

mgr inż. Maciej Ławryńczuk
Instytut Automatyki i Informatyki Stosowanej
Politechnika Warszawska
ul. Nowowiejska 15/19, 00 - 665 Warszawa
Lawrynczuk@ia.pw.edu.pl

Wykorzystanie sieci neuronowych do modelowania silnie nieliniowych procesów*

Zagadnienie modelowania jest jednym z najważniejszych etapów projektowania układu automatycznej regulacji. Do opisu silnych nieliniowości można wykorzystać sztuczne sieci neuronowe. Modele tego typu są często bardziej dokładne niż tradycyjne struktury (np. wielomianowe), natomiast ich architektura jest stosunkowo prosta. W referacie omówiono dwa praktyczne przykłady wykorzystania modeli neuronowych. Model statyczny umożliwia obliczenie niemierzalnych strat produkcji w oparciu o dostępne wielkości pomiarowe (pomiar pośrednie). Omówiono także metody wykorzystania odwrotnego modelu procesu dynamicznego w algorytmie regulacji. W celu szybkiej kompensacji mierzzonego zakłócenia zastosowano sterowanie w układzie otwartym. Dodatkowy regulator ogranicza wpływ pozostałych zakłóceń i błędów modelowania.

Modelling of highly nonlinear processes by means of neural networks

Modelling is the crucial issue in developing an automatic control system. Artificial neural networks are able to work as universal approximators. Neural models can outperform traditional structures (e.g. polynomial) while their architecture is relatively simple. In this paper two practical examples of applying neural models are presented. The unmeasured composition of the bottom product in a distillation column is calculated by means of a steady-state model. The inferential model is fed only by accessible measurements. In the control schemes of a nonlinear dynamic process an inverse model is employed. In an open-loop structure the influence of a measured disturbance is reduced. Additional linear controller is used to compensate modelling errors and others disturbances.

1. WSTĘP

Zadaniem układu automatycznej regulacji jest przeciwdziałanie towarzyszącym każdemu procesowi zakłóceniom i możliwie wiernie nadażanie za zmianami sygnału zadanego. Podstawą opracowania skutecznego algorytmu jest możliwie wierny model procesu. W praktyce przemysłowej stosuje się zazwyczaj modele liniowe, podczas gdy rzeczywiste procesy charakteryzują się występowaniem silnych nieliniowości [1], [2], [6]. W celu poprawy jakości regulacji, a tym samym pracy całej instalacji, należy opracować bardziej dokładny model danego procesu. Warto również dodać, że coraz więcej pracujących aktualnie w przemyśle in-

* Praca została wykonana w ramach działalności statutowej Politechniki Warszawskiej (nr 504G)

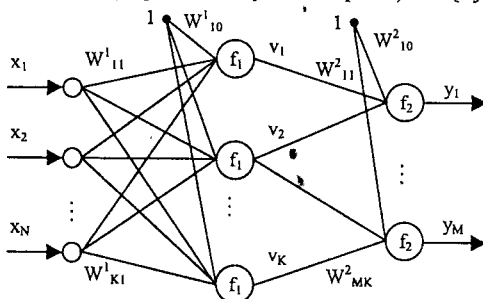
stalacji technologicznych wyposażonych jest w nowoczesne systemy sterowania, które umożliwiają implementację złożonych algorytmów regulacji.

Sformułowanie dokładnego modelu procesu w postaci analitycznej wymaga podania zależności opisujących poszczególne elementy aparatury. Jest to zazwyczaj trudne, natomiast modele tego typu (zwane fizykochemicznymi) nie nadają się do bezpośredniego wykorzystania w algorytmie regulacji. Wśród wielu metod opisu nieliniowych procesów na szczególną uwagę zasługują sztuczne sieci neuronowe [3], [7], [8], [9]. W ostatnich latach obserwuje się nie tylko gwałtowny rozwój badań teoretycznych, lecz także zastosowania praktyczne [1], [2], [4], [6]. Modele budowane w oparciu o sieci neuronowe umożliwiają aproksymację silnie nieliniowych zależności. W porównaniu z wieloma klasycznymi strukturami modeli (np. wielomianowymi) okazują się znacznie bardziej dokładne, natomiast ich zaletą jest stosunkowo prosta struktura.

