

106/2012

Raport Badawczy

RB/24/2012

Research Report

**Algorytmy modelowania,
optymalizacji i aproksymacji
wspomagające zarządzanie
miejską siecią wodociągową**

J. Studziński

**Instytut Badań Systemowych
Polska Akademia Nauk**

**Systems Research Institute
Polish Academy of Sciences**



POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
dr hab. inż. prof. nadzw. Jan Studziński

Warszawa 2012

ALGORYTMY MODELOWANIA, OPTYMALIZACJI I APROKSYMACJI WSPOMAGAJĄCE ZARZĄDZANIE MIEJSKĄ SIECIĄ WODOCIĄGOWĄ

JAN STUDZIŃSKI

Spis treści

1. Wstęp
2. Informatyzacja miejskiej sieci wodociągowej
3. Algorytmy modelowania
4. Algorytmy optymalizacji
5. Algorytmy aproksymacji
6. Uwagi końcowe

Literatura

1. Wstęp

Miejskie systemy wodno-ściekowe składają się zwykle z czterech podsystemów: ujęcia wody, sieci wodociągowej, sieci kanalizacyjnej i oczyszczalni ścieków (rys. 1), przy czym sieć wodociągowa wydaje się w tym systemie elementem kluczowym. Zmieniające się obciążenie sieci wodociągowej wpływa na pracę pompowni w stacjach ujęcia wody, decyduje o obciążeniu hydraulicznym sieci kanalizacyjnej i oczyszczalni ścieków, a poprawne zarządzanie siecią wodociągową decyduje o właściwej jakości i ilości wody dostarczanej do odbiorców. Dlatego jednym z podstawowych zadań w przedsiębiorstwie wodociągowym jest efektywne zarządzanie tą siecią. W ramach tego zadania należy realizować szereg podzadań, takich jak:

- pełne pokrycie potrzeb odbiorów na wodę pitną,
- dostarczanie do odbiorców wody o odpowiedniej jakości,
- dostarczanie wody pod odpowiednim ciśnieniem,
- zapobieganie awariom sieci wodociągowej,
- w przypadku wystąpienia awarii, szybka jej lokalizacja i usunięcie,
- planowanie i wykonywanie remontów sieci,
- modernizacja i rozbudowa sieci,
- zapobieganie stratom wody,
- planowanie ceny wody na kolejny rok kalendarzowy itp.



Rys. 1. Schemat miejskiego systemu wodno-ściekowego.

Dane pozyskiwane z monitoringu sieci wodociągowej ułatwiają rozwiązywanie problemów związanych z jej zarządzaniem. Uzasadnionym więc wydaje się budowa całych złożonych systemów informatycznych do zarządzania sieciami wodociągowymi, w których

monitoring jest jednym z kluczowych elementów systemu. Co więcej, problematyka właściwego zarządzania, czyli działania o charakterze jakościowym, jest związane z pewnymi obszarami wiedzy technicznej, czyli działaniami o charakterze ilościowym.

Zarządzanie wiedzą jest działaniem interdyscyplinarnym, łączącym w sobie techniki ilościowe i jakościowe i nie można spodziewać się skutecznego zarządzania opartego tylko na intuicji i doświadczeniu, jak również nie można w zarządzaniu ograniczać się jedynie do stosowania algorytmów obliczeniowych. Jednak połączenie obu rodzajów działania, jakościowego i ilościowego, chociaż wskazane, jest bardzo trudne, ponieważ jest związane z dużymi trudnościami organizacyjnymi i ponoszeniem wysokich kosztów a przede wszystkim z przełamaniem pewnych nawyków myślowych oraz przyzwyczajzeń, co wydaje się najtrudniejsze. Stąd potrzeba wskazywania ośrodkom decyzyjnym na zalety działań integracyjnych i celowość kompleksowej informatyzacji przedsiębiorstw wodociągowych.

2. Informatyzacja miejskiej sieci wodociągowej

Systemy informatyczne stosowane do kompleksowego zarządzania przedsiębiorstwami wodociągowymi, to w Polsce jeszcze rzadkość. Przy czym przez zarządzanie kompleksowe rozumiemy tutaj wykonywanie przez system zarówno zadań organizacyjno-finansowych (gospodarka magazynowa, księgowość, obsługa klientów przedsiębiorstwa itp.), jak i techniczno technologicznych (symulacja komputerowa realizowanego procesu, optymalizacja i sterowanie procesem itp.). Wykonujący takie zadania system informatyczny powinien składać się ze współpracujących ze sobą modułów odpowiedzialnych za funkcje administracyjne i techniczne.

Jednym z kluczowych elementów takiego zintegrowanego systemu informatycznego jest system monitoringu. System monitoringu sieci wodociągowej może oczywiście funkcjonować w przedsiębiorstwie wodociągowym jako system autonomiczny i jeżeli będzie odpowiednio dokładny, to jest w stanie realizować samodzielnie szereg zadań znacznie usprawniających zarządzanie siecią. Jednak rozwiązaniem znacznie korzystniejszym jest wdrożenie w przedsiębiorstwie kompleksowego systemu wspomaganie decyzji, w którym monitoring byłby jedynie jednym z jego elementów. Aby system monitoringu mógł być właściwie wykorzystany, powinien on współpracować z modelem hydraulicznym sieci wodociągowej. Dodając jako kolejny element mapę numeryczną sieci wodociągowej oraz modele prognozowania obciążenia sieci a także algorytmy optymalizacji i sterowania siecią

otrzymujemy strukturę takiego kompleksowego systemu wspomagania decyzji, będącego idealnym narzędziem zarządzania siecią wodociągową.

W ostatnich latach nastąpił gwałtowny rozwój automatyzacji w krajowym przemyśle. W większości przedsiębiorstw zostały wdrożone nowoczesne systemy automatyki. Stwarza to korzystny grunt do wprowadzania zaawansowanych technik nadzoru, sterowania, diagnostyki i wspomagania decyzji kadry zarządzającej, operatorów i inżynierów procesu. Pojawia się również zainteresowanie przedsiębiorstw wodociągowych takimi aplikacjami. Wynika ono z potencjalnie dużych korzyści ekonomicznych, które mogą przynieść wdrożenia.

Rozwiązania oferowane przez renomowane firmy są jednak bardzo drogie dla polskich przedsiębiorstw, nawet takich jak elektrownie lub rafinerie, a ponadto ich funkcjonalność jest zwykle ograniczona. Wdrożenia wiążą się najczęściej z koniecznością sprowadzania specjalistów z zagranicy.

Współczesne systemy automatyki są integrowane z systemami wspomagania decyzji oraz z systemami zarządzania przedsiębiorstwem. Z punktu widzenia kompleksowego podejścia, do systemów wspomagania decyzji należy więc te trzy grupy systemów rozpatrywać jako całość. W ogólności, zarządzanie przedsiębiorstwem można podzielić na następujące warstwy (rys. 2):

a) Warstwa realizacji pomiarów i oddziaływań sterujących. W strukturze sprzętowej jest to zbiór urządzeń pomiarowych i wykonawczych. Coraz częściej są to tzw. inteligentne urządzenia pomiarowe i wykonawcze, wyposażone w jednostki mikroprocesorowe, realizujące wiele funkcji wstępnego przetwarzania sygnałów, regulacji położenia elementów nastawczych oraz komunikacji z jednostkami sterującymi przez sieci polowe.

b) Warstwa sterowania bezpośredniego, której głównymi zadaniami jest regulacja i sterowanie binarne. Fizycznie w warstwie tej występują różnego rodzaju sterowniki i regulatory.

c) Warstwa sterowania nadrzędnego, w której realizowane są algorytmy regulacji, kompensacji zakłóceń, adaptacji oraz sterowania optymalnego poszczególnych węzłów technologicznych. Sygnały sterujące wypracowywane w tej warstwie nie oddziałują bezpośrednio na urządzenia wykonawcze, lecz stanowią wejścia zadane (wiodące) dla algorytmów w warstwie sterowania bezpośredniego. Algorytmy tej warstwy mogą być realizowane zarówno przez sterowniki dużej mocy (w systemach klasy DCS nazywane stacjami procesowymi), jak też komputery.

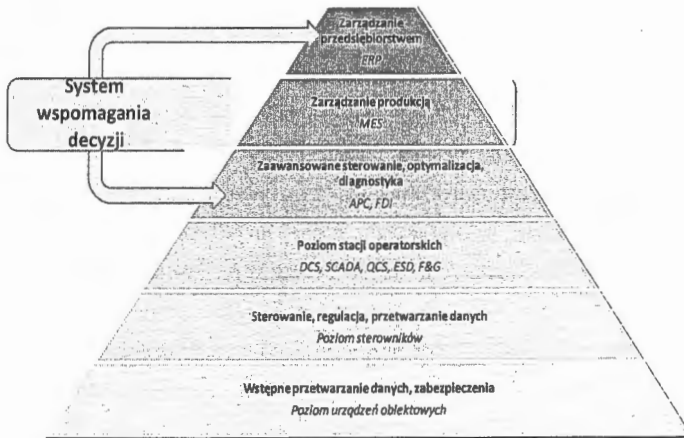
d) Warstwa nadzorowania, koordynacji i optymalizacji procesu. Nadzorowanie obejmuje zadania wykrywania, rejestrowania i sygnalizacji alarmów. Zadaniem sterowania realizowanymi dla procesów ciągłych w tej warstwie jest koordynacja strumieni materiałów i energii przepływających między różnymi częściami procesu (węzłami technologicznymi) oraz optymalizacja punktów pracy procesów. Dla procesów dyskretnych realizowane są algorytmy koordynacji pracy grup maszyn i urządzeń. Funkcje te realizują komputery (stacje operatorskie lub odrębne jednostki komputerowe).

