatorem odchylenia standardowego próby residuów r_i , uwzględnianym ze względu na niepewność związaną z niedokładnością modelu procesu [14,16].

3 Weryfikacja modelu

Weryfikacja modelu stanowi procedurę mającą na celu sprawdzenie jakości oszacowanego modelu, co sprowadza się do określenia z jakim błędem model naśladuje zachowanie się rzeczywistego procesu. Dla danych pomiarowych $\{(x_{t1}, x_{t2}, \dots, x_{tp}), y_t\}_{t=1}^n$, rejestrowanych w trakcie eksploatacji instalacji gazociągowej zdefiniujemy model addytywny w postaci

$$\hat{y}_t = \sum_{j=1}^p \hat{\phi}_j(x_{tj}) + \varepsilon_t, \quad (4)$$

gdzie $\hat{y}_t = \hat{y}_t - \alpha$, ε jest błędem oszacowania \hat{y}_t poprzez model addytywny oraz $\hat{\phi}_j$ są jednowymiarowymi funkcjami rzeczywistymi. W badaniach wykorzystano następujące wskaźniki jakości modelowania:

- średni kwadrat błędów (ang. *mean square error*)

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2, \quad (5)$$

- średni moduł błędów (ang. *mean-absolute deviation error*)

$$\text{MADE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|, \quad (6)$$

- średni bezwzględny błąd procentowy wyrażony w zakresie pomiarowym wartości wyjścia (ang. *mean-absolute percentage error*)

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{(y_{\max} - y_{\min})} \cdot 100\%, \quad (7)$$

- wariancja błędów (ang. *variance error*)

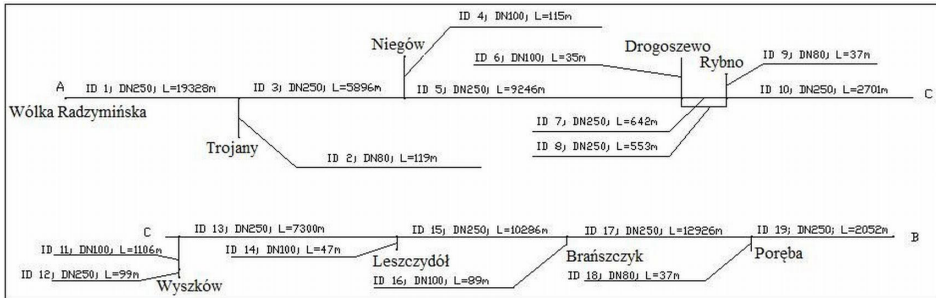
$$\text{VAR} = \frac{1}{n} \sum_{t=1}^n (y_t - \bar{\hat{y}}_t)^2, \quad (8)$$

gdzie y_t oznacza wartość zmiennej wyjściowej procesu w chwili t , \hat{y}_t oznacza estymowaną wartość zmiennej wyjściowej procesu w chwili t , y_{\max} i y_{\min} oznaczają odpowiednio największą i najmniejszą wartość zmiennej wyjściowej procesu, $\bar{\hat{y}}_t$ oznacza średnią wartość estymowanej zmiennej wyjściowej procesu w chwili t oraz n jest licznością próby.

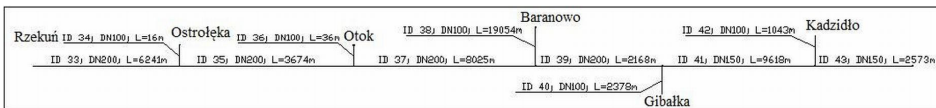
4 Addytywny model zmian ciśnień w sieci gazociągowej

Badania przeprowadzono dla fragmentu rzeczywistej sieci przesyłowej gazu znajdującej się na dużym obszarze Polski, przy użyciu programu R-project [4]. Wytypowany fragment sieci składa się z rurociągu głównego wraz z odnogami, które są zakończone osiemnastoma sta-

