

*Extreme Learning Machine, ELM,
modelowanie tarcia, regulacja adaptacyjna,
modelowanie rozmyte, silnik liniowy*

Marcin JASTRZĘBSKI*, Jacek KABZIŃSKI*,
Grzegorz WASIAK*, Rafał ZAWIŚLAK*

OCENA EFEKTYWNOŚCI TECHNIKI EXTREME LEARNING MACHINE (ELM) DO MODELOWANIA DWUWYMIAROWYCH NIELINIOWOŚCI W UKŁADACH NAPĘDOWYCH

Celem artykułu jest ocena przydatności techniki Extreme Learning Machine (ELM) do modelowania dwuwymiarowych nieliniowości występujących w układach napędowych. Jako przykład typowej nieliniowości wybrana została funkcja opisująca sumaryczną siłę oporów w napędzie liniowym. Do stworzenia neuronowego modelu oporu użyta została klasyczna technika ELM. Autorzy ocenili efekty zastosowania tej techniki i zaproponowali modyfikację modelu pozwalającą na zmniejszenie błędów modelowania. Przeprowadzone testy symulacyjne i eksperymenty pozwoliły autorom na sformułowanie ogólnych wniosków dotyczących możliwości stosowania techniki ELM do modelowania nieliniowych, dwuwymiarowych zależności na podstawie zaszumionych danych uzyskanych eksperymentalnie.

1. WPROWADZENIE

W układach napędowych wielowymiarowe modele matematyczne wykorzystuje się między innymi w zadaniach identyfikacji, kompensacji czy sterowania adaptacyjnego. Wiedza pochodząca z procesu identyfikacji obiektu może być wykorzystana w układach regulacji, zarówno do określenia struktury układu sterowania jak i do strojenia parametrów regulatorów. Dobrym przykładem może być kompensacja siły tarcia zależnej od położenia i prędkości [2]–[5] w układach sterowania pozycyjnego, czy też modelowanie kształtu strumienia magnetycznego zależnego od prądu i położenia, wykorzystywane w układach regulacji momentu. Satysfakcjonujący model

* Instytut Automatyki, Politechnika Łódzka, 90-924 Łódź, ul. B. Stefanowskiego 18/22, e-mail: marcin.jastrzebski@p.lodz.pl, jacek.kabzinski@p.lodz.pl, grzegorz.wasiak@p.lodz.pl, rafal.zawislak@p.lodz.pl

stosowany w układzie sterowania napędem powinien łączyć w sobie następujące cechy:

- dostateczną dokładność pozwalającą na wierne odtworzenie modelowanej zależności,
- dostateczną prostotę, pozwalającą na szybkie strojenie modelu i szybkie obliczanie jego wyjścia w czasie rzeczywistym,
- odporność na zaszumienia i błędy danych,
- elastyczność pozwalającą na szybką modyfikację w czasie rzeczywistym, np. poprzez adaptację jego parametrów.

Otrzymywane w wyniku eksperymentów dane pomiarowe z reguły wskazują na występowanie nieliniowości, a czasami także nieciągłości. Do modelowania zależności tego rodzaju można wykorzystywać różne techniki aproksymacji, także modelowanie rozmyte i modelowanie poprzez sztuczne sieci neuronowe. Powszechnie znaną zaletą systemów rozmytych jest duża zdolność aproksymacji nieliniowości oraz przydatność do rozwiązywania zadań klasyfikacji oraz stosunkowa łatwość interpretacji zgromadzonej w nich wiedzy. Do niewątpliwych zalet charakterystycznych dla sztucznych sieci neuronowych należy zaliczyć ich zdolność uczenia się. Połączenie tych koncepcji doprowadziło do powstania systemów rozmyto-neuronowych, a w dalszym etapie poprzez kolejne próby hybrydyzacji systemów ewolucyjno-rozmytych o strukturach strojonych przez algorytmy bakteryjne, rojowe, itp. Wadą tych rozwiązań są skomplikowanie algorytmy strojenia oraz stosunkowo długi czas poszukiwania wartości optymalnych ich parametrów.

Extreme Learning Machine (ELM) można zaliczyć do relatywnie nowych technik modelowania neuronowego. W literaturze przedmiotu można znaleźć szereg artykułów przeglądowych opisujących, charakteryzujących i promujących tę technikę, głównie ze względu na prostotę modelu oraz krótki czas jego uczenia [1, 11, 12]. Dość szeroki jest również opisywany zakres zastosowań ELM: od aproksymacji danych pomiarowych do problemów klasyfikacji, z technikami uczenia wsadowego bądź sekwencyjnego. Liczni autorzy publikacji jak również sami autorzy algorytmu podkreślają główną zaletę ELM, to jest prostą strukturę sieci – neuronowy model z jedną warstwą ukrytą oraz niemalże błyskawiczny czas nauki uzyskany poprzez:

- losowy wybór parametrów wag neuronów w warstwie ukrytej, które pozostają niezmiennione (są więc całkowicie niezależne od danych uczących)
- obliczenie wag neuronu wyjściowego z pojedynczej formuły analitycznej, bez konieczności iteracyjnego uczenia modelu.