W referacie omówiono dwa praktyczne przykłady wykorzystania modeli neuronowych. Pierwszy z modeli, statyczny, umożliwia obliczenie niemierzonej bezpośrednio wielkości strat produkcji. Skład chemiczny produktu ubocznego destylacji jest wyznaczany na podstawie dostępnych pomiarów, co jest realizacją koncepcji pomiarów pośrednich [5]. W dalszej części pracy przedstawiono metody wykorzystania odwrotnego modelu procesu w algorytmie regulacji. W celu szybkiej kompensacji mierzonych zakłóceń zastosowano otwarty układ sterowania. Aby przeciwdziałać nieuniknionym błędom modelowania, a także pozostałym zakłóceniom, wykorzystano dodatkowy regulator. W kontekście przykładów omówiono praktyczne problemy związane z modelowaniem neuronowym, a mianowicie wybór architektury sieci i argumentów modelu, dobór danych stosowanych podczas procesu uczenia oraz jego testowanie.

2. STRUKTURA SIECI NEURONOWEJ I JEJ UCZENIE

Na potrzeby badań teoretycznych opracowano wiele nowych struktur sieci neuronowych [3], [8], [9]. Niektóre z nich (np. sieci Elmana lub Jordana) mają dodatkowe sprzężenia zwrotne, co predysponuje je do modelowania obiektów dynamicznych. Z drugiej jednak strony, klasyczna sieć perceptronowa MLP (ang. Multi Layer Perceptron) wciąż jest szeroko stosowana.



Rys. 1. Dwuwarstwowa sieć neuronowa

Na rys. 1 pokazano strukturę dwuwarstwowej sieci jednokierunkowej, która posiada N węzłów wejściowych. Ilość neuronów w drugiej warstwie jest zdeterminowana przez ilość wyjść (M), natomiast ilość neuronów w warstwie ukrytej (K) jest parametrem decydującym o dokładności modelu. Sygnał wyjściowy i -tego neuronu ukrytego może być obliczony jako:

$$v_i = f_1 \left(\sum_{j=1}^N W_{ij}^1 x_j + W_{i0}^1 \right) = f_1 \left(\sum_{j=0}^N W_{ij}^1 x_j \right) \quad (1)$$

Jako funkcję aktywacji f_1 przyjmuje się zazwyczaj różniczkowalną funkcję sigmoidalną, przy czym może być ona w postaci bipolarnej lub unipolarnej:

$$f_1(u_1) = \operatorname{tg} h(\beta u_1) \text{ lub } f_1(u_1) = [1 + \exp(-\beta u_1)]^{-1} \quad (2)$$

Funkcja aktywacji drugiej warstwy (f_2) ma zazwyczaj postać liniową. Sygnał na k -tym wyjściu jest obliczany następująco:

$$y_k = f_2 \left(\sum_{i=0}^K W_{ki}^2 f_1 \left(\sum_{j=0}^N W_{ij}^1 x_j \right) \right) \quad (3)$$

Występujące w zależności (1) współczynniki W_{ij} noszą nazwę wag sieci. Ich wyznaczenie jest celem procesu uczenia sieci neuronowej. Jest on tożsamy z optymalizacją funkcji celu i ma na celu minimalizację błędów między wartościami żadanymi a aktualnie otrzymywanymi na wyjściu sieci. Opisany tryb uczenia nosi nazwę uczenia pod nadzorem lub uczenia z nauczycielem, ponieważ muszą być znane żadane wartości sygnałów wyjściowych oraz odpowiadające im wartości sygnałów wejściowych. Jeżeli szukany wektor wag zostanie oznaczony jako W , natomiast wektor wyjściowych wartości żadanych d , funkcja celu ma postać:

$$E(W) = \frac{1}{2} \sum_{j=1}^p \sum_{i=1}^m (y_i - d_i)^2 \quad (4)$$

symbol p oznacza licznosc zbioru uczacego.

