

Paweł Licznar, Janusz Łomotowski

Zastosowanie sztucznych sieci neuronowych Kohonena do prognozowania dobowego poboru wody

W ostatnich latach obserwuje się wzrost zainteresowania sztucznymi sieciami neuronowymi, zaliczanymi do podstawowych narzędzi *data-mining* (analizy dużych zbiorów danych), w prognozowaniu i modelowaniu wielu procesów naturalnych i technologicznych. W doniesieniach literaturowych krajowych [1,2] i zagranicznych [3–8] można odnaleźć informacje, że sztuczne sieci neuronowe mogą być wykorzystywane do prognozowania chwilowego, godzinowego i dobowego poboru wody.

Zastosowanie sztucznych sieci neuronowych w obrębie prognozowania poboru wody ograniczało się dotychczas najczęściej do najprostszych typów sieci, głównie sieci perceptronowych o pojedynczej warstwie ukrytej, i służyło tylko do prognozy poboru dobowego [1,2]. Otrzymywano przy tym wyniki porównywalne lub lepsze od modeli stochastycznych. Budowane sieci neuronowe stanowiły w rzeczywistości modele tzw. czarnej skrzynki, co uniemożliwiało przeprowadzenie analizy wpływu parametrów sieci na czynniki kształtujące pobór wody. Wagi poszczególnych neuronów sieci perceptronowych, ustalane w trakcie ich uczenia, nie były bowiem powiązane z fizycznymi cechami prognozowanego szeregu czasowego. Poznawcze aspekty perceptronowych sztucznych sieci neuronowych są nieporównywalnie mniejsze w stosunku do modeli stochastycznych opartych o analizę szeregów czasowych [9,10].

Celem przeprowadzonych badań było sprawdzenie możliwości prognozowania dobowego poboru wody za pomocą samoorganizującej sieci Kohonena połączonej z siecią perceptronową. Takie hybrydowe połączenie sztucznych sieci neuronowych pozwala nie tylko na prognozowanie poboru wody, ale pełni również rolę narzędzia poznawczego umożliwiającego identyfikację przyczyn obserwowanych zmian.

Materiał i metody

W badaniach poddano analizie szereg czasowy dobowego zużycia wody w latach 1996–2002 w jednym z większych polskich wodociągów. Szereg ten był już wcześniej wykorzystywany do prognozowania dobowego poboru wody metodą wygładzania wykładniczego [11] oraz prostych sztucznych sieci neuronowych, tj. liniowych, perceptronowych oraz o radialnych funkcjach bazowych [1]. Charakterystyczną cechą analizowanego szeregu była spadkowa tendencja zużycia wody, obserwowana w ostatnich latach w większości polskich

wodociągów, będąca efektem ekonomizacji gospodarki wodno-ściekowej oraz zmniejszenia zużycia wody przez przemysł.

Do prognozowania dobowego poboru wody postanowiono zastosować metodykę analogiczną do stosowanej w prognozowaniu obciążeń w systemach elektroenergetycznych [12, 13]. W pracy [13] wykazano, że sztuczne sieci neuronowe – wielowarstwowe i samoorganizujące – są skutecznym narzędziem do krótkoterminowej prognozy dobowego, a nawet godzinowego obciążenia elektroenergetycznego. Metoda prognozowania obciążeń godzinowych składa się z dwóch elementów [13]:

– wyznaczenia za pomocą sieci samoorganizującej, wytrenowanej przy użyciu algorytmu Kohonena, tzw. profilu godzinowego j -tej doby – $p(j,h)$, zdefiniowanego zależnością:

$$p(j,h) = \frac{P(j,h) - P_m(j)}{\sigma^2(j)} \quad (1)$$

w której:

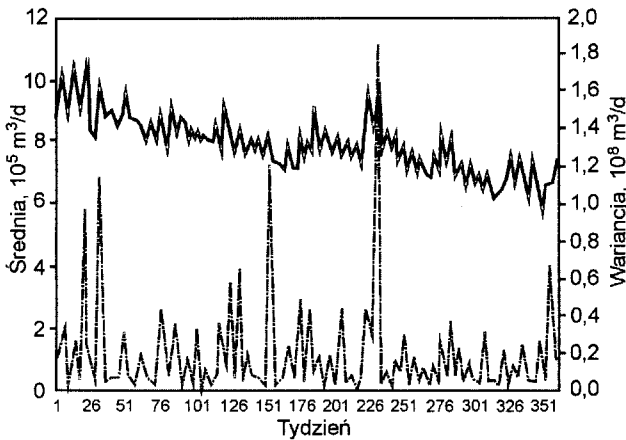
$h=1, 2, \dots, 24$ – godzina w dobie

$P(j,h)$ – rzeczywisty pobór mocy w systemie elektroenergetycznym w j -tej dobie o godzinie h , MW

