Wydział Elektroniki i Technik Informacyjnych

Politechnika Warszawska

Modelowanie i Identyfikacja

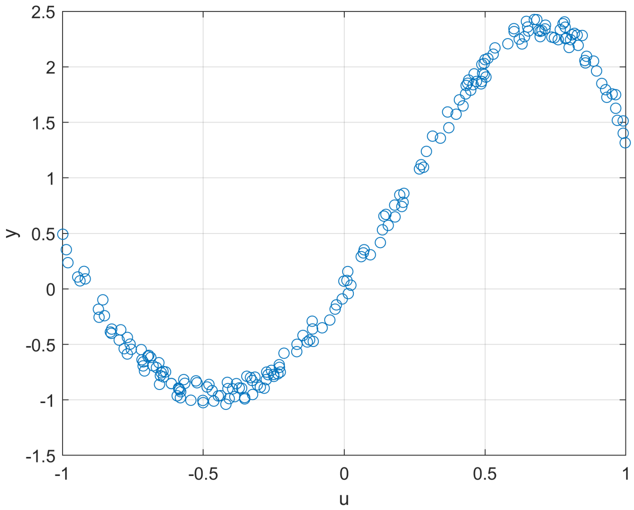
Sprawozdanie z projektu nr 2
zadanie nr 43

Mikołaj Wewiór
318407

Warszawa, 2023

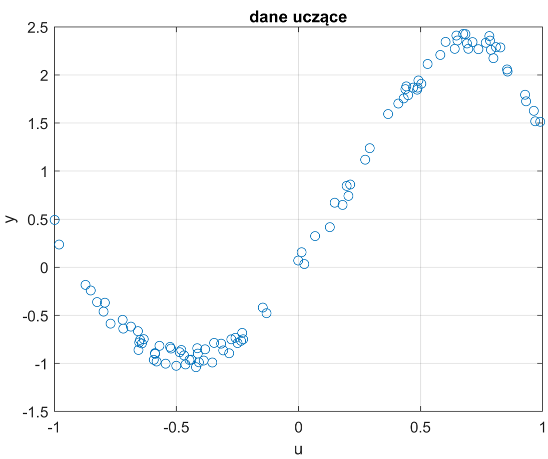
1. Identyfikacja modeli statycznych

a. Reprezentacja graficzna danych statycznych

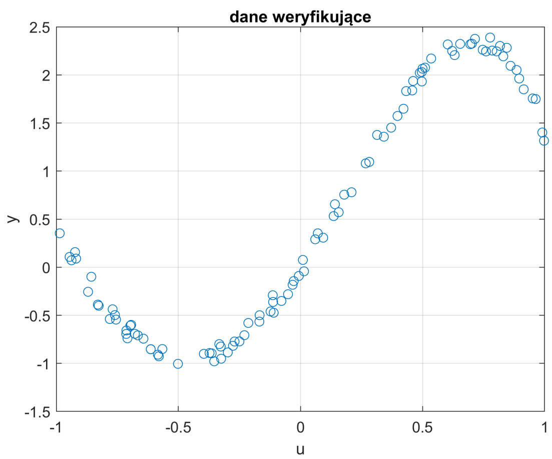


Rys. 1. dane statyczne

Aby podzielić wszystkie dane na dane uczące i weryfikujące co drugą próbkę przydzieliłem do każdego ze zbiorów. Dzięki temu zbiory są tak samo liczne oraz na pierwszy rzut oka przypominają kształtem wykres wszystkich danych.



Rys. 2. zbiór danych uczących



Rys. 3. zbiór danych weryfikujących

b. Statyczny model liniowy

Model ten dany jest wzorem:

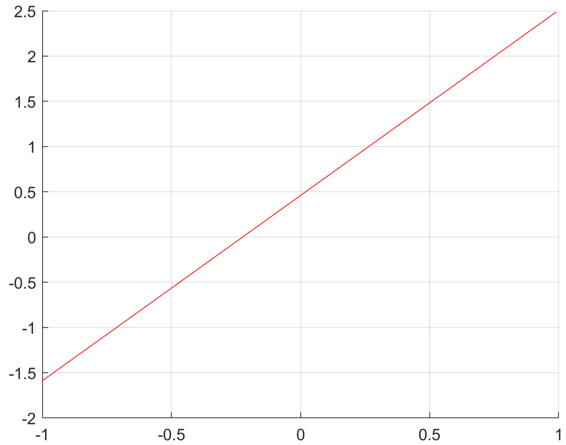
$$y(u) = a_0 + a_1 u$$

Do obliczenia współczynników a_0 oraz a_1 wykorzystamy metodę najmniejszych kwadratów i zastosujemy zapis wektorowo-macierzowy, dzięki któremu możemy w uniwersalny sposób zapisać zadanie:

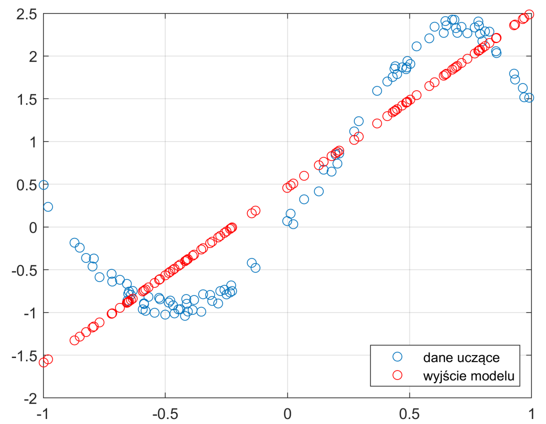
$$Y = M \cdot w$$

gdzie Y to wektor pionowy wartości wyjścia modelu w zależności od danych, M jest macierzą, gdzie jedną kolumną są dane uczące wejściowe, a drugą - jedynki, a w to wektor współczynników - w tym przypadku a_0 i a_1. Do wyliczenia elementów wektora w wykorzystałem tzw. lewe dzielenie: $w = M \setminus Y$ i prezentują się w następujący sposób:

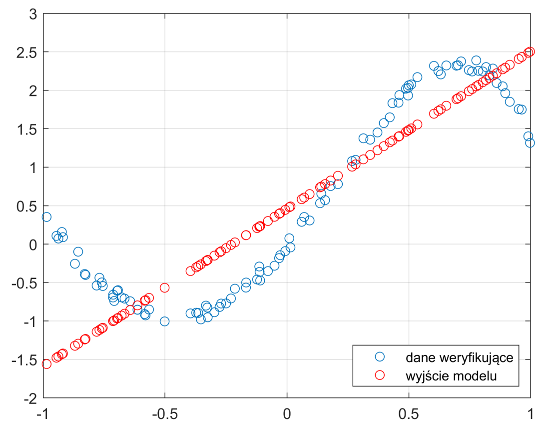
$$a_0 = 0,4593, \quad a_1 = 2,0475.$$



Rys. 4. charakterystyka modelu liniowego $y(u) = a_1 \cdot u + a_0$



Rys. 5. wyjście modelu na tle danych uczących
Błąd modelu: $E_{ucz} = 32,1536$



Rys. 6. wyjście modelu na tle danych weryfikujących
Błąd modelu: $E_{wer} = 39.2603$

Błędy zostałe obliczone według wzoru:

$$E = \sum_{i=1}^{P} (y^{mod}(i) - y(i))^2$$

Ewidentnie model liniowy nie jest w dobrym modelem dla danego zadania, co jednak nie powinno nas dziwić patrząc na charakterystykę danych.

c. Statyczne modele wielomianowe

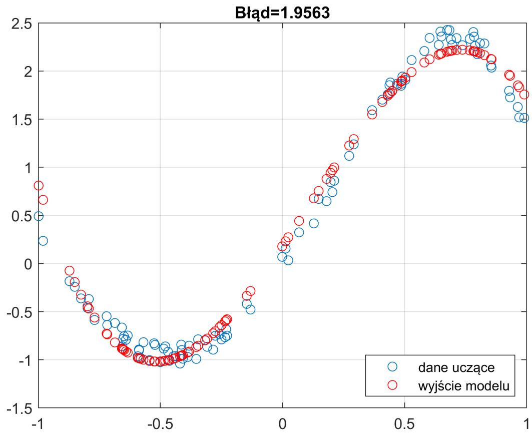
Ogólna postać modelu dla wielomianu stopnia N wygląda następująco:

$$y(u) = a_0 + \sum_{i=1}^{N} a_i u^i$$

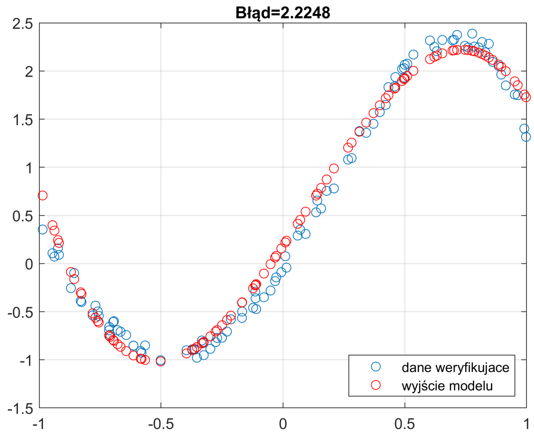
Do identyfikacji modelu ponownie wykorzystamy metodę najmniejszych kwadratów oraz wzór $w = M \setminus Y$. Wektor Y jak poprzednio zawiera wyjściowe dane, w również jest wektorem współczynników, tylko tym razem ma on długość N i kolejnymi jego elementami są: a_0, a_1, itd. aż do a_N. Macierz M zawiera kolejno kolumny: jedynek, danych wejściowych w pierwszej potędze, danych wejściowych w drugiej potędze, itd. aż do danych w potędze stopnia N.

Na pierwszy rzut oka wielomian w trzeciej potędze mógłby już całkiem dobrze odwzorować model, dlatego do testów wybrałem następujące stopnie wielomianów: 3, 5, 7 i 8.

