

# **PORTFOLIO:**

## **Koncepcja informatycznego systemu prognozy krótkoterminowej propagacji wezbrania na rzekach z kontrolowanym stanem**

*Autorzy: Tomasz Siuta, Mirosław Gajer, Zbigniew Handzel*

**Centrum Inteligentnych Systemów Informatycznych** Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 [www.isi.agh.edu.pl](http://www.isi.agh.edu.pl) [isi@agh.edu.pl](mailto:isi@agh.edu.pl)

# 1. Wprowadzenie

Odptyw ze zlewni, rejestrowany w przekroju zamykającym zlewnię i będący skutkiem opadu, jaki w tej zlewni wystąpił, jest jednym z ważniejszych elementów cyklu hydrologicznego, mieszczącym się w centrum zainteresowania inżynierów. Prognozy odpływu ze zlewni rzecznej mają podstawowe znaczenie dla prawidłowego funkcjonowania gospodarki wodnej, zwłaszcza w warunkach zagrożenia powodziowego. Na ich podstawie podejmowane są decyzje dotyczące sterowania odpływem ze zbiorników wodnych oraz pracą innych obiektów retencyjnych (poldery) i piętrzących (stopnie wodne). Ponadto prognozy odpływu chwilowego i średnich stanów wód z nim związanych są wykorzystywane przez ośrodki wczesnego ostrzegania w sytuacjach zagrożenia powodziowego.

W ostatnich latach coraz częściej spotykamy się z powodziami o gwałtownym przebiegu, których zasięg i wielkość trudno przewidzieć na podstawie samych tylko prognoz meteorologicznych dostarczających danych wejściowych do modeli hydrologicznych.

Wiarygodne scenariusze przebiegu wezbrań można uzyskać dopiero na podstawie udokumentowanej obserwacji reakcji zlewni na opad w czasie rzeczywistym. Z punktu widzenia poprawnego funkcjonowania systemów wczesnego ostrzegania interesuje nas nie tylko jak najdłuższy okres uprzedzenia, ale także odpowiednia dokładność prognozy. W hydrologii spotykamy się z wieloma modelami pozwalającymi na wyznaczenie wielkości odpływu zlewniowego. Wyróżniamy tutaj modele, które bazują na relacji opad – odpływ, oraz modele opisujące związki czasowe występujące pomiędzy dopływem i odpływem mierzonym (obliczonym) w różnych przekrojach kontrolnych sieci rzecznej, związki takie są efektem procesu transformacji fali wezbraniowej w korycie rzecznej. Podstawą opracowania modeli pochodzących z obu wspomnianych grup może być fizyczna analiza procesów hydrologicznych biorących udział w transformacji opadu w odpływ (takich jak: ewapotranspiracja, infiltracja, spływ powierzchniowy, spływ gruntowy) oraz interpretacja prawa zachowania pędu i masy w przypadku matematycznych modeli transformacji fali wezbraniowej w korycie rzecznej. Innym sposobem modelowania jest zastosowanie konceptualnych reprezentacji działania i budowy systemu zlewniowego oraz wykorzystanie metod statystycznych i stochastycznych.

Choć w ostatnich dziesięcioleciach opracowano wiele modeli pozwalających na symulację matematyczną procesów hydrologicznych towarzyszących transformacji opadu w odpływ, jak również złożone modele hydrauliczne przepływów nieustalonych w korycie rzecznej, to wciąż brakuje modeli przystosowanych do obliczeń operacyjnych w czasie rzeczywistym, stanowiących podstawę dla optymalnej gospodarki wodnej, szczególnie w warunkach zagrożenia powodziowego. Ponadto, jak wykazują liczne publikacje [Thirumalaiah, Deo, 2000], większość dostępnych dziś modeli ma charakter dystrybucyjny, tzn. prognozy są stawiane w wielu lokalizacjach w obrębie zlewni, co czasami wiąże się z pogorszeniem dokładności prognozy w pojedynczych wyróżnionych lokalizacjach. W sytuacjach, gdy wymagana jest jedynie prognoza w kilku ustalonych przekroju rzeki (np. na dopływach do zbiorników retencyjnych) leżących w zlewni, dla której brak adekwatnych danych meteorologicznych oraz topograficznych, nie zawsze celowe staje się stosowanie złożonych modeli wymagających dużej ilości precyzyjnych danych dotyczących systemu. W takich przypadkach pojawi się możliwość opracowania modeli typu czarnej skrzynki bądź modeli konceptualnych. W przypadku prognozy w czasie rzeczywistym wspomniane modele często wyposażone są

dodatkowo w procedury służące do uaktualniania parametrów modeli lub ich stanów oraz korekty błędów pojawiających się w danych wejściowych i wyjściowych. Prognoza odpływu ze zlewni, jak już wcześniej wspomniano, może być uzyskana na podstawie modeli transformacji opad – odpływ, mówimy wówczas o tzw. prognozie średnioterminowej, bądź w oparciu o związki czasowo – przestrzenne pomiędzy stanami zaistniałymi w danym systemie rzeczonym a stanem prognozowanym w przypadku prognozy krótkoterminowej [Lambor, 1962]. Wybór właściwego typu prognozy podyktowany jest przede wszystkim charakterem zlewni, w której prognoza odpływu jest prowadzona, a więc w zlewniach górskich, gdzie spływ powierzchniowy stanowi podstawę zasilania cieku głównego, na ogół stosowane będą modele transformacji opadu w odpływ, natomiast w zlewniach nizinnych i podgórskich często dobre wyniki obliczeń uzyskuje się na podstawie prognoz krótkoterminowych (opierającą się na stanach już zaistniałych w systemie) [Zealand i in., 1999]. Dla tego drugiego typu prognozy przedstawiono koncepcję operacyjnych modeli prognostycznych w ramach niniejszego opracowania. Zwrócono tutaj uwagę na przypadek graniczny, dla którego nie można przewidzieć a priori, który z wymienionych wcześniej typów prognozy będzie bardziej efektywny. W większości badanych zlewni występuje bowiem niekontrolowany dopływ boczny, który w istotny sposób przyczynia się do wielkości odpływu w przekroju zamykającym zlewnię, a dane opadowe nie są dostępne lub niekompletne. Pojawia się więc pytanie o możliwość skutecznego prognozowania natężeń przepływu w przekroju zamykającym wspomnianą zlewnię w oparciu o prognozę krótkoterminową.

Postęp w zakresie technologii komputerowej przyczynił się w ostatnim dziesięcioleciu do silnego rozwoju i praktycznego wykorzystania numerycznych technik obliczeniowych. Szczególnie intensywnie prowadzono badania w zakresie zagadnień tzw. sztucznej inteligencji (ang. Artificial Intelligence). Celem tej dziedziny informatyki jest symulacja, za pomocą maszyn cyfrowych, intelektualnych możliwości człowieka w procesie obliczeniowym [Tadeusiewicz, 1995]. Jednym z ważnych elementów rozwoju technik obliczeniowych z zakresu sztucznej inteligencji stały się sztuczne sieci neuronowe. Sieci neuronowe reprezentujące modele typu czarnej skrzynki wykorzystywane są często do modelowania systemów rzeczywistych, dla których postać transformaty nie jest znana, bądź system jest zbyt kompleksowy, aby można było uwzględnić wszystkie czynniki mające wpływ na jego stan. Z powodów, o których mowa, sieci neuronowe znalazły swoje zastosowanie również w hydrologii do opisu funkcjonowania naturalnych systemów wodnych [Jayawardena, 1995]. W ramach niniejszego opracowania przedstawiono możliwości zastosowania sztucznych sieci neuronowych do określenia związków czasowo – przestrzennych pomiędzy natężeniami/ stanami przepływu obserwowanymi w przekrojach wodowskazowych systemu rzeczego. Modele sztucznych sieci neuronowych (SSN), jak również modele typu ARMA XARMA mogą być zastosowane do symulacji prognozy natężeń przepływu w czasie rzeczywistym w sposób bezpośredni lub pośrednio, jako modele hybrydowe w połączeniu z modelami hydrodynamicznymi typu „szarej skrzynki”.

## 2. Modele stosowane w prognozie przepływów w czasie rzeczywistym

Zarówno prognozy średnioterminowe, jak i krótkoterminowe mogą ulegać modyfikacji w czasie operacyjnym, gdy dysponujemy na bieżąco pomiarami danych wejściowych do modeli

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl

prognostycznych takimi, jak: wysokości opadu mierzone na stacjach monitoringu w obrębie zlewni, stany wód lub natężenia przepływu mierzone w przekrojach wodowskazowych sieci rzecznej. W prognozach hydrologicznych w czasie rzeczywistym stosowane są zarówno modele opisujące relację opad – odpływ, jak i modele odwzorowują relację pomiędzy zmiennymi przepływu (stan wody lub natężenie przepływu) występującymi w górnych i dolnych przekrojach sieci rzecznej.

Ze względu na częstą konieczność przeprowadzenia procedury aktualizacji modeli w czasie operacyjnym, nie każdy model hydrologiczny wywodzący się ze wspomnianych dwóch grup modeli będzie spełniał wymogi aplikacji w czasie rzeczywistym, można tutaj wymienić takie kryteria kwalifikujące jak:

- wymagania, co do zakresu i jakości danych wejściowych do modelu,
- spełnienie założeń stosowalności modelu,
- zdolność do skutecznego odwzorowania przypadków ekstremalnych,
- łatwość kalibracji i aktualizacji.

