

Prof. dr hab. inż. Witold Pedrycz
członek zagraniczny PAN
Dept. of Electrical & Computer Engineering
University of Alberta, Edmonton, Canada
Instytut Badań Systemowych PAN

30 sierpnia 2017

Recenzja

pracy doktorskiej p. mgr Adama Marszałka

pt.

Skierowane liczby rozmyte w modelowaniu i symulacji finansowych szeregów czasowych

Recenzja pracy doktorskiej p. mgr Adama Marszałka przedstawiona jest w nawiązaniu do pisma Pana Sekretarza Rady Naukowej IPPT PAN Pana Prof. dr hab inż. Zbigniewa Ranachowskiego z dnia 30 czerwca 2017 powołującego mnie w roli recenzenta niniejszej pracy doktorskiej.

1. Tematyka rozprawy doktorskiej

Analiza szeregów czasowych i związane z nimi metody predykcji (prognozowania) stanowią jedno z centralnych zagadnień o charakterze teoretycznym i aplikacyjnym. Wraz z coraz większą liczbą danych eksperymentalnych i szerokim spektrum zastosowań obejmujących złożone procesy przemysłowe i ekonomiczne, pojawia się nagląca konieczność budowy zaawansowanych modeli predykcyjnych. Podczas konstrukcji tej klasy modeli rozważane są sposoby efektywnej reprezentacji danych tak aby osiągnąć

wymagany poziom abstrakcji, jakości predykcji oraz interpretowalności modelu. Ziarna informacji odgrywają tutaj istotną rolę. W tym sensie, użyty formalizm skierowanych liczb rozmytych stanowi interesującą i obiecującą alternatywę przetwarzania oryginalnych danych numerycznych szeregów czasowych.

Z powyższego punktu widzenia, wybór tematyki badawczej podjętej przez Doktoranta jest aktualny i w pełni uzasadniony.

2. Struktura rozprawy doktorskiej

Rozprawa doktorska liczy 124 strony, składa się z 8 rozdziałów, bibliografii, spisu rysunków i tabel oraz streszczenia (w języku polskim i angielskim). Rozdział 1 ma charakter wstępny i omawia cel i tezę pracy. Rozdział 2 dotyczy prezentacji podstawowych elementów teorii zbiorów rozmytych. Skierowane liczby rozmyte (definicje, operacje, reprezentacja numeryczna) stanowią treść Rozdziału 3. Reprezentacja danych szeregu czasowego realizowana za pomocą skierowanych liczb rozmytych omawiana jest w Rozdziale 4. Modele rozmyte szeregów czasowych (bazujące na klasycznych ujęciach wnioskowania rozmytego, rozmytej regresji, modelach Box-Jenkins) wraz z koncepcją skierowanych rozmytych szeregów czasowych omawiane są w Rozdziale 5. Aspekty symulacyjne rozmytych szeregów czasowych stanowią treść Rozdziału 6. Przykłady zastosowań skierowanych liczb rozmytych w problematyce analizy szeregów danych finansowych (wycena opcji europejskiej na indeks WIG20, dywersyfikacja portfela akcji, analiza techniczna rynków finansowych) zawarte są w Rozdziale 7. Podsumowanie i wnioski przedstawione są w Rozdziale 8.

Organizacja rozdziałów ma pełne uzasadnienie logiczne i strukturalizacja materiału jest właściwa.

Koncepcje rozwijane w poszczególnych rozdziałach są interesujące i mogą prowadzić do innowacyjnych rozwiązań modelowania i symulacji szeregów czasowych. Rozprawa jest pracą o dużym poziomie spójności przedstawiając

aktualny stan wiedzy w powyższym zakresie tematycznym jak również proponując nowe ujęcie wykorzystania skierowanych liczb rozmytych w analizie szeregów czasowych.

3. Oryginalność rozprawy

Głównymi elementami rozprawy o charakterze oryginalnym są:

- propozycja modeli szeregów czasowych o wysokiej częstotliwości z użyciem skierowanych zbiorów rozmytych,
- konstrukcja szczegółowych algorytmów predykcji wykorzystujących skierowane zbiory rozmyte,
- wprowadzenie koncepcji skierowanej rozmytej zmiennej losowej.

