

Zbigniew Siwoń, Wojciech Cieżak, Jan Cieżak

Modele neuronowe szeregów czasowych godzinowego poboru wody w osiedlach mieszkaniowych

Stochastyczne modele szeregów czasowych poboru wody z sieci wodociagowych w miastach i aglomeracjach miejsko-przemysłowych, a także w ich wydzielonych rejonach oraz przez określone grupy odbiorców wody, mogą być wykorzystywane między innymi do [1, 2]:

- kalibracji modeli przepływów w systemach dystrybucji wody oraz komputerowej symulacji działania tych systemów,
- optymalnego sterowania dystrybucją wody i procesami jej oczyszczania,
- wyznaczania optymalnych poleceń operacyjnych w procesach oczyszczania i dystrybucji wody.

Poznanie zmienności godzinowego poboru wody w poszczególnych dniach tygodnia umożliwia ustalenie wzorcowych histogramów godzinowego poboru wody w czasie całej doby przez określone grupy odbiorców oraz przez wszystkich odbiorców. Histogramy stanowią między innymi jeden z zasadniczych elementów procesów kalibracji i weryfikacji modeli hydraulicznych oraz dynamicznego modelowania przepływów w systemach dystrybucji wody zarówno na potrzeby ich bieżącej eksploatacji, jak i programowania ich modernizacji i rozbudowy. Są również niezwykle przydatne w wykrywaniu nieprawidłowości w bieżącym funkcjonowaniu wodociągów, gdyż umożliwiają diagnozowanie odstępstw od stanu uznawanego za normalny stan eksploatacyjny i w konsekwencji podejmowanie działań lub sygnalizowanie, że ma miejsce przypadek pracy systemu technicznego wymagający nadzoru lub natychmiastowej reakcji [1]. Wiedza ta nie jest doceniana przez eksploatatorów systemów wodociagowych. Należy jednakże zauważyć, że diagnozowanie z wykorzystaniem analizy szeregów czasowych będzie tylko wtedy możliwe, gdy na podstawie wstępnych obserwacji ustalą się stany uznawane za normalne stany eksploatacyjne oraz charakterystyczne stany wywołane zjawiskami losowymi lub awariami urządzeń technicznych. Praktyczne możliwości wykorzystywania modeli szeregów czasowych poboru wody zostały bardziej szczegółowo omówione w pracy [1].

Do analizy i prognozowania szeregów czasowych poboru wody z sieci wodociagowych w miastach i aglomeracjach miejsko-przemysłowych, a także w ich wydzielonych rejonach oraz przez określone grupy odbiorców wody, mogą być wykorzystane niestacjonarne modele stochastyczne. Opisują one rzeczywisty proces na podstawie

pomiarów wejścia i wyjścia z obiektu (obciążonych szumem) z wykorzystaniem metod interpolacyjnych, aproksymacyjnych, analizy szeregów czasowych, modeli neuronowych itp. Są to najczęściej modele tzw. czarnej skrzynki, które ustala się z wykorzystaniem metod parametrycznych ARIMA lub ARIMAX, metod wygładzania szeregów czasowych, sztucznych sieci neuronowych (SSN), modeli rozmytych oraz analizy falkowej.

Stosunkowo nowym narzędziem, które może być wykorzystane do opisu i prognozowania szeregów czasowych, są sztuczne sieci neuronowe. Ich atrakcyjność wynika przede wszystkim z możliwości aproksymacji dowolnych nieliniowości oraz dostrajania przyjętej struktury na podstawie danych eksperymentalnych. Co więcej, proces budowy modelu neuronowego polega na badaniu zbiorów danych i prowadzi do całkowicie automatycznego oszacowania na tej podstawie modelu opisującego stwierdzone prawidłowości i współzależności. Stosowanie modeli neuronowych nie wymaga w związku z tym znajomości postaci funkcji opisującej daną prawidłowość. W efekcie modele neuronowe mogą znaleźć zastosowanie wszędzie tam, gdzie nie jest znane dokładne prawo opisujące kształtowanie się badanych zależności. Dodatkową zaletą sztucznych sieci neuronowych jest łatwość ich adaptacji do zmiennych warunków środowiska. Cecha ta predestynuje je do opisu zależności zmieniających się w czasie. Ewolucja zmian dynamiki procesów poboru wody i odpływu ścieków oraz wpływ warunków zewnętrznych utrudniają stosowanie wszelkich metod obliczeniowych opartych na koncepcji modelu stacjonarnego. Tymczasem w przypadku sieci neuronowej zmiany zachodzące w modelowanym obiekcie nie stanowią zwykle przeszkody. W chwili pojawienia się nowych danych przeprowadzony może zostać zawsze proces douczenia sieci (lub uczenia od nowa), co umożliwia uwzględnienie w tworzonym modelu informacji zawartych także w najnowszych obserwacjach. Rezultaty badań nad możliwościami praktycznego wykorzystania sztucznych sieci neuronowych w analizie danych pochodzących z monitoringu systemów wodociagowo-kanalizacyjnych zostały omówione w pracy [3].

