

Zbigniew Siwoń, Wojciech Cieżak, Jan Cieżak

Stochastyczne modele godzinowego poboru wody w wybranym systemie wodociągowym

Modele prognostyczne poboru wody w obszarze prognoz bieżących i krótkoterminowych są wykorzystywane między innymi w układach sterowania szeroko pojętymi procesami zaopatrzenia w wodę, w tym procesami uzdatniania wody i jej dystrybucji. W systemach sterowania układ regulacji odbiera zmiany poboru wody lub zapotrzebowania na nią jako zakłócenia, które powinny być kompensowane zmianami wielkości sterującej. Z tego względu metody modelowania i prognozowania szeregów czasowych poboru wody określane są niekiedy jako metody prognozowania zakłóceń w systemach wodociągowych [1].

Zjawisko poboru (zużycia, rozbioru) wody i zapotrzebowania na nią przez jej konsumentów, a także zjawisko odpływu ścieków z miast i aglomeracji miejsko-przemysłowych, mają charakter złożony: deterministyczno-losowy. Podlegają one zatem pewnym prawidłowościom, na które nakładają się czynniki losowe. Ponieważ modele prognostyczne (w zakresie prognoz bieżących i krótkoterminowych dokonywanych z wyprzedzeniem od 1 do kilkudziesięciu godzin) są często wykorzystywane prawie natychmiast w algorytmach symulacji działania systemu dystrybucji wody i wyznaczania optymalnej strategii sterowania, muszą umożliwiać szybkie uzyskiwanie prognoz na podstawie możliwie małej liczby łatwo dostępnych predyktorów, powinny być odporne na nieujmowane przez nie zakłócenia danych oraz skonstruowane jako rekursywne, tj. w każdym kroku po napłynięciu nowych obserwacji powinny je uwzględniać i generować nową prognozę [2,3]. Z tego względu do prognozowania krótkoterminowego są wykorzystywane obecnie najczęściej modele stochastyczne. Za model stochastyczny uważa się opis rzeczywistego procesu na podstawie pomiarów wejścia i wyjścia z obiektu obciążonych białym szumem, z wykorzystaniem metod interpolacyjnych, aproksymacyjnych, analizy szeregów czasowych itp. Tak skonstruowane modele w postaci zalgorytmizowanej można wprowadzić do struktury zarządzania procesem sterowania. Optymalizacja klasy i struktury modelu wymaga odpowiednio dużej liczby danych o poborze wody, uzyskiwanych z ciągłych pomiarów.

Alternatywą dla wyżej wymienionych metod statystycznych są sztuczne sieci neuronowe, które wg doniesień literaturowych [4–6] mogą być ewentualnie wykorzystywane do prognozowania dobowego poboru wody. Analiza możliwości ich zastosowania do prognozowania szeregów czasowych godzinowego poboru wody będzie przedmiotem oddzielnych badań.

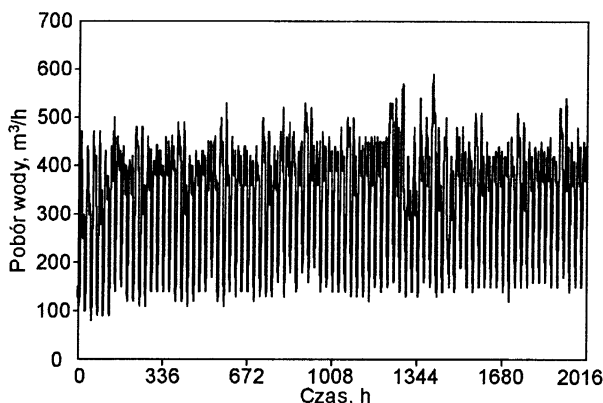
Dane przyjęte do analiz

Do identyfikacji i weryfikacji modeli prognostycznych wykorzystano wyniki pomiarów globalnego poboru wody w systemie wodociągowym Brzegu. System ten jest zasilany tylko z jednego źródła (tj. pompowni drugiego stopnia) i nie występują w nim zbiorniki sieciowe. Woda jest dostarczana do miejskiej sieci wodociągowej dwoma rurociągami tłocznymi, na których są zainstalowane wodomierze produkcji firmy PoWoGaz z ciągłym przekazem wskazań do dyspozytorni.

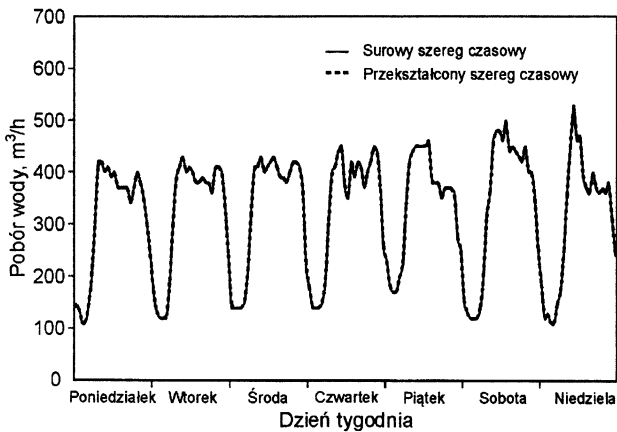
Chronologicznie uszeregowane wyniki pomiarów godzinowego poboru wody przez wszystkich jej konsumentów w okresie od 1 stycznia 2001 r. do 31 października 2002 r. stanowiły wyjściowy szereg czasowy poddawany analizie. Pierwszym krokiem w tej analizie był podział całego zbioru obserwacji na cztery szeregi czasowe grupujące wyniki pomiarów w czterech sezonach (tj. dwóch jesienno-zimowych i dwóch wiosenno-letnich), a także wyodrębnienie dni roboczych, sobót oraz niedziel wraz z dniami świątecznymi. Każdy analizowany szereg czasowy poddano badaniom statystycznym mającym na celu wyeliminowanie tzw. błędów grubych, spowodowanych awariami w systemie dystrybucji wody oraz błędami odczytów wodomierzy.

