

SUBMITTED ARTICLES

PROGNOZOWANIE DOPŁYWU ŚCIEKÓW SUROWYCH DO OCZYSZCZALNI MIEJSKIEJ ZA POMOCĄ ALGORYTMU MACD

Aleksy Barski

*Instytut Badań Systemowych
Polska Akademia Nauk
ul. Nowelska 6, 01-447 Warszawa
e-mail: alex@ibspan.waw.pl*

Streszczenie: *Przedstawiono program do prognozowania trendu dopływu ścieków surowych do oczyszczalni opracowany na podstawie algorytmu MACD (zbieżności/rozbieżności średniej kroczącej). Przewidywanie z odpowiednim wyprzedzeniem czasowym zmian dopływu ścieków surowych, spowodowanych deszczem lub roztopami, ma kluczowe znaczenie przy sterowaniu oczyszczalnią, umożliwiając odpowiednio wczesne skierowanie ścieków do kanału awaryjnego (ulgi). Skuteczność algorytmu MACD w tego rodzaju prognozach przetestowano na przykładzie rzeczywistych danych pomiarowych z oczyszczalni ścieków w Łądku-Zdroju.*

Słowa kluczowe: *Algorytm MACD, prognozowanie dopływu ścieków surowych do oczyszczalni, monitoring oczyszczalni ścieków, ekstrakcja wiedzy z danych pomiarowych.*

Forecasting the raw sewage inflow by means of the MACD algorithm

Abstract: *In the paper a computer program for forecasting the trends of the raw sewage inflows into wastewater purification plants is presented. In the program the algorithm MACD specialized to forecast the prices of shares at the stock market has been implemented. The exact forecasting of raw sewage inflows is of great importance by the control of sewage purification plants making possible some counteractions against the disturbances of the purification process realized there. The efficiency of the MACD algorithm has been tested using the real data from the sewage purification plant in Łądek-Zdrój in Poland.*

Keywords: *MACD algorithm, forecasting the raw sewage inflow into wastewater purification plant, wastewater flow monitoring, data mining.*

1. WSTĘP

Istotnym elementem sterowania oczyszczalnią jest utrzymanie jej funkcjonowania w trybie pracy ciągłej. Jednak często zdarza się, że przy zbyt dużym i gwałtownym dopływie ścieków surowych do oczyszczalni może nastąpić zaburzenie jej poprawnego działania poprzez przeciążenie hydrauliczne oczyszczalni względnie poprzez wypłukiwanie osadu czynnego, odpowiedzialnego za redukcję w ściekach zanieczyszczeń

organicznych i biologicznych. Aby temu zapobiec, przy zbyt dużych dopływach należy skierować przynajmniej część ścieków do tak zwanego kanału ulgi, nie przepuszczając ich przez oczyszczalnię i odprowadzając je bezpośrednio do odbiornika. W większości oczyszczalni takie przekierowywanie ścieków surowych może zająć nawet kilkanaście minut powodując, że jednak ich nadmierna ilość może dostać się do zbiorników oczyszczalni i zakłócić ich pracę. W związku z tym celowe staje się opracowanie sposobu przewidywania, kiedy dopływ ścieków może osiągnąć niebezpieczną granicę, aby zawnocześnie poczynić działania eliminujące zagrożenie.

Ciągłe pomiary dopływu ścieków, pozyskiwane z systemów monitoringu zainstalowanych w oczyszczalni, stanowią pewien szereg czasowy. Za pomocą średnich ruchomych i/lub filtrów regresywnych można uzyskać zarys trendu zmiany dopływu. Jeżeli da się określić trend szeregu czasowego z wyprzedzeniem czasowym do kilkudziesięciu minut, to zwykle wystarcza, aby przewidywać niebezpieczne poziomy dopływu i móc odpowiednio zareagować. praktycznych algorytmie MACD (ang. *Moving Average Convergence/Divergence*)¹.

