

Recenzja rozprawy doktorskiej

Anna Labijak-Kowalska

zatytułowanej:

New Directions in Robust Data Envelopment Analysis

1. Problem badawczy i jego znaczenie

Jaki jest najważniejszy problem rozważany w rozprawie?

Czy ma on charakter naukowy?

Czy ma on znaczenie praktyczne?

Data Envelopment Analysis (DEA) [5] jest metodą oceny efektywności jednostek decyzyjnych (DMU) zużywających wiele nakładów i w efekcie uzyskują wiele rezultatów. Od zaproponowania przez Charnes'a et al. (model CCR) [4] metoda DEA doczekała się wielu zastosowań obejmujących między innymi obszary i dziedziny takie jak bankowość, opieka zdrowotna, rolnictwo, edukacja, kultura itd. Z czasem rozwinęły się jej liczne rozszerzenia i modyfikacje, gdzie do najważniejszych należy zaliczyć model nadefektywności DEA (ang. super efficiency DEA) SE-CCR [2], nie-radialnej DEA (ang. non-radial DEA) NR-DEA [8], ich połączenie tj. SE-NR-DEA, czy model BCC [1]. W kontekście tej metody badacze zwykle nie podejmują w swoich pracach problemu odporności otrzymanych rezultatów, która w ostatecznym efekcie może być istotna przy podejmowaniu decyzji tj. wyborze najlepszego (najbardziej efektywnego) rozwiązania. Sytuacja taka tworzy możliwość wpisania się z nowymi pomysłami w tak rozumianą lukę badawczą.

W prezentowanej rozprawie, tj. cyklu siedmiu publikacji, autorka stawia sobie za główny cel zaproponowanie rozszerzeń metody DEA, których wspólnym zadaniem jest dostarczenie decydentowi dodatkowych informacji o badanych obiektach DMUs (od ang. decision-making units). Cel ten jest realizowany na wiele sposobów włączając w to:

- zaproponowanie nowej metodologii analizy odporności (ang. robustness) w metodzie DEA [P1, P2] oraz VDEA [P4]
- analizę działania różnych metod rankingowych używanych do oceny DMUs [P3]
- analizę odporności w kontekście zaproponowanych metod w sytuacji niepewności [P5, P6]
- i w końcu analizę odporności dla modelu, w którym wejścia i wyjścia (nakłady i rezultaty) tworzą strukturę hierarchiczną.

Jedną z własności metody DEA wymienianą w kontekście jej wad jest stosunkowo duża wrażliwość wyników na zaburzenia w wartościach opisujących obiekty wzorcowe (optymalnych) [7]. Analiza odporności stanowiąca główny temat rozprawy pozwala zmniejszyć niepożądane skutki tego zjawiska i aktywnie przed nimi się bronić. **Tym samym podjęcie tego tematu badawczego przez panią Annę Labijak-Kowalską trzeba uznać za ważne i w pełni uzasadnione.**

Analiza odporności to zagadnienie łączące w sobie zarówno elementy teoretyczne jak i praktyczne. Propozycja modelu, a w zasadzie modeli analizy odporności zaproponowane dla różnych wersji DEA ma charakter teoretyczny i naukowy. Modele jednak można w pełni ocenić tylko wtedy, gdy zostaną zasilone danymi obrazującymi ich działanie w praktyce. W prezentowanej rozprawie,

znajdujemy kilka obszernych przykładów wykorzystania zaproponowanej teorii dotyczących oceny efektywności specjalnych stref ekonomicznych, lekarzy oddziałów ratunkowych, polskich lotnisk czy bezpieczeństwa energetycznego na przykładzie wybranych krajów.

2. Wkład autora

Jaki jest najważniejszy wkład autora opisywane w rozprawie? Jeżeli niezbędne, recenzent może rozróżnić wkład deklarowany przez autora rozprawy i wkład, który sam recenzent uważa za najważniejszy. W tym wypadku należy podać powody, dla których recenzent nie zgadza się z twierdzeniami autora (np. ktoś wcześniej zaproponował już dany pomysł, lub jest on oryginalny ale niepoprawny z powodów opisanych w sekcji 3). Prosimy także o komentarze dotyczące praktyczności zaproponowanych rozwiązań (może być tak, że problem jest bardzo praktyczny, ale zaproponowane rozwiązanie nie). Jeżeli ma to zastosowanie, recenzent może odwołać się do innych wskaźników jakości (np. jakości publikacji, patentów autora, cytowań, wdrożonych zastosowań...).