Zgodnie z procedurą zaproponowaną w pracach [7,8], kolejnym etapem analizy, umożliwiającym minimalizację błędów modeli, było wygładzenie empirycznego szeregu czasowego $\{q_t, t=1,2,\dots,N\}$ za pomocą funkcji drgań harmonicznym (szeregu Fouriera) w celu dekompozycji empirycznego szeregu czasowego zawierającego składniki cykliczne na r funkcji sinusoidalnych (sinus i cosinus) o określonych długościach fal. Do wyrównania wartości periodogramów zastosowano transformację ważonej średniej ruchomej Hamminga.

Przykłady surowego i przekształconego szeregu czasowego pokazano na rysunkach 1 i 2.



Rys. 1. Surowy szereg czasowy godzinowego poboru wody w czasie od 1 listopada 2001 r. do 31 stycznia 2002 r.



Rys. 2. Przetworzony szereg czasowy godzinowego poboru wody w czasie od 19 do 25 listopada 2001 r.

Modele klasy ARIMA

Scałkowane modele autoregresji i średniej ruchomej (ARIMA) odwzorowują właściwości statyczne i dynamiczne szeregów stacjonarnych i pewnych klas szeregów niestacjonarnych, interpretowanych jako wynik przejścia białego szumu przez dyskretny filtr liniowy, skończenie wymiarowy. Charakteryzują się one różnymi właściwościami przy jednolitym zapisie formalnym oraz identycznych metodach estymacji parametrów dla różnych typów i podklas modeli. Z wielu przyczyn są wygodne w praktycznym stosowaniu. Nadają się szczególnie do bieżącego i krótkoterminowego prognozowania szeregów czasowych poboru wody. Sezonowe modele ARIMA mają ogólną postać [9]:

$$\Phi(B)\Phi(B^s)(w_t - \mu) = b(B)\Theta(B^s)\varepsilon_t \quad (1)$$

$$w_t = \nabla^d \nabla^{sD} q_t, \quad t=1,2,3,\dots,N \quad (2)$$

w których:

$\Phi(B)=1-\varphi_1(B)-\varphi_2(B^2)-\dots-\varphi_p(B^p)$ – niesezonowy operator autoregresji rzędu p ,

$\Phi(B^s)=1-\Phi_1(B^s)-\Phi_2(B^{s2})-\dots-\Phi_P(B^{sP})$ – sezonowy operator autoregresji rzędu P ,

$b(B)=1-b_1(B)-b_2(B^2)-\dots-b_g(B^g)$ – niesezonowy operator średniej ruchomej rzędu g ,

$\theta(B^s)=1-\theta_1(B^s)-\theta_2(B^{s2})-\dots-\theta_G(B^{sG})$ – sezonowy operator średniej ruchomej rzędu G ,

w_t – stacjonarny szereg czasowy poboru wody utworzony przez różnicowanie szeregu surowego,

N – liczba elementów szeregu czasowego,

$\nabla^d=(1-B)^d$ – niesezonowy operator różnicowania rzędu d ,

$\nabla^{sD}=(1-B^s)^D$ – sezonowy operator różnicowania rzędu D ,

ε_t – szereg czasowy reszt modelu (tzw. biały szum), gdzie $t=1,2,3,\dots,N$,

μ – wartość średnia szeregu czasowego,

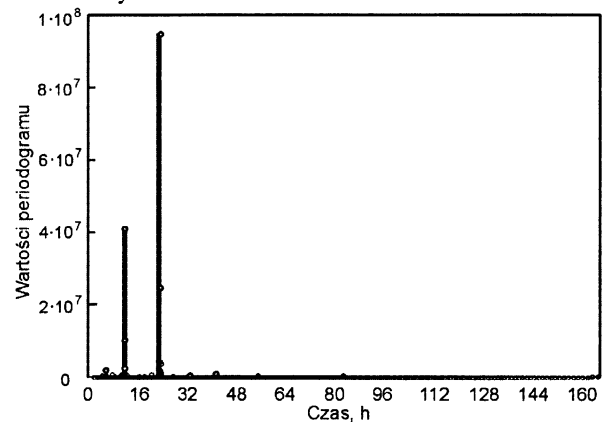
t – zdyskretyzowany czas, $t=1,2,3,\dots,N$,

B – operator przesunięcia wstecz,

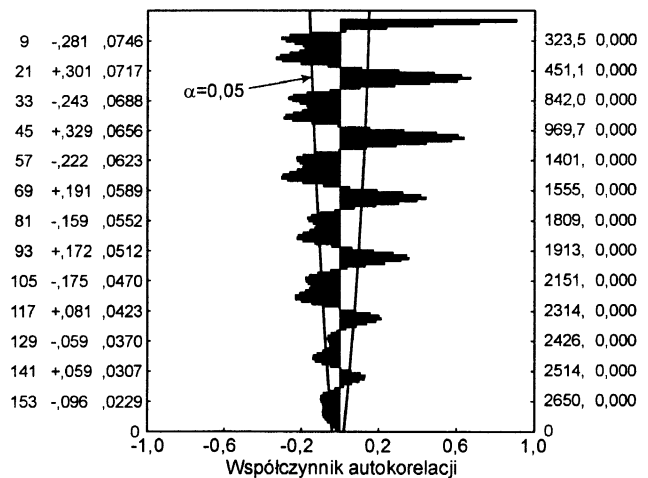
s – okres składowej okresowej modelu szeregu czasowego (tzn. obserwacje odległe od siebie o s jednostek czasu, np. o 24 godz., są statystycznie podobne),

