

Warsaw University of Life Sciences – SGGW
Institute of Economics and Finance
Department of Econometrics and Statistics

**QUANTITATIVE METHODS
IN ECONOMICS**

**METODY ILOŚCIOWE W BADANIACH
EKONOMICZNYCH**

Volume XXI, No. 4

Warsaw 2020

EDITORIAL BOARD

Editor-in-Chief: Bolesław Borkowski

Vice-Editor-in-Chief: Hanna Dudek

Managing Editor: Grzegorz Koszela

Theme Editors:

Econometrics: Bolesław Borkowski

Multidimensional Data Analysis: Wiesław Szczesny

Mathematical Economy: Zbigniew Binderman

Analysis of Labour Market: Joanna Landmesser

Financial Engineering: Monika Krawiec

Data Science: Michał Gostkowski

Theory of Risk: Marek Andrzej Kociński

Statistical Editor: Wojciech Zieliński

Technical Editors: Jolanta Kotlarska, Elżbieta Saganowska

Language Editor: Agata Cienkusz

Native Speaker: Yochanan Shachmurove

Editorial Assistant: Luiza Ochnio

SCIENTIFIC BOARD

Adnene Ajimi (University of Sousse, Tunisia)

Heni Boubaker (University of Sousse, Tunisia)

Vasily Dikussar (Doradnicyn Computing Centre of the Russian Academy of Sciences, Russia)

Peter Friedrich (University of Tartu, Estonia)

Paolo Gajo (University of Florence, Italy)

Agnieszka Gehringer (University of Göttingen, Germany)

Anna Maria Gil-Lafuente (University of Barcelona, Spain)

Jaime Gil-Lafuente (University of Barcelona, Spain)

Vasile Glavan (Moldova State University, Moldova)

Francesca Greselin (The University of Milano-Bicocca, Italy)

Ana Kapaj (Agriculture University of Tirana, Albania)

Jirawan Kitchaicharoen (Chiang Mai University, Thailand)

Yuriy Kondratenko (Black Sea State University, Ukraine)

Vassilis Kostoglou (Alexander Technological Educational Institute of Thessaloniki, Greece)

Karol Kukula (University of Agriculture in Krakow, Poland)

Kesra Nermend (University of Szczecin, Poland)

Nikolas N. Olenev (Doradnicyn Computing Centre of the Russian Academy of Sciences, Russia)

Alexander N. Prokopenya (Brest State Technical University, Belarus)

Yochanan Shachmurove (The City College of The City University of New York, USA)

Mirbulat B. Sikhov (al-Farabi Kazakh National University, Kazakhstan)

Marina Z. Solesvik (Nord University, Norway)

Ewa Syczewska (Warsaw School of Economics, Poland)

Achille Vernizzi (University of Milan, Italy)

Andrzej Wiatrak (University of Warsaw, Poland)

Dorota Witkowska (University of Lodz, Poland)

ISSN 2082-792X

e-ISSN 2543-8565

© Copyright by Department of Econometrics and Statistics WULS – SGGW
(Katedra Ekonometrii i Statystyki SGGW)

Warsaw 2020, Volume XXI, No. 4

The original version is the paper version

Journal homepage: qme.sggw.pl

Published by Warsaw University of Life Sciences Press

BADANIE PUNKTÓW ZWROTNYCH CEN ROPY NAFTOWEJ W LATACH 2015 - 2019

Monika Krawiec  <https://orcid.org/0000-0002-4765-244X>

Instytut Ekonomii i Finansów

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: monika_krawiec@sggw.edu.pl

Rafał Furman

e-mail: s178817@sggw.edu.pl

Streszczenie: Celem pracy była identyfikacja punktów zwrotnych cen ropy naftowej w latach 2015 – 2019, do czego zastosowano test Perrona. Następnie podjęto próbę wskazania wydarzeń o charakterze geopolitycznym, które mogły być przyczyną istotnych zmian w strukturze trendu. Wykorzystano w tym celu, publikowane na portalach biznesowych, komentarze analityków oraz raporty surowcowe. Na tej podstawie oceniono, że w większości przypadków na notowania ropy naftowej w badanym okresie najbardziej wpływały problemy z nadwyżką podaży ropy naftowej w Stanach Zjednoczonych. Raporty informujące o zwiększającej się liczbie funkcjonujących platform wiertniczych w Ameryce Północnej oraz niski popyt na rynku paliwowym często powodowały długotrwałe obniżki cen na giełdach New York Mercantile Exchange (NYMEX) i Intercontinental Exchange (ICE).

Słowa kluczowe: rynek ropy naftowej, punkt zwrotny, test Perrona, wydarzenia geopolityczne

JEL classification: C22, C58, Q31

WSTĘP

Ropa naftowa to, obok węgla kamiennego i gazu ziemnego, jeden z najważniejszych naturalnych surowców energetycznych, wydobywanych aktualnie na świecie. Wykorzystywana jest przede wszystkim do produkcji paliw płynnych, takich jak benzyna, paliwo lotnicze czy olej napędowy, ale także w produkcji asfaltu [Schofield 2007]. Ropa jest również strategicznym towarem, niezwykle istotnym dla bezpieczeństwa energetycznego większości krajów.

<https://doi.org/10.22630/MIBE.2020.21.4.17>

W ostatnim stuleciu ropa naftowa stała się wyjątkowo pożądanym produktem, niezbędnym dla rozwoju ekonomicznego i obecnie gospodarki wielu państw są istotnie uzależnione od zużycia tego surowca. W efekcie zmienność światowych cen ropy ma wpływ na wzrost gospodarczy i rozwój przemysłowy poszczególnych krajów [Chai i in. 2011].

Wśród czynników kształtujących ceny ropy naftowej, między innymi, należy wymienić poziom jej podaży, zapasy, rozmiary produkcji i konsumpcji energii, ryzyko geopolityczne, a także zmienność kursów walutowych, inwestycje spekulacyjne na rynku ropy, giełdowy handel instrumentami pochodnymi opartymi na cenach tego towaru, jak również zmienność cen na rynkach akcji [Płókarz 2013]. Istotny wpływ na ceny ropy naftowej ma też działalność inwestycyjna spółek wydobywczych i ich nakłady na rozwój nowych technologii, wolne moce produkcyjne i zdolności przerobowe koncernów naftowych, od których zależy podaż, a także wysokość marż rafineryjnych, zdolność magazynowania ropy, jak również koszty transportu i siły roboczej [Czech i in. 2019].

W ciągu kilku ostatnich dekad rynek ropy stał się największym światowym rynkiem towarowym, który przyciąga wielu uczestników. Są to nie tylko firmy zajmujące się wydobyciem ropy i producenci paliw, ale też instytucje finansowe (banki, fundusze inwestycyjne, fundusze emerytalne, firmy ubezpieczeniowe, fundusze hedgingowe) oraz inwestorzy indywidualni [Geman 2007]. W wyniku postępującego procesu finansjalizacji rynków towarowych, rynek ropy został zdominowany przez tzw. uczestników niekomercyjnych (*noncommercials*), utożsamianych z podmiotami osiągającymi zyski w wyniku spekulacji. Jednocześnie maleje znaczenie tzw. uczestników komercyjnych (*commercials*), angażujących się w transakcje na rynku towarowym ze względu na rzeczywistą potrzebę zabezpieczenia się przed ryzykiem zmiany ceny towarów [Falkowski 2011].