Ze względu na wymienione cechy, można wyróżnić następujące grupy modeli, które znajdują zastosowanie w prognozach przepływów w czasie rzeczywistym:

- modele empiryczne (np. związki wodowskazowe, modele typu SSN),
- modele opad – odpływ typu „czarnej skrzynki” (np. modele typu SSN),
- hydrologiczne modele propagacji fali,
- modele hydrodynamiczne.

W obrębie modeli opad – odpływ typu czarnej skrzynki często stosowane są w prognozach operacyjnych modele tzw. funkcji transformującej (ang. Transfer Function). Zgodnie z tym podejściem przepływ w chwili  $t$  (w jednostkach dyskretyzacji czasu) jest skorelowany z przepływem wcześniejszym poprzez  $m$  parametrów oraz z wielkością opadu poprzez  $n$  parametrów i czas opóźnienia (czas, po którym opad wywołuje odpływ w danym przekroju koryta rzeki). Liniowa funkcja transformująca dla ustalonego przekroju koryta rzeki ma postać:

$$(1) \quad Q_t = a_1 Q_{t-1} + a_2 Q_{t-2} + \dots + a_m Q_{t-m} + b_1 R_{t-\tau} + b_2 R_{t-1-\tau} + \dots + b_n R_{t-n-\tau}$$

gdzie:

- $Q_t$  – natężenie przepływu w chwili  $t$
- $R_t$  – natężenie opadu w chwili  $t$
- $a_i$  – parametry przepływu
- $b_i$  – parametry opadu
- czas opóźnienia

Parametry powyższej funkcji mogą być interpretowane na podstawie cech dynamicznych zlewni (takich jak proporcja pomiędzy wielkością szybkiego spływu powierzchniowego a wielkością wolnego odpływu podpowierzchniowego). Interpretacja ta uzyskiwana jest jednak w oparciu o wnioskowanie, a nie bezpośrednio na bazie związków fizycznych lub modeli konceptualnych [Young i Tomlin, 2000]. Przedstawioną funkcję (3) stosuje się najczęściej w wariacie tzw. fizycznie dopuszczalnej funkcji transformującej (PRTF) [Young, 2001] w celu uniknięcia problemów ze stabilnością wyników modelowania. Modyfikacja polega tutaj na

dołączeniu ograniczeń na strukturę i zakres wartości parametrów modelu, tak aby uzyskać dodatnią, nie oscylującą wartość na wyjściu z modelu. W prognozach operacyjnych stosowana jest również nieliniowa postać funkcji transformującej w połączeniu z liniową funkcją oddzielającą opad całkowity od opadu efektywnego. Funkcja przejścia typu opad – odpływ ma istotne zalety z punktu widzenia możliwości łatwej aktualizacji jej stanu i wartości parametrów w czasie rzeczywistym.

### 3. Wprowadzenie teoretyczne do budowy i działania sztucznych sieci neuronowych typu „backpropagation”

Podstawowe właściwości sztucznych sieci neuronowych

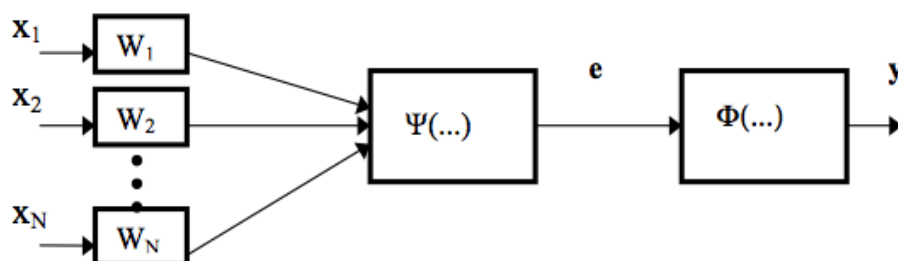
Sztuczne sieci neuronowe (SSN) stanowią skuteczne narzędzie wykorzystywane w wielu dziedzinach techniki. Idea sztucznych sieci neuronowych wywodzi się z badań neurofizjologicznych prowadzonych nad strukturą i działaniem sieci komórek nerwowych budujących ludzki mózg. Pomimo ciągłego doskonalenia technik SSN wykazują one jedynie niewielkie podobieństwo w działaniu i budowie w porównaniu z rzeczywistymi sieciami komórek nerwowych. Można jednak wykazać, że relatywnie proste struktury SSN są zdolne do realizacji praktycznych zadań z zakresu automatyki, statystyki czy optymalizacji. Do najczęstszych zastosowań należą:

- aproksymacja funkcji transformującej sygnały wejściowe na wyjściowe,
- prognozowanie przebiegów sygnałów wyjściowych analizowanego systemu,
- klasyfikacja danych wejściowych, poprzez przypisanie ich do zdefiniowanych wcześniej klas, na wyjściach z sieci pojawiają się wówczas odpowiedzi binarne 0 lub 1, które odpowiadają wykluczeniu lub przynależności do danej klasy,
- autoasocjacja polegająca na uzupełnianiu i korygowaniu zestawu niekompletnych danych lub danych obciążonych błędami (R. Tadeusiewicz, 1993).

Do chwili obecnej opracowano wiele różnych typów sztucznych sieci neuronowych takich, jak: sieci wielowarstwowe typu *backpropagation*, sieci Kohonena, Hopfielda, RBF i inne.

Sieci wielowarstwowe

Podstawowym elementem składowym wielowarstwowych sieci neuronowych jest nieliniowy neuron, którego schemat budowy został przedstawiony na rys. 1.



Rys. 1. Schemat budowy nieliniowego neuronu.

Sposób funkcjonowania neuronu opisany jest przez wartość wektora współczynników wagowych  $\langle w_1, w_2, \dots, w_N \rangle^T$  oraz przez postać dwóch funkcji:  $\sigma$ . Funkcja spełnia rolę agregatu sygnałów wejściowych, dając na wyjściu sygnał  $e$ , stanowiący łączne pobudzenie neuronu. Spośród wielu możliwych postaci realizacji funkcji, najczęściej wybierana jest funkcja, wyznaczająca iloczyn skalarny wektora reprezentującego sygnał wejściowy  $\langle x_1, x_2, \dots, x_N \rangle^T$  oraz wektora współczynników wagowych  $\langle w_1, w_2, \dots, w_N \rangle^T$ . Wartość sygnału  $e$  wyznaczana jest ze wzoru:

( 2 )

$$e = \sum_{i=1}^N w_i \cdot x_i + \Theta$$

Ważona suma sygnałów wejściowych jest często dodatkowo uzupełniana o stały składnik (ang. bias). Z kolei funkcja jest funkcją nieliniową pełniącą rolę przełącznika pomiędzy dwoma skrajnymi poziomami sygnału w rezultacie czego generowany jest sygnał wyjściowy  $y$ . Najczęściej stosowaną do celów praktycznych jest funkcja sigmoidalna, o postaci:

$$y = \frac{1}{1 + \exp(-x)}$$

( 3 )

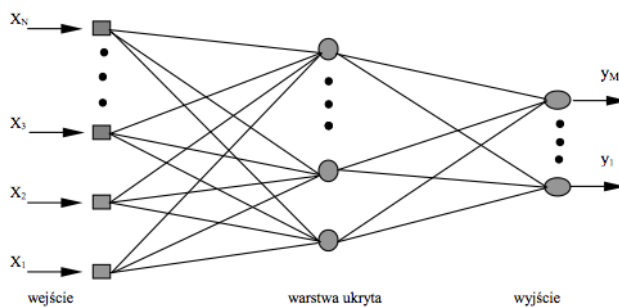
Pochodna rozważanej funkcji wyraża się wzorem:

$$y' = y \cdot (1 - y)$$

( 4 )

Omawiane nieliniowe neurony organizowane są w warstwy. Sieć neuronowa o najprostszym typie struktury może posiadać tylko jedną warstwę, jednakże sieć taka może być wykorzystywana jedynie do realizacji bardzo ograniczonej klasy zadań, sprowadzających się do liniowej dyskryminacji przestrzeni sygnałów wejściowych. Natomiast zastosowanie nieliniowej sieci neuronowej, o co najmniej trzech warstwach, pozwala na realizację, w dziedzinie sygnałów wejściowych, odwzorowań nieliniowych dowolnego typu. Przykładowa sieć neuronowa, zbudowana z trzech warstw nieliniowych neuronów, została przedstawiona na rys. 2.





Rys. 2. Struktura trójwarstwowej sieci neuronowej

Spośród warstw sieci można wyróżnić tzw. warstwę wejściową, do której podawane są, za pośrednictwem buforów, sygnały wejściowe. Następnie jedną lub więcej warstw ukrytych oraz warstwę wyjściową, której neurony wytwarzają wektor sygnałów wyjściowych  $\langle y_1, y_2, \dots, y_M \rangle^T$ , stanowiący odpowiedź sieci na wektor sygnałów wejściowych  $\langle x_1, x_2, \dots, x_N \rangle^T$ . Ponieważ rozważane są tylko sieci bez sprzężeń zwrotnych, propagacja sygnałów przebiega od warstw niższych ku warstwom wyższym. Przy czym wyjście każdego neuronu warstwy niższej łączone jest z odpowiednim wejściem każdego z neuronów warstwy wyższej [Tadeusiewicz, 1993].