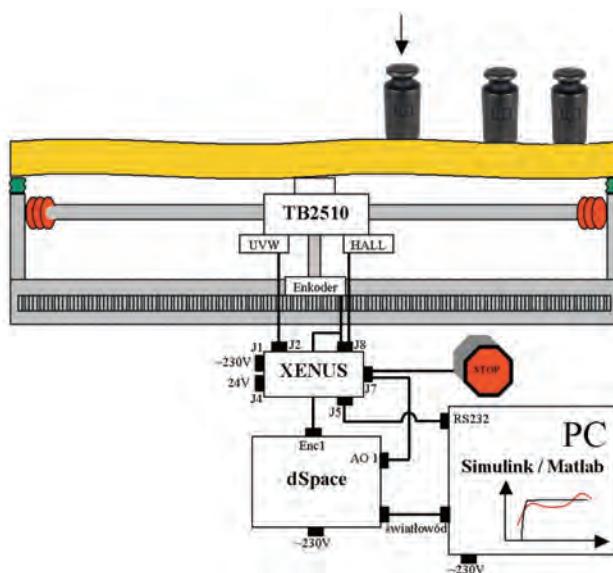
Z drugiej strony część autorów publikacji wskazuje jednak na możliwość powstania problemów numerycznych związanych np. ze złym uwarunkowaniem macierzy [10], nadmiarowością danych oraz trudnościami wynikającymi z zastosowania całkowicie losowych wartości wag.

Poniżej przedstawiono efekty wykorzystania techniki ELM do odtworzenia przebiegu zmienności siły oporów działających na silnik liniowy. Problem kompensacji siły

oporów jest tym trudniejszy, że zawiera ona składniki pochodzące z wielu niezależnych źródeł i może być silnie nieliniową, nieciągłą funkcją położenia silnika i jego prędkości.

2. DWUWYMIAROWY MODEL SIŁY OPORÓW

Na potrzeby testów techniki ELM zaproponowano modelowanie siły oporów w serwonapędzie liniowym, w postaci statycznej, nieliniowej zależności ww. siły od położenia i prędkości silnika. Dane do testów zebrane zostały w układzie laboratoryjnym wyposażonym w: silnik liniowy TB2510 o całkowitej masie 7,04 kg, długotrwałej sile 104 N, z enkoderem o rozdzielczości 1 μm i z tensometrycznymi czujnikami do pomiaru sił. Napęd był zasilany z falownika XTL-230-18 z wbudowanym regulatorem prądu, a całość algorytmów sterowania realizował modułarny zestaw kontrolno-pomiarowy oparty na karcie dSpace DS1106, z czasem próbkowania równym 50 ms. Wartości sił obciążenia mogły być zbierane z czujników tensometrycznych lub z wyjścia obserwatora zakłócenia [4, 5]. Schemat stanowiska laboratoryjnego pokazano na rysunku 1.



Rys. 1. Schemat układu laboratoryjnego

Dane zebrane za pomocą obserwatora ilustruje rys. 2. Analiza tych danych wykazała nieciągłość dla prędkości równej zero, której źródłem jest istnienie tarcia statycznego oraz okresową (o okresie równym podziałce biegunowej silnika 51,2 mm) zmienność siły oporu w funkcji położenia nałożoną na wzrost oporu, który odzwierciedla np. efekt narzędzia obrabiającego przedmiot o zmiennej twardości czy też pe-

netrację tkanki w przypadku robotów chirurgicznych. Efekty te uzyskano zmieniając obciążenie, kształt i materiał belki zaznaczonej na rys. 1.

W pierwszej fazie testów techniki ELM zastosowano dane wygenerowane na podstawie wzoru (1):

$$F_o(p, v) = 0,4 \sin\left(\frac{2\pi}{0,0512} p\right) + 2 \operatorname{sign}(v) \xi_t(100(p - 0,15)) + 3,62 \operatorname{sign}(v) + 24,81v \quad (1)$$