e) Warstwa wizualizacji i obsługi. Zadaniem tej warstwy jest współpraca systemu sterowania z operatorami, automatykami i innymi użytkownikami systemu. Służą do tego panele operatorskie, stacje operatorskie, stacje inżynierskie oraz stacje nadzorcze (informacyjne). Wyświetlać na nich można odpowiednio przetworzone i przygotowane dane pochodzące z innych warstw. Przebieg procesu prezentowany jest graficznie na schematach synoptycznych procesu, wykresach zmiennych procesowych, listach zdarzeń i alarmów itp. Obserwować i analizować można nie tylko dane bieżące, ale także archiwizowane dane z przeszłości. Raporty udostępniają obsłudze odpowiednio dostosowane do potrzeb użytkowników zestawienia danych. Stacje operatorskie umożliwia oddziaływanie na proces, stacje inżynierskie pozwalają na konfigurację sprzętową i programową (algorytmiczną) systemu oraz na wprowadzanie modyfikacji, natomiast stacje nadzorcze umożliwiają jedynie wizualizację przebiegu procesu i danych.

f) Warstwa zarządzania produkcją: systemy MES (*Manufacturing Execution Systems*). Systemy klasy MES, wykorzystując technologie informatyczne, oprogramowanie, urządzenia elektroniczne i elementy automatyki, umożliwiają efektywne zbieranie informacji w czasie rzeczywistym wprost ze stanowisk produkcyjnych i ich transfer na obszar biznesowy. Informacje o realizacji produkcji mogą być pobierane bezpośrednio z maszyn oraz przy udziale pracowników bezpośrednio-produkcyjnych. Dzięki funkcjonalności systemu można uzyskać natychmiastowy sygnał zwrotny o stopniu wykonania produkcji, podejmować na bieżąco właściwe decyzje i reagować na bieżąco na nieprawidłowości pojawiające się w czasie procesu produkcyjnego. Pozyskane dane z procesu produkcyjnego pozwalają na analizę kluczowych wskaźników wydajności na produkcji i uzyskanie prawdziwego obrazu wykorzystania zdolności produkcyjnych.

g) Warstwa zarządzania przedsiębiorstwem, realizowana wyłącznie w technice komputerowej. Zadania zarządzania realizują systemy ERP (*Enterprise Resource Planning*).

Wspomaganie to może obejmować wszystkie lub część szczebli zarządzania i ułatwia optymalizację wykorzystania zasobów przedsiębiorstwa oraz zachodzących w nim procesów.



Rys. 2. Warstwowa struktura systemów zarządzania i sterowania.

W Instytucie Badań Systemowych PAN opracowano koncepcję zintegrowanego systemu informatycznego do kompleksowego zarządzania miejską siecią wodociągową. Projektowany system wspomaganie decyzji mieści się głównie w grupie funkcji dotyczących zarządzania produkcją (jakość wody). Poniżej przedstawiono wybór algorytmów działających w strukturze tego systemu.

3. Algorytmy modelowania

3.1. Modelowanie sieci wodociągowej

Podstawowym narzędziem informatycznym ułatwiającym zarządzanie siecią wodociągową jest model hydrauliczny do obliczania przepływów i ciśnień wody w sieci. Model taki jest opisany liniowymi i nieliniowymi równaniami algebraicznymi, podobnymi do równań opisujących bilanse natężeń i napięć prądów w sieciach elektrycznych i wynikających z I i II prawa Kirchhoffa znanych z elektrotechniki. Dla sformułowania równań modelu należy zadać strukturę sieci wodociągowej oraz charakterystyki obiektów aparaturowych znajdujących się w sieci. Sieć składa się z węzłów i odcinków, czyli zadanie jej struktury

polega na podaniu współrzędnych przestrzennych węzłów, ich typów oraz długości i średnic odcinków.

Podstawowe typy węzłów, to:

- węzły zasilające sieć,
- węzły odbiorcze,
- węzły montażowe.

Z kolei podstawowe obiekty aparaturowe sieci wodociągowej to:

- zasuw,
- zawory zwrotne,
- reduktory ciśnienia,
- zbiorniki wyrównawcze, układy pompowe w węzłach zasilających na wejściu sieci i w hydroforniach umieszczonych wewnątrz sieci.

Komputerowe modele przepływów stanowią obecnie wysoce wydajne narzędzia analizy hydraulicznej zjawisk przepływów w warunkach ustalonych, wspomagające pracę inżyniera przy eksploatacji, modernizacji czy rozbudowie systemu wodociągów miejskich. Istota pracy z modelem hydraulicznym polega na odtworzeniu rzeczywistych warunków pracy sieci w stopniu możliwie najbliższym wiarygodnym. Zgodnie z definicją Wosiewicza (1996) „*model matematyczny jest teoretycznym opisem matematycznym odzwierciedlającym z żadaną dokładnością zachowanie się oryginału (obiektu procesu) w warunkach rzeczywistych*”. Model matematyczny w ostatecznym kształcie jest tylko podobny do oryginału. Z powodu złożoności zagadnienia konieczne jest wprowadzenie pewnych uproszczeń w modelu, co pozwala odwzorować tylko istotne cechy danego obiektu.

Metody używane do rozwiązywania równań ciągłości przepływu i strat opisujących stan hydrauliczny sieci w danym przedziale czasowym, mogą być ustalone hybrydową węzłowo-pętlową metodą iteracyjną. Todini i Pilati (1978), a później Salgado (1988) nazwali tę metodę Metodą Gradientową. Podobne podejście zostało opisane przez Hamam'a i Brameller'a (1971) i przez Osiadacz (1987). Jediną różnicą pomiędzy tymi metodami jest sposób, w jaki przepływy w połączeniach są uaktualniane po nowej próbie rozwiązania dla wysokości węzłów.

Zakładając, że sieć składa się z N złączy węzłowych i NF węzłów stałego stopnia (zbiorniki i rezerwuary), to zależność pomiędzy stratami przepływu w rurze pomiędzy węzłami i oraz j określa się jako:

$$H_i - H_j = h_{ij} = rQ_{ij}^n + mQ_{ij}^2$$

gdzie: H – wysokość ciśnienia w węźle, h – straty, r – współczynnik oporów, Q – natężenie przepływu, n – wykładnik przepływu, m – współczynnik strat miejscowych.

Drugim zestawem równań, które muszą być spełnione, są równania zapewniające utrzymanie ciągłości przepływu we wszystkich węzłach:

$$\sum_j Q_{ij} - D_i = 0 \quad \text{dla } i = 1, \dots, N$$

gdzie: D_i – rozbiór w węźle i , przy założeniu, że dopływ do węzła jest dodatni.

Mając dane wysokości w węzłach stałych szukamy rozwiązania dla wszystkich wysokości H_i i przepływów Q_{ij} spełniających równania (1) i (2).

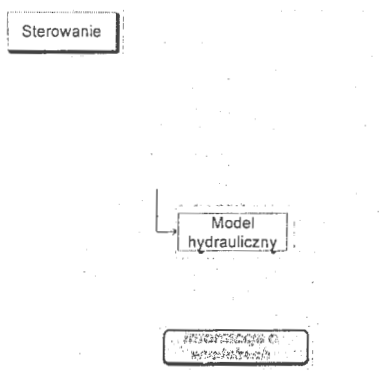
Rozwiązanie metodą gradientową rozpoczyna się od początkowego przybliżenia przepływu w każdej z rur, przy czym niekoniecznie zapewniają one ciągłość przepływu. W każdej kolejnej iteracji tej metody zostaje znajdowana nowa wysokość w węzłach dzięki rozwiązaniu macierzy równań:

$$\mathbf{AH} = \mathbf{F}$$

gdzie: \mathbf{A} – macierz Jacobiego ($N \times N$), \mathbf{H} – wektor nieznanych wysokości w węzłach ($N \times 1$), \mathbf{F} – wektor warunku prawostronnego ($N \times 1$).

3.2. Modele prognozujące obciążenie sieci

Obciążenie sieci zmienia się cyklicznie w zależności od pory roku, dnia tygodnia czy pory dnia. Te zmiany mają wpływ na sterowanie siecią, w szczególności na napelnianie i opróżnianie zbiorników wyrównawczych zlokalizowanych w sieci: przewidując większe obciążenie sieci powinno się odpowiednio wcześniej zbiorniki napelnić, podczas gdy można je opróżniać w przypadku mniejszego zapotrzebowania na wodę. Również do wykonywania dokładnych obliczeń hydraulicznych sieci należy znać godzinowe rozkłady obciążenia już nie tylko całej sieci, ale również jej poszczególnych węzłów lub grup węzłów. Aby takie modele obciążenia sieci względnie węzłów sieci wyznaczyć, należy dysponować odpowiednio długimi seriami pomiarów przepływu z zainstalowanego na sieci systemu monitoringu.



Rys. 3. Wyznaczanie modeli prognozujących obciążenie sieci.

Poniżej przedstawiono wyniki badań dotyczących możliwości prognozowania dobowego poboru wody na 1 dzień, na podstawie archiwalnych jej wartości, czyli wyznaczenia modelu postaci (rys. 3):

$$p_w(k+1) = f(p_w(k), p_w(k-1) \dots p_w(k-r))$$

gdzie: p_w – dobowa produkcja wody, k – dyskretna chwila czasowa, r – rząd modelu.

W ostatnich latach popularne stało się wykorzystywanie w tym celu sztucznych sieci neuronowych (SSN). Prognozowanie z użyciem SSN daje dobre rezultaty, charakteryzuje się jednak pewnymi wadami, do których można zaliczyć (Osowski, 2000):

- brak możliwości analizy struktury wyznaczonego modelu, zapis sztucznej sieci neuronowej zakodowany jest w formie niezrozumiałej dla człowieka,
- sztuczne sieci posiadają ograniczone zdolności ekstrapolacji poza zakres danych na postawie których zostały wyznaczone.