W oczywisty sposób uczestnicy transakcji na rynku ropy są zainteresowani monitorowaniem jej cen, które ekonomiści, poczynając od lat 70. XX wieku, postrzegają jako jedno z istotnych źródeł fluktuacji w gospodarce i impuls globalnych szoków cenowych. Wraz z ceną złota i kursami walut, cena ropy jest indykatorem najważniejszych procesów, zachodzących w światowej gospodarce. Informacje płynące z rynku ropy stanowią również podstawę krótko- i długoterminowych decyzji inwestycyjnych. Zmienność cen wpływa więc nie tylko na wzrost gospodarczy, ale również na kondycję firm branży paliwowej i energetycznej oraz dochody inwestorów. Dlatego tak ważne są badania funkcjonowania tego rynku. Istotnym ich aspektem jest próba wykrycia punktów zwrotnych. Można tego dokonać na przykład za pomocą testu Perrona [Perron 1989]. Celem niniejszej pracy była identyfikacja punktów zwrotnych cen ropy naftowej w latach 2015 – 2019 z wykorzystaniem właśnie tego testu. Dodatkowo podjęto próbę powiązania punktów zwrotnych, wskazanych przez test Perrona, z wydarzeniami o charakterze gospodarczym i politycznym, które mogły powodować istotne zmiany w strukturze trendu na rynku ropy.

METODYKA BADAŃ

Test Perrona [1989], który służy do badania pierwiastka jednostkowego i jest rozszerzeniem testu ADF (*augmented Dickey-Fuller test*), można wykorzystać do wykrywania istotnych zmian strukturalnych w trendzie deterministycznym szeregu czasowego. Wskazane momenty określa się mianem punktów zwrotnych. Badanie zmian w strukturze trendu liniowego może dotyczyć zmian stałej, współczynnika kierunkowego lub obu parametrów równocześnie, co zapisuje się w postaci odpowiednio sformułowanego modelu. Jeśli zakłada się zmianę wyłącznie wyrazu wolnego w funkcji trendu, to analizowany model ma postać:

$$y_t = \alpha_0 + \lambda_1 t + \lambda_2 D(U)_t + \lambda_3 D(TB)_t + \alpha_1 y_{t-1} + \sum_{i=1}^k c_i \Delta y_{t-i} + \epsilon_t, \quad (1)$$

gdzie: y_t – wartość obserwacji w okresie t ,

$\alpha_0, \alpha_1, \lambda_1, \lambda_2, \lambda_3, c_i$ – parametry modelu,

t – zmienna czasowa,

T_B – moment wystąpienia zmiany w strukturze trendu oraz

$$D(U)_t = \begin{cases} 1, & \text{dla } t > T_B \\ 0, & \text{dla } t \leq T_B \end{cases},$$

$$D(TB)_t = \begin{cases} 1, & \text{dla } t = T_B + 1 \\ 0, & \text{dla } t \neq T_B + 1 \end{cases}.$$

W tym przypadku testowaniu podlega następujący zestaw hipotez:

$$H_0: \lambda_1 = 0; \lambda_2 = 0; \lambda_3 \neq 0; \alpha_1 = 1$$

wobec

$$H_1: \lambda_1 \neq 0; \lambda_2 \neq 0; \lambda_3 = 0; \alpha_1 < 1.$$

Hipoteza zerowa zakłada istnienie pierwiastka jednostkowego ($\alpha_1=1$) oraz wystąpienie pojedynczej zmiany – tak zwanego „szoku”. Hipoteza alternatywna zakłada, że nie występuje pierwiastek jednostkowy (proces jest stacjonarny – $\alpha_1 < 1$) oraz, że zmiana nie ma charakteru szokowego, lecz dotyczy zmiany stałej w funkcji trendu.

Kiedy istnieje przypuszczenie, że uległ zmianie współczynnik kierunkowy w funkcji trendu, to testowaniu podlega model postaci:

$$y_t = \alpha_0 + \lambda_1 t + \lambda_2 D(U)_t + \lambda_4 D(T^*)_t + \alpha_1 y_{t-1} + \sum_{i=1}^k c_i \Delta y_{t-i} + \epsilon_t, \quad (2)$$

gdzie:

$$D(T^*)_t = \begin{cases} t - T_B, & \text{dla } t > T_B \\ 0, & \text{dla } t \leq T_B \end{cases}.$$

Dla tego modelu zespół hipotez formułuje się następująco:

$$H_0: \lambda_1 = 0; \lambda_2 \neq 0; \lambda_4 = 0; \alpha_1 = 1$$

wobec

$$H_1: \lambda_1 \neq 0; \lambda_2 = 0; \lambda_4 \neq 0; \alpha_1 < 1.$$

Hipoteza zerowa zakłada istnienie pierwiastka jednostkowego wobec H_1 , że proces nie posiada pierwiastka jednostkowego. Ponadto w hipotezie alternatywnej zakłada się, że zmiana współczynnika kierunkowego jest istotna.

Natomiast w celu sprawdzenia, czy w funkcji trendu zmianie ulegają jednocześnie współczynnik nachylenia oraz wyraz wolny, szacuje się równanie postaci:

$$y_t = \alpha_0 + \lambda_1 t + \lambda_2 D(U)_t + \lambda_3 D(T)_t + \lambda_4 D(TB)_t + \alpha_1 y_{t-1} + \sum_{i=1}^k c_i \Delta y_{t-i} + \epsilon_t, \quad (3)$$

$$\text{gdzie: } D(T)_t = \begin{cases} t, & \text{dla } t > T_B \\ 0, & \text{dla } t \leq T_B \end{cases}$$

W tym przypadku zespół hipotez stanowi połączenie hipotez sformułowanych dla modeli (1) i (2), czyli:

$$H_0: \lambda_1 = 0; \lambda_2 = 0; \lambda_3 = 0; \lambda_4 \neq 0; \alpha_1 = 1$$

wobec

$$H_1: \lambda_1 \neq 0; \lambda_2 \neq 0; \lambda_3 \neq 0; \lambda_4 = 0; \alpha_1 < 1.$$

Do weryfikacji prawdziwości hipotezy, że $\alpha_1=1$, Perron zaproponował sprawdzian testu postaci:

$$t^* = \frac{\hat{\alpha}_1 - 1}{S(\hat{\alpha}_1)}, \quad (4)$$

gdzie: $\hat{\alpha}_1$ – oszacowany parametr α_1 ,

$S(\hat{\alpha}_1)$ – standardowy błąd szacunku dla parametru α_1 .

Wartości krytyczne, podane przez Perrona, są odczytywane dla ustalonego poziomu istotności oraz dla parametru $\lambda = \frac{T_B}{T}$. Natomiast do weryfikacji istotności pozostałych parametrów wykorzystuje się wartości krytyczne z rozkładu t-Studenta. Przy tym istotną kwestią w teście Perrona jest wybór odpowiedniego momentu, w którym zakłada się występowanie zmiany strukturalnej. Można to ustalić na podstawie statystyki t-Studenta, oszacowanej dla zmiennej obrazującej moment zmiany: λ_2 dla pierwszego przypadku, λ_4 dla drugiego przypadku oraz λ_2, λ_4 – dla trzeciego przypadku. Należy wybrać ten punkt, dla którego wartość bezwzględna tej statystyki jest największa [Witkowska i in. 2012].

WYNIKI BADAŃ EMPIRYCZNYCH

Prezentowane w tej części wyniki badań empirycznych stanowią wycinek szerszych analiz, przeprowadzonych w ramach przygotowania pracy magisterskiej, napisanej przez R. Furmana pod kierunkiem M. Krawiec [Furman 2020]. Ich podstawę stanowiły dzienne notowania na zamknięcie kontraktów terminowych na ropę naftową West Texas Intermediate (WTI) i Brent w okresie od 02.01.2015 r. do 31.12.2019 r. w USD/baryłkę, publikowane przez serwis internetowy Stooq (www.stooq.pl). WTI to lekka ropa naftowa, wydobywana w Stanach Zjednoczonych. Charakteryzuje się wysoką jakością oraz niską zawartością siarki. Wykorzystywana jest do produkcji paliw i benzyny, przy czym uzyskuje się z niej więcej oleju napędowego i benzyny niż z ropy Brent. Z kolei Brent to lekka ropa wydobywana w Europie (na Morzu Północnym), o niskiej zawartości siarki, również wykorzystywana do produkcji paliw i benzyny. WTI jest notowana na New York Mercantile Exchange (NYMEX), podczas gdy Brent – na Intercontinental Exchange (ICE) z siedzibą w Londynie [Kowalik, Herczakowska 2010].