4. Punkty dyskusyjne

Nasuwać się tutaj następujące uwagi o charakterze krytycznym. Należy tutaj wyraźnie podkreślić, że nie mają one negatywnego wpływu na ogólnie pozytywną ocenę osiągnięć Doktoranta:

skierowane przedziały – skierowane liczby rozmyte stanowią modyfikacje liczb rozmytych. W kontekście badań przedstawionych w rozprawie, interesująca byłaby dodatkowa analiza skierowanych przedziałów wraz z ich rachunkiem.

propagacja ziarnistości - jest ogólnie wiadomym, że zasada rozszerzenia prowadzi do bardzo "konserwatywnych" wyników (tzn. o niskim poziomie szczegółowości). To samo zjawisko redukcji szczegółowości jest widoczne w przypadku analizy przedziałowej (*interval analysis*). W rzeczywistości, zasada rozszerzenia zaproponowana przez Zadeh w zbiorach rozmytych bazuje na koncepcjach przetwarzania danych przedziałowych. Wiadomym też jest że dane przedziałowe pojawiły się w związku z zastosowaniem obliczeń cyfrowych i opisem propagacji i kumulacji błędów w tych obliczeniach (np. praca Warmusa z lat 50tych). Opis kumulacji błędów ma charakter konserwatywny, co oczywiście prowadzi do bardzo szybkiej redukcji precyzji wyników (albo wzrostu poziomu

ziarnistości). Zbiory rozmyte nie dotyczą tej samej klasy problemów technicznych tak że propagacja ziarnistości nie musi podlegać tym samym zasadom, które stosują się do analizy przedziałowej (*interval mathematics*). Skierowane liczby rozmyte pozwalają na redukcje wzrostu poziomu ziarnistości tym niemniej nie mają wbudowanego mechanizmu kontroli i ewentualnej optymalizacji tego poziomu, który może być bezpośrednio związany z naturą rozwiązywanego problem i jakością uzyskanego rozwiązania.

Praca zawiera szereg aspektów oryginalnych, tym niemniej nie są one wyraźnie uwypuklone i czytelnik musi się w pewnym stopniu domyslać co stanowi oryginalne osiągnięcie Doktoranta.

Skierowane liczby rozmyte są ziarnami informacji – w tym sensie pojawia się interesujący aspekt kompresji (uogólnienia) danych numerycznych, a mianowicie jak ocenić jakość uogólnienia (abstrakcji) dostarczanego przez skierowane liczby rozmyte – np. w formie kryterium rekonstrukcji (dekompresji).

Eksperymenty – warte byłoby podkreślenie w sposób bardzo zwięzły co może oczekiwać użytkownik (nie będący zainteresowanym aspektami teoretycznym jak i algorytmicznymi). Jak w sposób zwarty przedstawić ocenę otrzymanych wyników vis-à-vis wyników otrzymanych z wykorzystaniem “standardowych” numerycznych modeli predykcyjnych?

Bibliografia jest obszerna, tym niemniej liczba cytowanych prac opublikowanych po roku 2010 jest bardzo ograniczona.

Praca przygotowana została z dużą starannością ze strony edycyjnej; drobne uchybienia są łatwe do eliminacji, np. podejście → ujęcie, czy też użycie niezbyt fortunnego nazewnictwa technicznego: defuzzifikacja, wyostrzenie → dekodowanie.

5. Podsumowanie

Podsumowując, uważam że główne cele badań sformułowane przez Doktoranta:

“możliwe jest przedstawienie danych finansowych wysokiej częstotliwości za pomocą skierowanych liczb rozmytych zachowując przy tym więcej informacji o danym szeregu czasowym”

oraz

“modele szeregów czasowych skonstruowane na bazie skierowanych liczb rozmytych pozwalają na efektywne modelowanie i symulacje finansowych szeregów czasowych oraz możliwe jest wykorzystanie ich w praktycznych zagadnieniach inżynierii finansowej”

zostały w bardzo dużym stopniu osiągnięte, a uzyskane wyniki są wartościowe, mają charakter oryginalny i stanowią ewidentny wkład w dziedzinę analizy i interpretacji szeregów czasowych.

Biorąc pod uwagę oryginalność rozprawy, w pełni osiągnięte cele badawcze, wszechstronną analizę, w tym też analizę o charakterze porównawczym oraz sposób prezentacji, uważam że spełnia ona wszystkie wymagania formalne określone art 13 ustawy o stopniach i tytule naukowym. Stawiam jednocześnie wniosek o wyróżnienie rozprawy doktorskiej ze względu na poziom oryginalności i dogłębność przeprowadzonych rozważań zarówno natury teoretycznej jak i eksperymentalnej. Chciałbym też podkreślić że Doktorant jest autorem publikacji w renomowanym czasopiśmie międzynarodowym (Information Sciences, Elsevier) oraz monografiach naukowych (Springer-Verlag).