Inną grupą modeli intensywnie badaną i rozwijaną w ostatnich latach są różnego typu modele oparte na logice rozmytej. Logika rozmyta posiada zdolność opisu obiektu w bardzo ogólnej formie zrozumiałej dla człowieka, co w połączeniu ze swobodą doboru nieliniowości oraz możliwościami opracowania dla niej algorytmów identyfikacyjnych daje interesujący typ modelu. Model rozmyty jest zapisywany w postaci reguł

Jeżeli <stwierdzenie rozmyte> to <stwierdzenie rozmyte>

gdzie stwierdzenie rozmyte określone jest na zbiorze funkcji przynależności zdefiniowanych na obszarze zmiennych fizycznych połączonych operatorami rozmytymi. Stwierdzenie występujące jako pierwsze nazywamy przesłanką bądź poprzednikiem, drugie nazywamy konkluzją bądź następnikiem.

W wykonanych badaniach sprawdzano możliwości wykorzystania modeli o strukturze Takagi Sugeno Kang'a, w których następnikiem jest klasyczny model liniowy uwzględniający dynamikę, co pozwala na określenie struktury jako "modele liniowe strefami rozmywane" (Takagi, Sugeno, 1985).

Tego typu model, w porównaniu do wcześniej wymienionych, charakteryzuje się interesującymi cechami:

- dzięki uwzględnieniu dynamiki obiektu w funkcji następnika modele TSK są w stanie dość dobrze naśladować proces nawet przy niewielkiej liczbie przedziałów rozmywania,
- następnikiem jest klasyczny model liniowy, co pozwala traktować nieliniowy model rozmyty jako zbiór lokalnych modeli liniowych. Daje to możliwość stosowania

bezpośrednio teorii sterowania opracowanej dla obiektów liniowych przy syntezie regulatora,

- bardzo prosta jest linearyzacja tego modelu w dowolnym punkcie pracy,
- modele liniowe są dobrze przebadane i szeroko znane, porównanie jego podstawowych właściwości (opóźnienia, stałe czasowe, rząd) z przesłankami wynikającymi z analizy teoretycznej identyfikowanego procesu (Ljung, 1987).

Obiekt badań

W przeprowadzonych badaniach wykorzystano dane pochodzące z sieci wodociągowej miasta Rzeszowa. Miejskie przedsiębiorstwo wodociągów i kanalizacji w Rzeszowie eksploatuje:

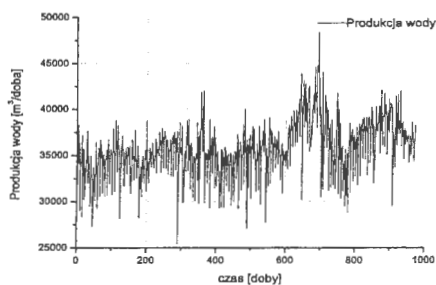
- sieć wodociągową o łącznej długości 864,1 km, rys. 4,
- 2 ujęcia powierzchniowe wody na rzece Wisłok o łącznej zdolności produkcyjnej 84,0 tys. m³/dobę,
- awaryjne ujęcie wgłębne przy ul. Mazowieckiej w Rzeszowie o wydajności 240 m³/dobę,
- lokalne ujęcie wody Osiedle Słocina w Rzeszowie o wydajności 465m³/dobę,
- 32 przepompownie wody (hydrofornie),
- 12 zbiorników wyrównawczych wody czystej (Krakowska Południe, Pobitno Słocina, Słocina Roch, Pustki) o łącznej pojemności 34,1 tys.m³,
- 187 studni publicznych, 12 źródeł ulicznych które znajdują się na majątku Gminy Miasta Rzeszów.



Rys. 4. Schemat magistrali wodociągowej miasta Rzeszowa.

Badane szeregi czasowe

Do przeprowadzenia badań związanych z modelowaniem wykorzystano szeregi czasowe zarejestrowanych wartości dobowych produkcji wody. Dane pozyskane zostały z przedsiębiorstwa wodociągowego w Rzeszowie, okres pomiarów obejmował okres od 1.07.2005r do 29.02.2008r, liczebność zbioru danych wynosiła 974 próbki. Zarejestrowane wartości modelowanego sygnału przedstawione zostały na rys. 5, natomiast jego charakterystyka w tabeli 1.

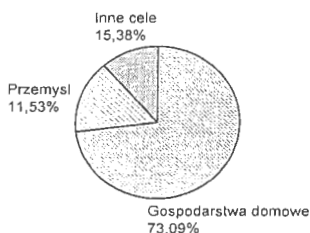


Rys. 5. Zarejestrowane wartości dobowej produkcji wody.

Tabela 1. Charakterystyka szeregu czasowego użytego w badaniach.

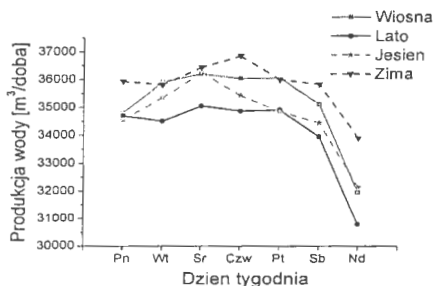
Liczba próbek	974
Wartość średnia	35788
Wartość minimalna	25433
Wartość maksymalna	48394
Odchylenie standardowe	2901,4

Zdecydowanie największa ilość wody sprzedawana jest w Rzeszowie gospodarstwom domowym (73,19%), tylko 11,53% wody używana jest przez przedsiębiorstwa przemysłowe (rys. 6). Pobór wody przez mieszkańców ma zazwyczaj charakter cykliczny, więc uznano, że możliwe jest wyznaczenie przyszłych jej wartości na podstawie zarejestrowanych sygnałów z chwil przeszłych.



Rys. 6. Podział na grupy odbiorców produkowanej wody.

Na podstawie zarejestrowanych danych porównano tygodniowe wartości produkcji wody w różnych porach roku. Z danych przedstawionych na rys. 7 wynika, że w ciągu całego badanego okresu najniższe pobory wody występowały w niedzielę. Największe sumaryczne pobory wody występowały w zimie, a najniższe w lecie. Produkcja wody jesienią i wiosną była porównywalna. Spadek zużycia wody zarówno w niedzielę, jak i w lecie, można tłumaczyć wyjazdami mieszkańców z miasta. (największe zużycie w Rzeszowie występowało w gospodarstwach domowych, rys. 6).



Rys. 7. Porównanie tygodniowego poboru wody dla różnych pór roku.

Modele rozmyte typu Takagi Sugeno Kang'a

Model TSK zalicza się do ogólnej grupy modeli rozmytych. Schemat przetwarzania takich modeli można w uproszczeniu opisać następująco (rys. 8): zmierzone wartości wejściowe są poddawane procesowi rozmywania. Następnie przeprowadzane jest wnioskowanie rozmyte w oparciu o zestaw reguł wchodzących w skład modelu. Na koniec następuje proces wyostrzenia i wyliczenie konkretnej, ostrej wartości wyjścia modelu.



Rys. 8. Schemat obliczania wyjścia modelu rozmytego.

Charakterystyką modelu TSK jest to, że w konkluzji każdej reguły występuje w miejsce zbioru rozmytego analityczna funkcja zmiennych wejściowych. Tak więc każda reguła modelu TSK przyjmuje następującą postać:

$$R_i : \text{if } [x_1 = A_i^1 \wedge \dots \wedge x_n = A_i^n] \text{ then } [y = f_i(u)]$$

gdzie: R_i to i -ta reguła wiedzy, x_j to j -te wejście rozmywane, A_j^i to j -ta partycja zapisana w i -tej regule modelu, n to liczba wejść rozmywanych, y wyjście modelu, f_i funkcja następnika (konkluzja) i -tej reguły, u to wektor wejść modelu liniowego - funkcji następnika.

Model składa się z M reguł. Rozmyte przesłanki tych reguł dzielą przestrzeń wejściową na określoną liczbę rozmytych regionów, natomiast funkcje następnika opisują zachowanie się modelu w tych regionach. Przesłanka jest zbudowana z jednego lub więcej stwierdzeń rozmytych połączonych rozmytym operatorem koniunkcji. Zmienne x_1, x_2, \dots, x_n tworzą n -wymiarowy wektor wejść rozmywanych. Wypadkowa (dla całego modelu) wartość wyjścia jest obliczana jako ważona suma wyjść poszczególnych reguł, gdzie wagą jest wypadkowy stopień przynależności danej reguły $\mu_{A'}(x)$:

$$y = \sum_{i=1}^M \mu_A^i(x) f_i(u)$$

gdzie $\mu_A^i(x)$, to spełnienie i -tej reguły określone jako iloczyn spełnienia poszczególnych jej przesłanek:

$$\mu_A^i(x) = \prod_{j=1}^n \mu_{A_j}(x_j)$$

Takie podejście upraszcza sposób obliczania wyjścia z modelu dzięki wyeliminowaniu procesu wyostrzania. Najczęściej wykorzystywaną postacią funkcji f_i jest wielomian 1-go stopnia składowych wektora \mathbf{u}

$$f_i(u) = a_0^i + a_1^i u_1 + \dots + a_L^i u_L$$

Wektor \mathbf{x} definiuje punkt pracy modelu, natomiast wektor \mathbf{u} zawiera informacje potrzebne do obliczenia jego dynamicznej odpowiedzi. Jako funkcje przynależności wykorzystane zostały funkcje trapezoidalne:

$$\mu(x) = \max\left(0, \min\left(1, \frac{x-a}{b-a}, \frac{d-x}{d-c}\right)\right)$$

gdzie: a, b, c, d – punkty charakterystyczne funkcji trapezoidalnej ($d > c, b > a$) (rys. 9):



Rys. 9. Trapezoidalna funkcja przynależności.