Zalozenie ciaglosci funkcji aktywacji neuronow umozliwia wykorzystanie gradientowych metod optymalizacji. Rozwinięcie funkcje $E(W)$ w szereg Taylora jest nastepujace:

$$E(W+p) = E(W) + [g(W)]^T p + \frac{1}{2} p^T H(W) p + \dots \quad (5)$$

gdzie g jest wektorem gradientu, H macierza drugich pochodnych (hesjanem), natomiast p wektorem kierunkowym. Zwykle wykorzystuje się tylko trzy pierwsze składniki rozwinięcia. Różnice między poszczególnymi algorytmami polegają na odmiennych metodach wyznaczania kierunku p oraz kroku η . Ogólny algorytm uczenia można zapisać następująco:

1. Inicjalizacja wektora wag W (najczęściej w sposób losowy).
2. Obliczenie wektora gradientu g_k , macierzy hesjanu H_k lub jej odwrotności V_k .
3. Wyznaczenie kierunku minimalizacji p_k .
4. Określenie kroku η_k .
5. Aktualizacja wektora wag:

$$W_{k+1} = W_k + \eta_k p_k(k) \quad (6)$$

6. Jeżeli spełnione jest kryterium stopu następuje koniec algorytmu, w przeciwnym przypadku powrót do punktu 2.

Współczynnik η może być stały lub też może być zmieniany adaptacyjnie wraz z postępowaniem optymalizacji. Rozwiązaniem najlepszym, związanym jednak ze znacznym wzrostem złożoności obliczeniowej, jest wyznaczanie tej wartości na podstawie minimalizacji kierunkowej. Jeżeli w rozwinięciu funkcji celu zostaną uwzględnione jedynie dwa pierwsze wyrazy (przybliżenie liniowe) kierunek poszukiwań będzie przeciwny do gradientu:

$$p_k = -g_k(W_k) \quad (7)$$

co oznacza metodę największego spadku. Jej podstawową wadą jest niewykorzystywanie informacji o krzywiznie funkcji zawartej w hesjanie. Dlatego też znacznie bardziej skuteczne są algorytmy, które wykorzystują kwadratowe przybliżenie funkcji celu (5). Z warunku zerowania się gradientu otrzymuje się następujący związek:

$$p_k = -[H_k(W_k)]^{-1} g_k(W_k) \quad (8)$$

Wyznaczenie nowego kierunku poszukiwań wymaga więc nie tylko znajomości gradientu, lecz także hesjanu. W praktycznych implementacjach algorytmów optymalizacji wyznacza się przybliżenie hesjanu lub też przybliżenie jego odwrotności. Jednym z najczęściej stosowanych sposobów iteracyjnego wyznaczania odwrotności hesjanu jest metoda BFGS:

$$V_k = V_{k-1} + \left[1 + \frac{r_k^T V_{k-1} r_k}{s_k^T r_k} \right] \frac{s_k s_k^T}{s_k^T r_k} - \frac{s_k r_k^T V_{k-1} + V_{k-1} r_k s_k^T}{s_k^T r_k}, \quad s_k = W_k - W_{k-1}, \quad r_k = g_k - g_{k-1} \quad (9)$$

Metoda największego spadku charakteryzuje się zbieżnością liniową, podczas gdy metoda zmiennej metryki zbieżnością superliniową. Jej wadą jest jednak znaczna złożoność obliczeniowa.