#### Proces trenowania sieci

Aby sieć neuronowa stała się zdolna do prawidłowego rozwiązywania postawionych przed nią zadań, musi zostać na wstępie poddana procesowi trenowania, zwanego również procesem uczenia sieci. W stosowanym, w przypadku sieci wielowarstwowych, wariantcie trenowania sieci, zwanym uczeniem z nauczycielem, programista dysponuje zbiorem sygnałów wejściowych  $X$ , przy czym z każdym wektorem wejściowym  $\langle x_1, x_2, \dots, x_N \rangle^T$  skojarzony jest wektor  $\langle z_1, z_2, \dots, z_M \rangle^T$  należący do zbioru  $Z$  znanych *a priori* odpowiedzi, jakich sieć powinna udzielić na prezentowane jej elementy, należące do ciągu uczącego. Ponieważ na początku procesu trenowania współczynniki wagowe  $w[i, j, k]$  (gdzie  $i$  oznacza numer warstwy sieci,  $j$  numer neuronu w warstwie oraz  $k$  numer wejścia neuronu) posiadają wybrane w sposób losowy wartości, sieć podczas prezentacji kolejnych elementów ciągu uczącego będzie generować błędy. Proces uczenia sieci jest procesem iteracyjnym, mającym za zadanie minimalizację wartości błędów popełnianych przez poszczególne neurony.

W przypadku warstwy wyjściowej błędy popełniane przez poszczególne neurony  $[i, j]$  mogą zostać wyliczone z następującego wzoru:

$$\delta[i, j] = z[i, j] - y[i, j]$$

( 5 )

Po wyznaczeniu wartości błędu, zmodyfikowane wartości współczynników wagowych  $w[i, j, k]$  mogą zostać wyliczone zgodnie ze wzorem:

$$w^{(n+1)}[i, j, k] = w^{(n)}[i, j, k] + \eta_1 \cdot \Delta w[i, j, k] + \eta_2 \cdot (w^{(n)}[i, j, k] - w^{(n-1)}[i, j, k])$$

( 6 )

gdzie:

$$\Delta w[i, j, k] = \delta[i, j] \cdot \frac{d\Phi(e)[i, j]}{de[i, j]} \cdot y[i, j]$$

( 7 )

We wzorze (33)  $n$  określa kolejny krok procesu uczenia, a współczynniki  $\eta_1$  i  $\eta_2$  odpowiedzialne są za szybkość procesu uczenia. Dla potrzeb procesu uczenia konieczne jest także wyznaczenie wartości pochodnej funkcji po zmiennej  $e$ . Trzeci składnik wzoru (33) nosi nazwę momentum i może być stosowany opcjonalnie. Jednakże jego występowanie posiada często duży wpływ na proces uczenia, gdyż zapobiega on utknięciu procesu treningu sieci w minimach lokalnych funkcji błędu [Tadeusiewicz, 1993].

W przypadku warstw ukrytych i warstwy wejściowej sieci programista nie dysponuje znanymi a priori wartościami prawidłowych odpowiedzi poszczególnych neuronów. Zatem wartości błędów popełnianych przez neurony warstwy ukrytej nie mogą zostać wyznaczone za pomocą wzoru (32). W tym celu stosowany jest tzw. algorytm wstecznej propagacji błędów (ang. backpropagation), gdzie błąd popełniany przez dany neuron warstwy ukrytej wyznaczany jest jako ważona suma błędów popełnianych przez każdy z neuronów warstwy wyższej. Zatem wartości błędów popełnionych przez neurony warstwy  $i+1$  mnożone są przez odpowiednie współczynniki wagowe  $w[i+1, j, k]$  i sumowane zgodnie ze wzorem:

( 8 )

$$\delta[i, k] = \sum_{j=0}^{N_i} w[i+1, j, k] \cdot \delta[i+1, j]$$

gdzie:

$N_i$  – liczba neuronów w  $i$ -tej warstwie

Sposób rzutowania błędów, popełnionych przez neurony warstwy  $i+1$  do neuronu  $k$ , należącego do warstwy  $i$ , został zobrazowany na rys. 4. Stosując algorytm wstecznej propagacji błędów konsekwentnie do coraz niższych warstw sieci, wyznaczone zostają wartości błędów popełnianych przez wszystkie neurony. W następnych etapach uczenie neuronów warstw ukrytych przebiega zgodnie z formułą wyrażoną wzorem (33).

Proces treningu wielowarstwowej sieci neuronowej wymaga zwykle wykonania dużej liczby iteracji, podczas których sygnały wejściowe należące do ciągu uczącego podawane są na wejście sieci w losowo ustalonej kolejności. Cykl obliczeń, po którym wszystkie elementy ciągu uczącego zostaną podane na wejścia sieci nazywany jest epoką.

Zakończenie procesu uczenia



Istnieje wiele kryteriów pozwalających na zakończenie obliczeń optymalizacyjnych. Kryteria te mogą być budowane na podstawie zbioru uczącego lub tzw. zbioru walidacyjnego. Ten ostatni zawiera znane *a priori* sygnały wejścia–wyjścia, które nie były dołączone do zbioru uczącego a więc nie brały bezpośrednio udziału w procesie optymalizacji wag sieci, a jego rola polega na bieżącym (w trakcie procesu uczenia) monitorowaniu wielkości błędu generowanego przez sieć [Bishop, 1994]. Najczęściej stosowaną miarą błędu jest średni błąd kwadratowy (Root Mean Square) o postaci:

$$RMS = \sqrt{\frac{\sum_{i=1}^K \sum_{j=1}^M (z_j - y_j)^2}{K}}$$

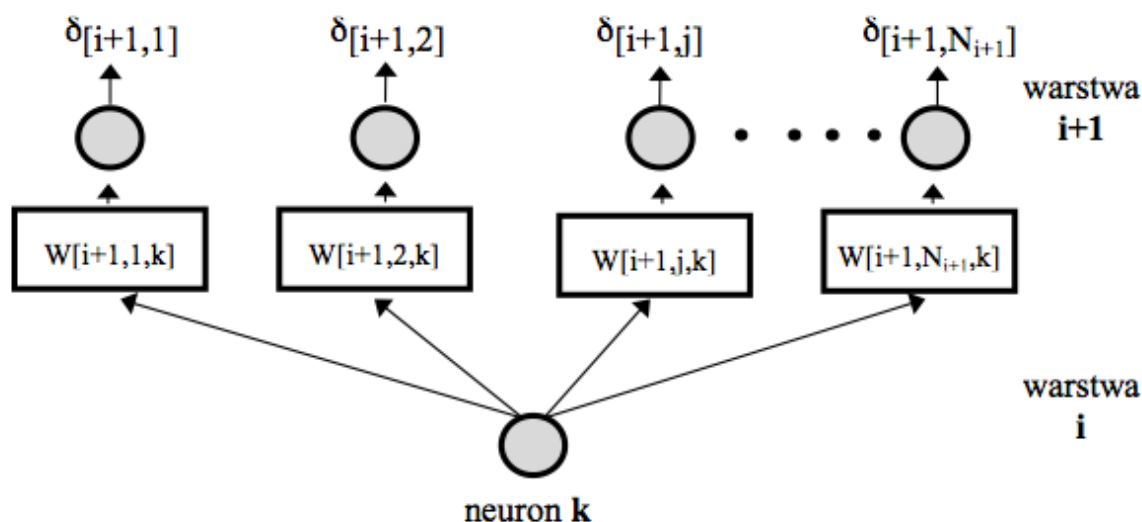
(9)

gdzie:

$\langle z_1, z_2, \dots, z_M \rangle^T$  – znana odpowiedź systemu na wymuszenie  $\langle x_1, x_2, \dots, x_N \rangle^T$

$\langle y_1, y_2, \dots, y_M \rangle^T$  – odpowiedź sieci na wymuszenie  $\langle x_1, x_2, \dots, x_N \rangle^T$

W momencie, gdy wartości błędu RMS osiągnie minimum proces uczenia w następnym kroku jest zatrzymywany, a wagi sieci, które minimalizują RMS uznawane są za optymalne. Inne lub dodatkowe kryteria mogą opierać się na maksymalnej liczbie iteracji lub na dopuszczalnej wartości błędu RMS obliczanego na podstawie zbioru uczącego.