Wyznaczanie modeli TSK

Identyfikacja modelu TSK jest procesem iteracyjnym, który można zawrzeć w dwóch fazach (Hellendorn, Driankov, 1997):

1. Określenie wszystkich parametrów rozmywania, tj. określenie zbioru sygnałów rozmywanych, parametrów poszczególnych funkcji przynależności itd. Zadanie to wykonane zostało przy wykorzystaniu algorytmów ewolucyjnych (Wnuk, 2005).
2. Estymacja współczynników a'_0, a'_1, \dots, a'_L wielomianu $f_i(u)$ dla każdej (i -tej) konkluzji z wykorzystaniem klasycznych metod obliczeniowych; wykorzystano metodę najmniejszej sumy kwadratów LS. Rozwiązanie zadania tego typu można przedstawić następująco:

$$\tilde{\Theta} = [V^T V]^{-1} V^T y$$

gdzie:

$$\tilde{\Theta} = [\Theta_1, \Theta_2, \dots, \Theta_M]^T$$

$$\Theta_i = [a'_0, a'_1, \dots, a'_L]$$

oraz:

$$V = \begin{bmatrix} \mu_A^1(x(1))u(1), \dots, \mu_A^M(x(1))u(1) \\ \dots \\ \mu_A^1(x(N))u(N), \dots, \mu_A^M(x(N))u(N) \end{bmatrix} \in \mathbb{R}^{N \times (M \times L)}$$

$$y = \begin{bmatrix} \underbrace{y(1), \dots, y(N)}_1, \dots, \underbrace{y(1), \dots, y(N)}_M \end{bmatrix}^T \in \mathbb{R}^{N \times M}$$

gdzie $u(k)$ - wektor wartości wejść modelu w dyskretnej chwili czasowej k , – wartość wyjścia modelu w dyskretnej chwili czasowej k .

Przy wyznaczaniu modelu rozmytego typu TSK, należy zwrócić szczególną uwagę na tzw. *poziom walidacji* poszczególnych modeli cząstkowych, tzn. liczebność zbioru danych (Hellendorn, Driankov, 1997), na podstawie których wyznaczona zostanie każda z konkluzji modelu:

$$\chi_i = \sum_{k=1}^N \mu_i(k)$$

Wartość ta nie może być zbyt mała, tzn. $\chi_i \geq L$ (L – liczba współczynników równania modelu), ponieważ w przeciwnym wypadku może zostać wyznaczony model o błędnie określonej dynamice. Zwykle przyjmuje się relację $\chi_i \geq (10 - 20)L$.

Zarejestrowane dane podzielono na dwa zbiory uczący (50%) i walidacyjny (50%). Współczynniki modelu wyznaczano na 60% losowo wybranych próbek ze zbioru uczącego, natomiast ocena modelu odbywała się z wykorzystaniem średniego błędu modelu MAE (ang. *Mean Absolute Error*), na wszystkich próbkach zbioru uczącego.

Struktury wyznaczanych modeli

Przez strukturę modelu TSK rozumie się sygnały rozmywające oraz ich podział na partycję oraz postać konkluzji modelu, czyli równania różnicowego. W przeprowadzonych rozważaniach praktycznie dostępne były trzy sygnały:

- produkcja wody, oznaczany dalej jako w ,
- dzień tygodnia, oznaczany dalej jako D_t , przy czym kolejnym dniom przypisano kolejne, całkowite wartości liczbowe, począwszy od poniedziałku (wartość 1) i skończywszy na niedzieli (wartość 7),
- pora roku P_r , przy czym kolejnym porom roku przypisano kolejne, również całkowite wartości liczbowe, począwszy od wiosny (wartość 1) a skończywszy na zimie (wartość 4).

Wpływ poszczególnych sygnałów na wartość produkcji wody oszacowano na podstawie wartości współczynnika korelacji r *Pearsona* poszczególnych sygnałów (w , D_t , P_r) z sygnałem produkcji wody w , tab. 2:

Tabela 2. Współczynnik korelacji r *Pearsona* dostępnych sygnałów z sygnałem produkcji wody.

Sygnal:	$x[k]$	$x[k-1]$	$x[k-2]$	$x[k-3]$	$x[k-4]$	$x[k-5]$	$x[k-6]$
w	1	0,416	0,171	0,058	0,013	0,052	0,013
D_t	-0,381	-0,318	0,017	0,09	0,261	0,286	0,259
P_r	0,04	0,05	0,06	0,07	0,08	0,08	0,09

Analizując wyniki, przedstawione w tab. 2 można zauważyć, iż wartości autokorelacji sygnału produkcji wody bardzo szybko maleją. Dla dyskretnej chwili czasowej $k-1$ wartość ta wynosi 0,416, co jest wartością porównywalną z wartością współczynnika r *Pearsona* dla korelacji dnia tygodnia z sygnałem produkcji wody w chwili k . Korelacja produkcji wody z porą roku oscyluje w praktyce w okolicach zera.

Biorąc pod uwagę wyznaczone wartości współczynników korelacji, przyjęto następującą strukturę wyznaczanych modeli:

- Sygnałami rozmywającymi były: wartość produkcji wody w chwili k (ostatnia dostępna pomiarowo wartość), i/lub indeks dnia tygodnia w chwili $k+1$ (indeks dnia tygodnia dla, którego wyznaczana jest prognoza).
- Równanie różnicowe konkluzji zawierało archiwalne dane produkcji wody i/lub indeks dnia tygodnia

Dyskusja otrzymanych wyników

Zbiór danych uczących zawierał 487 próbek. Aby zależność na liczebność zbioru danych została spełniona, należało wyznaczyć odpowiednią liczbę zbiorów rozmytych, tak aby łączna liczba współczynników modeli cząstkowych nie przekraczała 48 współczynników. Najlepsze wyniki modelowania uzyskano dla konkluzji zawierających wyraz wolny, wartość indeksu dnia tygodnia dla chwili $k+1$ oraz 6 archiwalnych wartości produkcji wody, co daje, maksymalnie, liczbę 8 współczynników w równaniu. Na tej podstawie określono, że maksymalna liczba zbiorów rozmytych może wynosić 6.

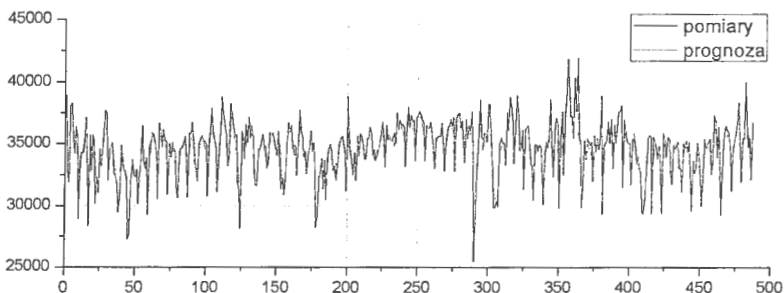
W badaniach wyznaczono modele dla różnych kombinacji rozmywania dostępnych sygnałów (tab. 3), przy czym należy zaznaczyć, iż wyniki uzyskane dla różnych modeli są

porównywalne. Nie są tutaj dyskutowane równania różnicowe poszczególnych modeli cząstkowych, gdyż jest ich zbyt dużo (4 do 6 w każdej wersji modelu).

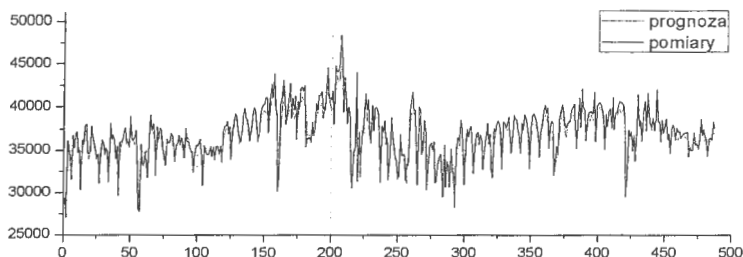
Komentując uzyskane wyniki modelowania wyniki, można uznać, że wyznaczone modele obniżają poziom reakcji na pobudzenia, co jest częstym efektem przy wykorzystaniu estymatora LS. Powodowane jest to zakłóceniami pomiarowymi, bądź wpływem innych zakłóceń o charakterze losowym. Wysoki poziom korelacji sygnału modelowanego i zarejestrowanego (0,87 dla danych uczących, 0,92 dla danych testowych; współczynnik korelacji jest wyższy dla danych testowych ze względu na większą zmienność sygnału produkcji wody w tym zbiorze) pozwala określić jakość modelowania jako zadowalającą. Średni błąd absolutny dla danych uczących wynosił ok. 850 m³/dobę, dla danych testowych 990 m³/dobę, co pozwala określić względny błąd absolutny na poziomach odpowiednio 2,4 % dla danych uczących i 2,7% dla danych testowych.

Tabela 3. Wskaźniki jakości i opis struktury wyznaczonych modeli.

Lp.	Struktura modelu			Wskaźniki jakości					
	Liczba zbiorów rozmytych przesłanki		Postać konkluzji	MAE [m ³ /dobę]		e_{wzg} [%]		r [-]	
	Dni tygodnia	Produkcja wody		uczenie	test	uczenie	test	uczenie	test
1	4	0	$y(k+1) = a_0^0 + a_1^1 D_t + \sum_{i=0}^6 a_2^i w(k-i)$	852	997	2,461574	2,69773	0,876	0,924
2	3	2	$y(k+1) = a_0^0 + \sum_{i=0}^6 a_2^i w(k-i)$	850	1026	2,455796	2,776199	0,871	0,916
3	6	0	$y(k+1) = a_0^0 + \sum_{i=0}^6 a_2^i w(k-i)$	855	1004	2,470242	2,716671	0,877	0,922
4	0	6	$y(k+1) = a_0^0 + a_1^1 D_t + \sum_{i=0}^6 a_2^i w(k-i)$	832	989	2,403791	2,676083	0,895	0,897
5	6	0	$y(k+1) = a_0^0 + a_1^1 D_t + \sum_{i=0}^6 a_2^i w(k-i)$	873	986	2,522247	2,667965	0,872	0,92



Rys. 10. Szeregi czasowe zarejestrowanych i prognozowanych wartości produkcji wody, dla danych uczących.