Operatory różnicowania (∇^d, ∇^{sD}) mają zapewnić odpowiednio niesezonową i sezonową stacjonarność d -tych i D -tych różnic wyrazów szeregu czasowego $\{q_t\}$. Do formalnego zapisu scałkowanego sezonowego multiplikatywnego modelu autoregresji i średniej ruchomej jest używana notacja ARIMA (p,d,g)(P,D,G) $_s$. W pierwszym nawiasie zostały zapisane rzędy operatorów

niesezonowych, a w drugim – rzędy operatorów sezonowych. Przykładowe wykresy periodogramu i funkcji autokorelacji surowego szeregu czasowego godzinowego poboru wody pokazano na rysunkach 3 i 4.



Rys. 3. Analiza widmowa szeregu czasowego obserwacji godzinowego poboru wody w czasie od 1 stycznia 2001 r. do 31 października 2002 r.



Rys. 4. Wykres funkcji autokorelacji szeregu czasowego obserwacji godzinowego poboru wody w czasie od 1 do 7 kwietnia 2001 r.

Szczegółowej analizie poddano 12 struktur modelu klasy ARIMA, spośród których wybrano 4 struktury charakteryzujące się najmniejszymi błędami, tj. ARIMA (1,0,0)(1,1,0) $_{24}$, ARIMA (1,0,1)(1,1,1) $_{24}$, ARIMA (1,0,2)(1,1,1) $_{24}$ oraz ARIMA (2,0,0)(1,1,0) $_{24}$. Analizy przeprowadzono dla surowych i przekształconych szeregów czasowych poboru wody w czterech okresach sezonowych, oddzielnie dla dni roboczych, sobót wolnych od pracy oraz niedziel wraz z dniami świątecznymi. Wybrane wyniki tych analiz zestawiono w tabeli 1, a przykładowy rozkład częstości reszt modelu na tle rozkładu normalnego pokazano na rysunku 5.

Przeanalizowano również liczebności zbiorów przeszłych obserwacji poboru wody w dniach roboczych, niezbędne do identyfikacji parametrów modelu prognostycznego, z uwzględnieniem okresów 2-, 3- i 4-tygodniowych. Ponieważ błędy prognoz okazały się porównywalne, jako okres minimalny przyjęto 2 tygodnie. Wyniki analizy jakości predykcji według czterech wybranych modeli zestawiono w tabeli 2. Dane w niej zamieszczone potwierdzają tezę podaną w pracach [2,7,8], że model o strukturze ARIMA (1,0,0)(1,1,0) $_{24}$ umożliwia uzyskiwanie prognoz godzinowego poboru wody, charakteryzujących się najmniejszymi błędami, w systemach wodociągowych miast o zasadniczo odmiennej liczbie mieszkańców i strukturze potrzeb wodnych (Łódź [7,8], Brzeg).

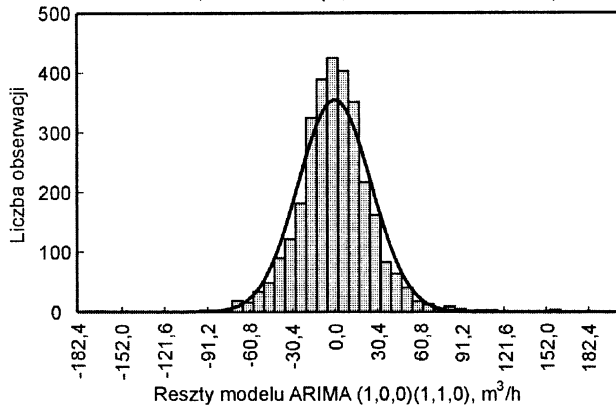
Tabela 1. Wyniki analizy efektywności różnych struktur sezonowych multiplikatywnych modeli ARIMA (od 1 listopada 2001 r. do 30 kwietnia 2002 r., dni robocze)

Struktura modelu	Bezwzględny średniokwadratowy błąd modelu m ³ /h		Średni pobór wody m ³ /h	Względny średniokwadratowy błąd modelu %	
	szereg surowy	szereg przekształcony		szereg surowy	szereg przekształcony
(0,1,1)(0,1,1) ₂₄	21,12	20,21	332,68	6,35	6,07
(0,1,2)(0,1,1) ₂₄	20,72	19,79		6,23	5,95
(1,0,0)(1,1,0) ₂₄	23,19	21,34		6,97	6,41
(2,0,0)(1,1,0) ₂₄	22,93	21,31		6,89	6,40
(1,0,1)(1,1,1) ₂₄	20,11	19,24		6,05	5,78
(2,0,1)(1,1,1) ₂₄	19,99	19,11		6,01	5,74
(1,0,2)(1,1,1) ₂₄	20,08	19,19		6,04	5,77
(1,1,1)(0,1,2) ₂₄	20,09	19,18		6,04	5,77
(1,1,1)(2,1,0) ₂₄	21,98	20,55		6,61	6,18
(1,1,1)(0,1,1) ₂₄	20,24	19,30		6,08	5,80
(1,1,1)(1,1,0) ₂₄	23,04	21,34		6,92	6,41
(1,1,1)(1,1,1) ₂₄	20,08	19,18		6,04	5,76