Rys. 3. Idea działania algorytmu wstecznej propagacji błędu

## 4. Zastosowania sztucznych sieci neuronowych w hydrologii

Sztuczne sieci neuronowe (SSN) są uznanym i użytecznym narzędziem stosowanym do modelowania kompleksowych nieliniowych systemów dynamicznych, dla których dysponujemy odpowiednio licznym (reprezentatywnym) zbiorem danych historycznych opisujących relację wejście – wyjście dla analizowanego systemu. Dane te w wielu przypadkach obciążone są błędem, co stanowi uzasadnienie poszukiwań nowych sposobów modelowania systemów naturalnych uwzględniających statystyczny charakter danych obserwowanych. SSN reprezentują grupę modeli typu „czarnej skrzynki”, na obecnym etapie wiedzy nie potrafimy podać interpretacji fizycznej struktury i współczynników wagowych sieci. Jest to jeden z powodów, dla którego zastosowanie SSN w hydrologii będzie ograniczone do przypadków, gdzie dysponujemy odpowiednio reprezentatywnym zbiorem danych historycznych. Prognozy krótkoterminowe

SSN stają się szczególnie użyteczne przy modelowaniu systemów zlewniowych w czasie rzeczywistym w sytuacji, gdy istnieje możliwość bieżącego monitorowania stanów systemu. SSN mogą wówczas adoptować się do aktualnych danych obserwowanych. Ta własność SSN może przynieść znaczące korzyści w czasie operacyjnej prognozy wezbrań [Openshaw i in., 1998]. Konwencjonalne modele stosowane w prognozie zdarzeń powodziowych bazują często na kompleksowym matematycznym opisie procesów fizycznych zachodzących w systemie. Modele te wymagają dokładnych specyfikacji parametrów oraz odpowiedniej jakości danych wejściowych. Niejednokrotnie istnieje potrzeba opracowania dodatkowej strategii korekty błędów obliczeń w czasie rzeczywistym szczególnie w czasie nieoczekiwanych i gwałtownych zmian w przebiegu wezbrań monitorowanych w systemie. SSN oferują realne perspektywy na tańszą, mniej wymagającą metodologię modelowania procesów powodziowych, które ze swej natury są bardzo złożone i wykazują cechy chaotyczne [Openshaw i in., 1998]. W sytuacji, gdy globalne zmiany w klimacie są przyczyną co raz częstszych ekstremalnych zdarzeń pogodowych, położenie nacisku na opracowanie „elastycznej” i jednocześnie skutecznej metodyki prognozy wezbrań staje się zadaniem priorytetowym.

Zastosowanie SSN w hydrologii nie oznacza rezygnacji z modeli matematycznych, które opisują w sposób szczegółowy procesy fizyczne zachodzące w danym systemie. Modele matematyczne na ogół nie wymagają tak licznego zbioru danych historycznych do kalibracji swoich parametrów, jaki jest wymagany do efektywnego wyniku procesu uczenia SSN. W przypadku systemów dobrze rozpoznanych pod względem ich szczegółowej budowy i procesów fizycznych w nich zachodzących, modele matematyczne mogą wykazywać większą skuteczność prognostyczną. Ponadto modele matematyczne są niezastąpione w obliczeniach symulacyjnych dostarczając i uzupełniając dane w sytuacji braku obserwacji wielkości wejścia – wyjścia dla danego systemu. Istnieje jeszcze jedna możliwość, która wydaje się szczególnie wartościowa z punktu widzenia praktycznych aplikacji modelowania systemów zlewniowych na potrzeby prognozy, jest to częściowa identyfikacja modeli SSN (ogólnie modeli statystycznych i konceptualnych) na podstawie interpretacji, poprzez model matematyczny, procesów fizycznych zachodzących w systemie. Współpraca ta może odbywać się na różnych płaszczyznach, np. znajomość podstaw fizycznych rządzących

procesami hydrologicznymi może być pomocna na etapie selekcji wektora danych wejściowych do SSN, przyczyniając się w ten sposób do zwiększenia skuteczności sieci.

Zastosowanie SSN w modelowaniu relacji opad – odpływ

Jednym z podstawowych problemów podejmowanym przez hydrologów i inżynierów jest modelowanie związku pomiędzy opadem a odpływem w analizowanej zlewni. Informacja na temat opadu i odpływu z nim związanego jest niezbędnym elementem poprawnego gospodarowania wodą w zbiornikach retencyjnych, a także stanowi podstawę niektórych projektów inżynierskich zabudowy hydrotechnicznej. Relacja opad – odpływ jest rozpoznawana jako silnie nieliniowa i kompleksowa. Na wielkość odpływu ze zlewni, poza natężeniem opadu, ma wpływ wiele innych czynników takich, jak: geomorfologia zlewni, początkowa wilgotność gleby, sposób użytkowania gleby, ewapotranspiracja, infiltracja, czas trwania opadu i inne. Tematykę modelowania relacji opad – odpływ za pomocą SSN podejmowało wielu badaczy. Jednym z pierwszych był Halff [1993], który opracował trójwarstwową SSN na podstawie obserwowanych hietogramów opadu i hydrogramów zarejestrowanych przez U.S. Geological Survey (USGS) w Bellvue w Waszyngtonie. Wyniki tej pracy stały się zachętą dla innych hydrologów do bardziej szczegółowych badań. Hjelmfelt i Wang [1993] opracowali sieć neuronową w oparciu o teorię hydrogramu jednostkowego. W tym celu autorzy dokonali pewnych modernizacji w klasycznej strukturze SSN wielowarstwowych. Aby sieć mogła realizować liniową superpozycję sygnału wejściowego, tj. rzędnych hietogramu opadu z rzędnymi hydrogramu jednostkowego, w dwóch pierwszych warstwach sieci występują jedynie pojedyncze połączenia pomiędzy  $i$ -tym neuronem warstwy wejściowej oraz  $i$ -tym neuronem drugiej warstwy. W warstwie ukrytej zastosowano pełne połączenie z pojedynczym neuronem wyjściowym. Ponadto zmodyfikowano klasyczną progową funkcję aktywacji tak, że na wyjściach neuronów warstwy ukrytej pojawiała się część wysokości opadu pomniejszona o wartość infiltracji. Rzędne hydrogramu odpływu obliczone były przez neuron wyjściowy w postaci superpozycji sygnałów generowanych w warstwie ukrytej. Autor wykazał większą dokładność obliczeniową opracowanej metody odwzorowania hydrogramu jednostkowego w porównaniu z klasyczną reprezentacją funkcji gama. Inne podejście do zagadnienia transformacji opad – odpływ z jednodniowym okresem dyskretyzacji zaprezentował Hsu i in. [1995]. W centrum jego uwagi było porównanie jakości i dokładności prognozy uzyskanej trójwarstwową SSN, statystycznym modelem liniowym ARMAX oraz modelem konceptualnym SAC – SMA. Wszystkie wymienione modele prowadzą do pewnego przeszacowania przepływów niskich. Pomimo tego SSN w przebadanych przypadkach generowały dokładniejsze wyniki obliczeń w porównaniu z analogicznymi uzyskanymi przez pozostałe modele. Fernando i Jayawardena [1996] zastosowali inny typ sztucznych sieci neuronowych opartych na tzw. funkcjach radialnych (Radial Basis Functions). Autorzy posłużyli się godzinowym opadem i odpływem ze zlewni o powierzchni 3.12 km<sup>2</sup> zarejestrowanym w czasie dużych wezbrań. Oprócz wspomnianej sieci RBF, opracowano dodatkowo trójwarstwową SSN i wykazano ich większą skuteczność w prognozie w porównaniu z modelem ARMAX.

Badania nad zastosowaniem SSN w hydrologii w Polsce nie znalazły jak dotąd tak dużego zainteresowania, jak miało to miejsce w krajach zachodnich. Spośród nielicznych prac można wymienić pracę dotyczącą modelowania relacji opad – stan dla zlewni Starej Rzeki z zastosowaniem sieci typu counter propagation [Ciszewski i Żelazny, 1996]. Na wejście do modelu podawany był stan początkowy w przekroju Łazy, suma dobowo opadu, średnia temperatura powietrza. Zastosowany tutaj model neuronowy w pierwszej fazie obliczeń dzielił przestrzeń zmiennych wejściowych na klasy, co jest efektem działania sieci Kohonena.

**Centrum Inteligentnych Systemów Informatycznych** Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl [isi@agh.edu.pl](mailto:isi@agh.edu.pl)

W następnym etapie sygnał wyjściowy z neuronów zwycięskich stanowił wejście do sieci Grossberga. Opracowany model był stosowany iteracyjnie, generując w ten sposób długoterminową prognozę stanów wód. Na podstawie uzyskanych wyników autorzy doszli do wniosku, że opracowany model neuronowy może skutecznie rywalizować z modelami tradycyjnymi.

Z kolei Robert Wójcik i Poul Torfs [1998] opracowali metodę selekcji zmiennych wejściowych do modelu w oparciu o algorytm genetyczny, którą zastosowali do modelu prognostycznego stanów wody na rzece Ren. Podstawą modelowania były dane opisujące stany w przekrojach wodowskazowych rzeki Ren i jej dopływów powyżej przekroju Lobith oraz dobowe sumy opadów. Spośród 155 możliwych danych wejściowych na drodze optymalizacji wyłoniono 13 zmiennych. Analiza korelacji liniowej tych zmiennych ze zmienną wyjściową (stanem w przekroju Lobith z dwudniowym opóźnieniem) wykazała dominujące znaczenie zmiennych wejściowych opisujących stany wód, podczas gdy zmienne opadowe odgrywają marginalną rolę. W efekcie końcowym opracowano czterowarstwową sieć neuronową, która dla zbioru testowego z okresu 3 lat generowała wyniki prognozy z dwu procentowym błędem.