2. ZASTOSOWANIE I OPIS ALGORYTMU MACD

Ponieważ dla wszelkiego rodzaju analiz dopływu ścieków powinno się uwzględniać trend dopływu i jego zmiany, to użycie wykładniczej średniej kroczącej pozwala na zastosowanie wskaźnika MACD *zbieżności/rozbieżności średniej kroczącej (ruchomej)*, skonstruowanego przez analityka Gerarda Appela [4]. W późniejszym okresie wskaźnik ten był bardzo dobrze przeanalizowany przez innych autorów [7]. Wskaźnik MACD bada zbieżności i rozbieżności średnich ruchomych. Jest on różnicą wartości długoterminowej i krótkoterminowej średniej wykładniczej. Początkowo wskaźnik ten był wykorzystywany do badania sygnałów i prognozowania kupna i sprzedaży akcji. Praktyczne zastosowania wskaźnika MACD, poza handlem papierami wartościowymi, nie są częste, jednak zdarzały się one w rozwiązywaniu następujących problemów:

Poprawa jakości połączenia oraz minimalizacja poboru mocy przez znajdującego się w ruchu użytkownika telefonii komórkowej [8].

Optymalizacji wykorzystania baterii słonecznych [9].

Przewidywanie powstania choroby serca na podstawie pomiarów wagi pacjenta [10].

W badaniach długich ciągów danych pomiarowych wskaźnik MACD jest wyznaczany na podstawie *krzywej sygnału*. Krzywą MACD tworzy się, odejmując średnią krocząca długookresową badanego sygnału od jego średniej kroczącej krótkookresowej a następnie wygładzając otrzymaną różnicę znowu za pomocą średniej kroczącej. Średnia krocząca obliczana jest z danych dotyczących pewnego okresu n wartości wstecz i jest on nazywany oknem lub okresem średniej kroczącej. Jako przykład mogą to być średnie z oknem 26 (długookresowa), z oknem 12 (krótkookresowa) i z oknem 9 (wygładzanie).

Jako średnie kroczące stosowane są: średnie ważone wykładnicze z wagami stanowiącymi ciąg geometryczny, średnie ważone z wagami stanowiącymi ciąg arytmetyczny oraz zwykłe średnie arytmetyczne. Ponieważ średnia wykładnicza znacznie szybciej reaguje na zmiany sygnału, to właśnie ona jest najczęściej używana w tym algorytmie.

Do wizualizacji uzyskanych wyników można zastosować *histogram MACD*, który składa się z pionowych słupków, odzwierciedlających wygładzoną różnicę między dwiema krzywymi. Histogram posiada własny poziom zero. W przypadku pozytywnego ułożenia krzywych okresowych (krzywa szybsza znajduje się ponad krzywą wolniejszą), histogram MACD przebiega ponad swoim poziomem zerowym. Przecięcia tego poziomu od dołu i od góry przez histogram zbiegają się z sygnałami wzrostu i spadku, wynikającymi z przecięć krzywych MACD.

Histogram pozwala określić moment, w którym odstęp między krzywymi zwiększa się lub zmniejsza. W chwili, gdy przebiegający powyżej poziomu zerowego histogram zaczyna opadać w kierunku zera, to trend wzrostowy badanego sygnału pomiarowego słabnie, i na odwrót, w sytuacji, gdy histogram znajduje się poniżej swojego poziomu zerowego, ale zaczyna rosnąć ku niemu, to trend spadkowy badanego sygnału traci impet. W ten sposób określa się nie tylko trend badanego sygnału, ale również przewiduje się jego zmianę.

3. ZASTOSOWANIE ALGORYTMU MACD DO BADANIA TRENDU DOPŁYWU ŚCIEKÓW

Dane pomiarowe pozyskiwane z systemu monitoringu i dotyczące dopływu ścieków surowych do oczyszczalni są w postaci szeregu czasowego. Jednym ze składników takiego szeregu jest składnik losowy, dlatego pierwszym etapem w procesie identyfikacji trendu jest wygładzanie danych pomiarowych. Wygładzanie wiąże się zawsze z pewnymi formami lokalnego uśredniania danych, tak że niesystematyczne składniki poszczególnych obserwacji znoszą się nawzajem. Najbardziej powszechną techniką wygładzania jest wygładzanie za pomocą średniej ruchomej, które polega na zastąpieniu każdego elementu szeregu przez zwykłą lub ważoną średnią n sąsiadujących wartości, gdzie n jest szerokością okna wygładzania.