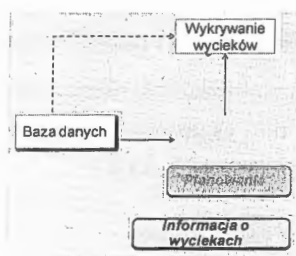
Rys. 11. Szeregi czasowe zarejestrowanych i prognozowanych wartości produkcji wody, dla danych testowych.

Wartości korelacji sygnału prognozowanego i modelowanego (przykładowe przebiegi dla modelu nr 2 (tab. 3) przedstawiono na rys. 10 i 11) dla danych testowych były niewiele mniejsze od wartości uzyskanych w (Licznar, Łomotowski, 2006) przy prognozowaniu z wykorzystaniem sieci neuronowych, z użyciem radialnych funkcji bazowych ($r = 0,93 - 0,94$). Należy pamiętać, że w badaniach wykorzystujących sieci neuronowe autorzy wykorzystywali zbiór 7-letniego szeregu czasowego, a nie jak w obecnych badaniach, zbiór 2-letniego szeregu.

3.3. Wykrywanie stanów awaryjnych oraz nieuzasadnionych poborów wody

W przypadku normalnej eksploatacji sieci mamy do czynienia z pewnym typowym rozkładem ciśnień w węzłach oraz z typowymi zmianami tego rozkładu. Wystąpienie awarii jest sytuacją nietypową, która wpływa na nietypową zmianę ciśnień węzłowych w obszarze awarii. Sygnalizacja tej zmiany przez system monitoringu jest informacją dla operatora sieci o możliwości wystąpienia awarii. Projektowany system informatyczny umożliwi dwa rodzaje wykrywania stanów awaryjnych: poprzez porównanie rozbiorów rzeczywistych zadeklarowanymi (z poziomu bazy danych, dla punktów w których istnieje monitorowanie), oraz poprzez porównanie rozbiorów zadeklarowanych, z obliczonymi na podstawie modelu (dla punktów, w których nie ma monitoringu) (rys. 12).

Często zdarza się praktyce, że kradnie się wodę z sieci wodociągowej. Oznacza to, że produkcja wody przekracza udokumentowany rozbiór wody. W przypadku braku systemu monitoringu można jedynie stwierdzić wystąpienie takiego zdarzenia w skali całej sieci. Natomiast jeżeli istnieje dokładny monitoring, to porównując rejestrowane odbiory wody przez użytkowników sieci z przepływami monitorowanymi przestrzennie można zlokalizować miejsca sieci, w których występują istotne nierówności między poborem i produkcją, a więc wskazać miejsca potencjalnych kradzieży względnie niejawnych wycieków wody.



Rys. 12. Wykrywanie wycieków.

Przykładowy algorytm wykrywania i lokalizacji wycieków ukrytych składa się z następujących kroków:

1. Wyznaczenie punktów wrażliwych sieci wodociągowej przy użyciu algorytmu do optymalnego planowania systemu monitoringu (Straubel, Holznel, 1999).

2. Wybór odpowiedniej liczby najbardziej wrażliwych punktów pomiarowych, z których pozyskuje się największy zasób informacji o występujących w sieci stanach awaryjnych.
3. Wykonanie obliczeń hydraulicznych sieci dla zadanego poziomu obciążenia, za pomocą wybranego i odpowiednio skalibrowanego modelu hydraulicznego.
4. Zapis wartości przepływów i ciśnień w punktach pomiarowych systemu monitoringu dla wykonanych obliczeń (zapis standardowego rozkładu ciśnień i przepływów).
5. Symulacja wycieków w kolejnych węzłach sieci wodociągowej za pomocą modelu hydraulicznego.
6. Zapis wartości przepływów i ciśnień w punktach pomiarowych dla wykonanych symulacji stanów awaryjnych (zapis awaryjnych rozkładów ciśnień i przepływów).
7. Utworzenie klasyfikatora lokalizującego awarie w postaci różnego typu sieci neuronowych i wybór klasyfikatora najlepszego według kryterium największej wrażliwości.
8. Rejestracja z systemu monitoringu pomiaru odbiegającego od standardu dla zadanego poziomu obciążenia sieci, czyli w przypadku wystąpienia awarii, znalezienie wśród modeli sieci neuronowych modelu takiego, który wskazuje dokładnie węzeł sieci z zaistniałym wyciekami.
9. Zmiana poziomu obciążenia sieci wodociągowej i powrót do punktu 3.
10. Porównanie i weryfikacja wyników dla różnych modeli hydraulicznych i różnych sieci neuronowych w celu najlepszego zamodelowania sieci wodociągowej i najdokładniejszego zlokalizowania awarii.

Wyznaczenie punktów wrażliwych sieci wodociągowej przy użyciu algorytmu do planowania systemu monitoringu

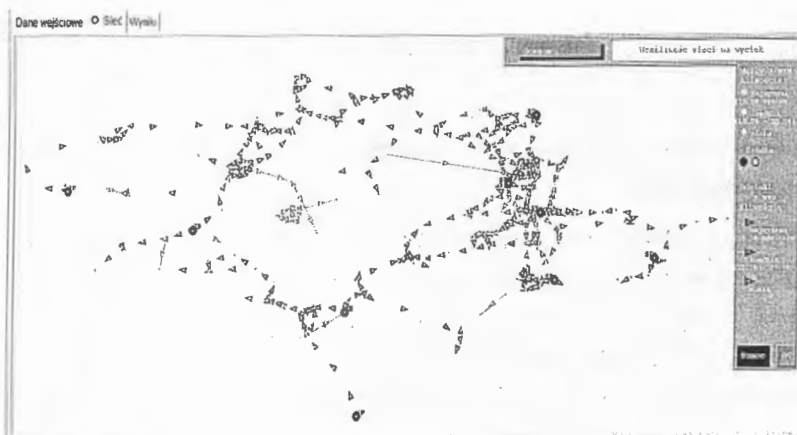
W celu znalezienia najlepszych lokalizacji czujników w punktach pomiarowych systemu monitoringu wyznacza się tzw. punkty wrażliwe sieci wodociągowej przy użyciu algorytmu do planowania systemu monitoringu (Straubel, Holznagel, 1999). Aby zmaksymalizować ilość informacji przekazywanych przez punkty pomiarowe, powinny one być tak wybrane, aby przekazywały wiedzę o stanie sieci nie tylko z danego punktu, ale również z jego dalszego

otoczenia. Takie punkty charakteryzujące się dużą wrażliwością na zmiany stanu sieci nazywa się punktami charakterystycznymi. Maksymalizacja przekazywanej informacji oznacza zwykle, w tradycyjnie planowanych systemach monitoringu, zwiększanie liczby punktów pomiarowych, a więc działanie przeciwne do opisanego powyżej.

Dlatego należy podkreślić, że odpowiedni wybór stosunkowo niewielkiej liczby punktów charakterystycznych może być równoważny, pod względem jakości i ilości przekazywanej wiedzy o sieci, większej liczbie punktów zlokalizowanych w mniej wrażliwych miejscach sieci.

Przy tym, jako miarę wrażliwości sieci na symulowane wycieki można stosować na przykład sumę różnic bezwzględnych między pomiarami ciśnień dla standardowego stanu pracy sieci i stanów awaryjnych dla przyjmowanych różnych kombinacji liczby i lokalizacji punktów pomiarowych, w przypadku, gdy w systemie monitoringu uwzględniamy jedynie pomiary ciśnienia.

Na rysunku 13 została przedstawiona sieć wodociągowa jednej dzielnicy Rzeszowa. Odpowiedni program obliczył wrażliwość tej sieci na wycieki.



Rys. 13. Wrażliwość sieci wodociągowej Rzeszowa na wyciek.

Następnym etapem algorytmu jest wprowadzenie wybranych punktów pomiarowych przepływu do planowanego systemu monitoringu. Punkty pomiarowe zostały wyznaczone w

miejscach o największej wrażliwości na wycieki. Eksperymenty zostały przeprowadzone dla dwóch przypadków. Pierwszy zawierał 10 punktów pomiarowych, a drugi 20. Zwiększenie liczby punktów pomiarowych spowodowało w konsekwencji zwiększenie dokładności lokalizacji awarii w sieci.

Na rysunku 14 zostały przedstawione wybrane punkty monitoringu uwzględnione w dalszych obliczeniach. Na początku obliczeń przyjmuje się, że wartość pomiaru jest równa 0.

Id. FM	Rodzaj pom.	Id. odcinka	Położenie	Stan	Wzrost pomiaru
FFF001	przepływ [l/s]	'3752'	na odcinku	aktywny	0.0
FFF002	przepływ [l/s]	'3758'	na odcinku	aktywny	0.0
FFF003	przepływ [l/s]	'2023'	na odcinku	aktywny	0.0
FFF004	przepływ [l/s]	'7460'	na odcinku	aktywny	0.0
FFF005	przepływ [l/s]	'5400'	na odcinku	aktywny	0.0
FFF006	przepływ [l/s]	'0689'	na odcinku	aktywny	0.0
FFF007	przepływ [l/s]	'0043'	na odcinku	aktywny	0.0
FFF008	przepływ [l/s]	'7056'	na odcinku	aktywny	0.0
FFF009	przepływ [l/s]	'0900'	na odcinku	aktywny	0.0
FFF010	przepływ [l/s]	'5839'	na odcinku	aktywny	0.0

Rys. 14. Punkty monitoringu – Rzeszów.