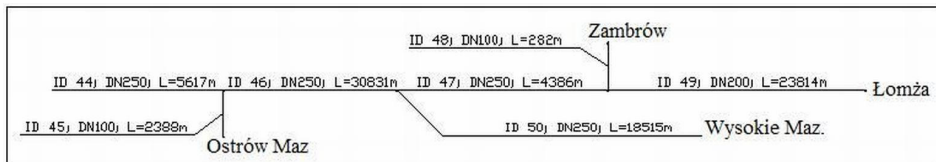
cjami redukcyjno-pomiarowymi. W pracy podjęto próbę oszacowania modelu procesu zmian ciśnień w sieci gazociągowej dla 3 punktów odbioru gazu, odpowiednio w trzech miejscowościach: Leszczydół, Ostrołęka, Wysokie Mazowieckie. Fragmenty schematów wybranych części instalacji przedstawiono na rys. 1-3.



Rys. 1. Fragment schematu instalacji w okolicy miejscowości Leszczydół



Rys. 2. Fragment schematu instalacji w okolicy miejscowości Ostrołęka



Rys. 3. Fragment schematu instalacji w okolicy miejscowości Wysokie Mazowieckie

Dane pobrane z czujników pomiarowych, podczas pracy systemu telemetrii, zostały zarejestrowane i przekazane dalej do bazy danych procesu, a następnie udostępnione w postaci plików na dysku, zawierających sygnały rejestrowane co dwie minuty. Wykorzystywane do modelowania zmienne pomiarowe zostały opisane w tab. 1.

Tabela. 1. Zmienne pomiarowe gazu użyte w modelowaniu

Ciśnienie	Przepływ	Stacja poboru gazu
Pwe_Drog	Qn_Drog	Drogoszewo
Pwe_Ryb	Qn_Ryb	Rybno
Pwe_Lesz	Qn_Lesz	Leszczydół
Pwe_Bran	Qn_Bran	Brańszczyk
Pwe_Rzek	Qn_Rzek	Rzekuń
Pwe_Ostr	Qn_Ostr	Ostrołęka
Pwe_Otok	Qn_Otok	Otok
Pwe_Bar	Qn_Bar	Baranowo
Pwe_Gib	Qn_Gib	Gibałka
Pwe_WysMaz	Qn_WysMaz	Wysokie Mazowieckie
Pwe_Zam	Qn_Zam	Zambrów
Pwe_Lom	Qn_Lom	Łomża

Aby dostępna baza danych była przydatna do celów eksploracji danych, dokonano jej wstępnej obróbki. Eksploracyjna analiza danych wykazała istnienie danych nieodpowiednich dla celów modelowania - ze względu na ich zły status. Rekordy z takimi danymi pominięto podczas analizy. Operacja ta nie doprowadziła do analizy obciążonego zbioru danych ani utraty informacji o zachowaniu się procesu.

W przypadku analizy danych pomiarowych pojawia się problem uwzględniania szumów pomiarowych, które mogą prowadzić do błędnych interpretacji zjawisk zachodzących w badanym procesie. Aby wyeliminować takie sygnały, zastosowano filtr FIR (ang. *finite impulse response*) (9) i (11) lub alternatywnie, w przypadku występowania impulsowych wartości ciśnienia w bazie danych, filtr ARMA (ang. *autoregressive moving average filter*) (10). Po wnikliwej analizie wywnioskowano, że do prawidłowego zamodelowania dynamiki sieci wystarczy zastosować filtry rzędu drugiego. Zwiększenie rzędu filtru nie skutkowało znaczną poprawą wyników modelowania w stosunku do wzrostu złożoności modelu.

Dla tak przygotowanych danych obliczany jest rozkład ciśnień w sieci na podstawie znanych, zmierzonych w rzeczywistej sieci wartości strumienia Q_n w chwilach poprzednich oraz wartości ciśnienia P_{we} w chwilach poprzednich. Niektóre zmienne pomiarowe przyjmowały wartości zerowe dla ciśnienia i/lub przepływu (np. przepływ dla Rybna jest zerowy). W takim przypadku nie wykorzystywano zmiennej do modelowania.