Wielomian trzeciego stopnia

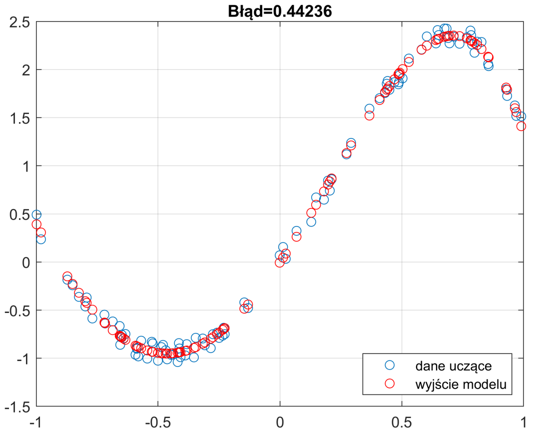


Rys. 7. wyjście modelu na tle danych uczących

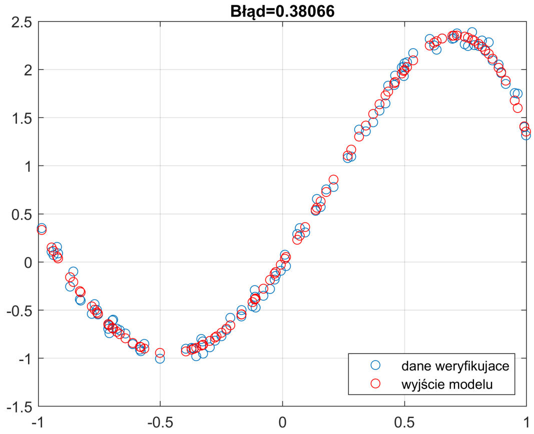


Rys. 8. wyjście modelu na tle danych weryfikujących

Wielomian piątego stopnia

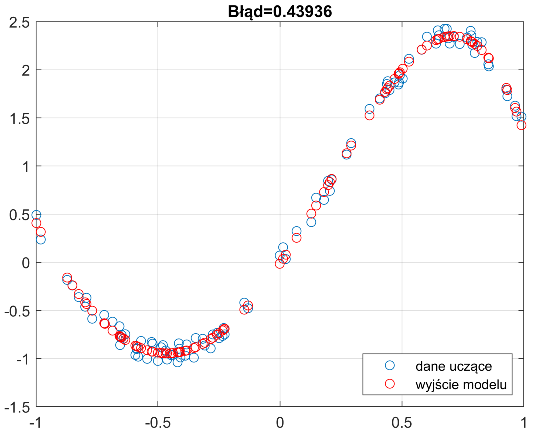


Rys. 9. wyjście modelu na tle danych uczących

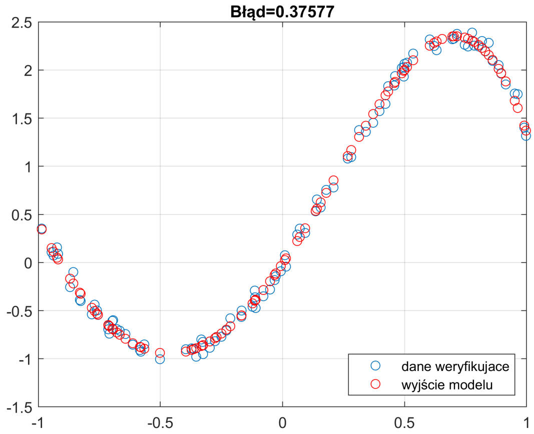


Rys. 10. wyjście modelu na tle danych weryfikujących

Wielomian siódmego stopnia

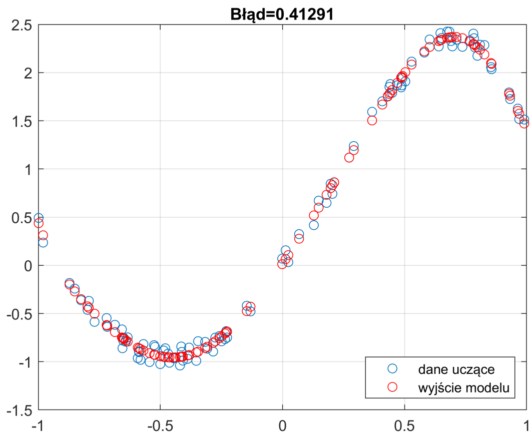


Rys. 11. wyjście modelu na tle danych uczących

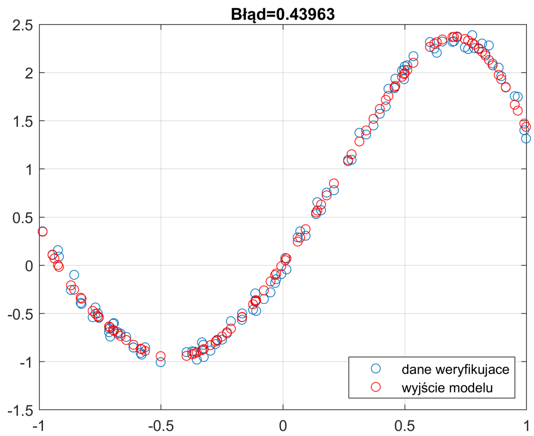


Rys. 12. wyjście modelu na tle danych weryfikujących

Wielomian ósmego stopnia



Rys. 13. wyjście modelu na tle danych uczących



Rys. 14. wyjście modelu na tle danych weryfikujących

stopień wielomianu	błąd dla danych uczących	błąd dla danych weryfikujących
3	1,9563	2,2238
5	0,4424	0,3807
7	0,4394	0,3758
8	0,4129	0,4396