Z doświadczeń własnych, a także z informacji literaturowych [4] wynika, że spośród znanych metod analizy i prognozowania statystycznych szeregów czasowych do badań i prognozowania procesów poboru wody i odpływu ścieków w praktyce stosuje się scałkowane modele autoregresji i średniej ruchomej (ARIMA i ARIMAX), metody wykładniczego wygładzania szeregów czasowych (zwłaszcza metodę wygładzania potrójnie wykładniczego oraz addytywny i multiplikatywny model Wintersa), modele regresyjne oraz sztuczne sieci neuronowe. Własne doświadczenia

w stosowaniu różnych metod prognozowania zostały zaprezentowane m.in. w pracach [2, 3, 5, 6]. Zamieszczone w nich informacje, a także doniesienia literaturowe wskazują, że w praktyce w większości przypadków najwyższą przydatność wykazują modele neuronowe, które mają szereg cech, dzięki którym stanowią przydatne narzędzie do analizy i prognozowania szeregów czasowych. W tym kontekście w niniejszej pracy zbadano i porównano przydatność sieci perceptronowych (MLP), sieci o radialnych funkcjach bazowych (RBG) oraz sieci liniowych do prognozowania dobowych histogramów godzinowego poboru wody.

Wyniki badań

Do modelowania i prognozowania dobowych profili godzinowego poboru wody przy użyciu sztucznych sieci neuronowych zostały wykorzystane szeregi czasowe wyników pomiarów poboru wody w osiedlach mieszkaniowych zaopatrywanych w wodę z wydzielonych rejonów sieci wodociągowej we Wrocławiu i Kłodzku. Osiedla te charakteryzują się zasadniczo odmienną liczbą mieszkańców. Wydzielony rejon sieci wodociągowej we Wrocławiu, z którego zaopatrywanych jest w wodę ok. 100 tys. mieszkańców osiedli „Nowy Dwór”, „Kozanów”, „Gądów” i „Muchobór Mały” (w przeważającej większości budynki 11-kondygnacyjne) wraz z kompletem usług, jest zasilany z jednego źródła (pompownia strefowa przy ul. Bystrzyckiej), a woda jest tłoczona dwoma rurociągami, na których są zamontowane przepływomierze elektromagnetyczne z ciągłą rejestracją wskazań. Wydzielony rejon sieci wodociągowej w Kłodzku obejmuje natomiast osiedle o zabudowie wielorodzinnej z budynkami 5- i 11-kondygnacyjnymi oraz osiedle z zabudową jednorodziną. Łączna liczba mieszkańców obu osiedli wynosi ok. 10 tys. Źródłem zasilania sieci jest pompownia przy ul. Dusznickiej. Pomiar sumarycznego poboru wody przez wszystkich jej odbiorców na terenie tych osiedli jest dokonywany za pomocą przepływomierza elektromagnetycznego oraz wodomierza z ciągłymi rejestracjami wskazań.

W analizach ilustrujących metodykę modelowania i prognozowania dobowych histogramów chwilowego poboru wody wykorzystano chronologicznie uszeregowane wyniki pomiarów godzinowego poboru wody w wydzielonych rejonach sieci wodociągowej w czasie od 25-04-2001 do 15-10-2001 we Wrocławiu oraz od 01-06-2006 do 11-02-2007 w Kłodzku. Zbiory wyników pomiarów podzielono na szeregi czasowe zawierające odpowiednio zbiory obserwacji obejmujące dni robocze, soboty oraz dni wolne od pracy (niedziele i święta) w czasie od 25-04-2001 do 16-09-2001 (Wrocław) i od 01-06-2006 do

31-12-2006 (Kłodzko). Dni od 17-09-2001 do 15-10-2001 (Wrocław) oraz od 01-01-2007 do 11-02-2007 (Kłodzko) zostały wyodrębnione w celu sprawdzenia jakości prognoz utworzonych przez wybrane optymalne modele prognostyczne.