W literaturze spotykane są próby wygładzenia badanego sygnału, w przypadku badania przepływów ścieków z dobowym krokiem próbkowania, za pomocą prostej średniej kroczącej (ruchomej). *Średnia krocząca (ruchoma) prosta* (ang. *Simple Moving Average, SMA*),

¹ Praca wykonana w ramach realizacji projektu badawczego NCN Nr N N519 6521 40.

jest to zwykła średnia arytmetyczna wartości badanego sygnału z ostatnich n okresów:

$$SMA = \frac{p_1 + p_2 + \dots + p_n}{n} \quad (1)$$

Oprócz tej metody pozbycia się przypadkowych wahań w wygładzanym sygnale, jest znana także ważona średnia krocząca. *Ważona średnia krocząca (ruchoma)* (ang. *Weighted Moving Average*, WMA) przypisuje różne wagi danym z poszczególnych kroków próbkowania (okresów). W analizie technicznej często stosuje się średnie ważone, w których wagi maleją w postępie arytmetycznym. Na przykład, w n -okresowej WMA ostatni okres ma wagę n , przedostatni $n-1$ itd.:

$$WMA = \frac{n \cdot p_1 + (n-1) \cdot p_2 + \dots + 2 \cdot p_{n-1} + 1 \cdot p_n}{n + (n-1) + \dots + 2 + 1} \quad (2)$$

Stosuje się również przy wygładzaniu pomiarów wykładniczą średnią krocząca. *Wykładnicza średnia krocząca (ruchoma)* (ang. *Exponential Moving Average*, EMA) jest odmianą średniej ważonej, w której znaczenie wartości badanego sygnału z coraz bardziej odległych w czasie okresów maleje w sposób wykładniczy.

$$EMA = \frac{\alpha^0 \cdot p_1 + \alpha^1 \cdot p_2 + \dots + \alpha^{n-1} \cdot p_n}{\alpha^0 + \alpha^1 + \dots + \alpha^{n-1}} \quad (3)$$

gdzie $\alpha = 1 + \frac{2}{n}$.

Metoda ważonej średniej kroczącej jest jedną z metod prognostycznych, która dotyczy analizy szeregów czasowych bez tendencji. Stosowana jest przy stałym poziomie zjawiska i znacznych waniach przypadkowych. Ta metoda znacznie szybciej reaguje na zmiany, niż średnia krocząca prosta.

4. OBIEKT BADAWCZY

Lokalizacja obiektu

Badania związane z pozyskiwaniem danych pomiarowych dopływu ścieków do oczyszczalni i ich analizą prowadzono na przykładzie oczyszczalni ścieków

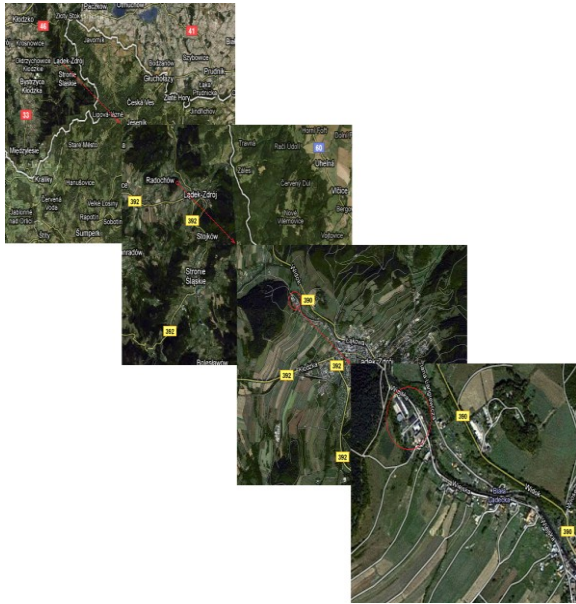
w Łądku-Zdroju. Lokalizację obiektu badawczego przedstawiono na rys. 1.

Schemat technologiczny oczyszczalni

Oczyszczalnia ścieków w Łądku-Zdroju, to oczyszczalnia mechaniczno-biologiczna o maksymalnej projektowanej przepustowości 8.000 m³/d (rys. 2). W rzeczywistości podana wartość znacznie przekracza realną maksymalną przepustowość obiektu. Część mechaniczna oczyszczalni składa się z dwóch naprzemiennie działających krat mechanicznych typu KUMP/600, dwóch zblokowanych piaskowników typu PP1-KPVS o przepustowości od 1.900 do 13.000 m³/d, oraz dwóch podłużnych płaskodennych osadników wstępnych typu OPP 6/2 ze zgarniaczem typu Miedera. Część biologiczna oczyszczalni składa się z dwóch złóż biologicznych typu ZBS o średnicy 20 m i czynnej wysokości 4 m. Ze złóż biologicznych ścieki dopływają kanałem otwartym do dwóch osadników wtórnych. Osad wtórny jest odpompowywany do dwóch komór fermentacyjnych osadów a po przefermentowaniu jest odprowadzany na trzy poletka osadowe.