Tab. 1.1. Porównanie błędów dla różnych stopni wielomianów

Z powyższych testów widać, że błąd wyliczony na danych weryfikujących maleje do wielomianu siódmego stopnia - jest tutaj najmniejszy. Dla wielomianu ósmego stopnia widoczne jest już przeparametryzowanie modelu, ponieważ błąd (weryfikujący) się zwiększa - jest to skutek zbytniego dopasowania modelu do danych uczących. Jednak wybrałem wielomian stopnia piątego ponieważ ma niewiele większy błąd niż ten otrzymany po zastosowaniu siódmego stopnia, jednak liczba parametrów jest o 2 mniejza, co sprawia że model jest prostszy. Ostatecznie wybrany przeze mnie model prezentuje się w następujący sposób:

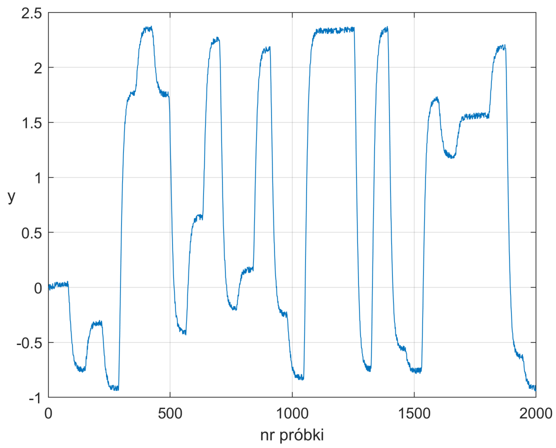
$$y(u) = 0,0034 + 3,7299u + 2,5267u^2 - 3,1359u^3 - 1,6557u^4 - 0,125u^5$$

Dodatkowe uwagi:
Dziwne wydawało się to, że błąd dla danych weryfikujących był mniejszy niż w przypadku danych uczących. Dlatego w ramach testów zamieniłem zbiór uczący z weryfikującym. Efekty tego doświadczenia były bardziej zgodne z intuicją - mniejszy błąd dla danych w porównaniu z weryfikującymi. Mam wrażenie, że skutkiem tego jest nierównomiern "wymieszanie" danych. Gdy spojrzymy na rys.2 i rys. 3 łatwo zauważyć, że dane uczące mają więcej danych skupionych w lokalnych ekstremach, a dane weryfikujące są bardziej równomiernie osadzone.

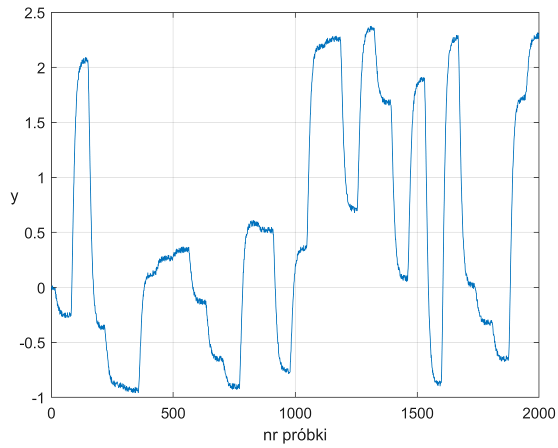
2. Identyfikacja modeli dynamicznych

a. Reprezentacja graficzna danych

wyjście modelu (kolejne próbki)

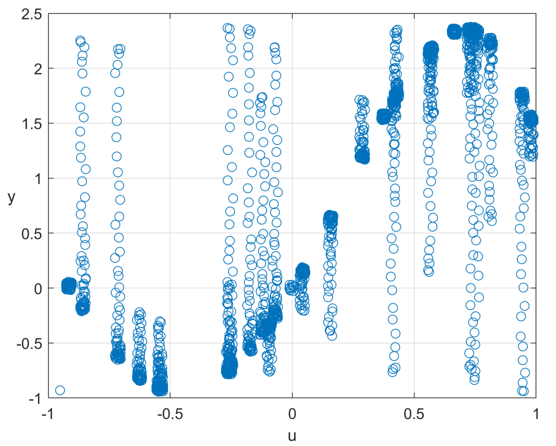


Rys. 15. dane uczące

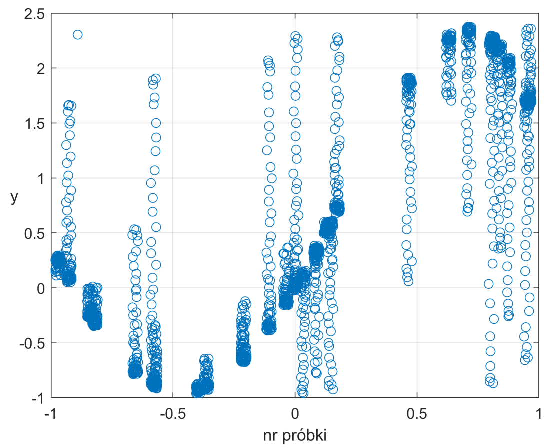


Rys. 16. dane weryfikujące

wyjście modelu w zależności od wejścia



Rys. 17. dane uczące



Rys. 18. dane weryfikujące

b. Liniowy model dynamiczny

Zadaniem jest wyznaczenie metodą najmnieszych kwadratów dynamicznego modelu liniowego postaci:

$$y(k) = \sum_{i=1}^{n_B} b_i u(k-i) + \sum_{i=1}^{n_A} a_i y(k-i)$$

rzędów pierwszego, drugiego i trzeciego. Testowane będą one w dwóch trybach: z rekurencją lub bez niej.

Wzory dla kolejnych modeli prezentują się w następujący sposób:

$$y(k) = 0,0634u(k-1) + 0,98y(k-1)$$

$$y(k) = 0,0089u(k-1) + 0,0328u(k-2) + 1,3833y(k-1) - 0,397y(k-2)$$

$$y(k) = 0,003u(k-1)+0,0416u(k-2)-0,0194(k-3)+1,1808y(k-1)+0,2954y(k-2)-0,4852y(k-3)$$

Błąd liczyłem podobnie jak w poprzednim przypadku, czyli metodą najmniejszych kwadratów. Ponieważ dla trzeciego rzędu dynamiki pierwszą próbką jest ta z indeksem 4, to błąd zacząłem liczyć dopiero od 4 próbki, ponieważ chciałem porównać wszystkie otrzymane modele dla tej samej liczby próbek. Wzór wygląda następująco:

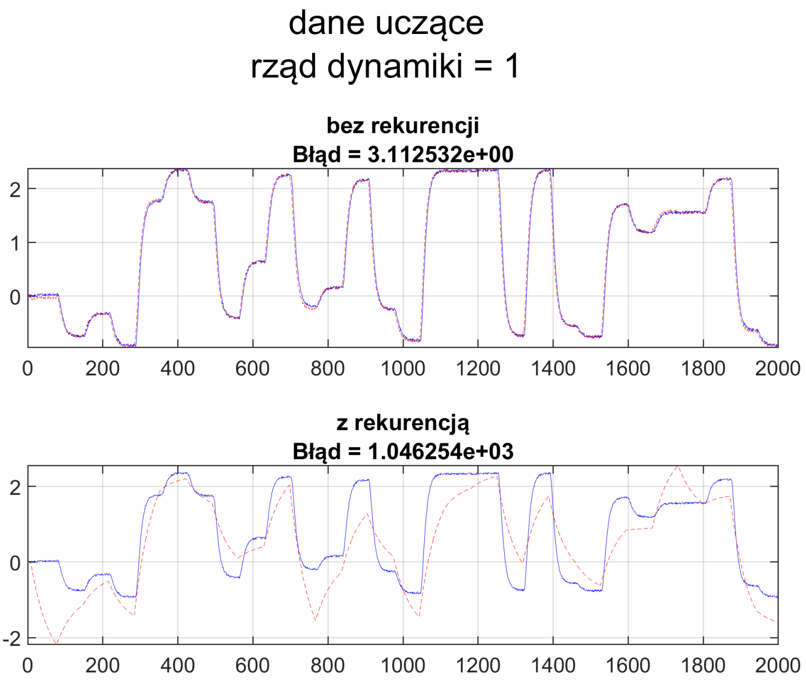
$$E = \sum_{i=4}^{P} (y^{mod}(i) - y(i))^2$$

	Zbiór uczący		Zbiór weryfikujący—	
rząd dynamiki	bez rekurencji	z rekurencją	bez rekurencji	z rekurencją
1	3,11	1046,25	3,04	1425,32
2	2,55	1021,04	2,52	1417,69
3	1,92	989,39	1,93	1393,46