W ramach wstępnej analizy graficznej przebiegu kształtowania się cen ropy naftowej WTI wskazano 35 okresów, które mogły zawierać punkty zwrotne, czyli momenty istotnie wpływające na zmianę trendu w szeregu. Dla ropy Brent otrzymano 37 hipotetycznych okresów, w których mogły wystąpić zmiany strukturalne. W większości przypadków przypuszczalne momenty zwrotne obu szeregów cen ropy (WTI i Brent) pokrywały się lub były przesunięte o kilka dni. Następnie, w celu wykrycia istotnych zmian w strukturze trendu rozpatrywanych szeregów czasowych, posłużono się testem Perrona. Dla każdego wyznaczonego okresu przeprowadzono test stacjonarności z uwzględnieniem zmiany trendu we wcześniej wskazanym hipotetycznym momencie zwrotnym oraz oszacowano modele, które miały na celu uzyskanie odpowiedzi na następujące pytania:

- Czy zmiana wyrazu wolnego w funkcji trendu jest istotna?
- Czy zmiana współczynnika kierunkowego funkcji trendu jest istotna?
- Czy zmiana wyrazu wolnego i współczynnika kierunkowego w funkcji trendu jest istotna?

W sytuacji, gdy w badanym okresie nie odrzucono żadnej z hipotez zerowych na poziomie istotności $\alpha=0,05$, podejmowano próbę wydłużenia lub skrócenia rozpatrywanego podokresu oraz testowano inne przypuszczalne momenty zwrotne, aby wykryć ich jak najwięcej. W testach przyjęto stałą liczbę opóźnień, równą 1.

Na podstawie przeprowadzonych testów wykryto 19 punktów zwrotnych w szeregu cen ropy naftowej WTI oraz 17 punktów zwrotnych w szeregu cen ropy naftowej Brent. Zaobserwowano, że jedynie 10 z hipotetycznych punktów zwrotnych, zidentyfikowanych we wstępnym etapie na podstawie analizy graficznej, wskazał test Perrona. W tabelach 1 i 2, odpowiednio dla ropy WTI i Brent, podano syntetyczną informację na temat analizowanych podokresów, wykrytych punktów zwrotnych i postaci testowanych modeli (zgodnie z oznaczeniami przyjętymi w

opisie metodyki badań). Ze względu na ograniczoną objętość artykułu nie zamieszczono szczegółowych wyników estymacji. Na rysunku 1 zilustrowano przebieg cen obu gatunków ropy w badanym okresie i wykryte punkty zwrotne.

Tabela 1. Wyniki testowania występowania zmian strukturalnych testem Perrona dla cen ropy WTI

Początek podokresu	Koniec podokresu	Punkt zwrotny	Testowany model
02.01.2015	17.02.2015	29.01.2015	(2)
29.01.2015	30.04.2015	19.03.2015	(2)
19.03.2015	23.06.2015	30.04.2015	(3)
30.04.2015	24.08.2015	23.06.2015	(3)
23.06.2015	08.10.2015	25.08.2015	(3)
25.08.2015	01.02.2016	08.10.2015	(3)
09.10.2015	21.03.2016	11.02.2016	(2)
11.02.2016	06.04.2016	21.03.2016	(2)
21.03.2016	08.06.2016	06.04.2016	(1)
06.04.2016	16.06.2016	08.06.2016	(2)
12.07.2016	23.08.2016	29.07.2016	(2)
23.05.2017	04.07.2017	21.06.2017	(3)
17.09.2018	30.10.2018	03.10.2018	(3)
03.10.2018	24.12.2018	29.11.2018	(1)
29.11.2018	23.04.2019	28.12.2018	(3)
28.12.2018	12.06.2019	24.04.2019	(2)
12.07.2019	16.09.2019	07.08.2019	(2)
07.08.2019	03.10.2019	16.09.2019	(3)
16.09.2019	31.12.2019	03.10.2019	(2)

Źródło: opracowanie własne

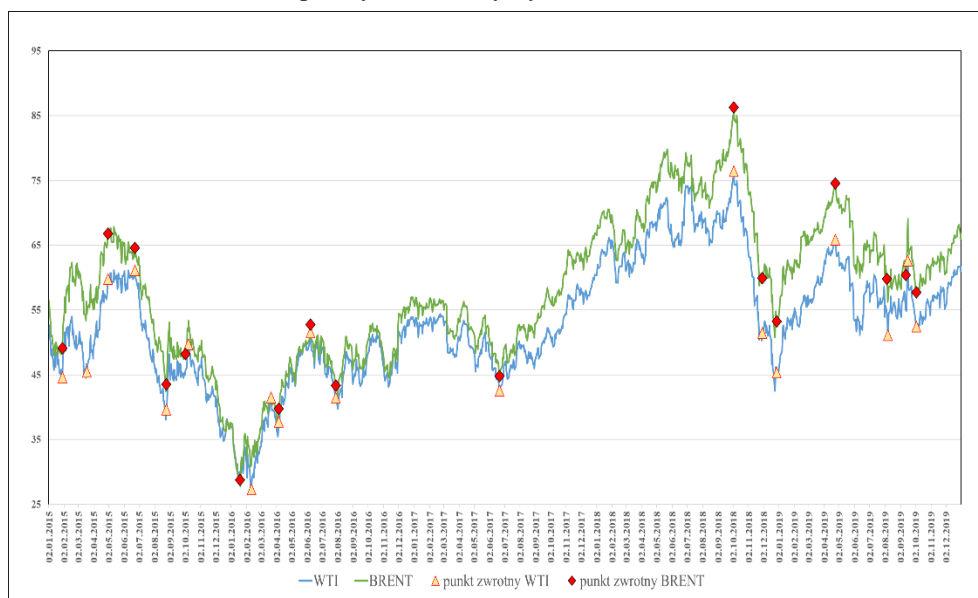
Tabela 2. Wyniki testowania występowania zmian strukturalnych testem Perrona dla cen ropy Brent

Początek podokresu	Koniec podokresu	Punkt zwrotny	Testowany model
02.01.2015	17.02.2015	29.01.2015	(2)
19.03.2015	23.06.2015	30.04.2015	(3)
30.04.2015	24.08.2015	23.06.2015	(3)
23.06.2015	08.10.2015	25.08.2015	(3)
25.08.2015	18.01.2016	02.10.2015	(3)
24.12.2015	01.02.2016	19.01.2016	(2)
21.03.2016	08.06.2016	06.04.2016	(1)
06.04.2016	16.06.2016	08.06.2016	(2)
12.07.2016	23.08.2016	29.07.2016	(2)
23.05.2017	04.07.2017	21.06.2017	(3)
17.09.2018	30.10.2018	03.10.2018	(3)
03.10.2018	21.12.2018	29.11.2018	(1)
29.11.2018	23.04.2019	28.12.2018	(3)
28.12.2018	10.06.2019	24.04.2019	(2)
12.07.2019	13.09.2019	05.08.2019	(3)
07.08.2019	03.10.2019	12.09.2019	(3)
16.09.2019	31.12.2019	03.10.2019	(2)

Źródło: opracowanie własne

Analizując szczegółowe wyniki testu Perrona, można było zauważyć, że tylko w czterech przypadkach zmiana strukturalna wyrazu wolnego w funkcji trendu była istotna. Jednakże zmiany te dotyczyły tych samych momentów zwrotnych (06.04.2016 r. i 29.11.2018 r.) dla obu badanych szeregów. Dla cen ropy naftowej WTI oszacowano 9 modeli, których zmiany dotyczyły wyłącznie współczynnika kierunkowego, a w przypadku cen ropy naftowej Brent uzyskano 6 takich punktów zwrotnych. Wszystkie punkty zwrotne, w przypadku których zmiana wyrazu wolnego i współczynnika kierunkowego w strukturze funkcji trendu była istotna (z wyjątkiem momentu zwrotnego w dniu 07.08.2019 r.), występowały równocześnie dla ropy amerykańskiej i europejskiej. W ramach przeprowadzonych badań podjęto również próbę identyfikacji wydarzeń, które spowodowały wystąpienie punktów zwrotnych w analizowanym okresie. W tym celu wykorzystano komentarze analityków rynkowych, raporty surowcowe i artykuły publikowane na portalach biznesowych.