$P_m(j)$ – wartość średnia obciążenia systemu w j -tej dobie, MW

$\sigma^2(j)$ – wariancja obciążenia systemu w j -tej dobie, (MW)²
– prognozowania wartości średniej i wariancji obciążenia systemu w poszczególnych dobach, za pomocą sieci perceptronowych szeregów czasowych o pojedynczych warstwach ukrytych.

We wstępnej fazie badań zastąpiono profil godzinowy profilem dobowym, odwzorowującym pobór wody w poszczególnych dniach tygodnia, a wartości średniej i wariancji w j -tej dobie przez ich wartości dla dobowego poboru wody w j -tym tygodniu. Osiągnięto przy tym zadowalające wyniki w zakresie wyznaczania profili poboru wody przy pomocy sieci samoorganizującej oraz w zakresie prognozowania średniego dobowego poboru wody w poszczególnych tygodniach. Niestety, nie było możliwe prognozowanie wariancji dobowego poboru wody w poszczególnych tygodniach. Podejmowano w tym zakresie próby zastosowania sieci perceptronowych o różnej architekturze, jednakże za każdym razem proces uczenia skończył się negatywnie. W świetle tych doświadczeń wydaje się, że szereg wartości wariancji dobowego poboru wody w poszczególnych tygodniach miał charakter tzw. białego szumu. Obserwacje te zdaje się potwierdzać wykres wartości wariancji szeregu czasowego o całkowicie odmiennym charakterze względem wykresu wartości średnich, z nagle występującymi skokami wartości maksymalnych (rys. 1).



Rys. 1. Szeregi czasowe średnich (linia ciągła) i wariacji (linia przerywana) dobowego rozbioru wody w poszczególnych tygodniach

W celu wyeliminowania tych trudności, w dalszej części badań postanowiono zmodyfikować i uprościć definicję profilu dobowego. Ostatecznie przyjęto prostą formułę profilu dobowego poboru wody (w kolejnych dniach tygodnia) w postaci:

$$q(i,j) = \frac{Q(i,j)}{Q_t(i)} \quad (2)$$

w której:

$j=1, 2, \dots, 7$ – indeks kolejnych dni tygodnia liczonych od niedzieli

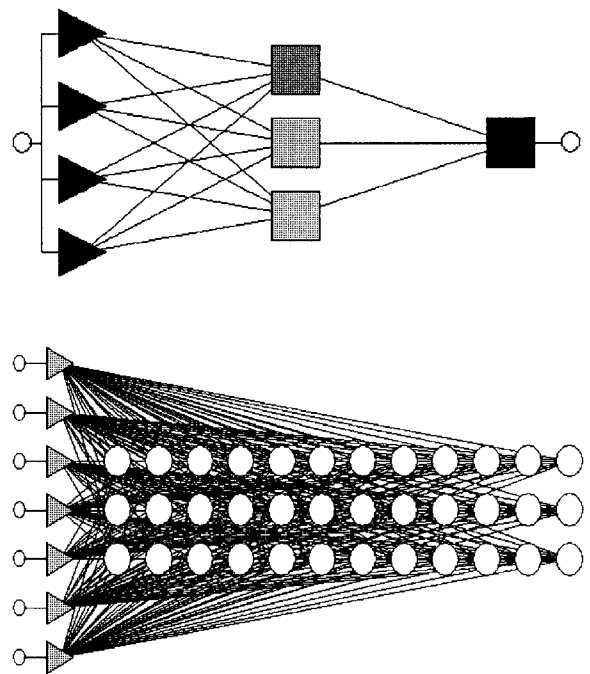
$Q(i,j)$ – rzeczywisty pobór wody w j -tej dobie i -tego tygodnia, m^3