Wybór zbiorów uczącego (50% ogólnej liczby obserwacji), walidacyjnego i testowego (odpowiednio po 25% liczby obserwacji), został przeprowadzony w taki sposób, aby zapewnić w każdym z nich dane z całego zakresu analizowanego szeregu czasowego. Zmienne przed podaniem na wejście sieci zostały przekonwertowane za pomocą metody „minimax” – wartości przeskalowano liniowo tak, że najmniejsza wartość ze zbioru uczącego przyjęła wartość 0, a największa 1. Model liniowy był reprezentowany przez sieć bez warstw ukrytych, zaś znajdujące się w warstwie wyjściowej neurony były w pełni liniowe. Wyboru optymalnych struktur sieci liniowej (tab. 1) dokonano analizując kilkadziesiąt różnych modeli oraz optymalizując liczbę wejść sieci (od 1 do 240).

Praktyczną przydatność modelowania i prognozowania przy wykorzystaniu liniowych sieci neuronowych ilustrują wyniki zamieszczone na rysunku 1. Względny średniokwadratowy błąd prognozy w poszczególnych dniach tygodnia wyniósł od 3,27% do 6,57% w Kłodzku oraz od 6,58% do 11,60% we Wrocławiu.

Perceptron wielowarstwowy (MLP) jest siecią jednokierunkową, składającą się z warstw wejściowej, ukrytych i wyjściowej. W sieci perceptronowej każdy z neuronów oblicza ważoną sumę wejść, a wyznaczony w ten sposób poziom pobudzenia staje się argumentem funkcji aktywacji (najczęściej jest to funkcja liniowa lub hiperboliczna), za pomocą której jest obliczana wartość wyjściowa neuronu. Następnie w każdej warstwie należy dobrać wartości wag i wartości progowe (bias) wszystkich neuronów. W obu rejonach sieci wodociągowej i w każdym z trzech określonych zbiorów (uczącego, walidacyjnego i testowego) przeprowadzono analizy mające na celu wstępne wyznaczenie opóźnienia (5 d, 10 d i 15 d), liczby warstw ukrytych (1 lub 2) oraz liczby neuronów w poszczególnych warstwach ukrytych (1÷25). W analizach wykorzystano jednoetapowy proces uczenia (Quasi-Newtona – BFGS) oraz zastosowano różne funkcje aktywacji w warstwach wejściowej i wyjściowej (liniową, logistyczną, tangens hiperboliczny, wykładniczą i sinusoidalną). Po przeprowadzeniu analiz każdego szeregu wybrano po jednej strukturze sieci perceptronowej, wykazującej wstępnie najmniejszy względny średniokwadratowy błąd modelu. Modele charakteryzujące się najmniejszym względnym średniokwadratowym błędem poddano dokładnym analizom, aby zminimalizować błąd prognozy. Rezultatem tych analiz było uzyskanie struktur sieci perceptronowej

Tabela 1. Wyniki analiz wyboru optymalnej struktury sieci liniowej w przypadku szeregów czasowych godzinowego poboru wody
Table 1. Results of analysis of the choice of an optimal linear network structure for the time series of hourly water demand

Dzień	Struktura modelu	Średni pobór wody, m ³ /h	Bezwzględny średniokwadratowy błąd modelu, m ³ /h	Względny średniokwadratowy błąd modelu, %	Współczynnik korelacji
wydzielony rejon sieci wodociągowej w Kłodzku					
Roboczy	liniowa s144 1:144-1:1	54,77	2,81	5,13	0,9898
Sobota	liniowa s30 1:30-1:1	58,34	3,34	5,73	0,9896
Niedziela	liniowa s60 1:60-1:1	51,86	3,71	7,15	0,9846
wydzielony rejon sieci wodociągowej we Wrocławiu					
Roboczy	liniowa s133 1:133-1:1	448,02	26,10	5,83	0,9894
Sobota	liniowa s47 1:47-1:1	501,17	25,86	5,16	0,9932
Niedziela	liniowa s108 1:108-1:1	466,38	37,78	8,10	0,9841

Metoda uczenia: Pseudoinwersji (PI) (optymalizacja liniowa najmniejszych kwadratów)

wykazujących najmniejszy średniokwadratowy błąd w modelowaniu i prognozowaniu dobowego profilu godzinowego poboru wody. W tabeli 2 zestawiono przykładowe struktury optymalnych sieci, a także metody ich uczenia i błędy średniokwadratowe. Do opisu struktury sieci zastosowano oznaczenia stosowane w programie STATISTICA.