Rysunek 1. Notowania ropy naftowej WTI i Brent w USD/baryłkę od 02.01.2015 r. do 31.12.2019 r. i punkty zwrotne wykryte testem Perrona



Źródło: Furman [2020], s. 47

Pierwszym wykrytym punktem zwrotnym był punkt w dniu 29.01.2015 r. W tym czasie największy wpływ na notowania cen ropy naftowej miała śmierć króla Arabii Saudyjskiej – Abdullaha. Podczas kolejnej sesji na giełdach towarowych odnotowano wzrosty cen ropy naftowej, ponieważ Arabia Saudyjska była wówczas największym producentem ropy w kartelu OPEC i przyszłe zmiany we władzy mogły wpłynąć na zmianę polityki produkcji ropy naftowej. Do zmian na rynku ropy

w tym okresie przyczyniły się również rosnące zapasy ropy w Stanach Zjednoczonych oraz mniejsza produkcja w Libii [Sierakowska 2015a].

Kolejny punkt zwrotny, wskazany przez test Perrona dla ropy WTI, przypadł na 19.03.2015 r. Przyczyną wzrostu notowań ropy okazało się oświadczenie Rezerwy Federalnej, dotyczące możliwości przyspieszenia terminu podwyżki stóp procentowych, które spowodowało spadek wartości dolara [Krishnan 2015]. Ponadto słabszy dolar podniósł atrakcyjność cen ropy naftowej dla posiadaczy innych walut.

Analizując rysunek 1, można zauważyć stagnację cen amerykańskiej ropy po 30.04.2015 r. Toczące się walki w Iraku i działania wojenne w Jemenie nie miały dużego wpływu na notowania cen ropy WTI, natomiast napięta sytuacja polityczna na Bliskim Wschodzie doprowadziła do spadku notowań ropy europejskiej Brent [Sierakowska 2015b].

Następnym momentem zmiany w strukturze trendu badanych szeregów, wykrytym testem Perrona, był dzień 23.06.2015 r. Za przyczynę spadku notowań ropy naftowej na kolejnych sesjach uznaje się:

- opublikowanie raportu przez Departament Energii Stanów Zjednoczonych, w którym podano, że zapasy ropy naftowej w USA zwiększyły się o 2,4 miliona baryłek [Sierakowska 2015c],
- pierwsze od 29 tygodni zwiększenie liczby czynnych platform wiertniczych w USA [Sierakowska 2015d].

Kolejny moment zwrotny na rynku ropy naftowej przypadł na dzień 25.08.2015 r. Wzrost cen ropy naftowej w kolejnych dniach był „technicznym odreagowaniem po wcześniejszej serii spadków”, która była wywołana paniką wśród inwestorów po spadku wartości chińskich indeksów akcji [Sierakowska 2015e]. Do wzrostu notowań ropy przyczyniły się również informacje o znaczącym spadku zapasów ropy w Stanach Zjednoczonych [Sierakowska 2015f].

Następnym momentem zmiany w strukturze trendu szeregu cen ropy naftowej, notowanej na nowojorskiej giełdzie, był 11.02.2016 r. Był to czas, kiedy państwa należące do kartelu OPEC oraz Rosja, w ramach porozumienia w Doha, uzgodniły, że przestaną wydobywać ropę naftową, aby zapobiec kolejnym spadkom notowań ropy [Dziennik Gazeta Prawna 2016a]. Ta decyzja miała się przyczynić do zmniejszenia podaży na rynku ropy naftowej.

W wyniku przeprowadzonych testów występowania zmian strukturalnych w rozpatrywanych szeregach, wskazano kolejny punkt zwrotny – dzień 21.03.2016 r., który jednak dotyczył tylko notowań ropy WTI. Po tym dniu, ceny, zarówno amerykańskiej, jak i europejskiej ropy, spadły. Przyczyny niższych cen upatrywano w reakcji inwestorów po wypowiedzi saudyjskiego księcia, który oznajmił, iż Arabia Saudyjska nie przestanie produkować ropy, dopóki Iran nie zgodzi się dołączyć do planu ograniczenia wydobycia ropy naftowej [Krishnan 2016]. Drugim powodem spadku notowań ropy była informacja, wynikająca z raportu Baker Hughes Inc., o zwiększeniu liczby funkcjonujących platform wiertniczych w Stanach

Zjednoczonych, co miało spowodować jeszcze większą nadpodaż na rynku paliwowym [Dziennik Gazeta Prawna 2016b].

Według Steve'a Chiavarone, jedną z przyczyn nagłego wzrostu cen ropy naftowej po kolejnym punkcie zwrotnym, przypadającym na dzień 06.04.2016 r., było osłabienie się dolara na skutek podniesienia stóp procentowych przez Rezerwę Federalną [Los Angeles Times 2016]. Innym powodem wzrostu cen ropy był optymizm inwestorów, uwarunkowany nadchodzącym spotkaniem w Doha [Sierakowska 2016a]. W przypadku zawarcia porozumienia w sprawie ograniczenia produkcji ropy naftowej, podczas posiedzenia w stolicy Kataru, inwestorzy spodziewali się wzrostu popytu na rynku paliwowym.

Następny punkt zwrotny, wskazany przez test Perrona, przypadł na dzień 08.06.2016 r. W tym dniu notowania cen ropy naftowej amerykańskiej i europejskiej po raz pierwszy w 2016 roku przekroczyły barierę 50 USD za baryłkę, ponieważ poziom zapasów ropy w USA był najniższy od 01.04.2016 r. [Dziennik Gazeta Prawna 2016c]. Jednakże obawy inwestorów związane z wystąpieniem Wielkiej Brytanii z Unii Europejskiej oraz informacje ujawnione w raporcie Amerykańskiego Instytutu Paliw o wzroście rezerw ropy naftowej w Stanach Zjednoczonych (gdy wcześniej zapowiadano ich spadek) spowodowały zniżki notowań ropy naftowej na kolejnych sesjach [Sierakowska 2016b].

Ostatnim momentem zwrotnym w 2016 roku był dzień 29 lipca. W pierwszych dniach sierpnia notowania ropy naftowej odbiły w górę, co wiązało się z opublikowaniem danych, przygotowanych przez amerykański Departament Energii, o większym spadku zapasów benzyny w Stanach Zjednoczonych niż oczekiwano [Sierakowska 2016c].

Wyniki przeprowadzonych badań pokazały, że jedyną zmianą w strukturze trendu badanych szeregów w 2017 roku był 21 czerwca. W tym czasie amerykańska Agencja Informacji Energetycznej (EIA) podała, że zapasy ropy naftowej w USA spadły o 2,5 mln baryłek w ciągu tygodnia, czyli o 400 tysięcy baryłek więcej niż wcześniej zakładano. Informacja ta była przyczyną zwyżek notowań ropy naftowej w kolejnych dniach [Reuters 2017].

Pierwszy w 2018 roku punkt zwrotny na rynku ropy naftowej, wskazany przez test Perrona, przypadł na 3 października. Tego dnia ropa naftowa na giełdzie w Nowym Jorku oraz w Londynie osiągnęła maksymalną cenę w całym okresie badania. Natomiast podczas kolejnych sesji notowania ropy naftowej, zarówno amerykańskiej, jak i europejskiej, zniżkowały. Analitycy stwierdzili, że przyczyną spadku wartości ropy było sekretne porozumienie Rosji i Arabii Saudyjskiej, na mocy którego wspomniane państwa zwiększyły wydobycie ropy w celu uzupełnienia luki produkcyjnej po Iranie, który musiał ograniczyć produkcję ropy ze względu na sankcje nałożone przez USA [Sierakowska 2018a]. Według analityka Rakuten Securities Inc. – Satoru Yoshidy, do zniżek cen ropy naftowej przyczyniły się również, nieustannie rosnące, rezerwy ropy w Stanach Zjednoczonych i zmniejszający się popyt [Dziennik Gazeta Prawna 2018].