Przed oczyszczalnią jest zainstalowana studnia przelewowa S-2 z zastawką mechaniczną KMP/600 i z kanałem bocznikowym ϕ 800 mm, który umożliwi zrzut ścieków, gdy ilość ścieków dopływających do oczyszczalni przekroczy dopuszczalną przepustowość hydrauliczną wynoszącą 4,5 m³/min. Do zrzutu ścieków dochodzi w czasie trwania deszczy nawalnych oraz intensywnych roztopów, gdy wody przypadkowe dopływają do oczyszczalni w dużej ilości i ze względu na grawitacyjny charakter sieci - w krótkim czasie.

Odbiornikiem ścieków oczyszczonych jest rzeka Biała Łądecka. Wylot ścieków oczyszczonych wybudowano jako boczny, żelbetowy, denny, o rzędnej dna kolektora wylotowego zrównanej z rzędną koryta rzeki równą 412,45 m.n.p.m. Rzędna terenu przy wylocie wynosi 415,30 m.n.p.m., a rzędna wody wysokiej wynosi 415,13 m.n.p.m. Ilość ścieków dopływających do oczyszczalni w Łądku-Zdroju jest mierzona za pomocą przepływomierza (rys. 3) zainstalowanego w korycie pomiarowym. Sygnał analogowy z przepływomierza jest doprowadzany do stacji przetwarzającej go na cyfrowy i następnie sygnał cyfrowy jest doprowadzany do komputera, na którym archiwizuje się i wizualizuje zmierzony przepływ ścieków.



Rysunek 1 Lokalizacja obiektu badawczego [www.maps.google.pl].

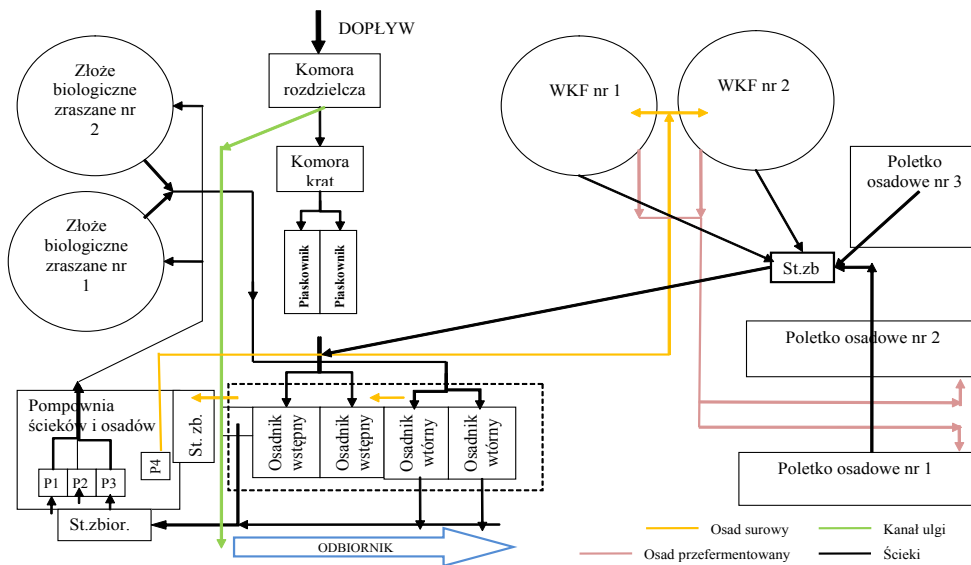
ścieków w Łądku-Zdroju badania przepływu ścieków zostały rozpoczęte 26 kwietnia 2007 roku i są nadal kontynuowane.



Rysunek 2 Urządzenie pomiarowe zainstalowane na oczyszczalni w Łądku-Zdroju.