Przedstawiona do oceny rozprawa składa się z siedmiu artykułów naukowych. W tekście pracy można je odnaleźć pod etykietami od P1 do P7. W większości prac udział autorski doktorantki jest deklarowany łącznie z udziałem promotora co uniemożliwia jednoznaczne określenie tzw. „wkładu autora”. Wyjątek stanowią prace oznaczone jako P2 i P4. W przypadku tej pierwszej wkład autorski pani Anny Labijak-Kowalskiej jest explicite wskazany jako:

- współautorstwo pomysłu zastosowania metod analizy odporności dla DEA w badaniu analizy odporności systemów elektroenergetycznych krajów,
- zastosowanie DEA opartej na proporcji w celu uzyskania zbiorów krajów efektywnych i nieefektywnych, projekcji na granice efektywności oraz (wyznaczenie) niezbędnych ulepszeń dla krajów (wskazanych) jako nieefektywne,
- zastosowanie metod analizy odporności do danych uwzględnionych w badaniu,
- wdrożenie oprogramowania niezbędnego w badaniu,
- autorstwo koncepcji i algorytmów służących do wyznaczania redukcji efektywności i kosztów,
- przygotowanie zaktualizowanych wyników dla trzech scenariuszy rozwoju krajów,
- konsultacje dotyczące przygotowania danych i interpretacji wyników,
- przeglądnięcie i korektę tekstu manuskryptu pod względem metodologii analizy

Wkład doktorantki w pracę P4 jest wprost wskazany jako:

- Współautorstwo pomysłu przystosowania zbioru metod i pojęć (z ang. framework'u) wykorzystywanych do analizy odporności na potrzeby modelu VDEA (DEA wykorzystującego funkcję wartości)
- autorstwo modeli matematycznych do analizy odporności dla modelu VDEA
- autorstwo procedury estymacji wskaźników akceptowalności z wykorzystaniem algorytmu HAR (Hit-And-Run) dla VDEA
- autorstwo proponowanych miar dla wielu scenariuszy analizy efektywności
- implementacja oprogramowania niezbędnego w badaniach opisanych w pracy P4,
- zastosowanie metod analizy odporności dla studium przypadku,
- analiza i dyskusja wyników,
- autorstwo pierwszego szkicu manuskryptu.

Wśród powyższego dla pracy P2 wysoko oceniam stworzenie algorytmu służącego do identyfikacji wszystkich reduktów wydajności oraz pokazanie zastosowania opracowanych rozwiązań

teoretycznych do przykładowych danych. Tu tymi danymi były wartości numeryczne wejść i wyjść do metody DEA wskazane w Tabeli 2 (strony 9 i 10 odnośnej pracy P2) zdefiniowane dla różnych krajów. Praca ta posiada również ciekawą dyskusję wyników której wnioski mogą mieć implikacje istotne z punktu widzenia polityki gospodarczej i społecznej wybranych krajów (np. zwiększenie odporności dostaw energii w Singapurze poprzez zwiększenie produkcji energii pochodzącej z paneli fotowoltaicznych itp.)

W pracy P4 na uwagę zasługuje zbudowanie modeli odporności dla modelu VDEA (Sekcja 2.2 Robustness analysis with mathematical programming). Budowa tych modeli wymagała dogłębnego zrozumienia samej metody VDEA, metod programowania liniowego oraz problemu badawczego związanego z analizą odporności.

W przypadku pozostałych prac, w których udział doktorantki należy rozpatrywać łącznie z pracą promotora na uwagę zasługuje między innymi praca P1 w której została zdefiniowana zintegrowana metoda (ang. framework) analizy odporności z wykorzystaniem modelu efektywności opartego o proporcję (ang. ratio). W publikacji P1 przeanalizowano wydajność 11 polskich lotnisk ocenianych na podstawie 4 danych wejściowych, takich jak roczna przepustowość terminala, maksymalna przepustowość, dynamiczna przepustowość płyty postojowej i obszar znaczenia biznesowego lotniska. Rozważane dane wyjściowe obejmują ruch pasażerski i roczną liczbę operacji (lądowań lub startów). W pracy zostały przedstawione wyniki analizy odporności w oparciu o zaproponowany model dla trzech różnych sytuacji: standardowej (tj. bez ograniczeń wagowych), liniowymi ograniczeniami wagowymi, oraz z lotniskami odstającymi (ang. outliers). Analiza wszystkich trzech konfiguracji problemu wskazała na przydatność zaproponowanej metody. Ponadto w pracach P3, P5 – P7 doktorantka wraz z promotorem mieli znaczący udział w powstaniu eksperymentalnego badania skupiającego się na porównaniu różnych (15 wybranych) wariantów metody DEA włączając w to efektywność krzyżową i super-efektywność, statystyki wielowymiarowe, analizę decyzji, benchmarking, wirtualne DMU itd.