Tabela 2. Wyniki analizy jakości predykcji dla dwutygodniowego czasu identyfikacji i wybranych modeli ARIMA oraz szeregu przekształconego

Dzień tygodnia	Data	Struktura modelu	Bezwzględny średniokwadratowy błąd prognozy m ³ /h	Średni pobór wody m ³ /h	Względny średniokwadratowy błąd prognozy %	Maksymalny błąd prognozy	
						(+) m ³ /h	(-) m ³ /h
Poniedziałek	2 IV 2001	(1,0,0)(1,1,0) ₂₄	39,68	397,92	9,97	84,05	-73,15
Wtorek	3 IV 2001		33,13	377,08	8,78	62,10	-91,12
Środa	4 IV 2001		20,85	398,75	5,23	70,37	-31,31
Czwartek	5 IV 2001		48,99	421,25	11,63	78,54	-38,31
Piątek	6 IV 2001		31,36	415,42	7,55	55,41	-55,29
Sobota	28 IV 2001		25,73	364,58	7,06	61,89	-43,89
Niedziela	29 IV 2001		23,02	302,50	7,61	59,29	-44,84
Poniedziałek	2 IV 2001	(1,0,1)(1,1,1) ₂₄	39,13	397,92	9,83	87,53	-60,79
Wtorek	3 IV 2001		29,87	377,08	7,92	47,15	-98,88
Środa	4 IV 2001		28,08	398,75	7,04	57,79	-38,82
Czwartek	5 IV 2001		56,68	421,25	13,46	90,84	-23,31
Piątek	6 IV 2001		22,22	415,42	5,35	46,40	-29,47
Sobota	28 IV 2001		30,86	364,58	8,47	34,90	-63,11
Niedziela	29 IV 2001		23,01	302,50	7,61	18,65	-51,44
Poniedziałek	2 IV 2001	(1,0,2)(1,1,1) ₂₄	39,13	397,92	9,83	87,53	-60,80
Wtorek	3 IV 2001		29,88	377,08	7,92	47,12	-98,93
Środa	4 IV 2001		27,40	398,75	6,87	57,41	-38,88
Czwartek	5 IV 2001		57,37	421,25	13,62	91,65	-23,25
Piątek	6 IV 2001		23,26	415,42	5,60	32,80	-41,73
Sobota	28 IV 2001		29,94	364,58	8,21	34,94	-61,61
Niedziela	29 IV 2001		23,84	302,50	7,88	17,22	-51,92
Poniedziałek	2 IV 2001	(2,0,0)(1,1,0) ₂₄	39,89	397,92	10,03	84,40	-73,03
Wtorek	3 IV 2001		33,01	377,08	8,75	62,39	-90,23
Środa	4 IV 2001		20,87	398,75	5,23	70,22	-31,84
Czwartek	5 IV 2001		49,01	421,25	11,63	78,48	-38,38
Piątek	6 IV 2001		32,07	415,42	7,72	51,21	-59,10
Sobota	28 IV 2001		29,30	364,58	8,04	67,14	-42,37
Niedziela	29 IV 2001		23,16	302,50	7,66	60,41	-44,73

d Kolmogorowa-Smirnowa 0,04788, $p < 0,01$, p Lillieforsa $< 0,01$
 Test chi-kwadrat=175,00364, $df=18$ (dopasow.), $p=0,00000$
 $Y=23164,8000 \cdot \text{Normal}(x; 693266E-1, 26,0774517)$



Rys. 5. Rozkład częstości reszt modelu ARIMA (1,0,0)(1,1,0)₂₄ godzinowego poboru wody w dniach roboczych w czasie od 1 maja do 31 października 2002 r. wraz z wynikami testów zgodności z rozkładem normalnym

Metoda wykładniczego wygładzania szeregów czasowych godzinowego poboru wody

Metody wygładzania wykładniczego szeregu czasowego umożliwiają wyznaczenie jego przyszłych wartości (prognoz) przez obliczenie pewnego rodzaju średniej ruchomej, gdzie bieżącym i bezpośrednio poprzedzającym obserwacjom przypisuje się większą wagę niż obserwacjom odpowiednio starszym. Do potrzeb bieżącego (z wyprzedzeniem od 1 do 24 godz.) prognozowania szeregów czasowych godzinowego lub o krótszym czasie trwania poboru wody, charakteryzujących się wahaniami w cyklach dobowych, przydatne są modele sezonowe addytywne i multiplikatywne, a w szczególności modele Wintersa. Równania addytywnej wersji modelu mają postać:

$$\hat{q}_t = F_{t-1} + S_{t-1} + C_{t-s} \quad (3)$$

$$F_{t-1} = \alpha(q_{t-1} - C_{t-1-s}) + (1 - \alpha)(F_{t-2} + S_{t-2}) \quad (4)$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1 - \beta)S_{t-2} \quad (5)$$

$$C_{t-1} = \gamma(q_{t-1} - F_{t-1}) + (1 - \gamma)C_{t-1-s} \quad (6)$$

a równanie na m kroków w przód:

$$\hat{q}_{t+m} = F_t + mS_t + C_{t+m-s} \quad (3a)$$

W wersji multiplikatywnej równania modelu mają następującą postać:

$$\hat{q}_t = (F_{t-1} + S_{t-1})C_{t-s} \quad (7)$$

$$F_{t-1} = \alpha \frac{q_{t-1}}{C_{t-1-s}} + (1 - \alpha)(F_{t-2} + S_{t-2}) \quad (8)$$

$$S_{t-1} = \beta(F_{t-1} - F_{t-2}) + (1 - \beta)S_{t-2} \quad (9)$$

$$C_{t-1} = \gamma \frac{q_{t-1}}{F_{t-1}} + (1 - \gamma)C_{t-1-s} \quad (10)$$

$$\hat{q}_{t+m} = (F_t + mS_t)C_{t+m-s} \quad (7a)$$

W celu zainicjowania procesu prognozowania należy przypisać następujące wartości początkowe (zarówno dla wersji addytywnej jak i multiplikatywnej):

$$S_0 = \frac{M_k - M_1}{(k - 1)s} \quad (11)$$

$$F_0 = M_1 - \frac{S_0 s}{2} \quad (12)$$

w których:

F_{t-1} – ocena wartości średniej w okresie $t-1$,

S_{t-1} – ocena przyrostu trendu w okresie $t-1$,

C_{t-1} – ocena wskaźnika sezonowości w okresie $t-1$,

α, β, γ – parametry gładzenia przyjmowane w przedziału $(0,1)$,

M_k – średnia z ostatniego cyklu sezonowego,

M_1 – średnia z pierwszego cyklu sezonowego,

k – liczba pełnych cykli sezonowych,

s – okres składowej okresowej modelu szeregu czasowego,

Wartości początkowe, odmienne dla wersji addytywnej oraz multiplikatywnej, dla składników sezonowych (C_{0j} dla $j=1, \dots, s$) oszacowano na podstawie danych przy pomocy klasycznej dekompozycji sezonowej szeregów czasowych (metoda Census I).

Minimalizacja błędów prognoz uzyskiwanych za pomocą metody wygładzania wykładniczego sprowadza się do optymalizacji parametrów gładzenia. W tym kontekście dla addytywnej i multiplikatywnej wersji modelu Wintersa przeanalizowano następujące warianty:

- parametry gładzenia równe 0,1; przeanalizowano szereg surowy,

- parametry gładzenia równe 0,1; przeanalizowano szereg przekształcony wg funkcji Fouriera drgań harmonicznnych,

- parametry gładzenia wyznaczone poprzez minimalizację błędu średniokwadratowego z wykorzystaniem procedury minimalizacji funkcji; przeanalizowano szereg surowy,

- parametry gładzenia wyznaczone poprzez minimalizację błędu średniokwadratowego z wykorzystaniem procedury minimalizacji funkcji; przeanalizowano szereg przekształcony.

Przeanalizowano również liczebności zbiorów przeszłych obserwacji poboru wody w dniach roboczych, niezbędne do identyfikacji parametrów modelu prognostycznego, z uwzględnieniem okresów 2-, 3- i 4-tygodniowych. Ponieważ błędy prognoz okazały się porównywalne, jako okres minimalny przyjęto 2 tygodnie. Analiza jakości predykcji dla czterech wyżej wyszczególnionych wariantów oraz dwóch wersji (tj. addytywnej i multiplikatywnej) modelu Wintersa wykazała, że w praktyce najlepszą efektywność opisu empirycznego szeregu czasowego uzyskuje się dla szeregu przekształconego za pomocą funkcji drgań harmonicznnych oraz optymalizowanych parametrach gładzenia. Wybrane wyniki tych analiz dla addytywnej wersji modelu Wintersa zestawiono w tabeli 3.

We wszystkich przeanalizowanych wypadkach wartości parametrów gładzenia wynosiły $\alpha \in (0,1)$, $\beta=0$, $\gamma=0$, natomiast wartości funkcji (5) i (9) oraz (6) i (10) wyniosły odpowiednio:

$$S_{t-1} = S_{t-2} = \dots = \text{const.} \quad (13)$$

$$\begin{aligned} C_{t-1} &= C_{t-1-s} \\ &\vdots \\ C_{t-s} &= C_{t-2s} \\ &\vdots \\ C_{t-(k-1)s} &= C_{t-ks} \end{aligned} \quad (14)$$

Oznacza to, że model wygładzania wykładniczego dla generowania prognoz godzinowego poboru wody charakteryzuje się stałym przyrostem trendu (S_0) oraz stałym wskaźnikiem sezonowym (C_{0j} dla $j=1, \dots, s$) dla każdego okresu sezonowego.

Tabela 3. Wyniki optymalizacji parametrów gładzenia addytywnych modeli wykładniczego wygładzania szeregów czasowych