Następnym dniem zmian w strukturze trendu rozpatrywanych szeregów był 29.11.2018 r. Kolejnego dnia doszło do spotkania przedstawicieli grupy G20, na którym Stany Zjednoczone oraz Chiny zobowiązały się nie podnosić ceł na import przez co najmniej 90 dni. Ceny ropy zwyżkowały w tym okresie, ponieważ inwestorzy oczekiwali na rezultaty obrad kartelu OPEC w Wiedniu [Sierakowska 2018b]. Państwa OPEC, wraz z Rosją, doszły do porozumienia w planowaniu zmniejszenia produkcji ropy naftowej. Jednak ustalenia z 6 grudnia negatywnie wpłynęły na notowania ropy, gdyż kilka dni później „minister energii Rosji, Aleksander Nowak, ogłosił, że Rosja planuje ograniczyć produkcję ropy naftowej w styczniu o około 50-60 tys. baryłek dziennie, co jest znacznie mniejszą liczbą niż ta zakładana w niedawnym porozumieniu z OPEC” [Sierakowska 2018c].

Przeprowadzone testy występowania zmian strukturalnych w trendzie badanych szeregów wskazały również punkt zwrotny w dniu 24.04.2019 r. Stwierdzono, że przyczyną zmiany trendu w tym okresie był zamiar zaostrenia przez rząd USA, od 1 maja, istniejących już sankcji, nałożonych na Iran. Prezydent Donald Trump chciał w ten sposób zredukować eksport ropy z Iranu do zerowego poziomu [Kozieł 2019]. Zapowiadane sankcje spowodowały zwiększenie importu ropy naftowej do Chin. Zaostrzający się konflikt handlowy między Stanami Zjednoczonymi i Państwem Środka, najwyższe rezerwy ropy w USA od września 2017 roku oraz wysokie wydobycie doprowadziły do spadku notowań ropy [Dziennik Gazeta Prawna 2019a].

W sierpniu 2019 roku test Perrona wskazał osobne punkty zwrotne dla dwóch rozpatrywanych szeregów. Dla cen ropy naftowej WTI moment zwrotny przypadł na 7 sierpnia, a dla cen ropy naftowej Brent – na 5 sierpnia. W tym okresie Arabia Saudyjska poinformowała o zamiarze ograniczenia produkcji ropy naftowej we wrześniu, a minister energii Zjednoczonych Emiratów Arabskich zakomunikował za pośrednictwem Twittera, że „jego kraj będzie popierał wszelkie inicjatywy zmierzające do ustabilizowania sytuacji na rynku ropy naftowej” [Sierakowska 2019].

Ostatnim wydarzeniem, stanowiącym punkt zwrotny na rynku ropy naftowej wykryty przez test Perrona, był atak dronów na saudyjskie rafinerie w Bukajk (Abqaiq) i Churajs (Khurais), który miał miejsce 14.09.2019 r. Natomiast dokładne daty zmian strukturalnych przypadły na 12.09.2019 r. (dla cen ropy europejskiej) oraz 16.09.2019 r. (dla cen ropy amerykańskiej). Na skutek zamachu doszło do rozległych pożarów, przez które Arabia Saudyjska musiała wstrzymać produkcję około 5,7 mln baryłek ropy dziennie. Powodem ataku było wspieranie przez Arabię Saudyjską i USA sił rządowych w Jemenie, które toczyły walki z rebelianckim ruchem Huti, odpowiedzialnym za ostrzał saudyjskich rafinerii [Dziennik Gazeta Prawna 2019b].

Dla kilku punktów zwrotnych, wykrytych przez test Perrona: 08.10.2015 r., 19.01.2016 r., 28.12.2018 r. oraz 03.10.2019 r., nie udało się wskazać wydarzeń o charakterze geopolitycznym lub ekonomicznym, wyjaśniających ich występowanie w badanych szeregach czasowych.

PODSUMOWANIE

Celem niniejszej pracy była identyfikacja punktów zwrotnych cen ropy naftowej w latach 2015 – 2019, do czego zastosowano test Perrona. Następnie podjęto próbę wskazania wydarzeń o charakterze geopolitycznym, które mogły być przyczyną istotnych zmian w strukturze trendu. Wykorzystano w tym celu, publikowane na portalach biznesowych, komentarze analityków oraz raporty surowcowe. Na tej podstawie oceniono, że w większości przypadków na notowania ropy naftowej w badanym okresie najbardziej wpływały problemy z nadwyżką podaży ropy naftowej w Stanach Zjednoczonych. Raporty informujące o zwiększającej się liczbie funkcjonujących platform wiertniczych w Ameryce Północnej oraz niski popyt na rynku paliwowym często powodowały długotrwałe obniżki cen na giełdach NYMEX i ICE.

Według komentarzy analityków giełdowych, istotnymi czynnikami, mającymi wpływ na notowania ropy naftowej w latach 2015 – 2019, były takie wydarzenia, jak:

- śmierć króla Arabii Saudyjskiej Abdullaha (23.01.2015 r.),
- sankcje nałożone na Iran przez USA w 2018 i 2019 roku,
- wojna handlowa między Stanami Zjednoczonymi a Chinami,
- obrady kartelu OPEC z Rosją w celu ograniczenia produkcji ropy,
- spotkanie grupy G20 (30.11.2018 r.),
- podnoszenie stóp procentowych przez Rezerwę Federalną.

Ewentualne dalsze badania w tym zakresie można by uzupełnić o notowania cen ropy naftowej z państw kartelu OPEC i Rosji. Ponadto zasadne byłoby wykorzystanie nowoczesnych narzędzi analitycznych w postaci zautomatyzowanych testów.

BIBLIOGRAFIA

- Chai J., Guo J.E., Meng L., Wang S.Y. (2011) Exploring the Core Factors and Its Dynamic Effects on Oil Price: An Application of Path Analysis and BVAR-TVP Model. *Energy Policy*, 39, 8028-8036.
- Czech K., Górská A., Koziół-Kaczorek D. (2019) Związki cenowe towarów w warunkach finansjeryzacji gospodarki na przykładzie cen ropy naftowej, złota i pszenicy. Wydawnictwo SGGW, Warszawa.
- Dziennik Gazeta Prawna (2016a) Katar: zamrożone wydobywanie ropy naftowej. 16.02.2016.
<https://finanse.gazetaprawna.pl/artykuly/923195,ropa-naftowa-zamrozone-wydobywanie-katar.html>