Praktyczną przydatność modelowania i prognozowania przy wykorzystaniu perceptronowych sieci neuronowych ilustrują wyniki zamieszczone na rysunku 2. Względny średniokwadratowy błąd prognozy w poszczególnych dniach tygodnia wyniósł od 2,77% do 5,47% w Kłodzku oraz od 6,01% do 10,39% we Wrocławiu.

Tabela 2. Wyniki analiz dotyczących wyboru optymalnej struktury sieci perceptronowej w przypadku szeregów czasowych godzinowego poboru wody

Table 2. Results of analysis of the choice of an optimal perceptron network structure for the time series of hourly water demand

Dzień	Struktura modelu	Średni pobór wody, m ³ /h	Bezwzględny średniokwadratowy błąd modelu, m ³ /h	Względny średniokwadratowy błąd modelu, %	Współczynnik korelacji
wydzielony rejon sieci wodociągowej w Kłodzku					
Roboczy	MLP s120 1:120-14-1:1	55,08	2,91	5,29	0,9888
Sobota	MLP s60 1:60-13-1:1	58,73	3,04	5,18	0,9907
Niedziela	MLP s120 1:120-11-1:1	51,86	3,53	6,80	0,9854
wydzielony rejon sieci wodociągowej we Wrocławiu					
Roboczy	MLP s120 1:120-9-1:1	447,59	25,68	5,74	0,9896
Sobota	MLP s60 1:60-6-1:1	505,03	25,59	5,07	0,9933
Niedziela	MLP s120 1:120-9-1:1	463,30	33,03	7,13	0,9893

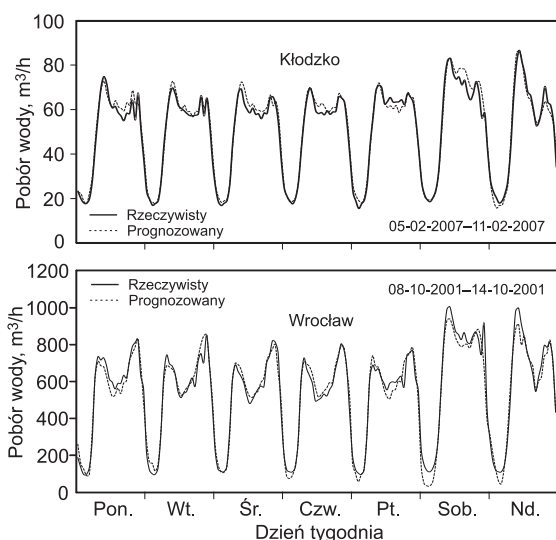
Metoda uczenia: Quasi-Newtona (BFGS) (w dni robocze we Wrocławiu: I etap – Quasi-Newtona (BFGS), II etap – gradienty sprzężone)

Tabela 3. Wyniki analiz dotyczących wyboru optymalnej struktury sieci o radialnych funkcjach bazowych w przypadku szeregów czasowych godzinowego poboru wody

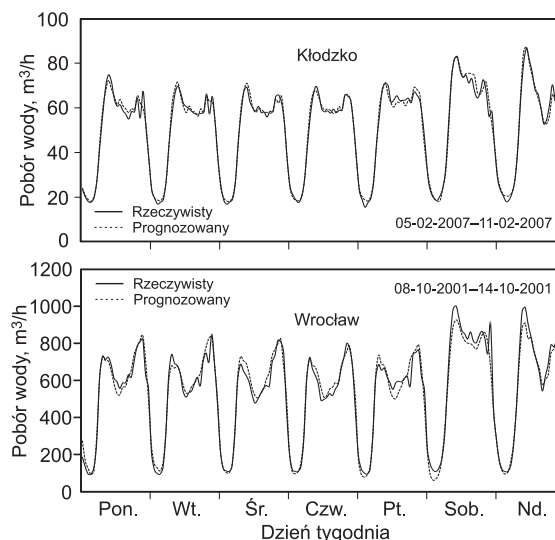
Table 3. Results of analysis of the choice of an optimal RBF network structure for the time series of hourly water demand