Wykonanie obliczeń hydraulicznych sieci dla zadanego poziomu obciążenia, za pomocą modelu hydraulicznego

Wykonanie obliczeń hydraulicznych sieci, zapis wartości przepływów i ciśnień w punktach pomiarowych systemu monitoringu dla wykonanych obliczeń (zapis standardowego rozkładu ciśnień i przepływów), symulacja wycieków w kolejnych węzłach sieci wodociągowej za pomocą modelu hydraulicznego w miejscach wrażliwych oraz zapis wartości przepływów i ciśnień w punktach pomiarowych dla wykonanych symulacji stanów awaryjnych (zapis awaryjnych rozkładów ciśnień i przepływów), służą do przygotowania plików uczących dla sieci neuronowych, które posłużą z kolei jako modele do lokalizacji awarii w sieci wodociągowej.

Wykonanie obliczeń hydraulicznych sieci powoduje pojawienie się określonych wartości pomiaru w punktach pomiarowych (rys. 15).

Wyniki dla punktów monitoringu

Id. pom.	Typ pom.	Id. odcinka	Polowanie	Odzn. [mSW]	Clm. obl.	Różnica	Prz. f. [l]	Prz. obl.	Różnica
FFF001	przepływ	'3752'	na odc.				0.0	-0.066	-0.066
FFF002	przepływ	'3758'	na odc.				0.0	-0.066	-0.066
FFF003	przepływ	'2023'	na odc.				0.0	0.358	0.358
FFF004	przepływ	'7480'	na odc.				0.0	-0.788	-0.788
FFF005	przepływ	'5400'	na odc.				0.0	0.035	0.035
FFF006	przepływ	'0639'	na odc.				0.0	0.166	0.166
FFF007	przepływ	'0843'	na odc.				0.0	0.166	0.166
FFF008	przepływ	'7056'	na odc.				0.0	2.655	2.655
FFF009	przepływ	'0900'	na odc.				0.0	0.856	0.856
FFF010	przepływ	'5239'	na odc.				0.0	-0.856	-0.856

Ekspozycja [Excel] z nagłówkami kolumn Zamiast

Rys. 15. Wartości pomiaru w punktach pomiarowych – stan normalny sieci.

W ten sposób dokonano zapisu standardowego rozkładu ciśnień i przepływów oraz wykonano szereg symulacji wycieków w różnych węzłach sieci wodociągowej za pomocą modelu hydraulicznego. Symulację awarii wykonano w następujący sposób:

- zwiększono kilkakrotnie wartość przepływu wody w wybranym węźle,
- obliczono hydraulikę sieci wodociągowej z symulowaną awarią,
- obliczono różnicę przepływów w sieci w stanie normalnym i w stanie awarii,
- określono punkt pomiarowy, który najbardziej zareagował na awarię węzła.

Następnie dokonano zapisu wartości przepływów i ciśnień w punktach pomiarowych dla wykonanych symulacji stanów awaryjnych.

Mając dane wartości przepływu wody generowane przez model hydrauliczny sieci, skalibrowany na podstawie rzeczywistych pomiarów, można przystąpić do budowy klasyfikatora w postaci sieci neuronowej, który będzie sygnalizował wystąpienie awarii w sieci wodociągowej oraz podawał jej lokalizację.

Utworzenie i wybór najlepszego klasyfikatora lokalizującego awarie w postaci sieci neuronowych typu MLP

Modele lokalizacji awarii sieci zostały utworzone przy wykorzystaniu sieci neuronowych typu MLP. Sieci MLP niezmiennie są najbardziej rozpowszechnionymi i uniwersalnymi sieciami neuronowymi stosowanymi do rozwiązywania różnych problemów.

W sieci neuronowej jednokierunkowej wielowarstwowej ze wsteczną propagacją błędów (MLP) dobór liczby neuronów w warstwie wejściowej jest uwarunkowany wymiarem wektora danych x . Model neuronu składa się z elementu sumacyjnego, do którego dochodzą sygnały wejściowe x_1, x_2, \dots, x_N , tworzące wektor wejściowy $x = [x_1, x_2, \dots, x_N]^T$ pomnożone przez przyporządkowane im wagi $w_{i1}, w_{i2}, \dots, w_{iN}$ tworzące wektor wag i -tego neuronu $w_i = [w_{i1}, w_{i2}, \dots, w_{iN}]^T$ oraz wartość w_{i0} zwaną progamiem. Sygnał wyjściowy sumatora oznaczono u_i , przy czym sygnał

$$u_i = \sum w_{ij} x_j + w_{i0}$$

jest podawany na blok realizujący nieliniową funkcję aktywacji $f(u_i)$. Najczęściej funkcja ta przyjmuje postać sigmoidalną lub tangensa hiperbolicznego. Postać funkcji stanowi ciągłe przybliżenie funkcji skokowej, przy której sygnał wyjściowy neuronu y_i przyjmuje dowolne wartości z przedziału $[0,1]$ dla funkcji unipolarnej.

Funkcja sigmoidalna unipolarna, to:

$$f_u(u) = \frac{1}{1 + \exp(-\beta u)}$$

Parametr β pozwala regulować pochyłość sigmoidy, co zmienia nieliniowość funkcji.

W ramach prowadzonych badań zostały wykonane eksperymenty z różnymi sieciami neuronowymi. Pierwszy etap badań został wykonany z wykorzystaniem 10 punktów pomiarowych oraz 36 węzłów sieci wodociągowej, w których symulowano awarie. Cała sieć wodociągowa zawiera 390 węzłów. W drugim etapie zwiększono liczbę punktów pomiarowych do 20 i węzłów z symulowaną awarią do 44.

Klasyfikator jest tworzony według opracowanej wcześniej metodyki i umieszczony w specjalnej bazie modeli (Rojek, 2010).

Eksperymenty związane z tworzeniem modeli sieci neuronowych typu MLP były parametryzowane dwoma parametrami: liczbą neuronów w warstwie ukrytej oraz liczbą epok uczących. W eksperymencie parametr 'liczba neuronów w warstwie ukrytej' przyjmował wartości od 5 do 25, natomiast drugi parametr 'liczba epok uczących' przyjmował wartości: 200, 500 i 1000. Po zakończeniu każdego eksperymentu przeprowadzone były testy, które dostarczały informacji na temat błędnie zakwalifikowanych decyzji. Liczba neuronów w warstwie ukrytej została dobrana eksperymentalnie. Wejścia sieci zawierają przepływy wody w wybranych węzłach sieci wodociągowej. Wyjście sieci wskazuje punkt pomiarowy, który lokalizuje awarię.

Na rys. 16 pokazano fragment danych, które posłużyły jako wejścia i wyjście sieci neuronowych.

22 1963	23 2048	24 2152	25 2186	26 2417	27 2740	28 2775	29 3029	30 6144	31 3587	32 3596	33 4138	34 4181	35 4250	36 4111	37 Punkt_pom
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF003
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF003
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF003
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF003
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF003
-7.82	-8.45	-5.74	-4.2	-4.83	-5.04	-94.46	-2.71	-2.79	-3.1	-8.43	-6.31	-2.2	-1.69	-4.76	FFF008

Rys. 16. Fragment danych pliku uczącego sieci neuronową.

Plik uczący zawiera 70% przykładów, plik testowy 15%, a plik walidacyjny kolejne 15%. Liczba przykładów wyniosła w pierwszym przypadku 304 dla 10 punktów pomiarowych i 360 w drugim przypadku.

Tabela 4 wyniki obliczeń dla różnych struktur sieci MLP w przypadku 10 punktów pomiarowych.

Tabela 4. Wyniki obliczeń dla sieci MLP dla 10 punktów pomiarowych.

Id sieci	Nazwa sieci	Jakość uczenie	Jakość testowanie	Jakość walidacja
1	MLP 36-9-11	97,66355	97,77778	95,55556
2	MLP 36-17-11	94,39252	97,77778	93,33333
3	MLP 36-20-11	97,19626	97,77778	97,77778

Natomiast tabela 5 przedstawia wyniki badań dla różnych struktur sieci MLP w przypadku 20 punktów pomiarowych.

Tabela 5. Wyniki obliczeń dla sieci MLP dla 20 punktów pomiarowych.

Id sieci	Nazwa sieci	Jakość uczenie	Jakość testowanie	Jakość walidacja
1	MLP 44-22-21	100,0000	100,0000	100,0000
2	MLP 44-10-21	85,3175	83,3333	81,4815
3	MLP 44-21-21	94,4444	88,8889	92,5926
4	MLP 44-17-21	90,8730	88,8889	94,4444
5	MLP 44-21-21	97,2222	100,0000	98,1481

Najlepszą siecią MLP dla 10 punktów pomiarowych okazała się sieć z 9 neuronami w warstwie ukrytej. Natomiast dla 20 punktów pomiarowych najlepsza jest sieć z 22 neuronami w warstwie ukrytej.

4. Algorytmy optymalizacji

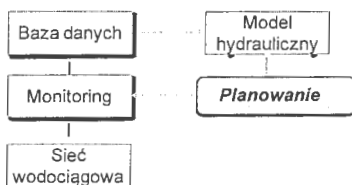
4.1. Planowanie punktów pomiarowych

Wybór odpowiednich punktów pomiarowych do monitorowania sieci wodociągowej jest nietrywialnym zadaniem, do rozwiązania którego można stosować różne algorytmy obliczeniowe. Zazwyczaj jest to optymalizacja wielokryterialna dla wyznaczenia liczby i lokalizacji punktów pomiarowych: Uwzględniając wszystkie możliwe kombinacje liczby i umiejscowienia punktów pomiarowych, wykonuje się obliczenia optymalizacji z następującymi trzema kryteriami celu:

- minimum liczby punktów
- maksimum wrażliwości lokalizacji wycieku
- minimum kosztów instalacji punktów pomiarowych.

Jako miarę wrażliwości lokalizacji wycieku można stosować sumę różnic między pomiarami ciśnień dla stanu standardowego pracy sieci i stanów awaryjnych dla przyjmowanych różnych kombinacji liczby i lokalizacji punktów pomiarowych.