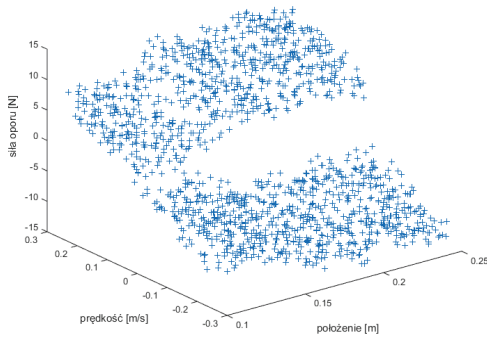
gdzie: $\xi_t(z) = 2/(1 + \exp(-2z)) - 1$,

p – wartość położenia [m],

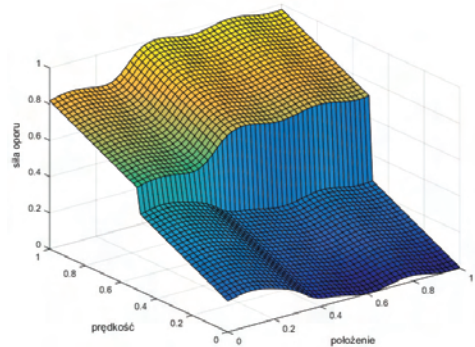
v – prędkość liniowa silnika [m/s],

$F_o(p, v)$ – sumaryczna siłę oporu [N].

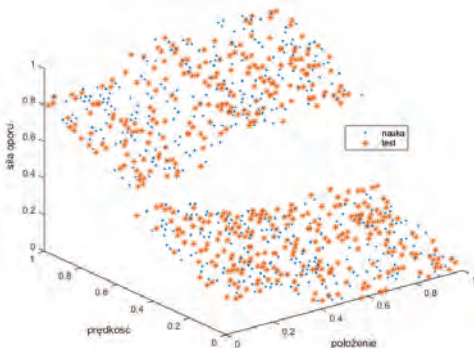
Wzór (1) generuje powierzchnię o właściwościach podobnych do danych przedstawionych na rys. 2. Wykorzystanie tej powierzchni pozwoliło na analizę właściwości metody i przetestowanie jej wariantów zanim zostanie zastosowana do danych z układu rzeczywistego. Na potrzeby symulacji wszystkie dane zostały znormalizowane do przedziałów $[0, 1]$. Ich przebieg (płaszczyzna wg wzoru (1) po normalizacji) przedstawiono na rys. 3.



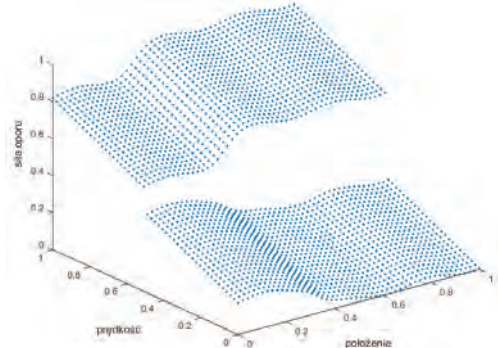
Rys. 2. Dane z obserwatora



Rys. 3. Model płaszczyzny



Rys. 4. Zaszumione dane do nauki i testów



Rys. 5. Dane do testów ze wzorca

Rysunki 4 i 5 przedstawiają wygenerowane dane służące do nauki i do testów poprawności modelu. Dane te zostały wygenerowane na podstawie wzoru (1), uwzględniając zarówno ich losowy rozkład w osiach p , v jak i losowo wygenerowany szum pomiarowy o wartościach z zakresu $[-0,1, 0,1]$ N.

Wprowadzone zostały średniokwadratowe miary jakości E_{train} , E_{test} i E_{Ttest} porównujące uzyskany metodą ELM model $F_m(p,v)$ odpowiednio z danymi użytymi do nauki i testów (zazsumionymi) oraz z danymi niezsumionymi uzyskanymi ze wzoru (1):

$$E_{\text{train}} = \sqrt{\frac{1}{N_n} \sum_{i=1}^{N_n} (F_m(p_{n_i}, v_{n_i}) - F_{sz}(p_{n_i}, v_{n_i}))^2}, \quad (2)$$

$$E_{\text{test}} = \sqrt{\frac{1}{N_t} \sum_{i=1}^{N_t} (F_m(p_{t_i}, v_{t_i}) - F_{sz}(p_{t_i}, v_{t_i}))^2}, \quad (3)$$

$$E_{\text{Ttrain}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (F_m(p_i, v_i) - F_o(p_i, v_i))^2}. \quad (4)$$

gdzie:

- N_n – liczba losowo wybranych próbek (p_{n_i}, v_{n_i}) służących do nauki modelu techniką ELM,
- N_t – liczba losowo wybranych próbek (p_{t_i}, v_{t_i}) służących do testowania otrzymanej powierzchni,
- N – liczba wszystkich punktów z równomiernej siatki, użytych do wygenerowania płaszczyzny z rys. 3,
- F_{sz}, F_o – odpowiednio zazsumione i niezsumione wartości funkcji oporu dla par (p, v) .