Biorąc pod uwagę strukturę sieci gazociągowej, zaproponowano następujące modele zmian ciśnień w sieci dla Leszczydołu (9), Ostrołęki (10) i Wysokie Mazowieckie (11):

$$\text{Pwe_Lesz} = \varphi(\text{Pwe_Drog}, \text{Qn_Drog}, \text{Pwe_Ryb}, \text{Pwe_Bran}, \text{Qn_Bran}), \quad (9)$$

$$\text{Pwe_Ostr} = \psi(\text{Pwe_Ostr}, \text{Pwe_Rzek}, \text{Pwe_Otok}, \text{Qn_Otok}), \quad (10)$$

$$\text{Pwe_WysMaz} = \chi(\text{Pwe_Zam}, \text{Qn_Zam}, \text{Pwe_Lom}, \text{Qn_Lom}), \quad (11)$$

gdzie $\varphi(\cdot), \psi(\cdot), \chi(\cdot)$ są pewnymi wielowymiarowymi funkcjami rzeczywistymi. Następnie dla modeli (9)-(11) zbudowano odpowiadające im modele addytywne (12)-(14), w następującej postaci:

$$\begin{aligned} \text{Pwe_Lesz}_t &= \hat{\text{Pwe_Lesz}}_t + \varepsilon_t \\ &= \varphi_1(\text{Pwe_Drog}_{t-1}) + \varphi_2(\text{Pwe_Drog}_{t-2}) + \varphi_3(\text{Qn_Drog}_{t-1}) + \varphi_4(\text{Qn_Drog}_{t-2}) \\ &\quad + \varphi_5(\text{Pwe_Ryb}_{t-1}) + \varphi_6(\text{Pwe_Ryb}_{t-2}) + \varphi_7(\text{Pwe_Bran}_{t-1}) + \varphi_8(\text{Pwe_Bran}_{t-2}) \\ &\quad + \varphi_9(\text{Qn_Bran}_{t-1}) + \varphi_{10}(\text{Qn_Bran}_{t-2}) + \varepsilon_t, \end{aligned} \quad (12)$$

$$\begin{aligned} \text{Pwe_Ostr}_t &= \hat{\text{Pwe_Ostr}}_t + \varepsilon_t \\ &= \psi_1(\text{Pwe_Ostr}_{t-1}) + \psi_2(\text{Pwe_Ostr}_{t-2}) + \psi_3(\text{Pwe_Rzek}_{t-1}) + \psi_4(\text{Pwe_Rzek}_{t-2}) \\ &\quad + \psi_5(\text{Pwe_Otok}_{t-1}) + \psi_6(\text{Pwe_Otok}_{t-2}) \\ &\quad + \psi_7(\text{Qn_Otok}_{t-1}) + \psi_8(\text{Qn_Otok}_{t-2}) + \varepsilon_t, \end{aligned} \quad (13)$$

$$\begin{aligned} \text{Pwe_WysMaz}_t &= \hat{\text{Pwe_WysMaz}}_t + \varepsilon_t \\ &= \chi_1(\text{Pwe_Zam}_{t-1}) + \chi_2(\text{Pwe_Zam}_{t-2}) \\ &\quad + \chi_3(\text{Qn_Zam}_{t-1}) + \chi_4(\text{Qn_Zam}_{t-2}) \\ &\quad + \chi_5(\text{Pwe_Lom}_{t-1}) + \chi_6(\text{Pwe_Lom}_{t-2}) \\ &\quad + \chi_7(\text{Qn_Lom}_{t-1}) + \chi_8(\text{Qn_Lom}_{t-2}) + \varepsilon_t, \end{aligned} \quad (14)$$

gdzie $t = 3, \dots, n$, $\varphi_1, \dots, \varphi_{10}, \psi_1, \dots, \psi_8, \chi_1, \dots, \chi_8$ jednowymiarowymi funkcjami poszczególnych sygnałów oraz ε_t są niezależnymi zakłóceniami losowymi.