W pracy P5 i P6 doktorantka wraz z promotorem zajmują się problemem pomiaru efektywności jednostek decyzyjnych za pomocą modelu opartego o proporcje (ang. ratio-based). W tym kontekście autorzy dostosowują koncepcję metody (ang. framework'u) analizy odporności, do różnego typu danych niepewnych takich np. jak przedziały liczbowe. W tym przypadku autorzy zajmują się metodami weryfikacji odporności wyników: efektywności, relacji preferencji efektywności i rang efektywności.

W pracy P7 doktorantka i promotor definiują strukturę hierarchiczną wejść i wyjść oraz nową metodologię dostępu. Takie podejście pozwala na zdekomponować dany problem na łatwe do zarządzania części i określić mocne i słabe strony analizowanych jednostek (DMUs). Otrzymane rozwiązanie jest solidne (ang. robust). Zaletą tego podejścia jest możliwość analizy węzłów na różnych stopniach hierarchii współczynników wejścia i wyjścia.

3. Poprawność

Czy stwierdzenia zawarte w rozprawie są godne zaufania? Czy uzasadnienia są poprawne? Wskaż zauważone słabości i błędy. Wskaż także te aspekty dotyczące poprawności, które są najbardziej wartościowe (elegancja dowodów, plan eksperymentów, analiza danych empirycznych, jakość prototypowego oprogramowania/sprzętu...).

Ponieważ przedstawiona rozprawa doktorska ma charakter zbioru artykułów opublikowanych albo przyjętych do publikacji to konieczność oceny poprawności rozprawy stawia recenzenta w niezręcznej sytuacji. Ponowne recenzowanie prac, które już zostały zrecenzowane w dobrych czasopiśmie ma wartość ograniczoną. Stąd oceniając poprawność pracy przede wszystkim skupię się na pierwszych

sześćdziesięciu stronach otrzymanego dokumentu, które stanowią coś w rodzaju streszczenia, wprowadzenia, podsumowania i być może uzupełnienia prac opublikowanych w czasopiśmie. Na potrzeby tej sekcji będę nazywał tę część rozprawy „wprowadzeniem”.

Wprowadzenie składa się z pięciu rozdziałów, gdzie oprócz wstępu (Rozdział 1), znajduje się wprowadzenie do metody DEA (Rozdział 2), opis zbioru narzędzi i metod (ang. framework) do analizy odporności w metodzie DEA (Rozdział 3), omówienie rozszerzeń zaproponowanych rozwiązań (Rozdział 4), oraz zakończenie (Rozdział 5).

W rozdziale 2 wprowadzającym w problematykę badawczą na uwagę i pochwałę zasługuje fakt, że doktorantka omawiając kolejno różne modele metody DEA trzyma się jednego, prostego przykładu demonstracyjnego. Takie podejście w istotny sposób ułatwia czytelnikowi zrozumienie istotnych różnic pomiędzy poszczególnymi modelami. Z dostrzeżonych niedociągnięć to omawiając metodę VDEA (str. 13) doktorantka wprowadza nowe oznaczenia w_q i u_q w miejsce dotychczasowych oznaczeń bez wyraźnego poinformowania o tym czytelnika. Czy nie można było zostawić dotychczasowych μ i v ?

Kolejny rozdział koncentruje się na opisie całości metod i pojęć związanych z analizą odporności. W tym rozdziale pojawia się szereg pojęć i oznaczeń wcześniej nie zdefiniowanych takich jak np. indeksy stochastyczne (ang. stochastic indices) (Str. 20), czy „pairwise efficiency outranking index”. Nawet zakładając, że jest to gdzieś później wyjaśnione w jednym z załączonych artykułów, w takim streszczeniu konieczne jest zamieścić choćby odpowiednią referencję z podaniem strony. Podobna uwaga dotyczy wprowadzania pozostałych pojęć takich jak np. Efficiency Acceptability Interval Index (str. 21) itd.

Przy okazji warto wspomnieć, że w trybie matematycznym LaTeX literały w rodzaju ERAI, PEOI itp. zamiast $\$ERAI\$$ lepiej zapisywać jako $\text{\textit{ERAI}}$ lub \textit{ERAI} , unikając w ten sposób zmiennych odstępów pomiędzy poszczególnymi znakami literału.