3. STANDARDOWY MODEL EML

Technika Extreme Learning Machine została zaproponowana w pracy Huang et al. [1] jako efektywny algorytm uczenia sieci neuronowej o jednej warstwie ukrytej. Głównym założeniem tej koncepcji jest fakt iż:

- wagi i biasy neuronów warstwy ukrytej są generowane losowo i pozostają stałe,
- wagi neuronu wyjściowego są wyznaczone z pojedynczej formuły analitycznej, bez konieczności iteracyjnego uczenia,
- wyjście modelu jest wyznaczone z formuły:

$$F_{ELM}(x) = \sum_{k=1}^M \beta_k h_k(x) \quad (5)$$

gdzie: x – n -wymiarowy wektor wejść, M – liczba neuronów w warstwie ukrytej, $h(x) = [h_1(x), \dots, h_M(x)]^T$ – wektor wyjść tych neuronów, $\beta(x) = [\beta_1(x), \dots, \beta_M(x)]^T$ – wektor wag neuronu wyjściowego.

Do konstrukcji modelu mogą być użyte różne funkcje aktywacji. Powszechnie stosowane są funkcje: sigmoidalne, Gaussowskie, skokowe, trójkątne, sinusoidalne. Najczęściej używana sigmoidalna funkcja aktywacji opisana jest wzorem:

$$h_k(x) = \frac{1}{1 + \exp(w_k^T(x) \cdot x + b_k)} \quad (6)$$

gdzie: $w_k^T(x) = [w_{k,1}(x), \dots, w_{k,n}(x)]$, $k = 1, \dots, M$ jest wektorem wag k -tego neuronu warstwy ukrytej, zaś b_k jego biasem.

Parametry w_k i b_k , są generowane losowo, z dowolnego rozkładu ciągłego, niezależnie od danych wejściowych i pozostają stałe. Jeśli:

$$\{(x_i, t_i), \quad i = 1, \dots, N_n\}, \quad x_i \in R^n \quad (7)$$

tworzą zbiór danych uczących, to N_n -wymiarowy wektor wyjść modelu F_m jest porównywany z wektorem wzorców F , gdzie:

$$F_m = [F_{ELM}(x_1), \dots, F_{ELM}(x_{N_n})]^T, \quad F = [t_1, \dots, t_{N_n}]^T \quad (8)$$

Z uwagi na to, że:

$$F_m = \begin{bmatrix} h_1(x_1) & \cdots & h_M(x_1) \\ \vdots & \ddots & \vdots \\ h_1(x_{N_n}) & \cdots & h_M(x_{N_n}) \end{bmatrix} \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_M \end{bmatrix} = H\beta, \quad (9)$$

wagi β można wyznaczyć minimalizując wskaźnik dopasowania:

$$E = \|H\beta - F\|^2 \quad (10)$$

Optymalną wartością jest:

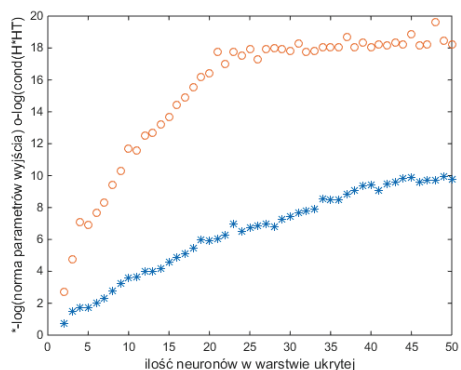
$$\beta_{opt} = H^+ F \quad (11)$$

gdzie H^+ jest odwrotnością Moore–Penrose’a macierzy H .

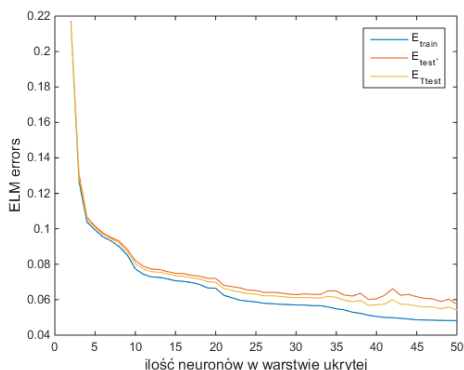
Do celów testowych w niniejszym artykule została użyta standardowa postać ELM z sigmoidalnymi funkcjami aktywacji o parametrach w_k i b_k , pochodzącymi z rozkładu jednostajnego w przedziale $[-1, 1]$.

Efekty uczenia sieci o różnej liczbie neuronów zostały przedstawione na rys. 6 i 7. Jak z nich wynika, zarówno wskaźnik uwarunkowania rozwiązywanego układu równań liniowych, jak i wartości parametrów wyjścia są bardzo duże. Użycie większej

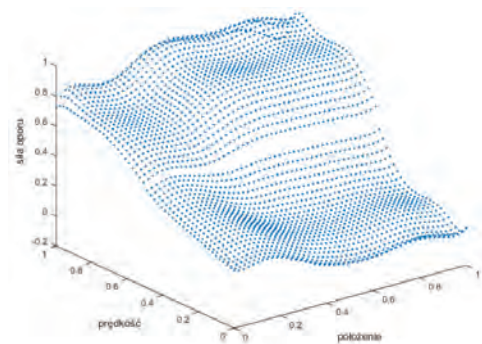
ilości neuronów w warstwie ukrytej nie poprawi więc jakości modelu. Analizując rys. 7 można zauważyć, że błędy również przyjmują spore wartości. Jak widać z rys. 8 i 9 model nauczony techniką ELM nie radzi sobie dobrze z odtworzeniem nieciągłości prędkości.