Obliczone ciśnienia $\hat{\text{Pwe_Lesz}}_t, \hat{\text{Pwe_Ostr}}_t, \hat{\text{Pwe_WysMaz}}_t$ są na bieżąco porównywane ze zmierzonymi w rzeczywistej sieci. Przedmiotem analizy są residua r_t , będące przybliżeniem błędów $\hat{\varepsilon}_t$, które moż-

na znormalizować do przedziału [-1.1] poprzez podzielenie jego wartości przez zakres zmienności sygnału:

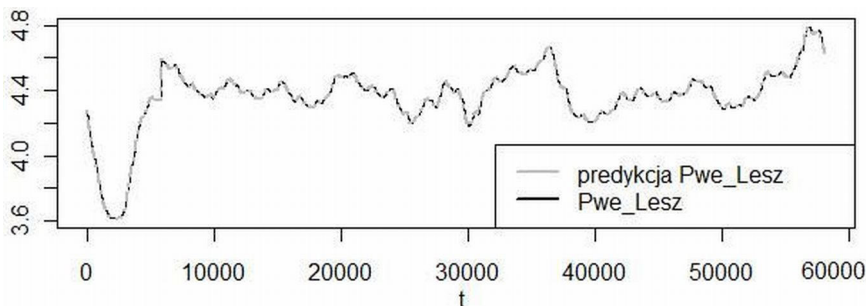
$$r_t := \frac{Pwe_t - \hat{P}we_t}{Pwe_{\max} - Pwe_{\min}}, \quad (15)$$

gdzie Pwe_{\max} i Pwe_{\min} są odpowiednio maksymalną i minimalną wartością zmiennej Pwe . Takie przeskalowanie umożliwi lepszą interpretację uzyskanych wartości residuów.

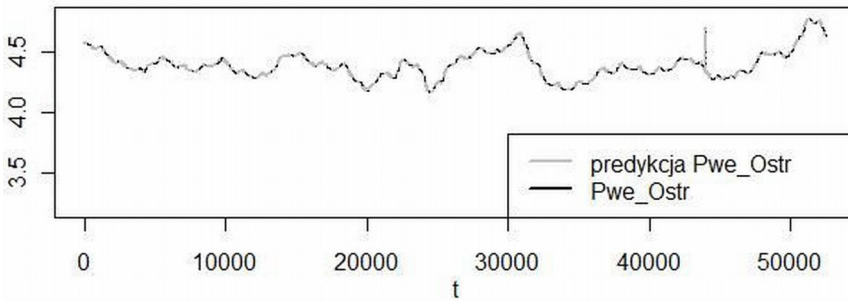
5 Wyniki modelowania

Zbiór danych pomiarowych został podzielony na zbiory uczące i zbiory testowe w zależności od modelowanego punktu odbioru gazu. W celu uniknięcia obciążenia danych trudnym do wychwycenia trendem wybrano dane uczące pochodzące z różnych okresów. Kryterium doboru była jak największa zmienność sygnału wyjściowego, reprezentująca najbardziej skrajne stany. Pozostałe dane wykorzystano do próby testowej. W celu estymacji modeli addytywnych (12)-(14), wykorzystano algorytm dopasowania wstecznego z naturalną kubiczną funkcją sklejaną z parametrem wygładzającym $df=4$, utożsamianym z liczbą stopni swobody [16].