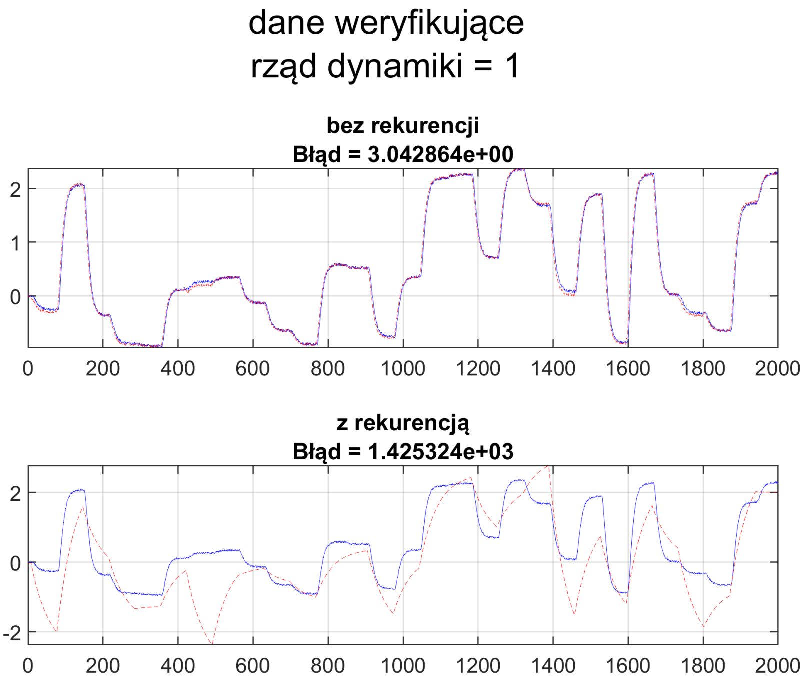
Tab. 2.1. Porównanie błędów dla różnych rzędów dynamiki

Jak widać z tabeli oraz rysunków poniżej model z rekurencją (OE) nie działa. Widzimy zależność, że wraz ze wzrostem dynamiki zmniejsza się błąd, co jest intuicyjne. Możliwe, że gdybyśmy go odpowiednio zwiększyli, to otrzymalibyśmy w pewnym momencie model dobrze odwzorowujący dany problem, jednak jest to tylko moje przypuszczenie. Patrząc na modele bez rekurencji widzimy, że naprawdę dobrze odwzorowują one dane - ich błędy wynoszą niewiele. Najlepszy okazał się model bez rekurencji (ARX) o rzędzie dynamiki równej 3.

Rząd dynamiki 1:

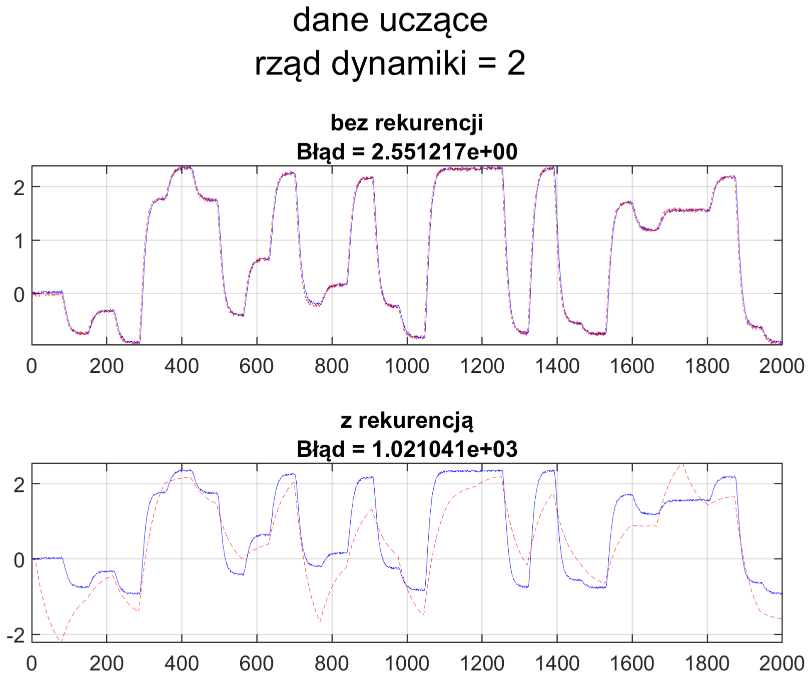


Rys. 19. dane uczace

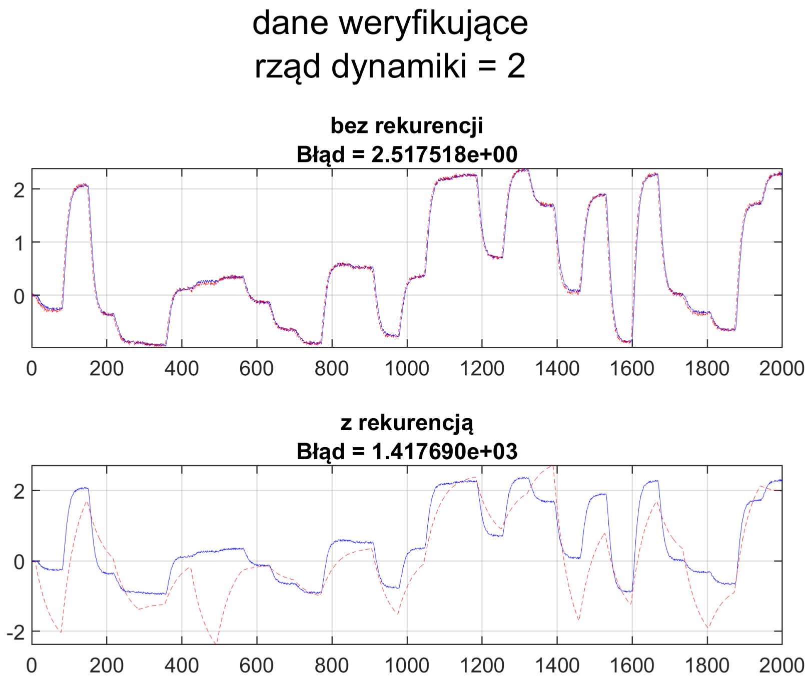


Rys. 20. dane weryfikujące

Rząd dynamiki 2:

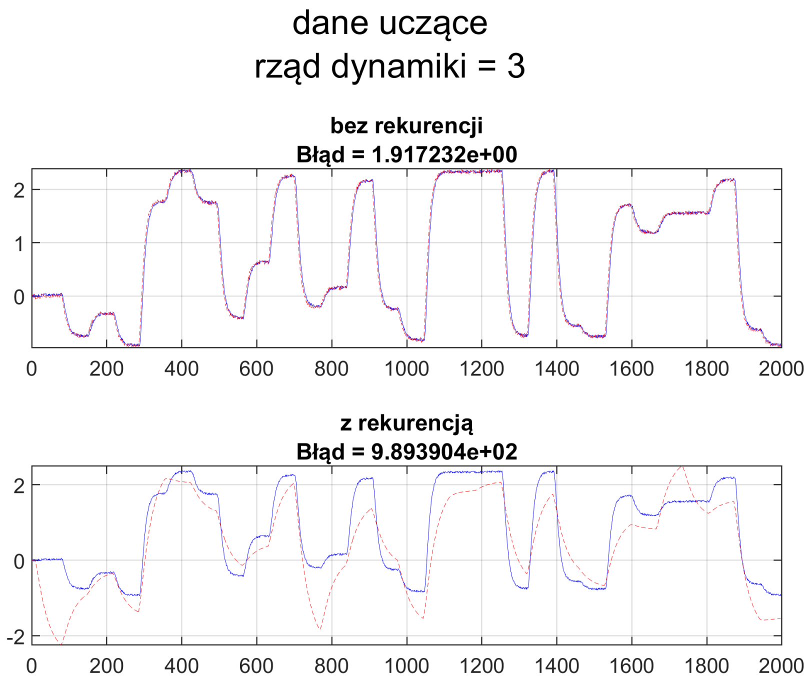


Rys. 21. dane uczace

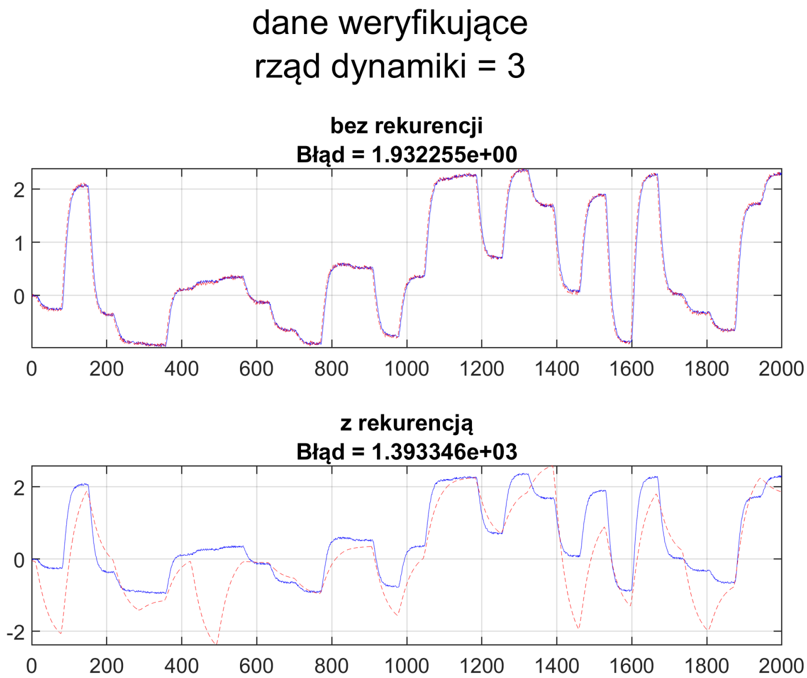


Rys. 22. dane weryfikujące

Rząd dynamiki = 3:



Rys. 23. dane uczace



Rys. 24. dane weryfikujące

c. Nieliniowy wielomianowy model dynamiczny

Zadaniem jest wyznaczenie metodą najmnieszych kwadratów dynamicznych modeli wielomianowych dynamicznych, czyli już z rodziny modeli nieliniowych. Badałem modele różnych rzędów dynamiki (od 1 do 3) oraz róznych stopni wielomianów (od 2 do 6), bez wyrazów mieszancyh.