3. MODEL STRAT PRODUKCJI

Podczas pracy wielu obiektów przemysłowych zachodzi konieczność wyznaczania wielkości, które nie są bezpośrednio mierzone. Przyczyny wykorzystywania pomiarów pośrednich są różne. Niektóre pomiary są kosztowne i stosunkowo kłopotliwe, podczas gdy wiadomo, że np. skład chemiczny można obliczyć na podstawie temperatury danej substancji. Konwencjonalne przyrządy mogą być precyzyjne, lecz ich wykorzystanie w regulacji jest kłopotliwe (np. ze względu na długi cykl pomiarowy). W warunkach przemysłowych dopuszcza się występowanie pewnych błędów pomiarów pośrednich, ponieważ nawet zniekształcona informacja dostarcza danych o pracy procesu. Dokładność pomiarów pośrednich jest określona przez jakość przygotowanego modelu danej wielkości.

Jako przykład ilustrujący wykorzystanie modeli neuronowych zostanie rozpatrzony model strat produkcji wysokociśnieniowej kolumny destylacyjnej [5]. Wielkość produkcji instalacji wynosi około 60 t/h, natomiast wielkość strumienia produktu ubocznego (wywaru) waha się w granicach od 10 do 12 t/h. Z punktu widzenia strat surowca (etylenu) jego zawartość w wywarze powinna być jak najmniejsza, lecz jest to związane ze zwiększonym zapotrzebowaniem na energię. Skład chemiczny wywaru jest wyznaczany na podstawie analizy laboratoryjnej, która wymaga pewnego nakładu pracy i wnosi istotne opóźnienie. Zastosowanie mechanizmu pomiarów pośrednich umożliwia więc bieżącą ocenę wielkości strat.

Korzystając z zestawu nieliniowych równań różniczkowych opracowano pełny model fizykochemiczny obiektu, jednak jego wykorzystanie w warunkach przemysłowych jest problematyczne. Zaproponowano więc następującą procedurę postępowania. Zweryfikowany pełny model procesu służy do badania wpływu zmian poszczególnych parametrów procesowych na interesującą wielkość wyjściową (skład wywaru). Z przeprowadzonych eksperymentów, a także z wiedzy ekspertów wynika, że skład x_w cieczy wyczerpanej można opisać następującą zależnością:

$$x_w = f(y_s, y_p, S, r) \quad (10)$$

Symbol y_s oznacza skład surowca, y_p skład fazy gazowej na tzw. półce kontrolnej kolumny destylacyjnej, S natężenie przepływu surowca, natomiast r jest współczynnikiem sterującym (stosunek przepływów). Wykorzystując pełny model procesu wykonano szereg symulacji, których celem było wyznaczenie wartości szukanej funkcji (10) dla różnych konfiguracji argumentów. Założono przy tym, że pozostałe parametry (ich wpływ na skład jest niewielki) miały wartość nominalną.

Zagadnienie opracowania modelu jest w istocie zadaniem aproksymacji funkcji (10) w przestrzeni \mathbb{R}^d . Wszystkie argumenty zostały przeskalowane do zakresu [0; 1]. Przygotowano trzy zbory danych, które służą do obliczania parametrów modeli i ich późniejszej weryfikacji. Dwa zbiory uczące (dane1, dane2) zawierały odpowiednio 2369 oraz 4742 punkty, natomiast zbiór testowy (dane3) 1730 punkty. Z uwagi na zmienną liczbę poszczególnych zbiorów jako wskaźnik jakości przyjęto znormalizowaną sumę kwadratów błędów (ang. Normalized