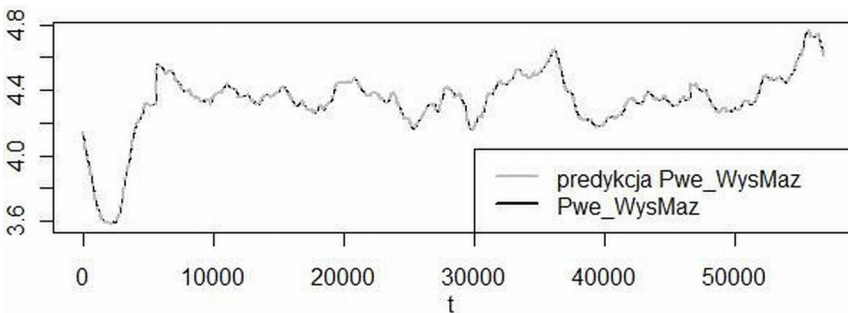
Na podstawie danych uczących, pochodzących ze stanu zdadności procesu, otrzymano estymowane wartości ciśnienia gazu (predykcja Pwe) wraz z rzeczywistymi wartościami ciśnienia danymi z procesu (Pwe) dla poszczególnych miejscowości, których wykresy przedstawiono na rys. 4 - 6.



Rys. 4. Przebieg sygnałów pomierzonego i modelowanego dla próby uczącej dla Leszczydołu



Rys. 5. Przebieg sygnałów pomierzonego i modelowanego dla próby uczącej dla Ostrołęki

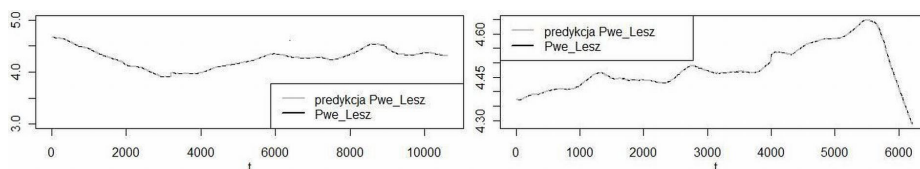


Rys. 6. Przebieg sygnałów pomierzonego i modelowanego dla próby uczącej dla Wysokie Mazowieckie

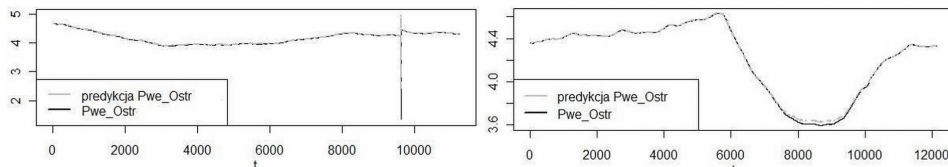
Przebiegi wyniku modelowania, niemalże idealnie nadążają za rzeczywistym procesem. Do badania algorytmu detekcyjnego (2), bazującego na otrzymanych modelach, wykorzystano próby testowe składające się z danych pomiarowych pochodzących tylko ze stanu zdatności, ze względu na aktualny brak danych w przypadku wystąpienia uszkodzeń. Kolejnym etapem prowadzonych badań będzie zastosowanie uzyskanych modeli dla potrzeb detekcji zasymulowanych nieszczelności instalacji gazociągowych.

Na rys. 7-12 przedstawiono wykresy przebiegów sygnału pomierzonego i modelowanego dla prób testowych. Na ich podstawie można wyraźnie zaobserwować brak istotnych odchyłek od stanu zdatności. Stąd zbudowane modele addytywne (12)-(14) dobrze oszacowują rozkład ciśnień w gazociągu.

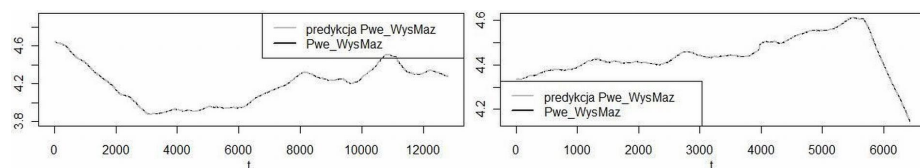
Zauważmy, że dla modelu zmian ciśnień w Ostrołęce pojawiają się znaczne skoki wartości ciśnienia, co powoduje, że residuum reaguje impulsowo na zaistniałe zmiany i sokowo odbiega od wartości zerowej oraz po chwili zanika. Jest to prawdopodobnie spowodowane tym, że niektóre wartości danych testowych wykraczają poza zakres wartości danych uczących.