$Q_t(i)$ – sumaryczny tygodniowy pobór wody w i -tym tygodniu, m^3

Profile dobowego poboru wody w tygodniu zostały wyznaczone osobno dla czterech okresów, tj. zimowego, wiosennego, letniego i jesiennego. Podział taki jest często stosowany w modelowaniu stochastycznym poboru wody w systemie wodociągowym [11]. Do opracowania wektorów profilu dobowego bazowano na obliczonych wartościach $q(i,j)$ dla pierwszych 4 lat szeregu czasowego (1996–1999). Stanowiły one wejście do sieci samoorganizującej celem wykrycia przez nią istniejących wzorców. Sieć zbudowana z neuronów ułożonych w węzłach kwadratowej siatki o wymiarach: 3×12 (rys. 2) została dobrana eksperymentalnie. Do uczenia sieci samoorganizującej przez konkurencję zastosowano klasyczny algorytm Kohonena. Szczegółowe informacje dotyczące sieci samoorganizujących, ich budowy, uczenia, funkcjonowania i możliwości potencjalnych zastosowań zawarte są w pracach [12, 14, 15].

Po etapie uczenia sieci, do wyznaczenia przewidywanego profilu obciążenia dobowego (rozkładu poboru wody w poszczególnych dniach tygodnia) dla danej pory roku, przyjęto uśrednione wartości wag neuronów zwycięzców w poszczególnych tygodniach analizowanego okresu. Przewidywany wektor profilu dobowego i -tego tygodnia wyodrębnionych pór roku obliczono ze wzoru:

$$\hat{q}_i = \frac{\sum_{j=1}^n k_{ij} w_j}{\sum_{j=1}^n k_{ij}} \quad (3)$$



Rys. 2. Struktura sieci perceptronowej o pojedynczej warstwie ukrytej typu MLP s4 1:4-3-1:1 (góra) oraz sieci samoorganizującej Kohonena typu SOFM 7:7-36:1 (dół)

w którym:

k_{ij} – liczba zwycięstw j -tego neuronu, odpowiadająca i -temu tygodniowi ($k_{ij}=0$, gdy j -ty neuron nigdy nie zwyciężył dla danej pory roku)

w_j – wektor wagowy neuronu j

Do prognozowania wartości sumarycznego poboru wody w kolejnych tygodniach opracowano prostą sieć perceptronową szeregu czasowego. Na jej wejściu podawane były wartości sumarycznego poboru wody dla czterech ostatnich tygodni, a na wyjściu uzyskano prognozowane sumaryczne zapotrzebowanie na wodę w następnym tygodniu. Struktura sieci była bardzo prosta, a jej pojedyncza warstwa ukryta składała się z trzech neuronów (rys. 2). Do uczenia sieci zastosowano klasyczny algorytm propagacji wstecznej, z podziałem całego szeregu czasowego na trzy podzbiory, tj. uczący, walidacyjny i testowy (liczące odpowiednio 50%, 25% i 25% łącznej liczby wektorów danych).

Prognozy dobowego poboru wody dokonano w oparciu o równanie:

$$\hat{Q}(j,k) = \hat{g}(j,h) \hat{Q}_t(j) \quad (4)$$

w którym:

$h=1, 2, \dots, 7$ – indeks kolejnych dni tygodnia liczonych od niedzieli

$\hat{Q}(j,h)$ – prognozowany pobór wody w h -tej dobie j -tego tygodnia, m^3

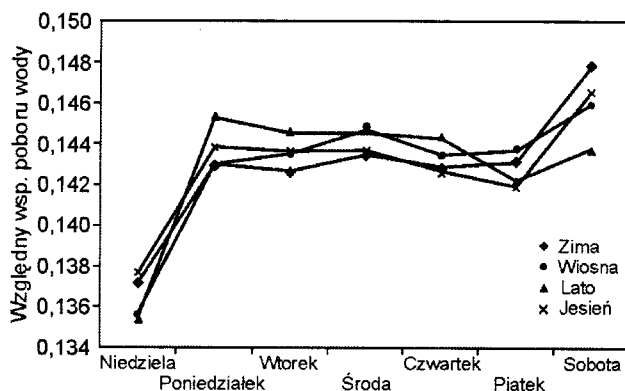
$\hat{q}(j,h)$ – przewidywany udział h -tej doby tygodnia w sumarycznym poborze wody w j -tym tygodniu

$\hat{Q}_t(j)$ – prognozowany pobór wody w j -tym tygodniu, m^3

Prognozy dobowe wykonane dla lat 2000–2002 porównano z rzeczywistymi poborami wody wodociągowej. Szereg czasowy z lat 2000–2003 nie był użyty do identyfikacji tygodniowych wzorców poborów wody. Przy poszukiwaniu parametrów sztucznych sieci neuronowych wykorzystano pakiet Sieci Neuronowe programu Statistica 6PL.