- Dziennik Gazeta Prawna (2016b) USA: ropa znów tanieje, bo rośnie liczba jej odwiertów. 21.03.2016.
<https://finanse.gazetaprawna.pl/artykuly/928979,ceny-ropy-usa.html>
- Dziennik Gazeta Prawna (2016c) Ropa w USA zakończy tydzień największym wzrostem ceny od 6 tygodni. 10.06.2016.
<https://finanse.gazetaprawna.pl/artykuly/950879,ropa-w-usa-zakoczy-tydzien-najwiekszym-wzrostem-ceny-od-6-tygodni.html>
- Dziennik Gazeta Prawna (2018) Na rynkach ropy wyprzedaż podobnie jak na globalnych rynkach akcji. 25.10.2018.
<https://finanse.gazetaprawna.pl/artykuly/1318265,na-rynkach-ropy-wyprzedaz.html>
- Dziennik Gazeta Prawna (2019a) Inwestorzy stracili wiarę, że możliwe są rajdy na rynku ropy. Surowiec silnie traci. 06.05.2019.
<https://finanse.gazetaprawna.pl/artykuly/1410726,inwestorzy-stracili-wiare-ze-mozliwe-sa-rajdy-na-rynku-ropy-surowiec-silnie-traci.html>
- Dziennik Gazeta Prawna (2019b) Spadki na nowojorskich giełdach po atakach na saudyjskie rafinerie. 16.09.2019.
<https://finanse.gazetaprawna.pl/artykuly/1430365,spadki-na-nowojorskich-gieldach-po-atakach-na-saudijskie-rafinerie.html>
- Falkowski M. (2011) Financialization of Commodities. *Contemporary Economics*, 5(4), 4-17.
- Furman R. (2020) Analiza punktów zwrotnych na rynku ropy naftowej w latach 2015-2020. Praca magisterska napisana pod kierunkiem dr hab. Moniki Krawiec, Wydział Zastosowań Informatyki i Matematyki, SGGW w Warszawie.
- Geman H. (2007) *Commodities and Commodity Derivatives*. John Wiley&Sons, Hoboken, Chichester, West Sussex.
- Kowalik S., Herczakowska J. (2010) Analiza i prognoza cen ropy naftowej na rynkach międzynarodowych. *Polityka Energetyczna*, 13(2), 253-262.
- Kozieł H. (2019) Ropa najdroższa od pół roku. Rzeczpospolita z dnia 23.04.2019.
<https://energia.rp.pl/ropa/art17990281-ropa-najdrozsza-od-pol-roku>
- Krishnan B. (2015) Oil Up Over 6 Percent As Dollar Falls on Worry of Slow U.S. Rate Hike. *Reuters Commodity News*. 18.03.2015.
<https://www.reuters.com/article/us-markets-oil/oil-up-over-6-percent-as-dollar-falls-on-worry-of-slow-u-s-rate-hike-idUSKBNOME03R20150318>
- Krishnan B. (2016) Oil Tumbles 4 Percent; Saudi Arabia Cools on Output Freeze. *Reuters Commodity News*. 31.03.2016. <https://www.reuters.com/article/us-global-oil-idUSKCN0WX00R>
- Los Angeles Times (2016) Stock Inch Higher as Oil Prices Jump, but Retailers Skid. 08.04.2016. <https://www.latimes.com/business/la-fi-0409-markets-20160408-story.html>
- Perron P. (1989) The Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Econometrica*, 57, 1361-1401.

- Płókarz R. (2013) Globalne rynki finansowe. Praktyka funkcjonowania. PWN, Warszawa.
- Reuters (2017) Corrected Oil-Prices Climb Off 10-Month Lows as U.S. Stockpiles Drop. Reuters Energy. 22.06.2017.
<https://www.reuters.com/article/global-oil-idUSL3N1JJ02V>
- Schofield N.C. (2007) Commodity Derivatives. John Wiley&Sons, Hoboken, Chichester, West Sussex.
- Sierakowska D. (2015a) Raport surowcowy DM BOŚ z dnia 26.01.2015.
https://info.bossa.pl/_gAllery/aid48478/20150126_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2015b) Raport surowcowy DM BOŚ z dnia 18.05.2015.
https://info.bossa.pl/_gAllery/aid50477/20150518_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2015c) Raport surowcowy DM BOŚ z dnia 02.07.2015.
https://info.bossa.pl/_gAllery/aid51429/20150702_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2015d) Raport surowcowy DM BOŚ z dnia 03.07.2015.
https://info.bossa.pl/_gAllery/aid51457/20150703_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2015e) Raport surowcowy DM BOŚ z dnia 25.08.2015.
https://info.bossa.pl/_gAllery/aid52506/20150825_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2015f) Raport surowcowy DM BOŚ z dnia 28.08.2015.
https://info.bossa.pl/_gAllery/aid52591/20150828_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2016a) Raport surowcowy DM BOŚ z dnia 13.04.2016.
https://info.bossa.pl/_gAllery/aid56979/20160413_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2016b) Raport surowcowy DM BOŚ z dnia 15.06.2016.
https://info.bossa.pl/_gAllery/aid58128/20160615_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2016c) Raport surowcowy DM BOŚ z dnia 04.08.2016.
https://info.bossa.pl/_gAllery/aid59123/20160804_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2018a) Raport surowcowy DM BOŚ z dnia 04.10.2018.
https://info.bossa.pl/_gAllery/aid73292/20181004_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2018b) Raport surowcowy DM BOŚ z dnia 03.12.2018.
https://info.bossa.pl/_gAllery/aid74321/20181203_raport_surowcowy_DMBOS.pdf

- Sierakowska D. (2018c) Raport surowcowy DM BOŚ z dnia 11.12.2018.
https://info.bossa.pl/_gAllery/aid74469/20181211_raport_surowcowy_DMBOS.pdf
- Sierakowska D. (2019) Raport surowcowy DM BOŚ z dnia 09.08.2019.
https://info.bossa.pl/_gAllery/aid78346/20190809_raport_surowcowy_DMBOS.pdf
- Witkowska D., Matuszewska-Janica A., Kompa K. (2012) Wprowadzenie do ekonometrii dynamicznej i finansowej. Wydawnictwo SGGW, Warszawa.

INVESTIGATION OF TURNING POINTS IN PRICES OF CRUDE OIL FROM 2015 TO 2019

Abstract: This paper aimed at identification of turning points in prices of crude oil from 2015 to 2019 through application of Perron test. It also attempted to detect geopolitical events that could have caused these changes in trend structure. This part of research was based on oil market analysts' comments and on market reports available at professional business websites. The research brings to conclusion that problems related to the oversupply of crude oil in the United States affected its prices the most. Reports informing about the increasing number of oil platforms working in North America and low demand in the fuel market often induced long-term price reductions at NYMEX and ICE.

Keywords: crude oil market, turning point, Perron test, geopolitical events

JEL classification: C22, C58, Q31

O PROBLEMACH STOSOWALNOŚCI MIERNIKÓW SYNTETYCZNYCH DO PORZĄDKOWANIA OBIEKTÓW

Zbigniew Binderman  <https://orcid.org/0000-0003-2917-4381>

Bolesław Borkowski  <https://orcid.org/0000-0001-6073-6173>

Instytut Ekonomii i Finansów

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: zbigniew_binderman@sggw.edu.pl; boleslaw_borkowski@sggw.edu.pl

Wiesław Szczesny  <https://orcid.org/0000-0002-8083-4624>

Instytut Informatyki Technicznej

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: wieslaw_szczesny@sggw.edu.pl

Rafał Zbyrowski  <https://orcid.org/0000-0001-8499-8899>

Katedra Metod Ilościowych, Uniwersytet Warszawski

e-mail: rzbyrowski@wz.uw.edu.pl

Streszczenie: Problem budowy stabilnego wskaźnika syntetycznego służącego do porządkowania obiektów opisanych za pomocą wielu wskaźników cząstkowych był i jest przedmiotem rozważań autorów. W naszych badaniach dążyliśmy do konstrukcji miernika, który byłby mało wrażliwy na dane wejściowe, do wyników, które nie zależałyby od sposobu normalizacji zmiennych, wyboru miary odległości (podobieństwa) i doboru cech. Nasze doświadczenie jest zgodne z wynikami prac wybitnego statystyka amerykańskiego L. Breimana. Pojedynczy klasyfikator (miernik syntetyczny) może być daleki od optymalnego, natomiast kombinacje wielu dają klasyfikator bliski optymalnemu i stabilny. Niestety w przypadku zastosowania „słabych” klasyfikatorów, w wyniku kombinacji można otrzymać klasyfikator jeszcze gorszy. W pracy przedstawiliśmy na praktycznym przykładzie gospodarstw rolniczych w krajach UE objętych badaniem FADN, konsekwencje wyboru wskaźnika syntetycznego przeznaczonego do porządkowania obiektów opisanych wieloma cechami.

Keywords: miernik syntetyczny, porządkowanie liniowe obiektów, klasyfikacja, wagi

JEL classification: C1, G2

WSTĘP

Podstawowym celem wielowymiarowej analizy danych jest dokonanie grupowania i porządkowanie obiektów (jednostek) będących elementami wielowymiarowej przestrzeni cech. Do klasyfikacji i grupowania stosowanych jest wiele metod [Binderman, Borkowski, Szczesny 2018, 2020; Hellwig 1968; Kukuła 2000; Młodak 2006; Gatnar, Walesiak 2009; Zeliaś 2000].