Przy wyznaczaniu punktów pomiarowych używa się modelu hydraulicznego do wykonania obliczeń symulacyjnych w celu określenia wrażliwości poszczególnych punktów na lokalizację wycieku. Wyznaczenie poprawnego modelu wymaga posiadania wiarygodnych danych pomiarowych z działającego już monitoringu. Oznacza to, że aby wyznaczyć model hydrauliczny sieci wodociągowej, należy dysponować systemem monitoringu, i aby zainstalować system monitoringu, należy dysponować modelem hydraulicznym. Problem ten rozwiązuje się zwykle metodą kolejnych przybliżeń, tzn. instaluje się na sieci wodociągowej system monitoringu w ograniczonym zakresie (z punktami pomiarowymi ustalonymi nie tyle w charakterystycznych, co w głównych punktach sieci, jak źródła i przepompownie), za pomocą którego wyznacza się zgrubny model hydrauliczny (opisujący jedynie główne magistrale sieci wodociągowej, tzn. uwzględniający odcinki sieci o największych średnicach), następnie za pomocą tego modelu rozszerza się system monitoringu o kolejne punkty, które służą z kolei do wyznaczenia bardziej szczegółowego modelu, i takie postępowanie można kontynuować aż do momentu uzyskania dokładnego modelu i efektywnego systemu monitoringu (rys. 17).



Rys. 17. Planowanie punktów pomiarowych.

Wybór odpowiednich punktów pomiarowych do monitorowania sieci wodociągowej można zrealizować na przykład za pomocą następującego algorytmu złożonego z czterech kroków (Straubel, Holznagel, 1999):

- Wykonanie obliczeń symulacyjnych przy użyciu modelu hydraulicznego sieci wodociągowej dla stanu standardowej eksploatacji sieci: Dla zadanego średniego obciążenia sieci wyznacza się ciśnienia węzłowe we wszystkich węzłach i przepływy we wszystkich odcinkach sieci.
- Wykonanie obliczeń symulacyjnych dla stanów awaryjnych sieci: Przy symulowanych kolejno w każdym węźle sieci wyciekach awaryjnych, istotnie różnych od obciążenia

średniego, wyznacza się nowe ciśnienia węzłowe we wszystkich węzłach i nowe przepływy we wszystkich odcinkach sieci.

- Sortowanie potencjalnych punktów pomiarowych m według malejącej wrażliwości S_P na zmiany ciśnienia i malejącej wrażliwości S_Q na zmiany przepływu:

$$S_{P_m} = \frac{\sum_{k \neq m} (\Delta p_{mk} / p_m) E_{km}}{\sum_{k \neq m} E_{km}}$$

$$S_{Q_m} = \frac{\sum_{k \neq m} (\Delta q_{mk} / q_m) E_{km}}{\sum_{k \neq m} E_{km}}$$

$$\Delta p_{mk} = |p_{mk} - p_m|, \quad \Delta q_{mk} = |Q_{mk} - q_m|$$

$$m = 1, 2, \dots, N$$

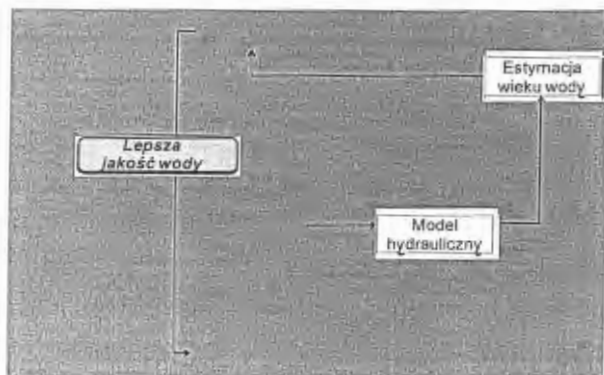
gdzie p_m, q_m oznaczają ciśnienie i przepływ w punkcie m dla średniego obciążenia sieci, p_{mk}, Q_{mk} - ciśnienie i przepływ w punkcie m dla wycieku awaryjnego w punkcie k , E_{km} - najkrótszą odległość liczoną wzdłuż odcinków sieci między punktami k i m , N oznacza liczbę wszystkich węzłów sieci.

- Optymalizacja wielokryterialna dla wyznaczenia liczby i lokalizacji punktów pomiarowych. Uwzględniając wszystkie możliwe kombinacje liczby i umiejscowienia punktów pomiarowych wykonuje się obliczenia optymalizacji z wymienionymi wcześniej trzema kryteriami celu.

4.2. Sterowanie prędkością przepływu wody w sieci

Problem dotyczy poprawiania jakości wody krążącej w sieci wodociągowej. Woda wprowadzana do sieci, po uzdatnieniu w stacji poboru wody, może tracić swoje walory smakowe i zapachowe, jeżeli przebywa w sieci odpowiednio długo i w dodatku nie krąży w niej z odpowiednio dużą prędkością. Szczególnie dotyczy to przypadków sieci starych o rurach stalowych, żeliwnych lub cementowych, które z upływem czasu rdzewieją lub zarastają glonami. Aby uniknąć pogarszania się jakości wody, należy w takich przypadkach

wymusić jej odpowiednio szybkie przepływy w sieci poprzez odpowiednie sterowanie pracą pomp w przepompowniach i stanem zasuw zainstalowanych w sieci (rys. 18). Pomiary z monitoringu są niezbędne do stwierdzania niekorzystnych przestojów wody i realizacji takiego sterowania.



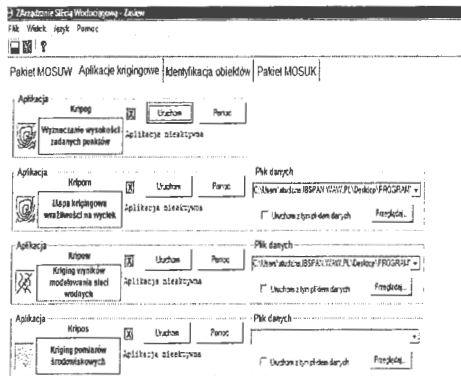
Rys. 18. Sterowanie prędkością wody w sieci.

5. Algorytmy aproksymacji

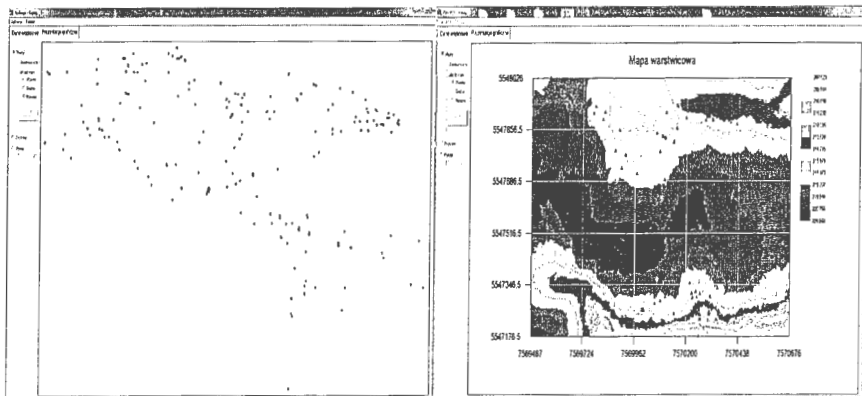
Projektowany system informatyczny w obecnej wersji składa się z 17 programów zlokalizowanych w czterech modułach. Modułem głównym jest oprogramowanie GIS, integrujące wszystkie programy systemu za pośrednictwem swojej bazy danych i odpowiedzialne za komputerową wizualizację sieci wodociągowej, tworzenie grafów sieci dla modelu hydraulicznego, obliczanie wysokości punktów węzłowych sieci oraz wykonywanie różnorodnych analiz tematyczno-przestrzennych dla sieci wodociągowej. Kolejne dwa moduły, to system SCADA monitoringu sieci wodociągowej i program iBOK odpowiedzialny za rejestrację danych o rozbiorach i użytkownikach sieci.

Modułem autorskim jest moduł zawierający programy specjalizujące się w wykonywaniu różnorodnych zadań zarządzania siecią wodociągową i wykorzystujące do tego celu dane zgromadzone w bazie systemu GIS. Programy te są zintegrowane ze sobą za pomocą wspólnego interfejsu, posiadającego trzy zakładki dla programów podzielonych na trzy grupy: grupę programów wykonujących obliczenia hydrauliczne i optymalizacyjne sieci

wodociągowej, grupę programów wykonujących obliczenia aproksymacyjne oraz grupę programów wykonujących obliczenia modelowania matematycznego.



Rys. 19. Grupa aplikacji aproksymacyjnych w projektowanym systemie informatycznym.

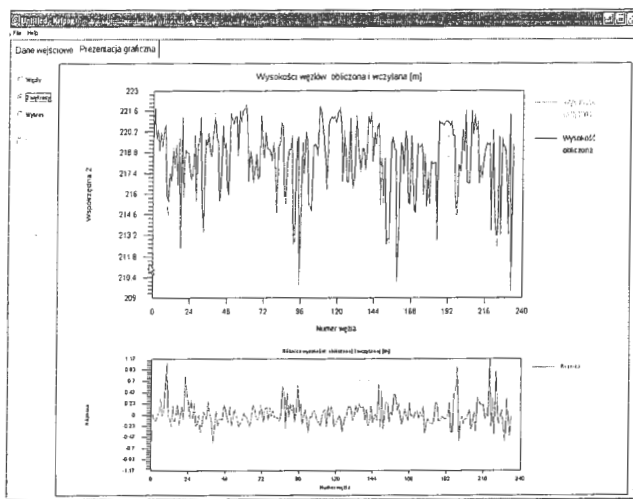


Rys. 20. Wizualizacja punktów węzłowych badanej sieci wodociągowej (rysunek z lewej strony) oraz wykreślona mapa wyznaczonych wysokości terenowych węzłów sieci.