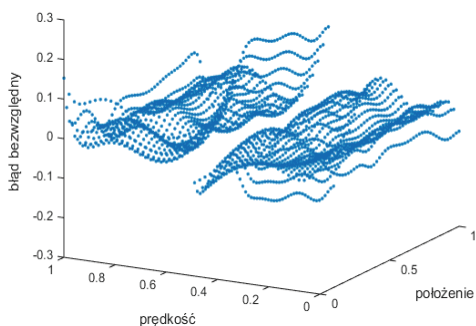
Rys. 6. Średnia wartość norm parametrów wyjścia oraz średnia wartość $\text{cond}(H*H^T)$



Rys. 7. Średnie wartości błędów z 20 eksperymentów w funkcji liczby neuronów w warstwie ukrytej



Rys. 8. Płaszczyzna otrzymana z modelu ELM zawierającego 50 neuronów dla punktów (p,v) ze zbioru Ttest

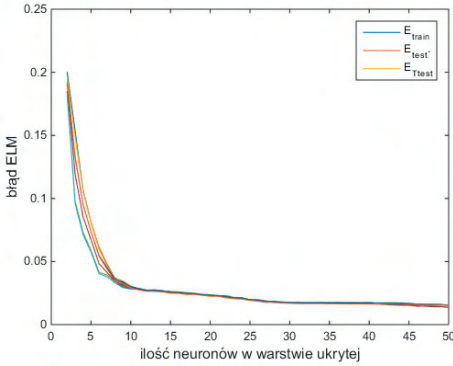


Rys. 9. Błąd bezwzględny modelu ELM zawierającego 50 neuronów

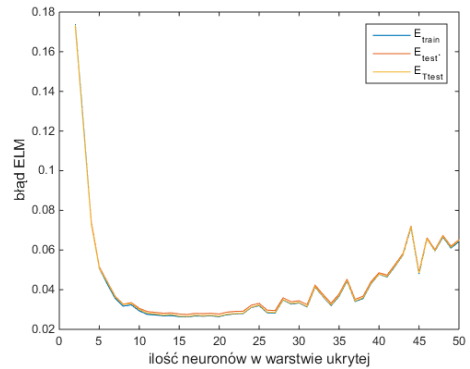
Jakkolwiek średnie wartości błędów pokazane na rys. 7 maleją w funkcji rosnącej liczby neuronów i osiągają minima $[E_{\text{train}}, E_{\text{test}}, E_{\text{Ttest}}] = [0,0488 \quad 0,0518 \quad 0,0534]$, to nie oznacza to zadowalającej dokładności modelu. Jak widać na rys. 9 błąd jest szczególnie duży w rejonie nieciągłości danych i na brzegach obszaru pokrytego danymi.

W literaturze opisano typowe modyfikacje metody ELM, które powinny poprawić jakość modelowania, czyli dokładność modelu, jego zdolność do uogólniania a przede wszystkim jego stabilność numeryczną. Przetestowano je w kolejnych eksperymentach.

W pierwszej kolejności rozważono wpływ zmiany funkcji aktywacji na jakość modelu. Zmiana typu funkcji aktywacji z sigmoidalnych na dowolne różniczkowalne funkcje przynależności, przy pozostawionej tej samej procedurze losowania wartości wag w warstwie ukrytej, nie wpłynęła znacząco na poprawę obserwowanych parametrów, co zilustrowano na rys. 10. Podobne wartości jak w poprzednich symulacjach przyjęły również $cond(H*H^T)$ i $\|\beta\|$.



Rys. 10. Błąd otrzymany dla różnych funkcji aktywacji (sigmoidalna, sinusoidalna, radialna)



Rys. 11. Błąd bezwzględny ELM dla $C = 50000$

Dobrze znanym remedium na problemy numeryczne spowodowane złym uwarunkowaniem macierzy $H^T H$ jest regularyzacja [1]. Zamiast minimalizować równanie (10) rozważmy wyrażenie:

$$E_c = \|\beta\|^2 + C\|H\beta - F\|^2 \quad (12)$$

gdzie C jest pewnym parametrem projektowym.