Rys. 7. Przebieg sygnałów pomierzonego i modelowanego dla 1 i 2 próby testowej dla Leszczydołu



Rys. 8. Przebieg sygnałów pomierzonego i modelowanego dla 1 i 2 próby testowej dla Ostrołęki



Rys. 9. Przebieg sygnałów pomierzonego i modelowanego dla 1 i 2 próby testowej dla Wysokie Mazowieckie

W celu sprawdzenia jakości oszacowanego modelu, dla próby uczącej i prób testowych dla poszczególnych punktów odbioru gazu, obliczono wartości miar dopasowania (5)-(8). Wyniki przedstawiono w tab. 2.

Tabela. 2. Wskaźniki jakości dopasowania dla poszczególnych modeli addytywnych

Miara	Próba ucząca			Próba testowa 1			Próba testowa 2		
	Lesz	Ostr	WysMaz	Lesz	Ostr	WysMaz	Lesz	Ostr	WysMaz
MSE	0	0	1e-06	7.876	0.001	2.687	3.937	1e-04	3.549
MADE	4e-04	2e-04	4e-04	0.005	0.003	4e-04	0.001	0.006	4e-04
MAPE[%]	0.032	0.031	0.033	0.063	0.09	0.053	0.133	0.554	0.093
VAR	1e-06	0	0	7.693	0.001	2.648	2.592	1e-04	3.499
Test[%]	0	0	0	0.009	0.071	0.009	0.016	0	0.186

Wykorzystując przebiegi wartości residuum w stanie normalnej eksploatacji zostały wyznaczone wartości progowe (3): $K_1 = -0.0149$ i $K_2 = 0.0145$ dla modelu (12), $K_1 = -0.0333$ i $K_2 = 0.0191$ dla modelu

(13) oraz $K_1 = -0.0036$ i $K_2 = 0.0045$ dla modelu (14). Następnie do otrzymanych residuów z poszczególnych prób testowych, zastosowano test (2), w celu oceny wrażliwości modelu na poszczególne uszkodzenia. Tab. 2 zawiera procentowe wyniki liczby wartości residuów przekraczających pasy zdatności (K_1, K_2) w zależności od modelowanego wyjścia. Podjęcie decyzji o wystąpieniu uszkodzenia można przeprowadzić za pomocą techniki progowania. W pracy zastosowano stały próg o wartości 0.05. Wartość progu została dobrana arbitralnie tak, aby uzyskać kompromis pomiędzy czułością detekcji, a liczbą fałszywych alarmów. Stąd jeśli wynik procentowy liczby wartości residuów przekraczających pasy zdatności (K_1, K_2) przekracza 5%, dany test wykrył uszkodzenie. W przeciwnym wypadku test nie wykrył uszkodzenia lub w przypadku próby zdatności procesu - próba nie zawiera uszkodzeń.

Otrzymane wyniki są zadowalające, gdyż na podstawie średniej procentowej liczby odchyłek od stanu normalnej eksploatacji, test nie wykrył istotnych nieprawidłowości w próbie zdatności. Stąd można sądzić, że uzyskane modele będą czułe na występowanie uszkodzeń, ale nie będą zbyt czułe dla danych pochodzących z procesu w stanie zdatności.

6 Podsumowanie

Możliwość generacji fałszywych symptomów w algorytmach detekcji zależy w dużym stopniu od poprawnego odzwierciedlenia dynamiki modelowanego procesu i tym samym od wartości przyjętych ograniczeń za dopuszczalne odchyłki od stanu prawidłowej pracy. Nieodpowiednia eksploracyjna analiza danych może spowodować znaczną zmianę wartości progowych i tym samym zmianę wyników testów diagnostycznych. Przeprowadzone eksperymenty pokazują, że addytywny model regresji może być efektywnie użyty do modelowania rozkładu ciśnienia w sieci gazociągowej. Otrzymane wyniki badań są zadowalające, gdyż zaprezentowane metody pozwoliły na skonstruowanie modeli dobrze odzwierciedlających dynamikę procesu, co pozwoli w dalszym etapie badań do ich wykorzystania w celach detekcji nieszczelności.