Dzień tygodnia	Data	Parametry modelu	Bezwzględny średniokwadratowy błąd prognozy m^3/h	Średni pobór wody m^3/h	Względny średniokwadratowy błąd prognozy %	Maksymalny błąd prognozy	
						(+) m^3/h	(-) m^3/h
Poniedziałek	4 II 2002	Szereg surowy (parametry równe 0,1)	21,88	344,17	6,36	43,13	-45,67
Wtorek	5 II 2002		25,46	354,17	7,19	64,80	-18,67
Środa	6 II 2002		21,99	351,67	6,25	9,14	-57,88
Czwartek	7 II 2002		20,85	336,25	6,20	17,93	-45,37
Piątek	8 II 2002		34,48	343,33	10,04	77,42	-21,67
Sobota	27 IV 2002		22,50	330,00	6,82	42,73	-13,84
Niedziela	28 IV 2002		21,30	277,08	7,69	39,29	-27,59
Poniedziałek	4 II 2002	Szereg przekształcony (parametry równe 0,1)	22,58	344,17	6,56	40,44	-46,10
Wtorek	5 II 2002		24,29	354,17	6,86	65,53	-22,16
Środa	6 II 2002		24,01	351,67	6,83	6,01	-58,71
Czwartek	7 II 2002		21,43	336,25	6,37	13,74	-43,85
Piątek	8 II 2002		32,52	343,33	9,47	72,39	-19,21
Sobota	27 IV 2002		20,33	330,00	6,16	37,33	-17,05
Niedziela	28 IV 2002		22,09	277,08	7,97	40,36	-31,00
Poniedziałek	4 II 2002	Szereg surowy (parametry zoptymalizowane)	21,01	344,17	6,10	50,21	-39,71
Wtorek	5 II 2002		28,98	354,17	8,18	74,68	-6,80
Środa	6 II 2002		16,47	351,67	4,68	20,33	-49,02
Czwartek	7 II 2002		24,28	336,25	7,22	14,92	-50,72
Piątek	8 II 2002		24,10	343,33	7,02	58,34	-32,78
Sobota	27 IV 2002		21,34	330,00	6,47	44,37	-24,53
Niedziela	28 IV 2002		23,83	277,08	8,60	51,76	-29,17
Poniedziałek	4 II 2002	Szereg przekształcony (parametry zoptymalizowane)	22,32	344,17	6,49	41,51	-44,63
Wtorek	5 II 2002		29,71	354,17	8,39	77,84	-8,65
Środa	6 II 2002		17,71	351,67	5,04	17,24	-50,55
Czwartek	7 II 2002		25,88	336,25	7,70	9,27	-50,83
Piątek	8 II 2002		22,29	343,33	6,49	53,72	-31,24
Sobota	27 IV 2002		18,03	330,00	5,46	35,12	-31,36
Niedziela	28 IV 2002		24,54	277,08	8,86	45,49	-34,31

Tabela 4. Analiza jakości predykcji dla dwutygodniowego czasu identyfikacji i optymalnego modelu ARIMA (1,0,0)(1,1,0)₂₄ oraz szeregu przekształconego

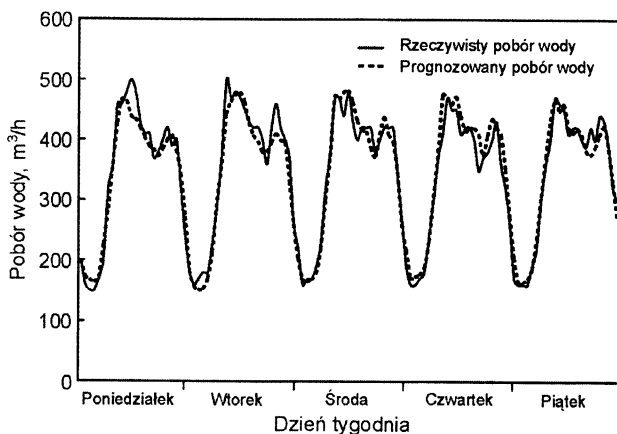
Dzień tygodnia	Data	Średni pobór wody m^3/h	Bezwzględny średniokwadratowy błąd prognozy m^3/h	Względny średniokwadratowy błąd prognozy %	Maksymalny błąd prognozy	
					(+) m^3/h	(-) m^3/h
Poniedziałek	4 II 2002	344,17	23,49	6,83	61,10	-31,58
Wtorek	5 II 2002	354,17	22,34	6,31	57,16	-20,97
Środa	6 II 2002	351,67	16,97	4,83	32,76	-39,29
Czwartek	7 II 2002	336,25	26,31	7,82	9,39	-65,44
Piątek	8 II 2002	343,33	18,10	5,27	44,15	-35,04
Sobota	27 IV 2002	330,00	21,32	6,46	41,77	-14,56
Niedziela	28 IV 2002	277,08	17,96	6,48	38,93	-25,46

Tabela 5. Analiza jakości predykcji dla dwutygodniowego czasu identyfikacji i optymalnego modelu wykładniczego (parametry zoptymalizowane) oraz szeregu przekształconego

Dzień tygodnia	Data	Średni pobór wody m ³ /h	Model addytywny				Model multiplikatywny			
			bezwzględny średnio-kwadratowy błąd prognozy m ³ /h	względny średnio-kwadratowy błąd prognozy %	maksymalny błąd prognozy		bezwzględny średnio-kwadratowy błąd prognozy m ³ /h	względny średnio-kwadratowy błąd prognozy %	maksymalny błąd prognozy	
					(+) m ³ /h	(-) m ³ /h			(+) m ³ /h	(-) m ³ /h
Poniedziałek	4 II 2002	344,17	22,32	6,49	41,51	-44,63	22,01	6,40	33,94	-50,60
Wtorek	5 II 2002	354,17	29,71	8,39	77,84	-8,65	33,37	9,42	84,14	-7,65
Środa	6 II 2002	351,67	17,71	5,04	17,24	-50,55	19,86	5,65	15,44	-60,10
Czwartek	7 II 2002	336,25	25,88	7,70	9,27	-50,83	28,07	8,35	12,90	-54,25
Piątek	8 II 2002	343,33	22,29	6,49	53,72	-31,24	21,91	6,38	54,59	-28,27
Sobota	27 IV 2002	330,00	18,03	5,46	35,12	-31,36	23,94	7,25	45,40	-19,22
Niedziela	28 IV 2002	277,08	24,54	8,86	45,49	-34,31	32,13	11,60	65,58	-40,99