Urządzenie pomiarowe

Wyniki pomiarów chwilowych ilości ścieków dopływających do oczyszczalni archiwizuje się za pomocą programu PRO-2000 firmy Mikrob. Na oczyszczalni



Rysunek 3 Schemat techniczny oczyszczalni ścieków w Łądku-Zdroju.

5. OCZYSZCZANIE DANYCH POMIAROWYCH

Udostępniane przez program pomiarowy PRO-2000 firmy Mikrob dane są rejestrowane w cyklu 10-sekundowym i po zagregowaniu archiwizowane w cyklu 1-minutowym. Są one obciążone pięcioma rodzajami błędów:

1. Brak danych (w ciągu danych gwiazdka zamiast wartości przepływu).
2. Dane zerowe (wartość zerowa przepływu).
3. Spadający przepływ (wartość przepływu spada do zera).
4. Powtarzające się dane (wartość przepływu nie zmienia się).
5. Dane chwilowe (dane minutowe nie są uśrednione z danych 10 sekundowych).

Po wykonaniu analizy dostarczanych pomiarów stwierdzono, że dwa pierwsze rodzaje błędów wynikają z zawodności zainstalowanego w oczyszczalni systemu pomiarowego i nie można ich wyeliminować bez zamiany tego systemu na bardziej niezawodny. Natomiast kolejne trzy rodzaje błędów są błędami powodowanymi przez wady oprogramowania zastosowanego w systemie monitoringu i również w sposób programowy zostały one wyeliminowane. W rezultacie powyższy wykaz błędów w pomiarach został zmodyfikowany następująco:

1. Brak danych (brak pomiaru, błąd urządzenia pomiarowego).
2. Dane zerowe (pomiar zerowy, błąd urządzenia pomiarowego).
3. Spadający przepływ (nie występuje, błąd wyeliminowany).
4. Powtarzające się dane (nie występują, błąd wyeliminowany).

Dane chwilowe (nie występują, błąd wyeliminowany po uśrednieniu pomiarów).

6. OPIS PROGRAMU PROGNOZOWANIA DOPŁYWU

Opracowany program analizy i prognozowania dopływu ścieków surowych do oczyszczalni (SCADA Analizer) jest całkowicie zautomatyzowany i aby go uruchomić, należy

ustalić jedynie kilka parametrów sterujących. Nazwy i funkcje tych parametrów są następujące:

- Dane – miejsce i nazwa pliku, skąd pozyskuje się dane pomiarowe do analizy.
- Wolna – długość okna średniej ruchomej długookresowej, w minutach.
- Szybka – długość okna średniej ruchomej krótkookresowej, w minutach.
- Różnica – długość okna uśredniania różnicy między średnią ruchomą długookresową i średnią ruchomą krótkookresową, w minutach.

Program ma dwa tryby:

- Tryb czasu rzeczywistego – przeznaczony do pracy w trybie on-line w warunkach eksploatacyjnych oczyszczalni ścieków; w przypadku, gdy wyznaczany na bieżąco trend badanego sygnału osiąga zadaną wartość krytyczną, rozlega się sygnalizacja dźwiękowa będąca ostrzeżeniem dla operatora oczyszczalni ścieków.
- Tryb przeglądania danych historycznych – przeznaczony do analizy zarchiwizowanych pomiarów w trybie off-line.

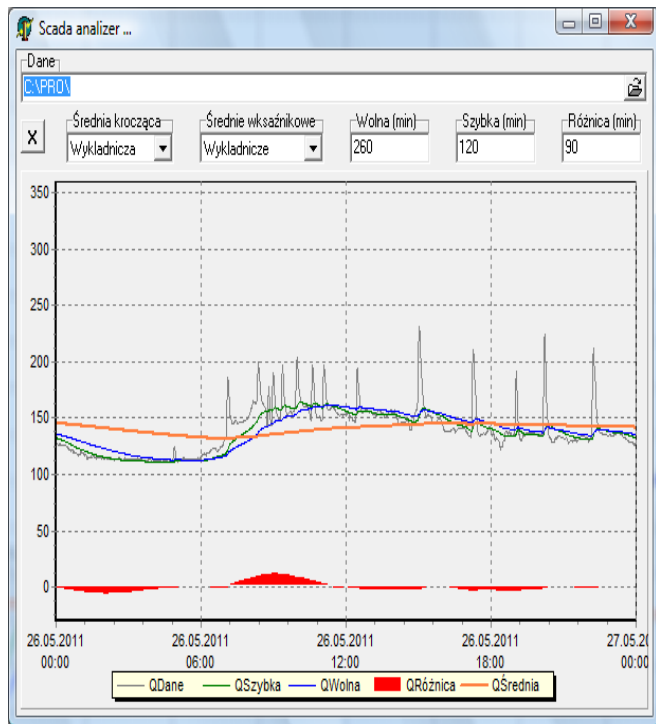
Jest możliwość wyboru rodzaju średnich używanych w algorytmie MACD, spośród następujących rodzajów:

- średnia wykładnicza
- średnia ważona
- średnia prosta.

Ekran główny opracowanego programu jest pokazany na rys. 4.

5. WYNIKI OBLICZEŃ

Badania testujące opracowanego programu wykonano na podstawie danych pomiarowych z oczyszczalni ścieków w Łądku Zdroju. Na rys. 5 widać, na ile szybciej reagują na zmianę przepływu wyznaczone średnie, agregujące rejestrowane pomiary chwilowe, to znaczy średnia krocząca wolna (oznaczona kolorem niebieskim) i średnia krocząca szybka (oznaczona kolorem zielonym), w porównaniu do średniej kroczącej dobowej (oznaczonej kolorem pomarańczowym). Przy tym średnie kroczące zostały w obliczeniach ustawione odpowiednio na 260 min dla średniej wolnej i 120 min dla średniej szybkiej.

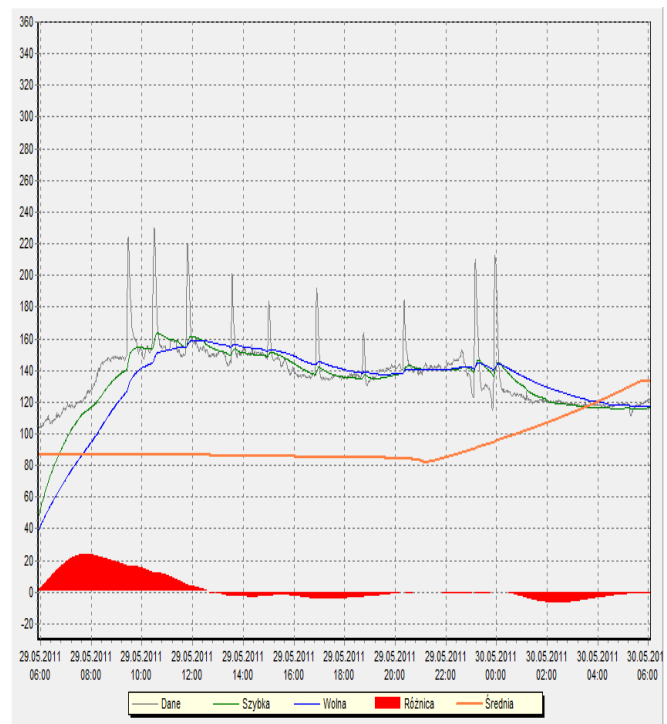


Rysunek 4 Ekran główny programu SCADA Analizer.

Wizualizujący wyniki w sposób bardziej pogłówny histogram (różnica średnich kroczących oznaczona kolorem czerwonym, z zadaną długością okna równą 90 min) pokazuje trend wzrostu względnie spadku przepływu. Można zauważyć, że z histogramu widać nadchodzącą zmianę trendu przepływu znacznie szybciej, niż wynika to z krzywych kroczących, co ma istotne znaczenie dla sterowania dopływem ścieków do oczyszczalni.

Ogólnie, interpretacja zmian trendu sygnału na podstawie wykreślonego histogramu jest następująca:

- jeżeli pole histogramu jest znacznie powyżej poziomu zerowego, oznacza to szybki trend wzrostu
- jeżeli pole histogramu jest nieznacznie powyżej poziomu zerowego, oznacza to mały trend wzrostu
- jeżeli pole histogramu jest nieznacznie poniżej poziomu zerowego, oznacza to mały trend spadku
- jeżeli pole histogramu jest znacznie poniżej poziomu zerowego, oznacza to szybki trend spadku.



Rysunek 5 Wykresy trendu wygenerowane za pomocą programu SCADA Analizer.