Literatura

- [1] Bilman L., Isermann R., Leak detection methods for pipelines, *Automatica*, vol.23, no. 3, pp. 381-385, 1987.
- [2] Friedman J.H., Stuetzle W., Projection pursuit regression, *Journal of the American Statistical Association*, 76, 1981, pp.817-823.
- [3] Gertler J., *Fault Detection and Diagnosis in Engineering Systems*, Marcel Dekker, Inc. New York - Basel - Hong Kong, 1998.

- [4] Good P.I., *Introduction to statistics through reasampling methods and R/S-PLUS*, Wiley, 2005.
- [5] Hand D., Mannila H., Smyth P., *Principles of Data Mining*.
- [6] Hastie T., Tibshirani R., *Generalized additive models*, Chapman and Hall, 1990.
- [7] Korbicz J., Kościelny J.M., Kowalczyk Z., Cholewa W.(red), *Diagnostyka procesów. Modele, metody sztucznej inteligencji, zastosowania*, WNT, Warszawa 2002.
- [8] Kościelny J.M., *Diagnostyka procesów przemysłowych*, EXIT, Warszawa 2001.
- [9] Kowalczyk Z., Gunawickrama K., *Detekcja i lokalizacja wycieków w rurociągach przemysłowych*, rozdział 21 pracy zbiorowej pod red. J. Korbicza i J. Kościelnego, WNT, Warszawa 2002.
- [10] Larose D.T., *Discovering Knowledge in Data: An Introduction to DATA MINING*. Wiley, 2005.
- [11] Łabęda Z.M., *Wykorzystanie addytywnego modelu regresji w eksploracyjnej analizie danych*, VI Sympozjum Modelowanie i Symulacja Komputerowa w Technice, Łódź 2008.
- [12] Łabęda Z.M., *Zastosowanie modelu addytywnego dla potrzeb detekcji uszkodzeń zaworów regulacyjnych*, III Konferencja Naukowo-Techniczna Doktorantów i Młodych Naukowców, Warszawa, 2008.
- [13] Łabęda Z.M., Additive model applications for the fault detection of actuators, *Pomiary, Automatyka i Kontrola*, Vol.55, pp. 152-155, (3)2009.
- [14] Łabęda Z.M., *Modele addytywne w układzie detekcji uszkodzeń w stacji wyparnej*, IX Międzynarodowa Konferencja Diagnostyka Procesów i Systemów, Gdańsk 2009, artykuł opublikowany jako rozdział w książce "Systemy wykrywające, analizujące i tolerujące usterki. Automatyka i Informatyka" Pomorskie Wydawnictwo Naukowo-Techniczne, rozdział 13, str. 113.
- [15] Łabęda Z.M., *The backfitting and marginal integration estimators for additive models*, IV Konferencja Naukowo-Techniczna Doktorantów i Młodych Naukowców, Warszawa, 2009, materiały konferencyjne, str. 307-314.
- [16] Łabęda Z.M., *Smoothing parameters selection in the additive regression models approach for the fault detection scheme*, IV Konferencja Naukowo-Techniczna Doktorantów i Młodych Naukowców, Warszawa, 2009, materiały konferencyjne, str. 299-306.

ADDITIVE MODEL OF PRESSURE DECOM- POSITION IN GAS NETWORK FOR THE LEAK DETECTION SCHEME

Summary – this paper presents methods based on additive regression models with knowledge discovery data to predict pressure values at determined nodes of gas network. The backfitting algorithm with nonparametric smoothness techniques has been used for estimating the additive model. The plan aim is to use the received models to design the leak detection scheme in gas pipeline systems. All research has been carried out based on the part of long range gas pipelines.

Keywords: additive model, gas pipelines, pressure prediction, leak detection, data mining.