Dzień	Struktura modelu	Średni pobór wody, m ³ /h	Bezwzględny średniokwadratowy błąd modelu, m ³ /h	Względny średniokwadratowy błąd modelu, %	Współczynnik korelacji
wydzielony rejon sieci wodociągowej w Kłodzku					
Roboczy	RBF s120 1:120-132-1:1	55,08	3,54	6,43	0,9841
Sobota	RBF s50 1:50-20-1:1	58,63	4,70	8,02	0,9795
Niedziela	RBF s72 1:72-21-1:1	52,03	4,50	8,65	0,9774
wydzielony rejon sieci wodociągowej Wrocławiu					
Roboczy	RBF s60 1:60-37-1:1	447,95	33,06	7,38	0,9829
Sobota	RBF s120 1:120-7-1:1	497,93	51,00	10,24	0,9734
Niedziela	RBF s120 1:120-15-1:1	463,30	51,42	11,10	0,9704

Metoda uczenia: K-średnich (KM), K-najbliższych sąsiadów (KN), Pseudoinwersji (PI)



Rys. 1. Rzeczywisty i prognozowany wg optymalnych sieci liniowych godzinowy pobór wody w wydzielonym rejonie sieci wodociągowej miasta (horyzont prognozy – 1 d)
Fig. 1. Real and forecasted (according to the optimal linear networks) hourly water demand within a separate water supply subsystem in the city (forecast horizon: 24 h)



Rys. 2. Rzeczywisty i prognozowany wg optymalnych sieci perceptronowych godzinowy pobór wody w wydzielonym rejonie sieci wodociągowej miasta (horyzont prognozy – 1 d)
Fig. 2. Real and forecasted (according to the optimal perceptron networks) hourly water demand within a separate water supply subsystem in the city (forecast horizon: 24 h)

Sieci o radialnych funkcjach bazowych (RBF) mają pojedynczą warstwę ukrytą z neuronami radialnymi i liniową warstwę wyjściową (z iloczynem skalarnym). Neurony warstwy ukrytej działają jako centra skupień podobnych przypadków (uczących), a warstwa wyjściowa tworzy funkcję dyskryminacyjną lub regresyjną (zależnie od problemu). Wyboru optymalnych struktur sieci RBF (tab. 3) dokonano analizując kilkadziesiąt różnych modeli i optymalizując:

- liczbę wejść sieci (od 1 do 240),
- liczbę neuronów w warstwie ukrytej (od 1 do 180).

Uczenie sieci przeprowadzono trzema metodami:

- KM: K-średnich – określanie wag neuronów radialnych,
- KN: K-najbliższych sąsiadów – określanie promienia (odchylenia),

- PI: Pseudoinwersji – optymalizacja liniowa najmniejszych kwadratów.

Praktyczną przydatność prognozowania przy wykorzystaniu sieci o radialnych funkcjach bazowych ilustrują wyniki analiz zamieszczone na rysunku 3. Wynika z nich, że przydatność tych sieci do prognozowania dobowych histogramów godzinowego poboru wody była ograniczona i mniejsza niż sieci liniowych oraz perceptronowych. Jedną z przyczyn może być fakt, że sieci tego typu są wrażliwe na dużą liczbę neuronów wejściowych. Ponadto wykresy pokazane na rysunku 3 wskazują, że prognoza „nie nadążała”

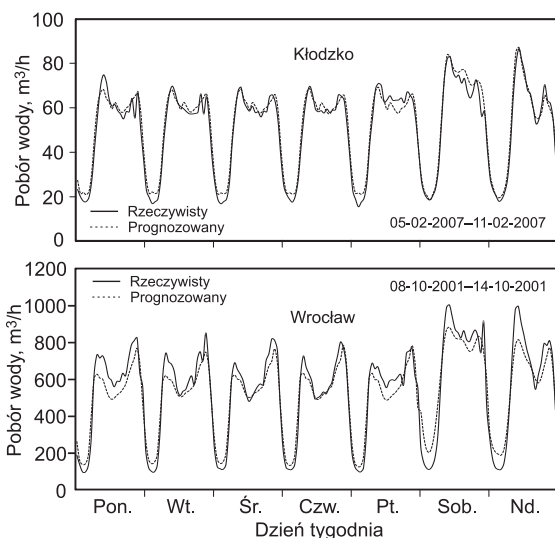
za szybkimi zmianami poboru wody w czasie. Względny średniokwadratowy błąd prognozy w poszczególnych dniach tygodnia wyniósł od 5,33% do 8,52% w Kłodzku oraz od 9,12% do 16,33% we Wrocławiu.