W kolejnym, czwartym rozdziale wprowadzenia, również autorka dość oszczędnie gospodaruje definicjami używanych pojęć. I tak na przykład w sekcji 4.2 zostają wprowadzone pojęcia „efficiency reducts” i „efficiency constructs”, których wyjaśnienie można znaleźć po przeszukaniu całości tekstu łącznie z publikacjami, ale nie w czytany właśnie wprowadzeniu. Podobnie skróty RDM i RAM (str. 41) powinny zostać rozwinięte i wyjaśnione w tekście wprowadzenia.

Oprócz RDM i RAM w rozdziale tym autorka posługuje się współczynnikiem korelacji Kendalla τ . Do jego wyznaczenia jest używana tzw. odległość Kendalla d_k (4.5). Ponieważ oryginalnie odległość Kendalla jest zdefiniowana w sposób dokładny dla rankingów bez remisów, a w przypadku pojawienia się takowych stosuje się zwykle odpowiednie aproksymacje i rozszerzenia oryginalnego pomysłu [3, 6], nie jest dla mnie jasne czy w tym miejscu albo charakter problemu wyklucza remisy w porównywanych rankingach, czy raczej jest tu zastosowane pewne uogólnienie oryginalnej odległości Kendalla.

Pewne usterki i niekonsekwencje, które można odnaleźć we „wprowadzeniu” do dysertacji, nie mają większego wpływu na ocenę całości przedstawionej pracy oraz osiągnięć pani Anny Labijak-Kowalskiej, która jest jednoznacznie pozytywna. Wyniki przedstawiane w pracy zostały w większości opublikowane w czasopiśmie z posiadającym oznaczony IF lub też zostały zaakceptowane do publikacji w czasopiśmie. Proponowanym rozwiązaniom teoretycznym towarzyszą przykłady numeryczne oraz przypadki studialne oparte o rzeczywiste dane. Analiza tych przykładów pozwala czytelnikowi z jednej strony lepiej zrozumieć proponowane rozwiązanie teoretyczne, z drugiej, zweryfikować zasadność oraz jego wartość praktyczną.

Przedstawione rozwiązania teoretyczne nie mogłyby się obyć bez odpowiedniego oprogramowania. W części przypadków oprogramowanie to zostało wykonane przez samą doktorantkę w języku R i udostępnione publicznie w serwisie github.com. Wchodzi ono również w skład otwartego projektu Decision Deck (<https://www.diviz.org>). Na uwagę zasługuje fakt opatrzenia większości procedur w języku R dużą ilością testów pozwalających kontrolować jakość powstałego oprogramowania.

Uwagi techniczne do „wprowadzenia”:

- Czy we fragmencie „the efficiency score with this weight (...)” (str. 21) nie chodziło raczej o “efficiency rank”?
- „propoerty” zmienić na „property” (str. 23).
- „when b_k is equal to 0” prościej „when b_k equals 0” albo jeszcze prościej „when b_k is 0” (str. 24)
- „required” zmienić na „required” (str. 24).
- „Having defined such permutation, we need to model the monotonicity of this permutation in mathematical programming.” (str. 28) – nie bardzo rozumiem na czym ma polegać tu monotoniczność permutacji.

4. Wiedza kandydata

Które z rozdziałów (lub sekcji w rozdziałach) rozprawy omawiają istniejący stan wiedzy i dzięki temu potwierdzają ogólny stan wiedzy kandydata w zakresie Informatyki? Jakie obszary tych dyscyplin zostały omówione w tych rozdziałach/sekcjach? Jaka jest opinia recenzenta o jakości tych rozdziałów sekcji? Jaka jest opinia recenzenta o bibliografii? Na ile bibliografia jest kompletna? Prosimy o podanie innych argumentów za lub przeciw stwierdzeniu, że kandydat posiada ogólną wiedzę w dyscyplinie Informatyka techniczna i telekomunikacja.