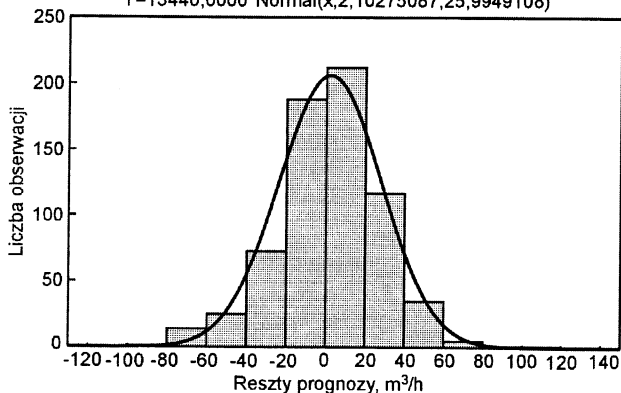
Analiza porównawcza efektywności modeli prognostycznych

Przykładowe rezultaty prognozowania godzinowego poboru wody na podstawie modeli ARIMA oraz addytywnej i multiplikatywnej wersji modelu Wintersa podano w tabelach 4 i 5, natomiast na rysunkach 6–9 pokazano rzeczywiste



Rys. 6. Rzeczywiste i prognozowane według modelu ARIMA (1,0,0)(1,1,0)₂₄ godzinowe pobory wody w czasie od 4 do 8 lutego 2002 r. w dniach roboczych

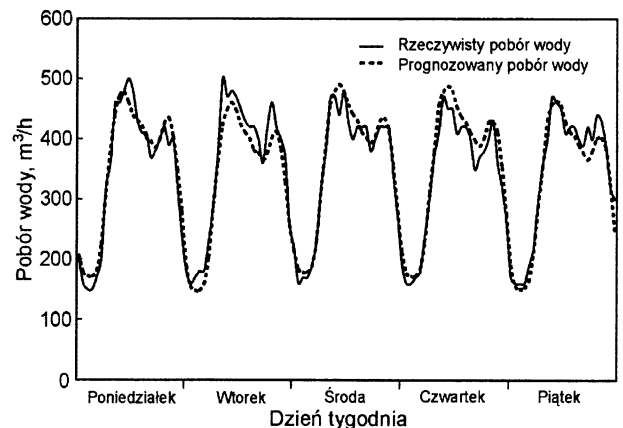
d Kolmogorowa-Smirnowa 0,04108, p=n.i., p Lillieforsa <0,01
Test chi-kwadrat=25,30509, df=5 (dopasow.), p=0,00012
Y=13440,0000*Normal(x;2,10275087;25,9949108)



Rys. 7. Rozkład częstości reszt dobowych prognoz godzinowego poboru wody wg modelu ARIMA (1,0,0)(1,1,0)₂₄ wraz z wynikami testów zgodności z rozkładem normalnym

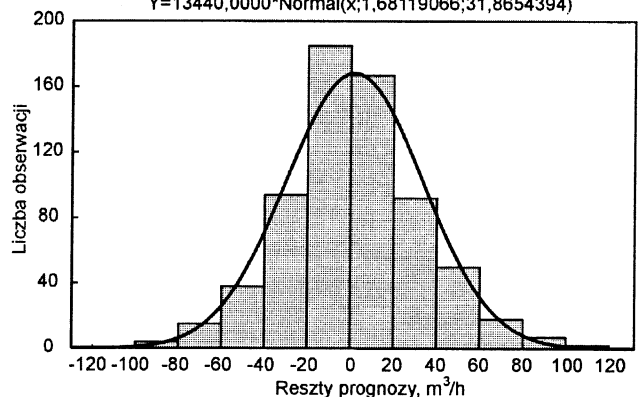
i prognozowane szeregi czasowe oraz rozkłady częstości reszt analizowanych modeli na tle rozkładu normalnego wraz z wynikami testów zgodności z rozkładem Gaussa.

Z przeprowadzonych analiz wynika, że efektywności prognozowania wg obu badanych metod były zbliżone, a szeregi reszt modeli prognostycznych podlegały procesowi białego szumu.



Rys. 8. Rzeczywiste i prognozowane według addytywnego modelu wykładniczego godzinowe pobory wody w czasie od 4 do 8 lutego 2002 r. w dniach roboczych

d Kolmogorowa-Smirnowa 0,04453, p<0,15, p Lillieforsa <0,01
Test chi-kwadrat=13,46855, df=5 (dopasow.), p=0,01936
Y=13440,0000*Normal(x;1,68119066;31,8654394)



Rys. 9. Rozkład częstości reszt dobowych prognoz godzinowego poboru wody dla addytywnego modelu wykładniczego wraz z wynikami testów zgodności z rozkładem normalnym

Podsumowanie

Omówione metody prognozowania szeregów czasowych godzinowego poboru wody (modele klasy ARIMA i metody wykładniczego wygładzania szeregów czasowych) spełniają wygodną dla potrzeb praktycznych zasadę łatwej dostępności danych wyjściowych do prognozowania, w tym predyktorów. Nie ujmują one żadnych zmiennych zewnętrznych, a bazują wyłącznie na chronologicznych ciągach obserwacji z bezpośredniej przeszłości. Do analizy efektywności poszczególnych metod prognozowania zostały wykorzystane wyniki pomiarów poboru wody w systemie wodociągowym Brzegu.