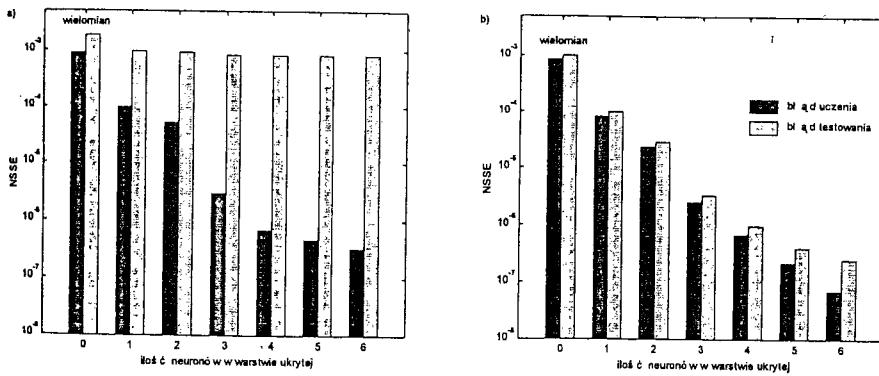
Sum of Squared Errors). Model ma jedno wyjście, a więc, na podstawie (4) otrzymuje się:

$$NSSE = \frac{1}{2N} \sum_{i=1}^p (x_{wi} - x_{wi}^{\text{mod}})^2 \quad (11)$$

W celu porównania z modelami neuronowymi poszukiwano również najlepszej aproksymacji zależności (10) w klasie wielomianów o stałych współczynnikach. Przyjęto model:

$$x_w^{\text{mod}} = a_1 y_s + a_2 y_p + a_3 S + a_4 r + a_5 y_s y_p + a_6 y_s^2 + a_7 r^2 + a_8 s^2 + a_9 r^3 \quad (12)$$

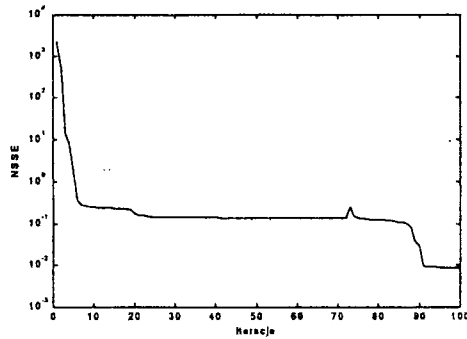
Podczas wyznaczania modeli neuronowych szczególną uwagę należy zwrócić na zdolności uogólniania sieci czyli generalizację danych wejściowych. Poprawny model jest w stanie podawać właściwe odpowiedzi nawet dla tych danych, które nie były bezpośrednio zawarte w zbiorze uczącym. Przygotowano dwie grupy modeli, przy czym różnica polegała na wykorzystaniu mniej lub bardziej liczego zbioru uczącego (dane1 lub dane2), natomiast do weryfikacji stosowano zawsze dane3. Parametrem zmiennym, mającym zasadniczy wpływ na dokładność modeli, jest ilość neuronów w warstwie ukrytej. Na rys. 2 graficznie przedstawiono otrzymane wskaźniki jakości w funkcji ilości neuronów ukrytych.



Rys. 2. Zależność błędów modeli w funkcji ilości neuronów ukrytych: a) ograniczony zbiór uczący, b) rozszerzony zbiór uczący

Wykorzystanie mniejszego zbioru danych (rys. 2a) nie pozwala na osiągnięcie przez model wymaganych zdolności generalizacji. Zwiększanie rozmiarów sieci prowadzi, co prawda, do ograniczenia błędu uczenia, natomiast błąd testowania pozostaje praktycznie stały. Zasadniczą poprawę przynosi rozszerzenie zbioru uczącego (rys 2b). Warto zwrócić uwagę, że nawet dla struktury złożonej z dwu neuronów odpowiednie błędy są o rząd wielkości mniejsze od błędów modelu wielomianowego.