Pomimo skoncentrowania się na problemie odporności w metodzie DEA rozprawa pani Anny Labiak-Kowalskiej jest wielowątkowa. W przedstawionym zbiorze prac doktorantka zajmuje się zarówno klasycznym modelem CCR jak i jego rozszerzeniami takimi jak podejście BCC, metodą addytywną czy modelem VDEA. W każdej z prac znajduje się solidna część bibliograficzna, w ramach której autorzy prezentują szereg prac opisujących zarówno podstawy metody DEA jak i specyficzne jej modyfikacje. Na tym tle wyróżnia się praca P3, która w swojej istocie ma charakter artykułu przeglądowego uzupełnionego o ocenę prezentowanych rozwiązań w oparciu o przeprowadzone eksperymenty numeryczne. Udział doktorantki w powstaniu tej pracy w sposób jednoznaczny pokazuje jej szeroką wiedzę w obszarze badań operacyjnych i wielokryterialnych metod podejmowania decyzji.

Pani Anna jest również, jak to zostało wspomniane wcześniej, autorką szeregu algorytmów zaimplementowanych w języku R. Implementacja jest uzupełniona o zbiór testów pozwalających sprawdzić i na bieżąco kontrolować jakość i poprawność powstającego oprogramowania. Powstałe oprogramowanie, staranność jego przygotowania oraz integracja opracowanych rozwiązań w ramach projektu Decision Deck potwierdza praktyczne umiejętności doktorantki związane z wytwarzaniem oprogramowania naukowego. Zawarty w artykułach pseudo-kod proponowanych algorytmów wraz z jego analizą sugeruje znajomość przez doktorantkę algorytmiki, która stanowi jedną z podstaw współczesnej informatyki. Podsumowując, wydaje się uzasadnionym twierdzenie, iż doktorantka posiada dobrze podbudowaną teoretycznie ogólną wiedzę w zakresie dyscypliny informatyka techniczna i telekomunikacja.

5. Inne uwagi

Przedstawiona praca (w postaci zbioru artykułów) jest bardzo rozbudowana i wielowątkowa. Gdyby nie udział promotora we wszystkich wskazanych pracach oraz pewna trudność w oddzieleniu

osiągnięć promotora od osiągnięć doktorantki to przedstawiony zbiór publikacji śmiało mógłby posłużyć jak podstawa otrzymania kolejnego stopnia naukowego tj. stopnia doktora habilitowanego.

6. Podsumowanie

Biorąc pod uwagę opinie zaprezentowane w poprzednich punktach i wymagania zdefiniowane przez art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (z późniejszymi zmianami)¹ moja ocena rozprawy pod względem trzech podstawowych kryteriów jest następująca:

A. Czy rozprawa zawiera oryginalne rozwiązanie problemu naukowego? (wybierz jedną opcję stawiając znak X)

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

B. Czy po przeczytaniu rozprawy zgadzasz się, że kandydat posiada ogólną wiedzę teoretyczną w dyscyplinie Informatyka techniczna i telekomunikacja?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

C. Czy kandydat posiada umiejętność samodzielnego prowadzenia pracy naukowej?

<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zdecydowanie TAK	Raczej TAK	Trudno powiedzieć	Raczej NIE	Zdecydowanie NIE

Ponadto, biorąc pod uwagę szerokość przeprowadzonych badań oraz ilość opublikowanych artykułów rekomenduję wyróżnienie rozprawy doktorskiej.



Podpis

Literatura

1. Banker, R.D. et al.: Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*. 30, 9, 1078–1092 (1984).
<https://doi.org/10.1287/mnsc.30.9.1078>.
2. Banker, R.D., Chang, H.: The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research*. 175, 2, 1311–1320 (2006).
<https://doi.org/10.1016/j.ejor.2005.06.028>.
3. Brandenburg, F.J. et al.: Comparing and Aggregating Partial Orders with Kendall Tau Distances, Department of Informatics and Mathematics, University of Passau, 2011.

¹ <http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000276>

4. Charnes, A. et al.: Measuring the efficiency of decision making units. *European Journal of Operational Research*. 2, 6, 429–444 (1978). [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8).
5. Cooper, W.W. et al. eds: *Handbook on Data Envelopment Analysis*. (2011). https://doi.org/10.1007/978-1-4419-6151-8_13.
6. Fagin, R. et al.: Comparing and aggregating rankings with ties. In: *Proceedings of the twenty-third ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*. ACM (2004). <https://doi.org/10.1145/1055558.1055568>.
7. Guzik, B.: *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*. Wydawnictwo Uniwersytetu Ekonomicznego (2009).
8. Thanassoulis, E., Dyson, R.G.: Estimating preferred target input-output levels using data envelopment analysis. *European Journal of Operational Research*. 56, 1, 80–97 (1992). [https://doi.org/10.1016/0377-2217\(92\)90294-J](https://doi.org/10.1016/0377-2217(92)90294-J).