Rozwiązaniem optymalnym będzie teraz [1]:

$$\beta_{opt} = \left(\frac{1}{C} I + H^T H \right)^{-1} H^T F \quad (13)$$

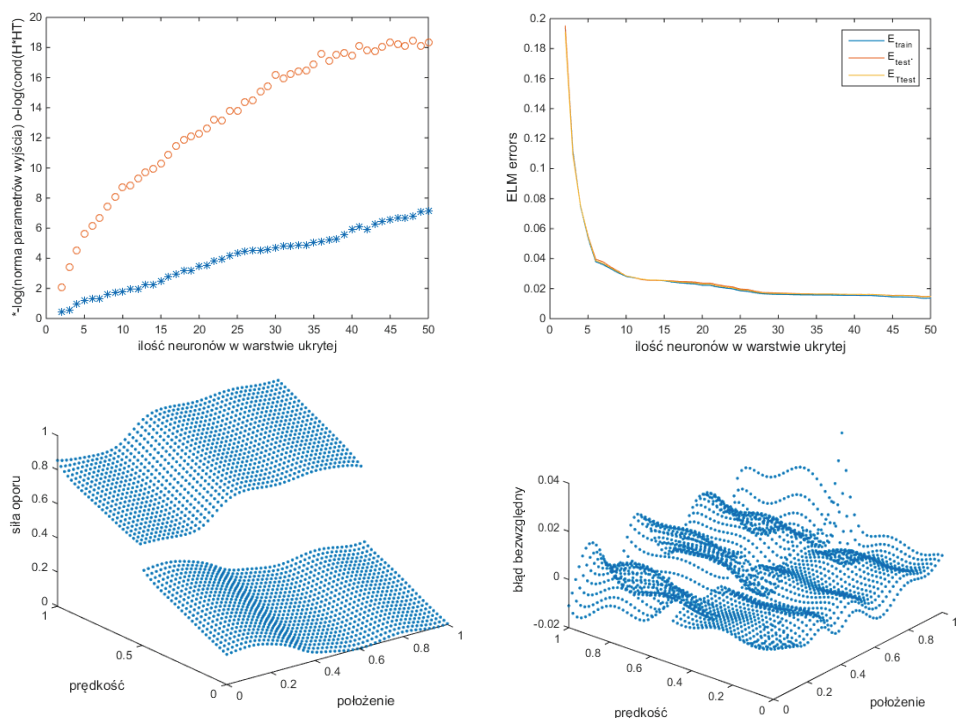
Zmniejszanie wartości C pozwala na ograniczenie wartości $cond\left(\frac{1}{C} I + H^T H\right)$.

Niestety równocześnie obserwujemy pogorszenie dokładności modelowania i jeszcze gorsze odtwarzanie nieciągłości. Przy dobranej wartości tego współczynnika trzeba się zatem kierować zdrowym rozsądkiem i wybrać rozwiązanie pośrednie. Dobór współczynnika C jest zależny od struktury problemu i z tego powodu trudno jest formułować ogólne wnioski. Na rys. 11 przedstawiony został przebieg błędów w układzie policzonym dla $C = 50000$. Wartość $\|\beta\|$ została znacznie ograniczona i pozostawała na poziomie akceptowalnym pomimo zwiększaniem liczby neuronów w warstwie

ukrytej. Jak wynika z rys. 11, ta sama wartość parametru C nie zapewnia dobrej dokładności modelowania dla żadnej liczby neuronów, a zwiększenie liczby neuronów powyżej 20 powodowało wręcz wzrost wszystkich błędów.

4. MODYFIKACJE MODELU EML

Jak wynika z przedstawionego przykładu, metoda ELM, mimo zalet intensywnie promowanych przez jej twórców zawodzi w przypadku modelowanej nieliniowości. Przyczyną jest niedostateczna zmienność funkcji bazowych przy narzuconym sposobie losowania parametrów. Sieć ELM nie pozwala na uczenie się cech z danych wejściowych, są one tylko rzutowane na przestrzeń cech skonstruowaną przez losowo wybrane funkcje aktywacji. Przestrzeń ta powinna być dostatecznie bogata by odwzorować wszystkie cechy danych, więc funkcje aktywacji powinny charakteryzować się dostateczną zmiennością wartości w swych dziedzinach. Można pokusić się o zwiększenie zmienności funkcji aktywacji poprzez modyfikację sposobu losowania parametrów, jak to proponowano w [7], lub też zapewnić lepszą ekstrakcję cech danych