Zastosowana metoda uczenia sieci neuronowej należy do grupy lokalnych metod optymalizacji. Znalezione minimum nie musi więc być wartością najmniejszą w sensie globalnym. W praktyce nie jest z góry znana wartość funkcji celu w optimum. Jeżeli po ukończeniu procesu uczenia uzyskany wynik nie jest zadowalający należy powtórzyć cały proces dla innych wartości początkowych wag lub też przyjąć inne parametry algorytmu. Niekiedy też łączy się klasyczne algorytmy uczenia sieci neuronowych z podejściami heurystycznymi. Jako przykłady metod tego typu można wymienić algorytmy genetyczne, strategie ewolucyjne czy też algorytm symulowanego wyżarzania [7], [9]. Podczas obliczeń rozważanego modelu strat produkcji zastosowano tzw. metodę *jog of weights*, która polega na losowej modyfikacji znalezionego rozwiązania i kontynuacji algorytmu, co często prowadzi do opuszczenia minimum lokalnego. Przykładowy przebieg procesu optymalizacji pokazano na rys. 3.



Rys. 3. Wpływ losowej zmiany wartości wag na proces uczenia sieci neuronowej

Po około 70 iteracjach algorytm doprowadził do minimum lokalnego. Obliczenia zostały przerwane, a wagi zmodyfikowane o pewien niewielki współczynnik losowy. Pozwoliło to na opuszczenie strefy przyciągania danego rozwiązania lokalnego. Po wykonaniu kolejnych 20 iteracji osiągnięto znacznie lepsze rozwiązanie. Wadą opisanego metody jest konieczność eksperymentalnego doboru wielkości zaburzenia wag: wartość zbyt mała nie daje pożądanego rezultatu, wartość zbyt duża może znacznie wydłużyć cały algorytm.

4. WYKORZYSTANIE MODELU NEURONOWEGO W REGULACJI

Modelowanie neuronowe obiektów dynamicznych można sformułować jako zadanie podobne do problemu aproksymacji funkcji statycznych, z tą tylko różnicą, że należy dodatkowo uwzględnić wartości odpowiednich sygnałów z poprzednich chwil czasowych. W przypadku modelu jednowymiarowego aktualna wartość wyjścia może być funkcją liniową dwu wartości sygnału wejściowego u oraz wyjściowego y (model ARX):

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + b_1 u(k-1) + b_2 u(k-2) \quad (13)$$

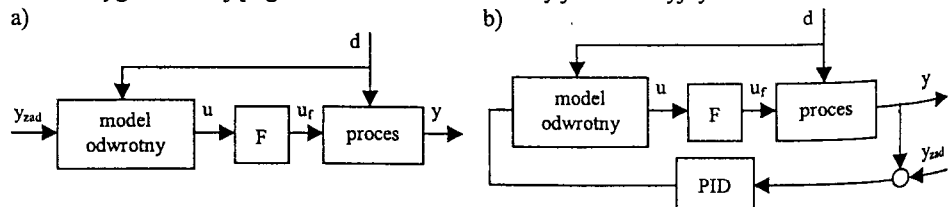
Wykorzystanie sieci neuronowej umożliwia modelowanie silnie nieliniowych obiektów. Model neuronowy, odpowiadający zależności (13), można opisać nieliniową funkcją f_0 :

$$y(k) = f_0(y(k-1), y(k-2), u(k-1), u(k-2)) \quad (14)$$

Typowym zadaniem regulacji przemysłowej jest kompensacja zakłóceń procesu. Zakłada się, że model obiektu (kolumny destylacyjnej [5]) o postaci (14) jest następujący:

$$y(k) = f(y(k-1), y(k-2), u(k-1), u(k-2), d(k-1)) \quad (15)$$

gdzie d oznacza zakłócenie. Jeżeli wielkość ta jest mierzona, można ją wykorzystać w algorytmie regulacji. Na rys. 4a pokazano strukturę sterowania w układzie otwartym. Zakłócenie d oddziałuje na proces oraz jednocześnie na jego model odwrotny. W rezultacie sygnał sterujący u zależy od wartości sygnału zadanego y_{zad} i zakłócenia. W celu eliminacji szkodliwych skoków sygnału sterującego u zastosowano dodatkowy jednoinercyjny filtr F .