Dyskusja wyników badań

Najistotniejszym – pod względem poznawczym – rezultatem przeprowadzonych badań było wyznaczenie profilu obciążeń dobowych dla poszczególnych pór roku (rys. 3). Od poniedziałku do piątku dobowy pobór wody wynosił około 14,3% tygodniowego zużycia wody w mieście. We wszystkich porach roku zaobserwowano wzrost zużycia wody w sobotę (w zimie dochodził on do 14,8% poboru tygodniowego) i spadek w niedzielę (do 13,5% poboru tygodniowego w lecie). Z danych przedstawionych na rysunku 3 wynika, że w końcu tygodnia (sobota–niedziela) wartości udziałów w poborze całotygodniowym były najniższe wiosną i latem. Wydaje się, że mogło to wynikać z wyjazdów mieszkańców w ciepłe weekendy poza miasto. W ostatnich latach, krótkie wyjazdy na wypoczynek od piątku do niedzieli, zwłaszcza w okresie letnim, stały się bardzo popularne. Mogłoby to być również wytłumaczeniem znacznego spadku krzywej profilowej w okresie letnim, już w piątek, kiedy mieszkańcy opuszczali miasto.



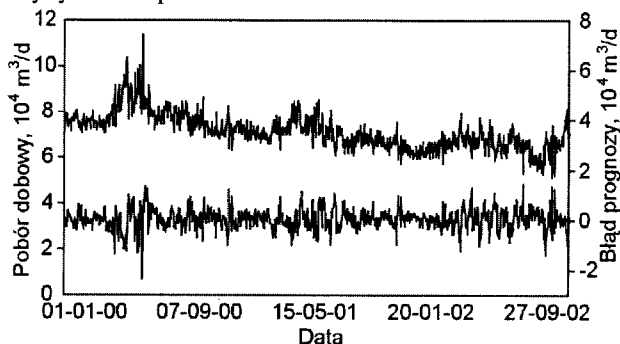
Rys. 3. Profile tygodniowego poboru wody (względne współczynniki poboru w poszczególnych dniach tygodnia) dla różnych pór roku

Podsumowanie wyników prognozowania tygodniowego poboru wody za pomocą prostej sieci perceptronowej przedstawiono w tabeli 1. Wysokie wartości współczynnika korelacji prognozowanego i obserwowanego tygodniowego poboru wody, wynoszące 0,9 we wszystkich podzbiorach, oraz stosunkowo niskie wartości ilorazów odchyień w zakresie od 0,3 do 0,4, pozostawały w pełnej zgodności z odpowiadającymi im wartościami uzyskanymi przez autorów przy prognozowaniu dobowego poboru wody za pomocą sztucznych sieci neuronowych dla tego samego szeregu czasowego [1]. Opracowana sieć neuronowa odwzorowywała spadkową tendencję zużycia wody.

Tabela 1. Statystyki regresji rzeczywistego i prognozowanego tygodniowego poboru wody przez sztuczną sieć neuronową typu MLP s4 1:4-3

Statystyka	Podzbiór		
	uczący	walidacyjny	testowy
Średnia	547 727,9	544 876,4	539 318,5
Odch. standardowe	59 101,9	58 175,8	66 348,4
Średni błąd	280,3	-2 261,9	-355,6
Odchylenie błędu	19 323,7	19 850,1	23 807,4
Średni błąd bezwzgl.	13 983,8	14 632,8	18 343,9
Iloraz odchyień	0,3	0,3	0,4
Korelacja	0,9	0,9	0,9