W grupie aplikacji aproksymacyjnych znajdują się programy (rys. 19): obliczania współrzędnych wysokościowych węzłów sieci wodociągowej i wykreślenia map dla

wyznaczonych wysokości terenowych węzłów (rys. 20), wykreślenia map rozkładów przepływów i ciśnień wody w sieci, wykreślenia map wrażliwości sieci na zakłócenia przy planowaniu systemu monitoringu, wykreślenia map wieku wody oraz wykreślenia rozkładów stężeń chloru wprowadzonego do sieci wodociągowej.

Obliczenia aproksymacji w wymienionych programach są wykonywane za pomocą algorytmów kringingowych (Studziński, 2011). W przypadku obliczania współrzędnych wysokościowych węzłów sieci wodociągowej stosuje się w systemie informatycznym, do celów weryfikacji poprawności obliczeń, dwa algorytmy: algorytm kringingowy oraz algorytm aproksymacji powierzchni sieci wodociągowej trójkątami, określonymi na miejskich punktach geodezyjnych (rys. 21).



Rys. 21. Porównanie wyników obliczeń wysokości węzłów sieci za pomocą algorytmu aproksymacji kringowej (*kolor niebieski*) i algorytmu aproksymacji trójkątami (*kolor zielony*).

6. Uwagi końcowe

Projektowany system informatyczny ma na celu wspomaganie kompleksowego zarządzania przedsiębiorstwami wodociągowymi. (symulacja komputerowa realizowanego procesu, optymalizacja i sterowanie procesem itp.). Kluczowym elementem takiego narzędzia jest baza danych z informacjami o sieci wodociągowej, wykorzystywanymi następnie przez

odpowiednie modele matematyczne do prognozowania stanów sieci i generowania scenariuszy sterowania operacyjnego siecią. Z kolei źródłem potrzebnych danych jest odpowiednio zaprojektowany system monitoringu sieci wodociągowej.

Takich systemów nie ma jeszcze w krajowych przedsiębiorstwach wodociągowych, jednak wydaje się, że w tym właśnie kierunku powinien iść rozwój informatyzacji tych przedsiębiorstw. Z kolei kolejnym poziomem informatyzacji byłoby obejmowanie zakresem działania systemu wspomagania decyzji wdrożonego dla potrzeb sieci wodociągowej również innych układów eksploatowanych w przedsiębiorstwie wodociągowym, takich jak sieć kanalizacyjna, oczyszczalnia ścieków oraz pompownie źródłowe tłoczące wodę do sieci wodociągowej.

Literatura

- Bogdan L., Karczmarska D., Studziński J. (2005) *Computerization of waterworks in Poland – current state and perspectives*. In: Applications of Informatics in Environment Engineering and Medicine (Studziński J., Drelichowski L., Hryniewicz O., Eds.) PAS SRI, Series Systems Research, Vol. 42, Warsaw 2005, pp. 157-169.
- Farmani R., Ingeduld P., Savic D., Walters G., Svitak Z., Berka J. (2007). *Real-time modeling of a major water supply system*. International Conference on Computing and Control for the Water Industry, N°8, Exeter, ROYAUME-UNI, vol. 160, no 2, 103-108, 2007.
- Hamam, Y.M. & Brameller, A. 1971. *Hybrid method for the solution of piping networks*. Proc. IEE, Vol. 113, No. 11, pp. 1607-1612.
- Hellendoorn H. i Driankov D. (1997). *Fuzzy Model Identification: Selected Approaches*, Springer-Verlag 1997.
- Licznar P., Łolmotowski J. (2006). *Zastosowanie sztucznych sieci neuronowych Kohonena do prognozowania dobowego poboru wody*. Ochrona Środowiska 2006, Vol. 28, Nr 1, 45-48.
- Ljung L. *System Identification: Theory for the User*, Prentice Hall, Englewood Cliffs 1987.
- Łolmotowski J., Siwoń Z. (2010). *Metodyka analizy danych pochodzących z monitoringu systemów wodociągowo-kanalizacyjnych*. Gaz, Woda i Technika Sanitarna 2010, nr 3, 16–20.
- Osiadacz, A.J. (1987). *Simulation and Analysis of Gas Networks*. E. & F.N. Spon, London.

- Oowski S. (2000). *Sieci neuronowe do przetwarzania informacji*. Oficyna Wydawnictwa Politechniki Warszawskiej, Warszawa 2000.
- Rojek I. (2010). *Wspomaganie procesów podejmowania decyzji i sterowania w systemach o różnej skali złożoności z udziałem metod sztucznej inteligencji*. Wyd. Uniwersytetu Kazimierza Wielkiego, 212 stron, Bydgoszcz 2010.
- Rojek I., Studziński J. (2011). *Algorytmy lokalizacji awarii w sieci wodociągowej przy użyciu sieci neuronowych*. Studia i Materiały Polskiego Stowarzyszenia Zarządzania Wiedzą, Tom 8, Bydgoszcz 2011, 146-156.
- Rossman L.A. and Boulos P.F. (1996). *Numerical methods for modeling water quality in distribution systems: A comparison*. J. Water Resour. Plng. And Mgmt, Vol. 122, No. 2, 137-146.
- Salgado R., Todini E., & O'Connell P.E. 1988. *Extending the gradient method to include pressure regulating valves in pipe networks*. Proc. Inter. Symposium on Computer Modeling of Water Distribution Systems, University of Kentucky, May 12-13.
- Siwoń Z., Cieżak W., Cieżak J. (2011). *Modele neuronowe szeregów czasowych godzinowego poboru wody w osiedlach mieszkaniowych*. Ochrona Środowiska 2011, Vol 33, Nr 2, 23-26.
- Siwoń Z., Cieżak W., Cieżak J. (2006) *Zastosowanie sztucznych sieci neuronowych do prognozowania szeregów czasowych krótkotrwałego poboru wody w wybranych systemach wodociągowych*. Ochrona Środowiska 2006, Vol. 28, Nr 1, 39-44.
- Siwoń Z., Cieżak W., Cieżak J. (2005). *Stochastyczne modele godzinowego poboru wody w wybranym systemie wodociągowym*. Ochrona Środowiska, 2005, nr 1, 7-13.
- Stachura M., Fajdek B., Studziński J. (2012). *Model based decision support system for communal water networks*, Industrial Simulation Conference ISC 2012, Brno.
- Straubel R., Holznagel B. (1999) *Mehrkriteriale Optimierung fuer Palmung und Steuerung von Trink- und Abwasser-Verbundsystemen*. In: Wasser•Abwasser, 140, No. 3, 1999, 191-196.
- Studziński J. (2010). *Innowacje XXI wieku – nowoczesne techniki informacyjne wspomaganie zarządzania w przedsiębiorstwach sieciowych*. W: Innowacyjne Mazowsze (A. Straszak, red.) Wyd. SWPW w Płocku, Płock 2010.

- Studzinski J. (2011). *Application of kriging algorithms for solving some water nets management tasks*. Enviroinfo 2011. Ispra 2011.
- Studzinski J., Bogdan L. (2006). *Informatyczny system wspomagania decyzji do zarządzania, sterowania operacyjnego i planowania miejskiego systemu wodno-ściekowego*. W: Studziński J., Drelichowski L., Hryniewicz O. (Red.): *Rozwój i zastosowania metod ilościowych i technik informatycznych wspomagających procesy decyzyjne*, IBS PAN, Seria Badania Systemowe 49, Warszawa 2006.
- Takagi T. i Sugeno M. (1985). *Fuzzy identifications of system and its applications to modeling and control*. IEEE Trans. on System Man and Cybernetic.
- Todini, E., Pilati, S. (1987). *A gradient method for the analysis of pipe networks*. In: International Conference on Computer Applications for Water Supply and Distribution, Leicester Polytechnic, UK, September 8-10.
- Waterworth G. 2002 *Efficiency modelling in the cost reduction of water pump maintenance*. In: Proceedings of QRM'2002, Ed. G. J. McNulty, PEP London, 141-144.
- Wnuk P. (2005). *Evolutionary optimization for fuzzy TSK model identification*. AI-METH 2005 Recent developments in artificial intelligence methods 2005.

the same way, the β parameter is estimated by the following equation:

$$\beta = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (1)$$

where \bar{x} and \bar{y} are the mean values of the independent and dependent variables, respectively.

Thus, the regression equation for the dependent variable y in terms of the independent variable x is:

$$\hat{y} = \alpha + \beta x \quad (2)$$

where \hat{y} is the predicted value of the dependent variable y for a given value of the independent variable x .

The regression equation (2) is used to estimate the value of y for a given value of x .

The regression equation (2) is also used to estimate the value of x for a given value of y .

The regression equation (2) is also used to estimate the value of α and β .

The regression equation (2) is also used to estimate the value of the standard error of the estimate.

The regression equation (2) is also used to estimate the value of the coefficient of determination.

The regression equation (2) is also used to estimate the value of the F-statistic.

The regression equation (2) is also used to estimate the value of the t-statistic.

The regression equation (2) is also used to estimate the value of the p-value.

The regression equation (2) is also used to estimate the value of the confidence interval.

The regression equation (2) is also used to estimate the value of the prediction interval.

The regression equation (2) is also used to estimate the value of the standard error of the regression.

The regression equation (2) is also used to estimate the value of the standard error of the estimate.

The regression equation (2) is also used to estimate the value of the coefficient of determination.

The regression equation (2) is also used to estimate the value of the F-statistic.

The regression equation (2) is also used to estimate the value of the t-statistic.

The regression equation (2) is also used to estimate the value of the p-value.

The regression equation (2) is also used to estimate the value of the confidence interval.

The regression equation (2) is also used to estimate the value of the prediction interval.

The regression equation (2) is also used to estimate the value of the standard error of the regression.