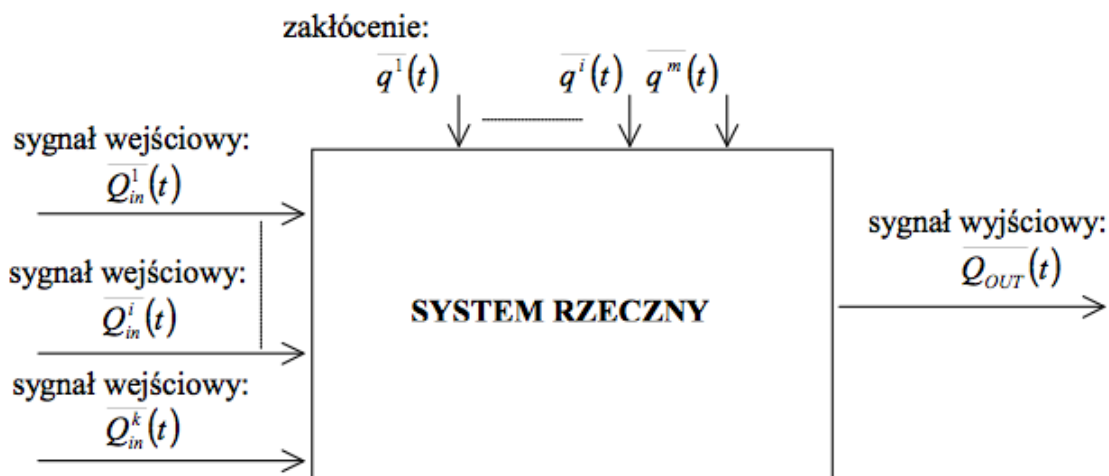
## 5. Proponowana koncepcja krótkoterminowej prognozy przepływów/stanów w czasie rzeczywistym z zastosowaniem SSN

Celem prognoz jest oszacowanie dyskretnych wartości natężeń przepływu pojawiających się w przekroju wyjściowym (zamykającym zlewnię) z przyjętym czasem uprzedzenia  $T_U$ . Sposób realizacji tego celu wygodnie jest przedstawić odwołując się do pojęć z zakresu teorii systemów. Teoria systemów dostarcza wiele metod analitycznych, które często wykorzystywane są przy modelowaniu zjawisk hydrologicznych.

**System** definiowany jest jako zbiór obiektów wyróżnionych przez badacza wraz z relacjami istniejącymi pomiędzy tymi obiektami, a także pomiędzy ich własnościami [Ozga–Zielińska, 1994]. Zbiór obiektów nienależących do systemu, których własności oddziałują na ten system nazywamy otoczeniem systemu. **Wejściami do systemu** nazywane są miejsca oddziaływania otoczenia na system, natomiast **wyjścia z systemu** to miejsca oddziaływania systemu na otoczenie. Nie wszystkie wejścia występujące w rzeczywistym systemie podlegają odwzorowaniu na jego modelu. Wejścia, które z różnych przyczyn są pomijane, a które wywierają wpływ na działanie systemu, określane są jako **zakłócenia wejścia**.

Na potrzeby analiz przeprowadzonych w ramach niniejszego opracowania użyto pojęcia systemu rzecznego, rozumianego, jako fragment sieci rzecznej obejmujący odcinek ciekę głównego oraz jego dopływy skupione, na których znajdują się czynne posterunki wodowskazowe. Zasilanie systemu rzecznego schematycznie przedstawiono na rys. 4. W systemie tym możemy wyróżnić **przekroje wejściowe** zlokalizowane na odcinku ciekę głównego oraz na jego monitorowanych dopływach powyżej przekroju zamykającego system zlewniowy, który nazywany jest również **przekrojem wyjściowym**. Sygnałem wejściowym (rys. 4) dla tak zdefiniowanego systemu są serie czasowe dopływów obserwowanych (wyrażone w postaci średnich natężeń przepływu) w przekrojach wejściowych. Sygnałem wyjściowym (rys. 4) jest seria czasowa odpływów obserwowanych w przekroju wyjściowym (zwanym również przekrojem zamykającym zlewnię). System reprezentowany jest w tym wypadku poprzez

relację czasoprzestrzenną zachodzącą pomiędzy samymi przepływami bez uwzględnienia bezpośredniej informacji na temat wartości i rozkładu czasoprzestrzennego opadu. Podstawą fizyczną funkcjonowania zdefiniowanego powyżej systemu rzeczno jest zatem proces transformacji fal wezbraniowych pojawiających się w przekrojach wejściowych systemu. Sam proces transformacji fali nie tłumaczy w pełni relacji pomiędzy rozważanymi przepływami, ponieważ na odcinku ciekę głównego w obrębie wybranego systemu występuje niekontrolowany dopływ boczny ze zlewni własnej w postaci skupionej i rozłożonej. Ten właśnie dopływ, który nie jest bezpośrednio mierzony należy traktować, jako **zakłócenie wejścia**. W niniejszym opracowaniu stosowane będzie pojęcie **dopływu niekontrolowanego** rozumianego, jako suma dopływu rozłożonego i dopływów skupionych leżących w obrębie danego systemu, dla których nie prowadzi się pomiarów wodowskazowych.



Rys. 4. Schemat systemu rzeczno

Metoda prognozy odpływu ze zlewni opisana w ramach niniejszego opracowania została przeanalizowana pod kątem jej zastosowania w czasie rzeczywistym w trakcie przebiegu wezbrania w analizowanym systemie rzeczno. Termin „czas rzeczywisty” zaczerpnięty jest z teorii systemów, gdzie stosowany jest powszechnie w odniesieniu do sterowania systemem na bieżąco tzn. decyzje sterowania zmieniające się w czasie podejmowane są na podstawie aktualnych danych wejściowych do systemu. W odniesieniu do prognozy termin „czas rzeczywisty” będzie wskazywał na pojawienie się zmiennego w czasie ciągu danych prognozowanych, którego ewolucja uzależniona będzie od aktualnych stanów pojawiających się na wejściu do systemu [Young, 2002].



Wobec przedstawionej wcześniej klasyfikacji prognoz hydrologicznych, metody prognozy opracowane w ramach niniejszej pracy opierające się na związkach czasowo-przestrzennych pomiędzy natężeniami przepływu, należy zaliczyć do **prognoz krótkoterminowych**.

Dla zdefiniowanego systemu rzeczno sformułowano następujące zadanie prognozy: poszukiwana jest funkcja  $F$  aproksymująca związek pomiędzy natężeniem przepływu w chwili  $t$  obserwowanym w przekroju wyjściowym systemu (rys. 12), a natężeniami przepływu obserwowanych w chwili  $t - T_U$  i chwilach wcześniejszych w przekrojach wejściowych i wyjściowym systemu rzeczno w następującej postaci:

$$\overline{Q}_{OUT}(t - T_U, n) = \left| Q_{OUT}(t - T_U), Q_{OUT}(t - T_U - 1 \cdot \Delta t) \dots Q_{OUT}(t - T_U - n \cdot \Delta t) \right| \quad (10)$$

gdzie:

$$Q_{OUT}(t) = F \left( \overline{Q}_{OUT}(t - T_U, n), \overline{Q}_{in}^1(t - T_U, m(i)), \overline{Q}_{in}^2(t - T_U, m(i)), \dots, \overline{Q}_{in}^i(t - T_U, m(i)), \dots, \overline{Q}_{in}^k(t - T_U, m(i)) \right) \quad (11)$$

$$\overline{Q}_{in}^i(t - T_U, m(i)) = \left| Q_{in}^i(t - T_U), Q_{in}^i(t - T_U - 1 \cdot \Delta t) \dots Q_{in}^i(t - T_U - m(i) \cdot \Delta t) \right| \quad (12)$$

- natężenie przepływu w przekroju wyjściowym systemu rzeczno,
- natężenie przepływu w  $i$ -tym przekroju wejściowym systemu rzeczno,
- $\Delta t$  – czas pomiędzy kolejnymi pomiarami przepływów,
- $T_U$  – czas uprzedzenia
- $k$  – liczba wodowskazowych przekrojów wejściowych
- $n$  – liczba zmiennych wejściowych (w postaci serii natężeń przepływu dla przekroju wyjściowego),
- $m(i)$  – liczba zmiennych wejściowych (w postaci serii natężeń przepływu dla  $i$ -tego przekroju wejściowego).

Zasady określenia optymalnego okresu uprzedzenia

Okres uprzedzenia  $T_U$  definiowany jest jako okres czasu, po którym spodziewamy się realizacji postawionej prognozy. Wybór długość okresu uprzedzenia ściśle wiąże się z wielkością błędu prognozy, dlatego ważnym staje się sposób jego określania. Można wprowadzić pojęcie optymalnego okresu uprzedzenia  $T_{Uopt}$ , dla którego korelacja liniowa pomiędzy natężeniem przepływu obserwowanym w chwili  $t - T_{Uopt}$  w górnym przekroju wejściowym (którym przepływa największa ze składowych fal wezbraniowych), a natężeniem przepływu obserwowanym w przekroju wyjściowym w chwili  $t$ , jest najsilniejsza (współczynnik korelacji liniowej ma największą wartość). Model prognostyczny będzie

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl



najefektywniej wykorzystywał dane pomiarowe z górnych przekrojów wejściowych pod względem dokładności prognozy natężeń przepływu w przekroju wyjściowym, jeżeli przyjęty okres uprzedzenia będzie miał optymalną wartość. W praktyce przyjęty okres uprzedzenia nie zawsze będzie równy optymalnemu ze względu na graniczne dopuszczalne wartości błędów prognozy. Wartość wspomnianego wcześniej współczynnika korelacji pomiędzy przepływami nie zawsze jest wystarczająco duża. Główną przyczyną takich sytuacji są niekontrolowane dopływy boczne będące istotną przyczyną błędów prognozy. Określenie optymalnego okresu  $TU_{opt}$  daje podstawę do wyznaczenia, na ogół na drodze empirycznej, najdłuższego okresu uprzedzenia, dla którego błędy symulacji prognozy nie przekraczają dopuszczalnych wartości.