8. PODSUMOWANIE I WNIOSKI

Na podstawie przeprowadzonych obliczeń i ich analizy można sformułować następujące wnioski:

- Powstały w trakcie pracy program komputerowy SCADA Analizer pozwala w szybki i prosty sposób na wcześniejsze ostrzeżenie operatora oczyszczalni przed gwałtownymi i nieoczekiwanymi zmianami dopływu ścieków surowych do obiektu. Odpowiednio wczesne wygenerowanie sygnału alarmowego pozwala na uruchomienie odpowiednich działań zapobiegających awarii oczyszczalni.
- Program może być stosowany w trybie rzeczywistym on-line w warunkach eksploatacyjnych oczyszczalni oraz w trybie off-line przeznaczonym do prowadzenia badań.
- Do zbierania danych pomiarowych należy stosować niezawodne systemy monitoringu. Z wykonanych obliczeń testowych wynika, że jeżeli dane z czterech kolejnych lat nie zawierają

ani jednego miesiąca bez błędów w pomiarach, to takie dane nie gwarantują wiarygodnych wyników analizy wykonanej za pomocą opracowanego programu. Tego rodzaju dane bardzo utrudniają wykonanie okresowych analiz statystycznych, które miałyby wartość aplikacyjną.

- Dla wszelkiego rodzaju analiz przepływu poleca się używanie wykładniczej średniej kroczącej ważonej z okresem kilkugodzinnym, zamiast powszechnie stosowanej średniej kroczącej prostej dobowej. Pozwala to na lepsze dopasowanie wynikowej średniej do trendu zmiany przepływu oraz na powiększenie ilości dostępnych danych, ponieważ średnia dobowa ma sensowne wyniki dopiero po dobie bezbłędnych danych wejściowych, natomiast średnia kilkugodzinna pozwala na uzyskanie sensownych wyników już po kilku godzinach.

Należy przy tym zauważyć, że w przypadku prognozowania dopływu ścieków surowych do oczyszczalni, co jest ważnym zagadnieniem z punktu widzenia efektywnego sterowania oczyszczalnią, były już prowadzone prace dotyczące prognozowania wartości chwilowych dopływu, na przykładów za pomocą modeli szeregów czasowych, sieci neuronowych czy zbiorów rozmytych. Natomiast bardzo nieliczne są prace zajmujące się badaniem trendu dopływu ścieków, co ma dla operatora oczyszczalni znacznie większe znaczenie, niż przewidywanie jego wartości chwilowej. Jednocześnie wydaje się, że przedstawione w tym artykule wyniki są pierwszymi, gdzie do prognozowania trendu dopływu ścieków stosuje się z powodzeniem algorytm MACD.

Literatura

1. *Elektronic Statistics Textbook*, StatSoft, 1984-2011.
2. Box i Jenkins, „*Time Series Analysis: Forecasting and Control*”, 1976.
3. Velleman i Hoaglin, „*The ABC's of EDA: Applications, Basics, and Computing of Exploratory Data Analysis*”, 1981.
4. Appel G., „*The moving average convergence-divergence trading method*”, Traders Press, Toronto, Kanada 1985.
5. Pring M., „*Introduction to Technical Analysis*”, McGraw-Hill, New York 1997.
6. Reuters Ltd., „*An Introduction to Technical Analysis*”, Wiley, Singapore, 1999.
7. Klinker F., „*Exponential moving average versus moving exponential average*”, *Mathematik in*

Forschung und Anwendung, Math Semesterber. 58, 2011, 97-107.

8. Chen Y.-M., Chung T.-Y., Lai M.-Y., Hsu C.-H., „MACD-based motion detection approach in heterogeneous networks”, *EURASIP Journal on Wireless Communication and Networks*, 2008.
9. Rowlands S.F., Livingstone J., Lund C.P., „Optical modelling of thin film solar cells with textured interfaces using the effective medium approximation”, *Solar Energy*, Vol: 76, Issue 1-3, January - March, 2004, 301-307.
10. Zhang, J.; Goode, K.; Cuddihy, P.; Cleland, J.G., „Predicting hospitalisation due to worsening heart failure using daily weight measurement: An analysis of the Trans-European Network-Home-Care Management System (TEM-HMS)”, *European Journal of Heart Failure*, Supplement Vol. 7, Issue 1, June 2008.