Zaletą modeli klasy ARIMA – przy ich oszczędnej parametryzacji – jest to, że proces prognozowania poboru wody może być zainicjowany przy stosunkowo małej liczbie wyrazów szeregu czasowego. W praktyce już dwutygodniowy ciąg obserwacji godzinowego poboru wody stwarza możliwość zainicjowania procesu prognozowania. Strukturę modelu ARIMA $(1,0,0)(1,1,0)_{24}$ można uznać za uniwersalną, nadającą się do prognozowania godzinowych poborów wody w miejskich systemach wodociągowych. Przyjęcie tej tezy, zwłaszcza, że pokrywa się ona z wynikami badań autorów [7,8], może uprościć automatyzację procesu prognozowania, bowiem umożliwi eliminację z niego etapu poszukiwania optymalnej struktury modelu. Wyniki prognozowania ulegają poprawie przez opis i przekształcenie szeregu surowego za pomocą szeregu Fouriera.

Metody prognozowania oparte na algorytmach wygładzania wykładniczego są łatwe do zastosowania i nie wymagają założenia o stacjonarności szeregu czasowego. W analizowanym systemie wodociągowym prognozy o najmniejszych błędach średniokwadratowych generuje addytywna wersja modelu Wintersa.

Alternatywą dla opisanych metod prognozowania mogą być ewentualnie sztuczne sieci neuronowe. Jako przykład ich efektywności można podać, że względne średniokwadratowe błędy prognoz godzinowego poboru wody w Brzegu (w przedziale czasu od 4 do 6 lutego 2002 r.) wynosiły od 6,16% do

6,69%. Wyniki analiz efektywności prognozowania przy użyciu sztucznych sieci neuronowych będą przedmiotem oddzielnej publikacji.

LITERATURA

1. J. T. DUDA, M. KLEMIATO: Klasyfikacja i nieklasyczne algorytmy prognozowania zakłóceń w systemach wodociągowych. Mat. konf. „Technologia i automatyzacja systemów wodociągowych i kanalizacyjnych – TiASWiK '99”, Stawiska 1999, ss. 79–86.
2. Z. SIWOŃ: Stochastyczne modelowanie procesu zużycia wody i prognozowanie zapotrzebowania na wodę w miastach. Prace Naukowe Instytutu Inżynierii Ochrony Środowiska Politechniki Wrocławskiej nr 56, Monografie nr 25, Wrocław 1986.
3. Z. SIWOŃ: Problemy krótkoterminowego prognozowania poboru wody w miejskich systemach wodociągowych. Mat. konf. „Aktualne zagadnienia w uzdatnianiu i dystrybucji wody”, Szczyrk 2003, ss. 323–332.
4. A. JAIN, L. E. ORMSBEE: Short-term water demand forecast modeling techniques – conventional methods versus AI. Journal AWWA, 2002, No. 7, pp. 64–72.
5. P. LICZYNAR, J. ŁOMOTOWSKI: Prognozowanie dobowych rozbiorów wody przy wykorzystaniu sztucznych sieci neuronowych. Mat. konf. „Zaopatrzenie w wodę, jakość i ochrona wód”, tom II, PZITS, Poznań 2004, ss. 175–183.
6. T. SŁOŃSKI, K. DUZINKIEWICZ, M. KIESIELEWICZ, D. TRAWICKI: Analiza możliwości zastosowania sztucznych sieci neuronowych do prognozowania zapotrzebowania na wodę w miejskich sieciach wodociągowych. Mat. konf. „Technologia i automatyzacja systemów wodociągowych i kanalizacyjnych – TiASWiK '99”, Stawiska 1999, ss. 235–242.
7. Z. SIWOŃ, J. STANISŁAWSKI: Über Prognose von Zeitreihen des Stundenwasserbedarfs in Städten und Regionen. Das Gas- und Wasserfach, 1991, 132, Nr. 6, S. 322–329.
8. Z. SIWOŃ, J. STANISŁAWSKI: Modelowanie i prognozowanie szeregów czasowych godzinowego poboru wody w miastach i regionach wiejskich. PAN, Instytut Budownictwa Wodnego, Rozprawy Hydrotechniczne, 1993, ss. 3–32.
9. G. E. P. BOX, G. M. JENKINS: Analiza szeregów czasowych. PWN, Warszawa 1983.

Siwoń, Z., Cieżak, W., Cieżak, J. Stochastic Models of Water Demand in a Chosen Water-Supply System. *Ochrona Środowiska* 2005, Vol. 27, No. 1, pp. 7–13.

Abstract: The study reported on in this paper involved two models, a model of ARIMA class and the model developed by Winters, which are used to forecast an hourly water demand time series. The models have been verified by analyzing the sets of continuous observations of the water demand in the water-supply system of Brzeg. Both the methods meet the requirement of an easy availability of the input data for the needs of forecasting. They do not include any external variables and are based solely on the preceding chronological water demand time series. Characterized by a reasonable parametrization, the ARIMA class models offer the potentiality for initiating the forecasting of water demand with a comparatively small number of points in the time series. In practice, 14-day hourly observations will suffice to initiate the forecasting process. The structure of the ARIMA $(1,0,0)(1,1,0)_{24}$ model can be considered universal and suitable

for forecasting the water demand in municipal water-supply systems. Accepting this thesis, as well as taking into account the fact that it is consistent with the results of our previous researches, we can notably simplify the automation of forecasting, because in this way the stage of search for the model structure is eliminated from the analysis of the time series. It is worth noting that the transformation of a raw time series into a Fourier series upgrades the quality of the forecasts and that seasonal multiplicative models are suitable for forecasting the water demand. Forecasting methods based on algorithms for an exponential smoothing of the time series are easy to use and do not require the assumption of a stationary input sequence. The additive Winters' model generates forecasts with the smallest errors. Our study has produced the following findings: sophisticated methods do not always generate optimal forecasts; the models presented in this paper rank high in the hierarchy of the methods used for forecasting of water demand.

Keywords: water-supply systems, water demand, forecasting.