Z naszych dotychczasowych prac wynika, że nie ma jednego uniwersalnego klasyfikatora, dlatego zalecamy stosowanie mierników będących funkcją różnych klasyfikatorów. W naszych badaniach dążyliśmy do konstrukcji miernika, który byłby mało wrażliwy na dane wejściowe, do wyników, które nie zależałyby od sposobu normalizacji zmiennych, wyboru miary odległości (podobieństwa), doboru cech i miernika syntetycznego.

Czytając prace z zakresy taksonomii, zauważyliśmy, że w wielu z nich uzyskuje się rezultaty, które są wynikiem zastosowania jednego doboru cech, jednego sposobu normalizacji zmiennych i jednego klasyfikatora. Skłoniło nas to podjęcia próby konstrukcji nowych mierników, jak również do poprawy używanych mierników. W tych konstrukcjach „wypadkowego” miernika syntetycznego, staraliśmy się wykorzystać głównie różne sposoby normalizacji cech, różne miary podobieństwa i różne „pojedyncze” mierniki syntetyczne. Rozważane mierniki, miały charakter funkcji użyteczności w warunkach niedosytu. Podobne podejście zaprezentowali autorzy w swoich pracach poświęconych analizie struktur [Binderman, Borkowski, Szczesny 2008, 2012, 2013, 2018a; Borkowski i in. 2020].

Taksonomiczny miernik rozwoju zaproponowany został przez Zdzisława Hellwiga w pracy [Hellwig 1968]. Zastosowanie tego miernika pozwala na przeprowadzenie porządkowania regionów, a następnie ich grupowania. Wartości taksonomicznego miernika rozwoju stanowią wypadkową poziomu zmiennych, dotyczących różnych aspektów badanego zjawiska i pozwalają na jego syntetyczny opis. Zaproponowany wówczas miernik wykorzystywał jeden wzorec oraz standaryzację zmiennych.

Nasze doświadczenia są zgodne z wynikami prac wybitnego statystyka amerykańskiego L. Breimana. Jak podaje L. Breiman w pracy [Breiman 1994] pojedynczy klasyfikator (miernik syntetyczny) może być daleki od optymalnego, natomiast kombinacje wielu dają klasyfikator bliski optymalnemu i stabilny. Niestety w przypadku zastosowania „słabych” klasyfikatorów, w wyniku kombinacji można otrzymać klasyfikator jeszcze gorszy [Breiman 1996, 1998].

Rozważając problemy ze statystycznej analizy danych należy dążyć, aby rozważane zagadnienia były poprawnie postawione. W naukach ekonomicznych zagadnienie poprawnie postawione jest różnie rozumiane [Binderman 2015]. Według określeń podanych w serii prac D.M. Jacksona [Jackson 1970, 1969, 1969a] zagadnienie jest poprawnie postawione, jeżeli:

- w wyniku zastosowanego algorytmu otrzymujemy jeden wynik,

- otrzymana klasyfikacja (uporządkowanie) jest stabilna. (tj. otrzymana klasyfikacja nie może być „rażąco odmienna” od klasyfikacji otrzymanych przy małych zmianach dokonanych w danych wejściowych),
- użyty algorytm musi być niezmienniczy względem permutacji, zarówno zmiennych i nazw obiektów, które mają być klasyfikowane. Oznacza to, że algorytm musi być niezależny od etykietowania zmiennych i obiektów,
- użyty algorytm musi być niezależny od skali. Oznacza to, że algorytm musi być niezmienniczy przez mnożenie macierzy podobieństwa, przez stałą dodatnią, różną od zera.

Warunki te są wyraźnie nie wystarczające, ponieważ pomijają całkowicie problem nierozwiązanych kwestii adekwatności konfiguracji poszczególnych klas w stosunku do szczegółowych związków między obiektami, czy do szczególnych potrzeb.

Do porządkowania obiektów stosowanych jest wiele metod. Najczęściej stosowane metody to:

1. Metody bezwzorcowe

- Metoda sum (np. metoda unitaryzacji zerowanej),
- Metoda średnich rang,
- Metody wykorzystujące wykresy radarowe wektorów.

2. Metody wzorcowe

- Metody wykorzystujące jeden wzorec (metoda Hellwiga),
- Metody wykorzystujące dwa lub więcej wzorców (metoda TOPSIS)

Metody wykorzystujące dwa wzorce stosowane były przez wielu autorów (poza metodą TOPSIS, głównie przez A. Binderman i autorów niniejszej pracy: [Binderman A. 2006, 2007, 2001; Binderman, Borkowski, Szczesny 2008, 2015, 2018; Borkowski i in. 2020; Binderman Z. 2010]. Przykład metody wykorzystującej trzy wzorce został podany w pracy autorów [Binderman, Borkowski, Szczesny 2020].

Do mierników „bezwzorcowych” zalicza się mierniki tworzone w oparciu o interpretacje geometryczne. Przykładem takich mierników są tzw. **radarowe mierniki** indukowane przez *s*-normy (*por.*[Borkowski, Wiliński, Szczesny, Binderman Z. 2020; Binderman Z. 2011]). Miary te, jako mierniki bez wzorców, wykorzystują pola wykresów radarowych, indukowanych przez wektory wierszowe macierzy obserwacji zawierającej dane po unormowaniu.

W naszej pracy rolę miernika rozwoju spełnia miernik syntetyczny, zbudowany z mierników wzorcowych jedno i dwuwzorcowych, jak również z miernika bez wzorca. Syntetyczny miernik umożliwił liniowe uporządkowanie rozważanych obiektów i ich podział na klasy.

Przeprowadzona w pracy procedura badawcza składała się z siedmiu, następujących etapów:

1. Dobór zmiennych diagnostycznych.
2. Określenie wag wybranych zmiennych.
3. Normalizacja zmiennych.
4. Przekształcenie destymulant w stymulanty.
5. Obliczenie wypadkowego miernika syntetycznego, jako miernika rozwoju.
6. Wizualizacja otrzymanych wartości wskaźnika syntetycznego.
7. Interpretacja otrzymanych wyników.

Doboru zmiennych diagnostycznych dokonano w oparciu o kryteria ocen merytoryczno-formalnych własności zmiennych oraz kryteria wartości informacyjnej zmiennych. Procedurę wyznaczania wag wybranych zmiennych diagnostycznych, przeprowadzono w dwóch wariantach. W wariacie pierwszym, ustalono, że wybrane zmienne mają podobny wpływ na badane zjawisko, tj. w obliczeniach przyjęto, że zmienne mają identyczne wagi. W drugim wariacie, wybór wag został oparty o współczynniki korelacji i miary rozproszenia. Normalizacja zmiennych polegała na zastosowaniu metody unitaryzacji zerowanej oraz standaryzacji [Kukuła 2000; Zeliaś 2000].

OPIS ZASTOSOWANYCH METOD PORZĄDKOWANIA OBIEKTÓW.

W dalszej części bez straty ogólności rozważań założymy, że dane zjawisko jest opisane przez zmienne będące stymulantami. Osiągnąć to można np. poprzez zmianę zwrotu osi współrzędnych określającej położenie destymulant oraz przez odpowiednie przekształcenie nominant.