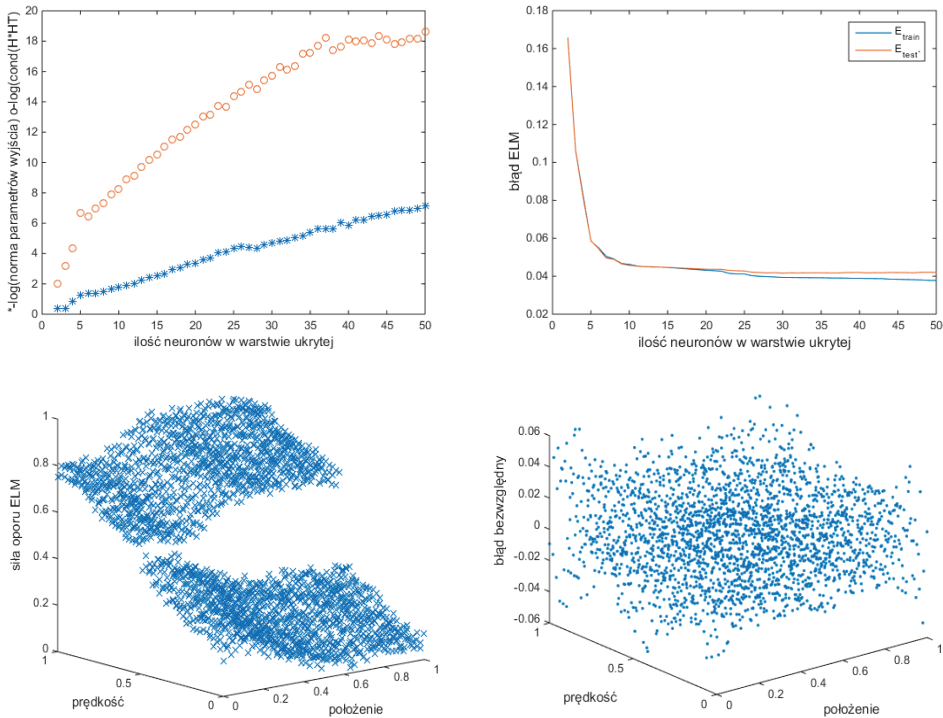
Rys. 12. Przebiegi uzyskane dla modelu ELM z trzema wejściami [położenie, prędkość, znak prędkości] i z 50 neuronami w warstwie ukrytej

wejściowych poprzez zwiększenie liczby wejść, lub dekompozycje danych na obszary, w których są one bardziej jednorodne.

Uwzględniając specyfikę nieciągłości powierzchni siły oporu, dla rozważanego przypadku związaną z przejściem prędkości przez zero, można zaproponować modyfikację modelu poprzez dodanie do niego kolejnego, trzeciego wejścia – sygnału znaku prędkości $sign(v)$. Uzyskane wyniki dla tak skonstruowanego modelu prezentuje rys. 12.

Uzyskane w wyniku symulacji wartości $[E_{train}, E_{test}, E_{Ttest}]$ wynoszą odpowiednio $[0,0136 \ 0,0154 \ 0,0155]$ i jak pokazano na rys. 12, błąd modelowania jest równomierny w całej dziedzinie. Z analizy tego rysunku wynika że, ten prosty zabieg pozwolił na znaczne zmniejszenie błędu aproksymacji modelowanej powierzchni.

Powierzchnia odtwarzana przez układ nauczony techniką ELM bez problemu radzi sobie z nieciągłością prędkości już przy 10–15 neuronach w warstwie ukrytej. Zaszumienie wartości próbek nie pozwala na zmniejszenie błędu dla niewielkiej liczby neuronów, z kolei zwiększanie ich liczby prowadzi do złego, z punktu widzenia numerycznego, uwarunkowania zadania.



Rys. 13. Przebiegi uzyskane dla modelu z trzema wejściami [położenie, prędkość, znak prędkości] uczonego wg. danych z obserwatora

Kolejnym etapem testów było uruchomienie procedury modelowania dla danych uzyskanych z eksperymentu (rys. 2). Analogicznie jak poprzednio z całego zbioru punktów (N) opisujących siłę oporu (p, v, f) wybrano podzbiór, który podzielono na część przeznaczoną do nauki modelu (N_n) i drugą do testów (N_t). Wyniki symulacji przedstawione zostały na rys. 13.

Jak wynika z rys. 13, układ uczony techniką ELM nadal dobrze sobie radzi w przypadku danych pochodzących z układu rzeczywistego. Nieco większe błędy wynikają z większego zaszumienia danych wejściowych.