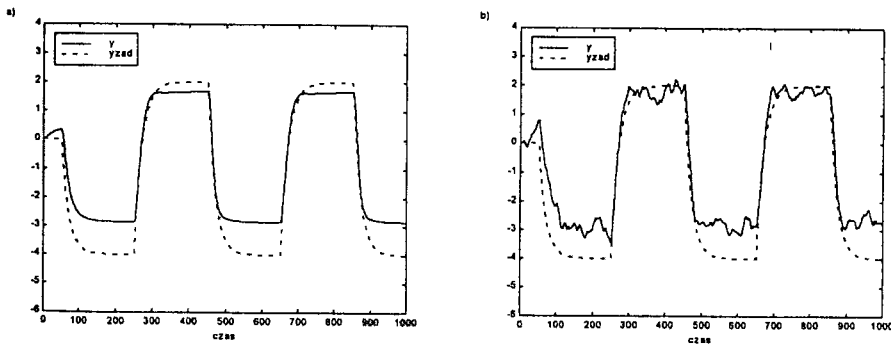
Rys. 4. Struktura sterowania: a) układ otwarty, b) układ zamknięty

Na podstawie zależności (15) można sformułować postać modelu odwrotnego procesu:

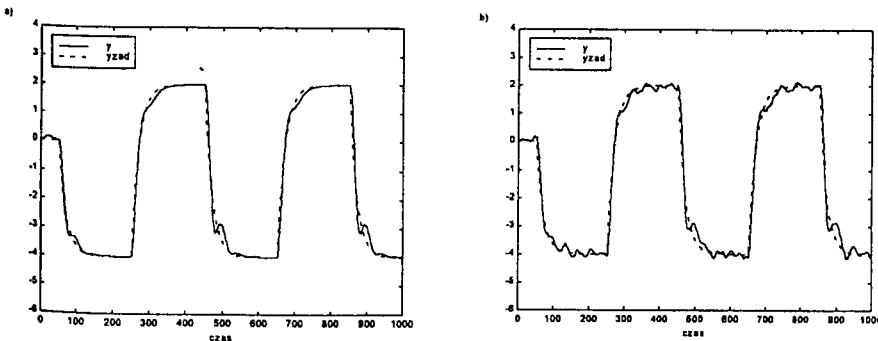
$$u(k) = g(y_{zad}(k-1), y_{zad}(k), y_{zad}(k+1), u(k-1), d(k)) \quad (16)$$

Model odwrotny jest realizowany przez sieć neuronową, która zawiera 10 neuronów ukrytych. Wyniki symulacji otwartego układu regulacji pokazano na rys. 5. Jeżeli brak jest jakichkolwiek zakłóceń procesu (sytuacja idealna) to niezgodność wyjścia obiektu z sygnałem wartości zadanej wynika z niedoskonałości modelu odwrotnego (rys. 5a). W warunkach przemysłowych parametry obiektu mogą ulegać zmianie w czasie. Wyniki symulacji w warunkach występowania zakłóceń pokazano na rys. 5b. Model odwrotny kompensuje (z ograniczoną dokładnością) zakłócenia mierzone, natomiast niemierzone nie są redukowane.

Do istniejącej struktury otwartej dołączono liniowy regulator PID, którego zadaniem jest modyfikacja sygnału wejściowego modelu odwrotnego (rys. 4b). Umożliwia to osiągnięcie zerowego uchybu ustalonego w stanie ustalonym (rys. 6a). Podobnie jak w strukturze otwartej, mierzone zakłócenia są redukowane przy pomocy modelu odwrotnego. Oprócz tego, w naturalny sposób regulator przeciwdziała wszystkim pozostałym czynnikom zakłócającym (rys. 6b).