Prognozy dobowego poboru wody dla trzech ostatnich testowych lat szeregu czasowego zostały przedstawione na tle wartości błędu prognozy (rys. 4). Uzyskano dużą zbieżność wartości prognozowanych i obserwowanych. Współczynnik korelacji pomiędzy tymi wartościami dla lat 2001–2002 wyniósł 0,84. Wartość ta była niewiele mniejsza od wartości współczynnika korelacji uzyskanych przez autorów w prognozowaniu z użyciem prostych sieci perceptronowych, liniowych oraz o radialnych funkcjach bazowych ($r=0,93+0,94$) [1]. Należy jednak pamiętać, że wyższa wartość współczynnika korelacji dotyczyła prognoz na jedną dobę w przód dla całego 7-letniego szeregu czasowego, a nie tylko znacznie szerszego zbioru 3-letniego szeregu oraz prognozy na cały tydzień w przód.



Rys. 4. Szeregi czasowe (linia góra) i błąd bezwzględny prognozy (linia dolna) dobowego poboru wody w latach 2000–2002

Dobre wyniki prognozowania znalazły potwierdzenie w sünkowo niskich wartościach średniego błędu bezwzględnego, wynoszącego $304 \text{ m}^3/\text{d}$, w porównaniu z średnią wartością dobowego poboru wody w analizowanym trzyleciu równą $71 653 \text{ m}^3/\text{d}$. Było to powiązane w oczywisty sposób z bardzo niską wartością średniego błędu względnego, wynoszącą jedynie 0,7%. Warto przypomnieć, że dla prognoz *ex ante* poboru wody dla całego szeregu czasowego, przy użyciu metody wykładniczego z sezonowością tygodniową i roczną, autorzy w pracy [11] otrzymali wartości średniego błędu względnego równe odpowiednio 3,97% i 4,43%, a w pracy [1] – stosując proste sieci neuronowe – w zakresie 3,09+3,56%.

Największe przeszacowanie prognozy zanotowano 20 września 2002 r. – 27,8% rzeczywistego poboru wody. Było ono mniejsze od otrzymanego przez autorów w pracy [11] najwyższego przeszacowania prognozy dla modelu z sezonowością tygodniową, wynoszącego 33,61% rzeczywistego poboru wody, oraz najwyższego przeszacowania 34,3% dla prognozy z użyciem prostych sieci neuronowych [1]. Prognozy dla wszystkich pozostałych dni tygodnia były dokładniejsze. W pozostałej części zbioru danych jeszcze tylko dla trzech dni stwierdzono przeszacowanie prognozy o ponad 20%, a 21 czerwca 2000 r. niedoszacowanie prognozy o ponad 20% wartości poboru rzeczywistego. Otrzymane rezultaty były lepsze od uprzednio uzyskanych przez autorów w pracach [1, 11]. Należy zwrócić uwagę, że cytowane wyniki poprzednich badań dotyczyły analizy 7-letniego szeregu czasowego, a przedstawione w tej pracy okresu 3-letniego (2000–2002).

Podsumowanie

Uzyskane wyniki wskazują, że prognozowanie dobowego poboru wody w dużym systemie wodociągowym może być wykonane dwuetapowo. Pierwszym etapem może być prognoza

tygodniowego poboru wody przy użyciu bardzo prostej sieci perceptronowej w oparciu o znajomość czterech ostatnich tygodniowych poborów wody. Drugi etap to rozdział tygodniowego poboru wody na poszczególne dni tygodnia, zgodnie z wzorcami rozpoznanymi dla poszczególnych pór roku przez samoorganizującą strukturę sieci Kohonena. Taki sposób prognozowania cechuje się lepszą jakością, w porównaniu do prognoz opartych o wygładzanie wykładnicze lub proste sieci neuronowe. Jego zaletą jest poszerzony horyzont prognozy (cały tydzień) oraz estymowanie tygodniowych profili dobowego poboru wody w różnych porach roku. Uzyskane wyniki wskazują na duży potencjał hybrydowego prognozowania poboru wody w systemach wodociągowych przez połączenie technologii sieci perceptronowych z sieciami samoorganizującymi. Celowe jest prowadzenie dalszych badań w tym zakresie, z wykorzystaniem szeregów czasowych poborów wody z innych wodociągów o zróżnicowanej produkcji wody.

Dr inż. Paweł Licznar pragnie podziękować Fundacji na Rzecz Nauki Polskiej za wsparcie badań w ramach stypendium krajowego dla młodych naukowców.