5. WNIOSKI

Metoda ELM potwierdziła swoje zalety – krótki czas uczenia, liniową parametryzację modelu przez wagi neuronu wyjściowego, dobrą właściwość filtracji szumu, odporność na izolowane błędy danych wejściowych oraz dobre właściwości generalizacji. Metoda w klasycznej postaci nie radzi sobie z aproksymacją danych z nieciągłościami, o dużej zmienności czy różnorodności przebiegu w podzbiorach dziedziny. Metoda nie pozwala na dowolne zwiększanie liczby neuronów w warstwie ukrytej, co teoretycznie powinno poprawiać dokładność modelu, z powodu ograniczeń numerycznych: rosnącego wskaźnika uwarunkowania macierzy używanej przy obliczaniu pseudoodwrotności oraz rosnących wartości wag neuronu wyjściowego. Przyczyną tych kłopotów jest nieefektywna ekstrakcja cech z danych wejściowych przez funkcje aktywacji, które w klasycznych warunkach losowania parametrów mogą charakteryzować się zbyt małą zmiennością [7], [8]. Zaproponowaną metodą pokonania tych trudności jest poprawa ekstrakcji cech poprzez wprowadzenie dodatkowych wejść charakteryzujących obszary jednorodnych danych – w prezentowanym przykładzie był to podział na obszar dodatniej i ujemnej prędkości (dodatkowym wejściem był znak prędkości). Przy takiej modyfikacji metoda ELM pozwala na osiągnięcie jakości modelu porównywalnej do uzyskanej techniką modelowania rozmytego z nieliniowymi następnikami [6], [9].

Przedstawione badania pokazują, że technika Extreme Learning Machine może być rozważana jako interesująca alternatywa modelowania rozmytego, choć nie jest tak uniwersalna jak ją przedstawiają w literaturze przedmiotu.

LITERATURA

- [1] GAO H., GUANG-BIN, HUANG, SHIJI SONG, KEYOU YOU, *Trends in extreme learning machines: A review*, Neural Networks, 2015, Vol. 61, 32–48.
- [2] GAO Z., *Active disturbance rejection: A paradigm shift in feedback control system design*, Proc. IEEE American Control Conference, 2006, 2399–2405.
- [3] HAN ME K., SEONG IK H., JONG SHIK K., *Precision position control of servo systems using adaptive backstepping and recurrent fuzzy neural networks*, Journal of Mechanical Science and Technology, 2009, Vol. 23, 2873–2890.

- [4] JASTRZĘBSKI M., *On-line parameter tuning of discontinuous fuzzy friction compensator in linear drive*, XXIV Inter. Conf. Information, Communication and Automation, Piscataway, USA, 2013, 139–144.
- [5] KABZIŃSKI J., *Experimental identification of disturbance forces effecting permanent magnet linear motors*, Proc. Inter. Conf. Power Electronics and Intelligent Control for Energy Conservation, 2005.
- [6] KABZIŃSKI J., *Fuzzy Friction Modeling for Adaptive Control of Mechatronic Systems. Artificial Intelligence Applications and Innovations*, IFIP Advances in Information and Communication Technology, Springer, Berlin Heidelberg, 2012, Vol. 381, 185–195.
- [7] KABZIŃSKI J., *Is Extreme Learning Machine Effective for Multisource Friction Modeling?*, Artificial Intelligence Applications and Innovations. IFIP Advances in Information and Communication Technology, Springer, Bayonne, France, 2015, Vol. 456, 318–333.
- [8] KABZIŃSKI J., *One-Dimensional Linear Local Prototypes for Effective Selection of Neuro-Fuzzy Sugeno Model Initial Structure*, IFIP WG 12.5 Inter. Conf. Artificial Intelligence Applications and Innovations, Springer, IFIP Series, Larnaca, Cyprus, Berlin 2010, 62–69.
- [9] KABZIŃSKI J., KACERKA J., *TSK Fuzzy Modeling with Nonlinear Consequences*, [in:] *Artificial Intelligence Applications and Innovations IFIP Advances in Information and Communication Technology*, Springer, Berlin–Heidelberg, 2014, 498–507.
- [10] PARVIAINEN E., RIIHIMÄKI J., *Connection between Extreme Learning Machine and Neural Network Kernel*, [in:] *Knowledge Discovery, Knowledge Engineering and Knowledge Management*, Springer, Berlin–Heidelberg, 2013, 122–135.
- [11] SHAOBO LIN, XIA LIU; JIAN FANG; ZONGBEN XU, *Is Extreme Learning Machine Feasible? A Theoretical Assessment (Part II)*, IEEE Trans. Neural Networks and Learning Systems, 2015, Vol. 26, 21–34.
- [12] XIA LIU, SHAOBO LIN, JIAN FANG, ZONGBEN XU, *Is Extreme Learning Machine Feasible? A Theoretical Assessment (Part I)*, IEEE Trans. Neural Networks and Learning Systems, 2015, Vol. 26, 7–20.

APPRAISAL OF EXTREME LEARNING MACHINE EFFECTIVE FOR 2D FRICTION MODELING

The aim of this article is to discuss suitability of extreme learning machine (ELM) approach for modeling multisource 2D friction for motion control purposes. The article describes a method of obtaining data used to modulate the resistance forces, taking into account specific to the case and machine external conditions. The features of multisource friction in mechatronic systems are defined, the main aspects of friction modeling by a Standard ELM are investigated and some modifications are proposed to make it more suitable for specific demands of the discussed task. This allows to formulate some general remarks concerning properties of ELM for function approximation.