Błąd liczyłem identycznie jak w poprzednim przypadku, czyli metodą najmniejszych kwadratów. Ponieważ badane układy nie przekraczały trzeciego rzędu dynamiki, to pierwszą próbką jest ta z indeksem 4. Wzór wygląda następująco:

$$E = \sum_{i=4}^{P} (y^{mod}(i) - y(i))^2$$

rząd dynamiki = 1

	Zbiór uczący		Zbiór weryfikujący—	
stopień wielomianu	bez rekurencji	z rekurencją	bez rekurencji	z rekurencją
2	2,81	$3,66 \cdot 10^3$	2,55	$2,53 \cdot 10^3$
3	1,88	$3,50 \cdot 10^3$	1,73	$2,44 \cdot 10^3$
4	1,85	$3,48 \cdot 10^3$	1,67	$2,43 \cdot 10^3$
5	1,84	$3,48 \cdot 10^3$	1,66	$2,43 \cdot 10^3$
6	1,84	$3,48 \cdot 10^3$	1,67	$2,43 \cdot 10^3$

Tab. 2.2. Porównanie błędów dla różnych stopni wielomianu przy rzędzie dynamiki 1

rząd dynamiki = 2

	Zbiór uczący		Zbiór weryfikujący—	
stopień wielomianu	bez rekurencji	z rekurencją	bez rekurencji	z rekurencją
2	2,40	$3,71 \cdot 10^3$	2,26	$2,58 \cdot 10^3$
3	1,39	$3,47 \cdot 10^3$	1,43	$2,42 \cdot 10^3$
4	1,36	$3,44 \cdot 10^3$	1,37	$2,40 \cdot 10^3$
5	1,35	$2,44 \cdot 10^3$	1,38	$2,40 \cdot 10^3$
6	1,35	$3,45 \cdot 10^3$	1,38	$2,40 \cdot 10^3$

Tab. 2.3. Porównanie błędów dla różnych stopni wielomianu przy rzędzie dynamiki 2

rząd dynamiki = 3

	Zbiór uczący		Zbiór weryfikujący—	
stopień wielomianu	bez rekurencji	z rekurencją	bez rekurencji	z rekurencją
2	1,85	$3,77 \cdot 10^3$	1,82	$2,62 \cdot 10^3$
3	1,17	$3,41 \cdot 10^3$	1,22	$2,38 \cdot 10^3$
4	1,13	$3,38 \cdot 10^3$	1,15	$2,36 \cdot 10^3$
5	1,11	$3,37 \cdot 10^3$	1,15	$2,35 \cdot 10^3$
6	1,10	$3,38 \cdot 10^3$	1,16	$2,35 \cdot 10^3$

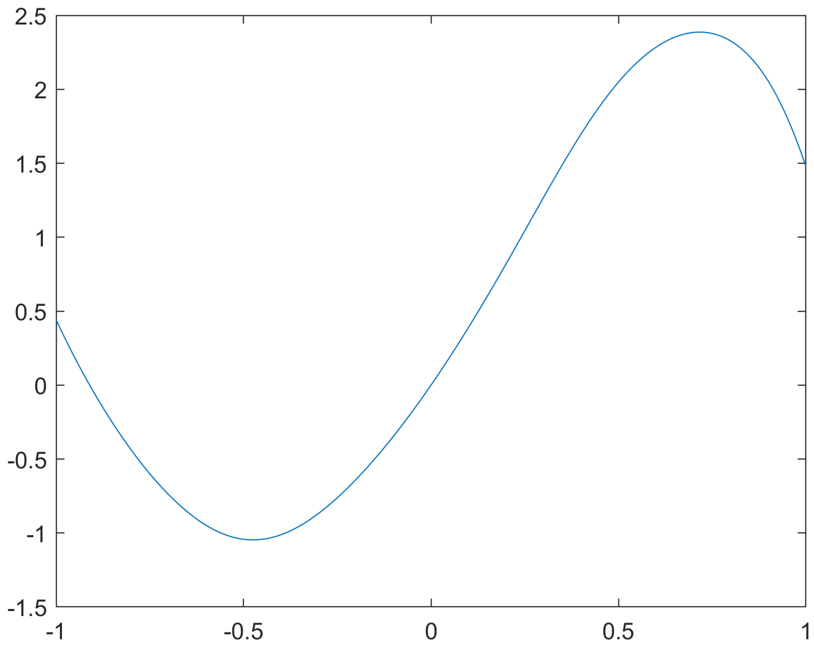
Tab. 2.4. Porównanie błędów dla różnych stopni wielomianu przy rzędzie dynamiki 3

Z powyższego doświadczenia wynika, że dla badanych danych optymalnym modelem jest wielomian 4 stopnia o rzędzie dynamiki równym 3. Dlatego dla niego zostanie przeprowadzona dalsza część projektu.

d. Nieliniowy model statyczny

Gdy mówimy o modelu statycznym, to jest to stan, gdy sygnały nie zmieniają się już w czasie. Możemy wtedy przyjąć, że wszystkie sygnały $u(k-i)$ są równe stałej wartości wartości u, analogicznie z sygnałami $y(k-1)$, które można przyjąć jako wartość y.

Do narysowania charakterystyki statycznej wykorzystałem funkcję `fsolve`. Jej wynikiem była następująca charakterystyka:



Rys. 25. charakterystyka statyczna wielomianu 4 stopnia o rzędzie dynamiki 3

Widać, że powyższa charakterystyka przypomina charakterystyki z rysunków 17 i 18, co wskazuje, że powyższy model wielomianowy zadziałał właściwie.