Optymalny okres uprzedzenia można utożsamić ze średnim czasem propagacji fali wezbraniowej z górnych przekrojów wejściowych. Czas propagacji fali definiowany jest często jako okres potrzebny do przejścia przepływu kulminacyjnego od przekroju wejściowego do przekroju zamykającego zlewnię. Ustalenie czasów propagacji fal dla poszczególnych przekrojów wodowskazowych (nawet na drodze empirycznej) na ogół nie jest zadaniem łatwym z powodu efektów nakładania się na siebie fal dopływowych z falą występującą w korycie głównym systemu rzecznej. W przypadku gdy dopływy boczne nie mają dużego wpływu na formowanie się fali wezbraniowej w przekroju zamykającym, czas propagacji fali w korycie głównym można utożsamić z różnicą pomiędzy czasem wystąpienia kulminacji w przekroju zamykającym a czasem kulminacji w górnym przekroju wodowskazowym cieku głównego. Powyższy sposób postępowania na ogół nie jest właściwy w przypadku określania czasu propagacji fal pochodzących z cieków dopływowych do cieku głównego. W przypadku dobrego rozpoznania geometrii koryta cieku głównego oraz współczynników oporu ruchu można stosować hydrodynamiczne modele symulacyjne pozwalające na oszacowanie czasu przebiegu fali.

Proponowane rodzaje modeli typu SSN

Proponowana jest ocena skuteczności krótkoterminowej prognozy natężeń przepływu/stanów w czasie rzeczywistym, obserwowanych w przekrojach zamykających analizowaną zlewnię różnicową rzeki. Cele ten może zostać osiągnięty poprzez opracowanie:

1. liniowych modeli typu SSN,
2. liniowych dwuwarstwowych sieci rekurencyjnych,
3. nieliniowych trójwarstwowych modeli typu SSN.

Zbiór danych uczących i testowych

Zbiór danych uczących ma postać hydrogramów historycznych zarejestrowanych (lub obliczonych za pomocą modeli matematycznych) w przekrojach wejściowych i wyjściowym systemu rzecznej. Zbiór uczący należy tak wybrać spośród dostępnych danych historycznych, aby był jak najbardziej reprezentatywny dla modelowanego systemu, tzn. obejmował jak najszerszy przedział zmienności sekwencji stanów pomierzonych. Na podstawie hydrogramów tworzone są rekordy serii czasowych stanów wielowymiarowej zmiennej wejściowej wraz z odpowiedzią systemu w postaci stanu prognozowanego z ustalonym czasem uprzedzenia reprezentującego zmienną wyjściową.

Zaleca się przygotowanie trzech zestawów danych historycznych:

- $Z_i$  reprezentującego przypadki, w których dane prognozowane są wynikiem interpolacji funkcji transformującej,
- $Z_e$  reprezentującego przypadki, w których dane prognozowane są wynikiem ekstrapolacji,

- $Z_{update}$  reprezentującego przypadki, w których część danych testowych jest dołączana do danych kalibracyjnych (uczących) w celu uaktualnienia parametrów modelu. W celu uzyskania lepszej reprezentatywności modeli SSN zaleca się ich kalibrację na odpowiednio liczny zbiór danych, dotyczy to głównie modeli nieliniowych. W przypadku modeli opracowanych na podstawie zbioru uczącego typu  $Z_e$ , można ocenić efektywność prognozy w przypadku ekstrapolacji zmiennej prognozowanej, kiedy jej wartość wykracza poza zakres zmienności danych historycznych użytych do kalibracji modelu. Dane historyczne, które nie były użyte do kalibracji modeli SSN, stanowią materiał do ich niezależnej weryfikacji.

Dobór wektora zmiennych wejściowych do modeli SSN

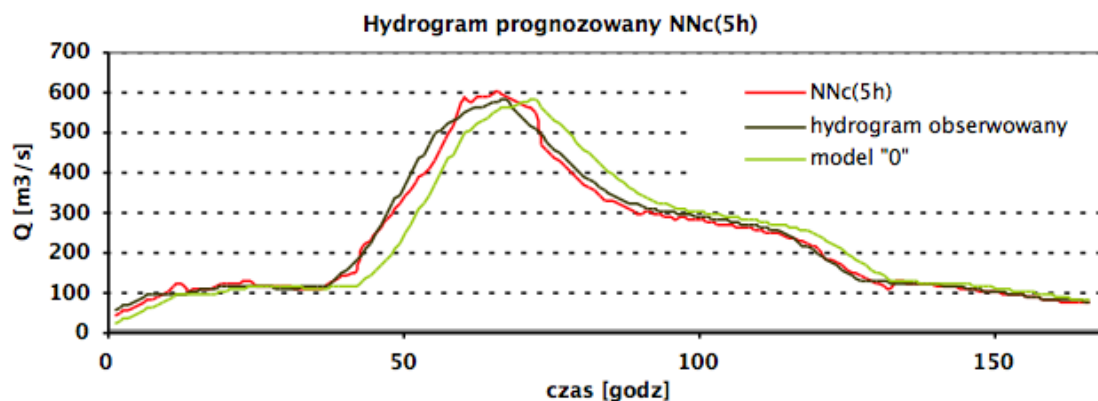
Celem opracowania modeli SSN jest aproksymacja relacji czasowo – przestrzennej pomiędzy natężeniami przepływu/stanami w wybranych przekrojach sieci rzecznej. Poszukiwanie zmiennych wejściowych w tym wypadku sprowadza się do ustalenia liczby wyrazów godzinowych serii natężeń przepływu/stanów mierzonych (obliczanych) w przekrojach wejściowych i wyjściowym systemu rzecznej, które mają istotny wpływ na dokładność odwzorowania prognozowanych wartości. Dobór wektora zmiennych wejściowych ściśle związany jest z przyjętym czasem uprzedzenia ( $T_U$ ), który określa prawostronną granicę poszukiwań zmiennych wejściowych w obrębie wspomnianych serii czasowych przepływów. Wektory zmiennych wejściowych do poszczególnych modeli prognostycznych określono w sposób uproszczony, w oparciu o analizę dokładności odwzorowania prognozowanych natężeń przepływu poprzez liniowy model regresyjny (dwuwarstwowa sieć neuronowa) stosowany dla różnych wariantów zmiennych wejściowych. Począwszy od jednoelementowych serii czasowych przepływów mierzonych w przekrojach kontrolnych systemu rzecznej badany był wpływ liczebności godzinowej serii natężeń przepływu na dokładność odwzorowania wartości prognozowanych natężeń przepływu poprzez liniowy model regresyjny. Przy wyborze zmiennych wejściowych należy kierować się zasadą doboru jak najmniejszej liczby zmiennych wejściowych, pozostawiając te, które w znaczącym stopniu są skorelowane ze zmienną wyjściową. Zbyt duża ilość zmiennych wejściowych zmniejsza zdolność do generalizacji zmiennej wyjścia oraz wydłuża czas kalibracji [Ignar i in. 1999]. Przykładowo, dla modelu prognozującego z 5-godzinnym czasem uprzedzenia, relacja czasowo – przestrzenna pomiędzy natężeniami przepływu może mieć postać:

$$Q_{out}(t) = F(Q_{out}(t-6), Q_{out}(t-5), Q_s(t-6), Q_s(t-5), Q_{in}(t-6), Q_{in}(t-5))$$

( 13 )

gdzie:

- t – czas w godzinach,
- $Q_{out}$  – natężenie przepływu prognozowane w przekroju wyjściowym rzeki,
- $Q_s$  – natężenie przepływu obserwowane w przekroju na dopływie,
- $Q_{in}$  – natężenie przepływu obserwowane w przekroju wejściowym rzeki,
- F – funkcja transformująca (definiująca model).



Rys. 5. Przykładowe wyniki testowe prognozy pięciogodzinnej natężenia przepływu w przekroju wyjściowym uzyskane w oparciu o model typu NN<sub>c</sub>(5h)

#### Struktura modeli SSN i wybór algorytmu uczącego

Strukturę modeli neuronowych determinuje ilość warstw ukrytych w sieci oraz liczba neuronów w każdej z warstw. W myśl zasady zwanej „brzytwą Occama” liczba parametrów modelu probabilistycznego powinna być jak najmniejsza pod warunkiem utrzymania zadowalającej dokładności odwzorowania zmiennej estymowanej w oparciu o dane testowe (służące do weryfikacji modelu).

Istnieje dowód, z którego wynika, że trójwarstwowa nieliniowa sieć neuronowa jest uniwersalnym aproksymatorem dowolnej rzeczywistej funkcji ciągłej [Tadeusiwicz, 1995]. Uwzględniając ten fakt zaleca się zastosowania takiej właśnie sieci w przypadkach modelowania nieliniowego.

Liczba neuronów w warstwie ukrytej może być różna w zależności od charakteru odwzorowania i podlega optymalizacji. Liczbę neuronów można dobrać w modelach hydrologicznych na drodze empirycznej, wykorzystując miarę średniego błędu kwadratowego jako funkcji kryterialnej. W wersji początkowej liczbę neuronów w warstwie ukrytej przyjęto jako równą wymiarowi wektora zmiennych wejściowych, a następnie stopniowo ją modyfikować, aż do momentu, kiedy średni błąd kwadratowy modelu liczony na podstawie zbioru testowego przestanie maleć.