Rys. 5. Praca otwartego układu regulacji: a) brak zakłóceń, b) obecność zakłóceń mierzonych oraz niemierzonych



Rys. 6. Praca zamkniętego układu regulacji: a) brak zakłóceń, b) obecność zakłóceń mierzonych oraz niemierzonych

Model odwrotny procesu dynamicznego można formułować jedynie do pewnej klasy obiektów minimalnofazowych. Jeżeli w procesie występuje opóźnienie, to w pierwszej kolejności należy określić jego model, a następnie obliczyć odwrotność tylko tej części, która nie zawiera opóźnienia.

5. PODSUMOWANIE

W referacie omówiono przykłady ilustrujące wykorzystanie sztucznych sieci neuronowych do modelowania silnie nieliniowych procesów. Podstawową zaletą struktur tego typu jest możliwość aproksymacji dowolnych odwzorowań nieliniowych. Umożliwia to więc rozwiązanie wielu trudnych problemów, przy których podejścia klasyczne (np. modele wielomianowe) nie dają wystarczająco dokładnych rezultatów. W ogólnym przypadku struktura sieci nie ogranicza ilości wejść i wyjść modelu. Uczenie i adaptacja parametrów modelu odbywa się na podstawie danych eksperymentalnych, a więc sieć neuronowa może również uwzględniać zmienność charakteru danego procesu w czasie. Z praktycznego punktu widzenia istotna jest również stosunkowo prosta struktura otrzymanych modeli.

Jako przykład omówiono statyczny model strat produkcji wysokociśnieniowego bloku destylacji. Sieci neuronowe mogą być również wykorzystywane w celu modelowania nieliniowych procesów dynamicznych. Przedstawiono dwie proste struktury sterowania z wykorzystaniem neuronowego odwrotnego modelu procesu. Sterowanie w układzie otwartym jest w rzeczywistych warunkach przemysłowych mało skuteczne. Dlatego też zastosowano dodatkowy regulator PID. W rezultacie układ sterowania umożliwia skutecznie kompensowanie wpływu mierzonych i niemierzonych zakłóceń oraz nadążanie za zmianami sygnału wartości zadanej.

6. PODZIĘKOWANIE

Autorem programu do symulacji wysokociśnieniowej kolumny destylacyjnej w stanach ustalonych jest dr Paweł Oracz (Uniwersytet Warszawski). Za jego udostępnienie składam tą drogą podziękowanie.

LITERATURA

- [1] R.Baratti, G.Vacca, A. Servida: *Neural network modeling of distillation columns*. Hydrocarbon Processing, no. 7, 1995, 35-38.
- [2] N. Bhat, T. J. McAvoy: *Use of neural nets for dynamic modeling and control of chemical processes*. Computers Chem. Engng., vol. 14, no. 4/5, 1990, 573-583.
- [3] K. J.Hunt, D.Sbarbaro, R.Zbikowski, P. J. Gawthrop : *Neural networks for control systems — A survey*. Automatica, vol. 28, no. 6, 1992, 1083-1112.
- [4] J. Keeler: *The Process Perfector: The next step in multivariable control and optimization*. Pavilion Technologies Inc. Technical Report, 1996.
- [5] M. Ławryńczuk, J. Pułaczewski: *Idea i przykłady wykorzystania pomiarów pośrednich w przemyśle*. Automation 99, 1999, 413-420.
- [6] J. C. MacMurray, D. M. Himmelblau: *Modeling and control of a packed distillation Column using artificial Neural Networks*. Computers Chem. Engng., vol. 19, no. 10, 1995, 1077-1088.
- [7] T. Masters: *Practical neural networks recipes in C++*. Academic Press, 1993.
- [8] K. S.Narendra, K. Pathasarathy: *Identification and control of dynamical systems using neural networks*. IEEE Transactions on Neural Networks. vol. 1, no. 1, 1990, 4-27.
- [9] S. Osowski: *Sieci neuronowe w ujęciu algorytmicznym*. WNT, Warszawa 1996.