LITERATURA

1. P. LICZNAR, J. ŁOMOTOWSKI: Prognozowanie dobowych rozbiórów wody przy wykorzystaniu sztucznych sieci neuronowych. Mat. konf. „Water Supply and Water Quality”, PZITS, Poznań 2004, ss. 175–183.
2. T. SŁONIMSKI, K. DUZINKIEWICZ, M. KWIESIELEWICZ, 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.
3. L.M. CAMARINHA-MATOS, F.J. MARTINELLI: Application of machine learning in water distribution networks. *Intelligent Data Analysis*, 1998, No. 2, pp. 311–332.
4. N. LERTPALANGSUNTI, C.W. CHAN, R. MASON, P. TONTI-WACHWUTHIKUL: A toolset for construction of hybrid intelligent forecasting system: Application for water demand prediction. *Artificial Intelligence in Engineering*, 1999, 13, pp. 21–42.
5. J. LIU, H.H.G. SAVENIJE, J. XU: Forecast of water demand in Weinan City in China using WDF-ANN model. *Physics and Chemistry of Earth*, 2003, 28, pp. 219–224.
6. A. MUKHOPADHYAY, A. AKBER, E. AL-AWADI: Analysis of freshwater consumption patterns in the private residences in Kuwait. *Urban Water*, 2001, No. 2, pp. 53–62.
7. S.L. ZHOU, T.A. MCMAHON, A. WALTON, J. LEWIS: Forecasting daily urban water demand: A case study of Melbourne. *Journal of Hydrology*, 2000, 236, pp. 153–164.
8. S.L. ZHOU, T.A. MCMAHON, A. WALTON, J. LEWIS: Forecasting operational demand for an urban water supply zone. *Journal of Hydrology*, 2002, 259, pp. 189–202.
9. Z. SIWON: Stochastyczne modelowanie procesu zużycia wody i prognozowanie zapotrzebowania na wodę w miastach. *Prace Naukowe Inst. Inż. Ochr. Środ. Politechniki Wrocławskiej, Seria Monografie nr 34*, Wrocław 1986.
10. Z. SIWON, W. CIEŻAK, J. CIEŻAK: Stochastyczne modele godzinowego poboru wody w wybranym systemie wodociągowym. *Ochrona Środowiska*, 2005, nr 1, ss. 7–13.
11. J. ŁOMOTOWSKI: Prognozowania rozbiórów wody metodą wygładzania wykładniczego. Mat. konf. „Zastosowanie technik informatycznych w zarządzaniu systemami wodno-kanalizacyjnymi”, PZITS, Poznań 2003, ss. 19–25.
12. S. OSOWSKI: Sieci neuronowe do przetwarzania informacji. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa 2000.
13. K. SIWEK: Prognozowanie obciążeń w systemie elektroenergetycznym przy wykorzystaniu sztucznych sieci neuronowych. Rozprawa doktorska, Politechnika Warszawska, Warszawa 2001 (praca nie publikowana).
14. M. KRAWCZAK: Multilayer Neural Systems and Generalized Net Models. Akademyka Oficyna Wydawnicza EXIT, Warszawa 2003.
15. R. TADEUSIEWICZ: Sieci neuronowe. Akademyka Oficyna Wydawnicza, Warszawa 1993.

Licznar, P., Łomotowski, J. Application of Kohonen Artificial Neural Networks to the Prediction of Daily Water Consumption. *Ochrona Środowiska* 2006, Vol. 28, No. 1, pp. 45–48.

Abstract: The objective of the study was to develop a hybrid tool for predicting daily water consumption by the combined use of the perceptron and Kohonen artificial neural networks. The investigations included a 7-year time series of total daily water consumption in the time span of 1996 to 2002, coming from one of Poland's largest water distribution systems. The prediction process was a two-stage one. At the first stage, the Self-Organizing Feature Map (SOFM) was made in order to

establish the weekly water distribution patterns that are typical for each season of the year. At the second stage, a simple single hidden layer perceptron network was built to enable the prediction of total weekly water consumption. Owing to the combined use of the perceptron and Kohonen artificial neural networks it was possible to work out high-quality daily water consumption predictions and to identify typical seasonal patterns of weekly water consumption.

Keywords: Daily water consumption, prediction, artificial neural networks, perceptron networks, Self-Organizing Feature Map (SOFM).