#### Weryfikacja modeli

Najczęściej stosowaną miarą błędów, która często pełni również rolę funkcji kryterialnej w procedurach optymalizacji (np. w procesie uczenia SSN) jest błąd średni kwadratowy, definiowany w następujący sposób:

$$BSK = \sqrt{\frac{1}{n} \sum (Q_{obl} - Q_{obs})^2} \quad (14)$$

W celu konfrontacji błędów obliczeniowych z wartościami obserwowanymi natężeń przepływu, wprowadzono miarę błędu średniego kwadratowego względnego zdefiniowanego w następujący sposób:

$$BSKW = \sqrt{\frac{1}{n} \sum \left( \frac{Q_{obl} - Q_{obs}}{Q_{obs}} \right)^2} \quad (15a)$$

W niektórych przypadkach chwilowa wartość błędu prognozy jest znacząco większy od wartości średniej, z tego względu zdefiniowano następującą miarę błędu maksymalnego:

$$BMAX = \max \left| \frac{Q_{obl} - Q_{obs}}{Q_{obs}} \right| \quad (15b)$$

Ponieważ w prognozie przepływów wezbraniowych szczególny nacisk kładziemy na prawidłowe oszacowanie wartości kulminacyjnych natężeń przepływu, dlatego w prowadzono dodatkowo miarę błędu względnego kulminacji definiowanego w następujący sposób:

$$BWK = \left| \frac{Q_{obl}(t_{kul}) - Q_{obs}(t_{kul})}{Q_{obs}(t_{kul})} \right| \quad (16)$$

gdzie:

- n – liczba punktów hydrogramu
- $Q_{obl}$  – natężenie przepływu obliczone
- $Q_{obs}$  – natężenie przepływu obserwowane
- $t_{kul}$  – czas wystąpienia przepływu kulminacyjnego

W celu określenia wielkości asynchronizmu przepływu kulminacyjnego prognozowanego względem obserwowanego w czasie, zdefiniowano następującą miarę błędu:

$$BTK = T_k^p - T_k \quad (17)$$

gdzie:

- $T_k$  – czas wystąpienia przepływu kulminacyjnego obserwowanego w przekroju wyjściowym
- $T_k^p$  – czas wystąpienia przepływu kulminacyjnego prognozowanego w przekroju wyjściowym

W oparciu o wartości zdefiniowanych powyżej błędów będzie możliwa ocena skuteczności opracowanych modeli neuronowych pod względem praktycznych ich zastosowań do celów prognostycznych.

W celu określenia miary efektywności opracowanych modeli neuronowych, wprowadzono pojęcie „Model 0”, które odnosi się do modelu referencyjnego, z udziałem którego przeprowadzona będzie konfrontacja jakości wyników obliczeń uzyskanych w oparciu o opracowane modele neuronowe. „Model 0” można przedstawić jako procedurę prognozy, według której zakłada się, że wartość przepływu występującego w danej chwili w przekroju wyjściowym podsystemu rzeczowego pozostaje stała w pewnym przedziale czasu. Taki model jest najprostszym z możliwych modeli prognostycznych i często nie ma on żadnego znaczenia praktycznego, niemniej jednak może stanowić punkt odniesienia dla innych zaawansowanych technik prognostycznych. Błędy, jakie generuje tak określony „Model 0”, odpowiadają przyrostom wartości przepływów, jakie następują w określonym przedziale czasu (równym czasowi uprzedzenia TU). Zrozumiałym jest fakt, że wszelkie zaawansowane modele prognostyczne, aby mogły mieć znaczenie praktyczne, nie mogą generować błędów, których wartości byłyby większe od wspomnianych przyrostów wartości przepływów w określonym czasie uprzedzenia, a więc większe od wartości błędów generowanych przez zdefiniowany „Model 0”.

W celu oceny poprawy wyników obliczeń uzyskanych w oparciu o opracowane modele neuronowe względem wyników generowanych przez „Model 0”, zdefiniowano współczynnik efektywności:

$$WE = \left( \frac{BSKW("Model\ 0") - BSKW(Model\_testowany)}{BSKW("Model\ 0")} \right) * 100\% \quad (18)$$

gdzie:  $BSKW(nazwa\ modelu)$  – wartość błędu BSKW obliczona dla modelu którego nazwę podano w nawiasach

## 6. Koncepcja multi-modelu

Istnieje jeszcze jedna możliwość polegająca na zastosowaniu kilku modeli jednocześnie, w taki sposób, że informacja pochodząca od każdego z nich ma wpływ na wynik tej samej wartości prognozowanej. Takie podejście prowadzi do opracowania tzw. multi-modelu, który na ogół wykazuje większą dokładność prognostyczną niż każdy z modeli składowych z osobna [Abrahart i See, 2002]. Istnieje wiele różnych algorytmów służących do fuzji danych pochodzących z różnych źródeł, można tutaj wymienić: wnioskowanie Bayesa, teorię Dempstera-Shafera, systemy wnioskowania eksperckiego, sztuczne sieci neuronowe [Hall Llinas, 1997]. Oprócz wymienionych stosowane są również proste techniki, takie jak średnia arytmetyczna danych pochodzących z modeli składowych lub strategia stosowania modelu bardziej wiarygodnego, którego wybór odbywa się na podstawie wartości błędu prognozy liczonego w chwili poprzedniej. Zastosowanie średniej arytmetycznej jako sposobu łączenia danych, może przynieść pożądany efekt tylko wtedy, jeżeli modele składowe generują rezysta przeciwnych znaków. Bardziej ogólny efekt można uzyskać stosując teorię Bayesa



na podstawie której ustalane są wartości parametrów warunkowego rozkładu prawdopodobieństwa zmieniającego się w czasie. Jeżeli założymy, że dane obserwowane są w pełni wiarygodne (nie obciążone błędem), wówczas funkcja rozkładu prawdopodobieństwa warunkowego może być określana niezależnie dla każdej chwili czasowej. Jeżeli poczynimy dodatkowe założenia, że zmienna losowa będąca różnicą pomiędzy wartością modelowaną a obserwowaną podlega rozkładowi prawdopodobieństwa Gaussa [Mayback, 1979], to parametry rozkładu prawdopodobieństwa warunkowego  $f_{x(t)|z_1(t),z_2(t)}(x)$ , określonego na podstawie dwóch źródeł danych-  $z_1(t)$  i  $z_2(t)$  estymujących zmienną  $x(t)$ , można wyznaczyć wg. następujących wzorów:

$$\mu = \left[ \frac{\sigma_{z_2}^2}{(\sigma_{z_2}^2 + \sigma_{z_1}^2)} \right] z_1 + \left[ \frac{\sigma_{z_1}^2}{(\sigma_{z_2}^2 + \sigma_{z_1}^2)} \right] z_2 \quad (*)$$

$$1/\sigma^2 = 1/\sigma_{z_1}^2 + 1/\sigma_{z_2}^2 \quad (**)$$

gdzie:

- $z_1$  – wariancja zmiennej  $xz_1$
- $z_2$  – wariancja zmiennej  $xz_2$
- $z_1$  – wartość oczekiwana zmiennej  $xz_1$
- $z_2$  – wartość oczekiwana zmiennej  $xz_2$
- wariancja zmiennej  $xz_1, z_2$
- wartość oczekiwana zmiennej  $xz_1, z_2$

Mając w ten sposób scharakteryzowaną funkcję rozkładu prawdopodobieństwa warunkowego można określić najlepsze z możliwych oszacowań wartość  $x(t_1)$  na podstawie wyników symulacji tej wartości  $-z_1, z_2$ , które jest równe wartości oczekiwanej ( $x'(t_1) =$ ). Analizując równanie (19) można zauważyć, że w przypadku gdy wartość wariancji  $z_1$  jest równa wartości  $z_2$ , tzn. wtedy, gdy wiarygodność prognostyczna obu modeli składowych jest taka sama, wartość optymalnego oszacowania zmiennej  $x$  odpowiada średniej arytmetycznej  $z_1$  i  $z_2$ . Z drugiej strony, jeżeli  $z_1$  jest większa od wartości  $z_2$ , a więc wtedy, gdy niepewność związana z wartością (oszacowania  $x$ )  $z_1$  jest większa niż z wartością  $z_2$ , wówczas zgodnie z formułą średniej ważonej wartość  $x'(t_1)$  jest bliższa wartości  $z_2$ . Według równania (76), wartość wariancji jest mniejsza od wartości  $z_1$  nawet jeśli wartość  $z_2$  byłaby bardzo duża - wynika stąd wniosek, że nawet dane złej jakości dostarczają pewnej informacji na temat prognozowanej zmiennej  $x$ , dlatego powinny w sensie statystycznym przyczyniać się do zwiększenia precyzji oszacowania zmiennej  $x$  [Maybeck, 1979].