Podsumowanie

Analiza bieżącego prognozowania dobowych profili godzinowego poboru wody w wydzielonych rejonach miejskich sieci wodociagowych w Kłodzku i we Wrocławiu wykazała względnie dobrą jakość predykcji za pomocą neuronowych sieci liniowych i perceptronowych, porównywalną lub lepszą od jakości predykcji wg modeli klasy ARIMA i metod wykładniczego wygładzania szeregów czasowych [3]. Perceptron wielowarstwowy charakteryzował się prostotą struktur z jedną warstwą ukrytą, przez co proces jego douczania lub uczenia od nowa nie wymagał długotrwałych obliczeń. Stwierdzono, że w procedurach doboru struktury sieci perceptronowej można ograniczyć opóźnienie tego samego typu do 5 d (dni robocze, soboty oraz niedziele i święta), liczbę warstw ukrytych (do 1) oraz liczbę neuronów w warstwie ukrytej (do 15). Jakości prognoz otrzymanych przez zastosowanie sieci liniowej i perceptronowej były zadowalające, natomiast sieci o radialnych funkcjach bazowych dawały większy błąd prognozy.

LITERATURA

1. J. ŁOMOTOWSKI, Z. SIWOŃ: Metodyka analizy danych pochodzących z monitoringu systemów wodociagowo-kanalizacyjnych. *Gaz, Woda i Technika Sanitarna* 2010, nr 3, ss. 16–20.
2. Z. SIWOŃ, W. CIEŻAK, J. CIEŻAK: Modele neuronowe szeregów czasowych krótkotrwałego zużycia wody w osiedlach mieszkaniowych. *Mat. konf. „Zaopatrzenie w wodę, jakość i ochrona wód”*, PZITS Oddział Wielkopolski, Poznań–Gniezno 2008, tom 1, ss. 79–92.
3. Z. SIWOŃ, J. ŁOMOTOWSKI, W. CIEŻAK, P. LICZNAR, J. CIEŻAK: Analizy i prognozowanie rozbiorów wody w systemach wodociagowych. *Komitet Inżynierii Ładowej i Wodnej PAN, Instytut Podstawowych Problemów Techniki, Studia z Zakresu Inżynierii nr 61*, Warszawa 2008.
4. A. JAIN, L. ORMSBEE: Evaluation of short-term water demand forecast modeling techniques: Conventional methods versus AI. *Journal American Water Works Association* 2002, Vol. 94, No. 7, pp. 64–72.
5. Z. SIWOŃ, W. CIEŻAK: Wybrane problemy kalibracji modeli przepływów w systemach dystrybucji wody. *Instal* 2008, wydanie specjalne, ss. 79–85.
6. W. CIEŻAK, Z. SIWOŃ, J. CIEŻAK: Zastosowanie sztucznych sieci neuronowych do prognozowania szeregów czasowych krótkotrwałego poboru wody w wybranych systemach wodociagowych. *Ochrona Środowiska* 2006, vol. 28, nr 1, ss. 39–44.



Rys. 3. Rzeczywisty i prognozowany wg optymalnych sieci o radialnych funkcjach bazowych godzinowy pobór wody w wydzielonym rejonie sieci wodociagowej miasta (horizont prognozy – 1 d)

Fig. 3. Real and forecasted (according to the optimal RBF networks) hourly water demand within a separate water supply subsystem in the city (forecast horizon: 24 h)

Siwoń, Z., Cieżak, W., Cieżak, J. Neural Network Models of Hourly Water Demand Time Series in Housing Areas. *Ochrona Środowiska* 2011, Vol. 33, No. 2, pp. 23–26.

Abstract: The paper outlines the results of modeling and forecasting the water demand time series for the optimal control of water supply processes in municipal water supply systems. The results of verification of the artificial neural network models have been presented for a separate water supply subsystem in Kłodzko and in Wrocław. Analysis of the performance of artificial neural networks when used to develop current predictions of the time series for hourly water demand has revealed that the optimal

structures of perceptron and linear networks are not very complicated, which facilitates the process of additional training or re-training. Practically, it has been found that forecasting produces comparable or smaller errors when focused on multilayer perceptron neural networks and linear neural networks than when based on the use of ARIMA models and exponential smoothing of the time series. Applicability of neural networks of radial base functions (RBF) to forecasting daily water demand histograms is limited, and lesser than that of linear and perceptron networks.

Keywords: Artificial neural networks, time series, forecasting, water demand, water supply system.