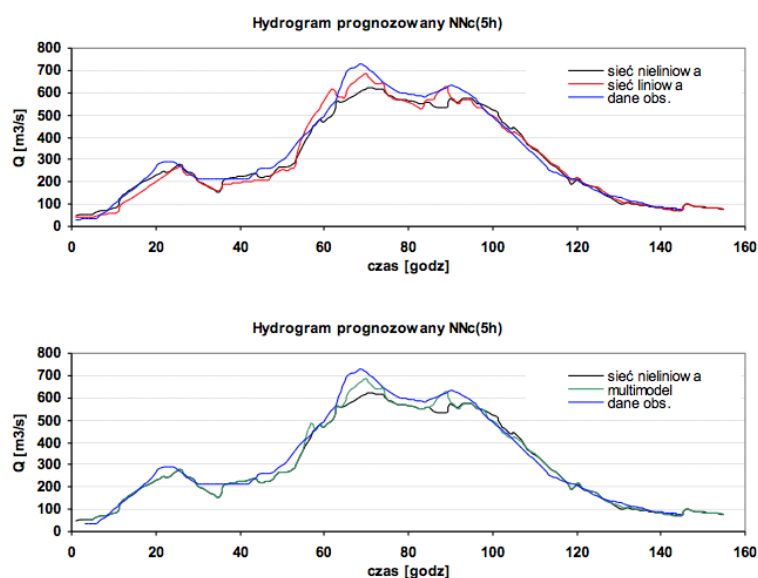
Przedstawiona powyżej teoria, nie zawsze przynosi praktyczne korzyści w prognozie przepływów w czasie rzeczywistym, szczególnie w przypadku wezbrań krótkotrwałych o dużej dynamice. Warunkiem istotnym, od którego zależy efektywności metody, jest spełnienie założenia o nieskorelowanym w czasie sygnale błędu prognozy (biały szum). W przypadkach, gdy mamy do czynienia z pewnym błędem skorelowanym, można wcześniej zastosować jedną z metod analizy serii czasowych (np. model ARMA).

Wyciąganie wniosków statystycznych na podstawie mało licznej próby losowej może być przyczyną ich negatywnej weryfikacji, dlatego w czasie prognozy przepływów w czasie rzeczywistym z procedurą aktualizacji modelu niezbędnym staje się określenie dostatecznie długiego okresu synoptycznego (czasu, w którym gromadzone są dane dotyczące wartości prognozowanych i obserwowanych, w oparciu o które często ustalane są parametry modelu aktualizującego). Skuteczna implementacja strategii aktualizacji modelu będzie silnie

**Centrum Inteligentnych Systemów Informatycznych** Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl [isi@agh.edu.pl](mailto:isi@agh.edu.pl)



uzależniona od czasu trwania wezbrania. W przypadku fali wezbraniowej, dla której czas koncentracji w przekroju kontrolnym rzeki jest krótki (sięgający kilkadziesiąt godzin), strategia prostej ekstrapolacji wartości błędu w pewnym przedziale czasu może okazać się skuteczniejsza od stochastycznych metod filtracji danych.



Rys. 6. Przykładowy wynik fuzji danych wyjściowych pochodzących z dwóch typów modeli SSN.

Również prosta strategia multimodelu polegająca na zastosowaniu do prognozy w chwili  $t$ , tego modelu, dla którego wartość błędu prognozy w chwili  $t-1$  (bądź wartość średniego błędu z pewnego przedziału czasu) jest mniejsza, może okazać się we wspomnianym przypadku skuteczna.

## 7. System informatyczny realizujący prognozę krótkoterminową za pomocą modeli typu SSN

System informatyczny niezbędny do realizacji postawionego zadania będzie obejmował następujące procedury:

- pomiaru bieżącego i akwizycji danych opisujących stan lub przepływ w przekrojach wodowskazowych systemu rzecznoego w czasie wezbrania,
- opracowania modeli nieliniowych SSN na podstawie danych historycznych typu  $Z_i$ ,
- opracowania modeli liniowych SSN na podstawie danych historycznych typu  $Z_e$ ,

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków  
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl

- wyboru i zastosowania jednego z opracowanych modeli w czasie rzeczywistym w zależności od tego, czy jest to przypadek „interpolacji”, czy „ekstrapolacji”,
- uaktualnienia modelu liniowego w momencie zarejestrowania odpowiednio liczebnej serii czasowej stanów pomierzonych,
- potencjalnego zastosowania multimodelu poprzez fuzję modelu nieliniowego i liniowego typu SSN.

## 6. Podsumowanie

W niniejszym opracowaniu przedstawiono koncepcję zastosowania sztucznych sieci neuronowych do prognozy krótkoterminowej przepływów/stanów wezbraniowych w przekroju wyjściowym zamykającym system zlewniowy rzeki. Zwrócono uwagę na aspekty fizyczne procesu propagacji fali w korycie rzeki, które mogą przyczynić się do poprawnego określenia tzw. czasu uprzedzenia prognozy (w teorii systemów nazywanym często czasem opóźnienia), jak również optymalnej identyfikacji zmiennych wejściowych (serii czasowych opóźnionych wartości stanu). Przedstawiono również ideę tzw. multimodelu, czyli koncepcji pozwalającej określić wartość prognozowaną w sposób bardziej wiarygodny na podstawie wyników prognozy uzyskiwanych przez dwa lub więcej modele w oparciu o charakterystyki ich statystycznej wiarygodności. Istota tej metody jest możliwość łączenia modeli typu SSN między sobą np. model liniowy z nieliniowym, jak również łączenie modeli SSN z hydrologicznymi modelami propagacji fali, których parametry podlegają optymalizacji statystycznej.

Praktyczna realizacja przedstawionych koncepcji dla fizycznego systemu monitoringu stanów w sieci rzecznej (ciek główny wraz z istotnymi dopływami skupionymi) wymaga informatycznego opracowania modeli typu SSN dostosowanego do fizycznego opisu procesu propagacji fali w korycie rzeczonym. Aby uzyskać narzędzie pozwalające na prognozę w trybie on-line, konieczne jest uwzględnienie przez programistów umożliwienie uaktualniania zbiorów danych uczących i testowych dla SSN, jak również danych wejściowych napływających z systemu monitoringu bieżących stanów wód. W przypadku realizacji multimodelu istnieje dodatkowo konieczność uwzględnienia w oprogramowaniu analizy statystycznej pozwalającej na określenie poziomu zaufania dla poszczególnych modeli składowych multimodelu.

### LITERATURA

- Bae Deg-Hyo, 1995, Georgakakos K., Nanda s., Operational forecasting with real-time databases, Journal of Hydraulic Engineering, Vol. 121, No.1, January.
- Bishop C.M., 1994, Neural networks and their applications, Review of Scientific Instruments, Vol. 65, No. 6, June.
- Campolo M., Andreussi P., Saldati A., 1999, River flood forecasting with a neural network model, Water Resources Research, Vol.35, No.4, April.
- Ciszewski S., Żelazny M., 1996, Modelowanie relacji opad stan wody dla zlewni Starej Rzeki przy pomocy sieci neuronowej CP (counter propagation) [w:] B. Więżik (red.), Modelowanie matematyczne w hydrologii, Politechnika Krakowska, PAN, Uniw. Bałt. w Uppsali, Kraków, 17-27.

- Crissman R., Chao-Lin Chiu, Weizu Yu, Kazumasa Mizumura, Ion Corbu, 1993, Uncertainties in flow modeling and forecasting for Niagara River, Journal of Hydraulic Engineering, Vol. 119, No.11, November.
- Cunge J.A., 1969, On the subject of the flood propagation computation method (Muskingum method), Journal of Hydraulic research, vol. 7, no 2, 205-230.
- Dibike Y. B., Abbot M. B., 1999, Application of artificial neural networks to the simulation of two dimensional flow, Journal of Hydraulic Research, Vol. 37, No.4.
- Dibike Y. B., Solomatine D., Abbot M. B, 1999, Encapsulation of numerical-hydraulic models in artificial neural network, Journal of Hydraulic Research, Vol. 37, pp.149-161.
- Georgakakos Konstantine P., Bras Rafael L., 1982, Real-time, statistically linearized, adaptive flood routing, Water Resources Research, v 18, n3, Jun, p 513-524.
- Halff, A. H., Halff, .H. M, Azmoodeh, M., 1993, Predicting runoff from rainfall using neural networks, Proceedings Engineering Hydrology, ASCE, New York, pp. 760-765.
- Hjelmfelt, A. T., and Wang, M., 1993, Artyficial neural networks as unit hydrograph applications, Proceedings Engineering Hydrology, ASCE, New York, pp. 754-759.
- Ignar S., Wójcik R., Torfs P., 1999, Selekcja zmiennych wejścia w przypadku modeli czarnej skrzynki, W ramach: Zastosowanie sztucznych sieci neuronowych w hydrologii, Warsztaty hydrologiczne, Inst. Geografii Uniwersytetu Jagiellońskiego, Kraków.
- Jayawardena A. W., Fernando D.A. K., 1996, Comparison of multi-layer perceptron and radial basis function network as tools for flood forecasting, Proceedings North America Water and Environment Conference, ASCE, New York, 457- 458.
- Jayawardena A.W., Fernando D.A.K., 1995, Artyficial neural networks in hydro-meteorological modelling, Developments in Neural Networks and Evolutionary Computing for Civil and Structural Engineering, 115-20.
- Lambor J., 1962, Metody prognoz hydrologicznych, Wydawnictwo komunikacji i łączności, Warszawa.
- Maybeck P.S., 1979, Stochastic models estimation, and control, volum1, Academic Press, New York.
- Ozga-Zielińska M., Brzeziński J., 1994, Hydrologia stosowana, Wydawnictwo Naukowe PWN Warszawa.
- Siuta T., Gajer M, 2002, Zastosowanie SSN do aproksymacji modelu matematycznego transformacji fali wezbraniowej w kanale otwartym, //W: Inżynieria Środowiska, AGH Uczelniane Wydawnictwa naukowo-dydaktyczne. Kraków, t. 7 z.2, s.187-201.
- Tadeusiewicz, R., 1995, Sieci neuronowe, Akademicka Oficyna Wydawnicza PLJ, Warszawa.