Rozważmy teraz problem polegający na porządkowaniu $m \in \mathbb{N}$ obiektów $\mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_m$ opisanych za pomocą $n \in \mathbb{N}$ zmiennych (cech). Niech wektor $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in})$, $i=1, 2, \dots, m$, opisuje i -ty obiekt \mathbf{Q}_i . To jest

$$\mathbf{x}_i \sim \mathbf{Q}_i, \quad i=1, 2, \dots, m.$$

Jeżeli $x_{ik} > x_{jk}$ ($x_{ik} \geq x_{jk}$) dla $k=1, 2, \dots, n$, to pisać będziemy $\mathbf{x}_i > \mathbf{x}_j$, ($\mathbf{x}_i \geq \mathbf{x}_j$), gdzie $i, j \in [1, m]$. Jeżeli $\mathbf{x}_i > \mathbf{x}_j$ i $\mathbf{x}_i \neq \mathbf{x}_j$ to naturalnym jest nazywać obiekt \mathbf{x}_i (\mathbf{Q}_i) lepszym (czyli wyżej ocenianym) od obiektu \mathbf{x}_j (\mathbf{Q}_j). Oznacza to, że żadna ze składowych wektora \mathbf{x}_i nie jest mniejsza od odpowiednich składowych wektora \mathbf{x}_j , a przynajmniej jedna z nich ma wartość większą, tj. istnieje takie $k \in [1, n]$, że $x_{ik} > x_{jk}$.

Przyjmijmy ponadto następujące oznaczenia:

$$x_{m+1,k} \stackrel{\text{def}}{=} \max_{1 \leq i \leq m} x_{ik}, \quad x_{0,k} \stackrel{\text{def}}{=} \min_{1 \leq i \leq m} x_{ik}$$

$$\mathbf{x}_0 := (x_{0,1}, x_{0,2}, \dots, x_{0,n}) \sim \mathbf{Q}_0, \quad \mathbf{x}_{m+1} := (x_{m+1,1}, x_{m+1,2}, \dots, x_{m+1,n}) \sim \mathbf{Q}_{m+1}.$$

Oczywistym jest, że tak określone obiekty $\mathbf{x}_0, \mathbf{x}_{m+1}$ (być może fikcyjne) są nie lepsze, nie gorsze od pozostałych $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$, tj.

$$\mathbf{x}_{m+1} \geq \mathbf{x}_i \text{ oraz } \mathbf{x}_i \geq \mathbf{x}_0 \text{ dla każdego } i = 1, \dots, m.$$

W przypadku gdy obiekty \mathbf{x}_0 i \mathbf{x}_{m+1} są różne od rozważanych obiektów $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$, to obiekty te spełniają rolę **obektu najlepszego, obiektu najgorszego**, odpowiednio i będą traktowane w ramach tej pracy, jako **wzorcy**. Otrzymano w ten sposób $m+2$ obiektów $\mathbf{Q}_0, \mathbf{Q}_1, \mathbf{Q}_2, \dots, \mathbf{Q}_m, \mathbf{Q}_{m+1}$ z których każdy opisany został przez n cech X_1, X_2, \dots, X_n . Wartości przyjętych zmiennych diagnostycznych dla rozważanych m obiektów, utworzyły macierz

$$\mathbf{X} := [x_{i,j}]_{i \times j}, \quad i=1, 2, \dots, m, j=1, \dots, n,$$

o wymiarach $m \times n$. W macierzy tej wiersze reprezentują obiekty, natomiast kolumny, uporządkowany zestaw zmiennych tych obiektów.

Na podstawie wartości przyjętych zmiennych diagnostycznych, dla m obiektów stworzono wyżej dwa stałe (statyczne), hipotetyczne obiekty: minimalny \mathbf{Q}_0 oraz maksymalny \mathbf{Q}_{m+1} , opisane odpowiednio za pomocą najmniej i najbardziej korzystnego zestawu wartości cech. Obiekty te będą więc reprezentowane przez wektory wzorcowe \mathbf{x}_0 oraz \mathbf{x}_{m+1} o n składowych każdy. Uwzględniając te obiekty (czyli wstawiając do macierzy \mathbf{X} obserwacji, wektory \mathbf{x}_0 i \mathbf{x}_{m+1} , jako wiersz $m+1$ i $m+2$, odpowiednio) otrzymujemy macierz dołączoną \mathbf{X}^* .

Wygodnie jest w dalszej części rozważań rozpatrywać zbiór m obiektów, jako zbiór opisany przez znormalizowaną macierz obserwacji \mathbf{U} o wymiarach $m \times n$, *indukowaną przez macierz \mathbf{X} oraz macierz indukowaną przez macierz dołączoną \mathbf{X}^** o wymiarach $(m+2) \times n$ (czyli np. po dokonaniu standaryzacji lub unitaryzacji zerowanej poszczególnych kolumn macierzy \mathbf{X} oraz analogicznie znormalizowanych tych dwu dodatkowych wierszy dołożonych odpowiadających \mathbf{x}_0 i \mathbf{x}_{m+1} , odpowiednio). Taką znormalizowaną, dołączoną macierz, przez analogię do \mathbf{X}^* , oznaczać będziemy symbolem

$$\mathbf{U}^* = [u_{i,j}]_{i \times j}, \quad i=1, 2, \dots, m, 0, m+1; j=1, \dots, n.$$

W dalszej części rozważań znormalizowane wektory wierszowe \mathbf{x}_0 i \mathbf{x}_{m+1} oznaczać będziemy przez

$$\mathbf{u}_0 \text{ i } \mathbf{u}_{m+1},$$

odpowiednio.

Niech d oznacza dowolną metrykę Minkowskiego:

$$d(\mathbf{x}, \mathbf{y}) = \left[\sum_{j=1}^n |x_j - y_j|^p \right]^{\frac{1}{p}}, \quad \mathbf{x} = (x_1, x_2, \dots, x_n), \mathbf{y} = (y_1, y_2, \dots, y_n) \in \mathfrak{R}^n, \quad 1 \leq p < \infty.$$

W polskiej literaturze, do porządkowania i klasyfikacji obiektów wykorzystuje się następujące wzory mierników syntetycznych (por. np. [Hellwig 1968; Młodak 2006]):

1. Miernik wzorcowy Z. Hellwiga

$$m(\mathbf{u}_i) = 1 - \frac{d_i}{\bar{d} + 2\sigma(D)}, \quad i=0,1,\dots,m,m+1.$$

gdzie D oznacza dodatkową kolumnę, która w każdym wierszu ma wartość równa odległości Minkowskiego dla parametru $p=2$, od wzorca pozytywnego \mathbf{u}_{m+1} , a \bar{d} oraz $\sigma(D)$ oznaczają, odpowiednio średnią i odchylenie standardowe z kolumny D dla wierszy $i=1,\dots,m$.

2. Bezwzorcowy miernik m określony wzorem :

$$m(\mathbf{u}_i) = \frac{1}{m} \sum_{j=1}^m u_{ij}, \quad i=0,1,\dots,m,m+1; j=1,2,\dots,n,$$

znany, jako miernik obiektu M. Cieślak (K. Kukuły), gdzie macierz $\mathbf{U}=[u_{ij}]_{m \times n}$ powstała poprzez normalizację macierzy obserwacji \mathbf{X} , nazywaną unitaryzacją zerowaną (por. [Cieślak 1974, 1993; Kukuła K. 2000]). Natomiast w przypadku, gdy zastosujemy, jako normalizację podstawową przekształcenie ilorazowe (czyli podzielenie przez średnią kolumnową z każdej kolumny macierzy \mathbf{X}), to wskaźnik określony tym wzorem występuje pod nazwą miernika Pluty [Pluta 1976].

3. Miernik wykorzystujący dwa wzorce TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [Hwang, Yoon 1981]:

$$m(\mathbf{u}_i) = \frac{d(\mathbf{u}_0, \mathbf{u}_i)}{d(\mathbf{u}_0, \mathbf{u}_i) + d(\mathbf{u}_{m+1}, \mathbf{u}_i)}, \quad i=1,2,\dots,m;$$

gdzie d oznacza tu metrykę Euklidesa (metryka Minkowskiego z $p=2$).

Rekomendujemy aby do porządkowania i grupowania obiektów wykorzystywać jednocześnie dwa wzorce „najgorszy”- \mathbf{Q}_0 i „najlepszy” \mathbf{Q}_{m+1} , jak również funkcje użyteczności w warunkach niedosytu. Wyniki naszych dotychczasowych badań i A. Binderman [Binderman A. 2006, 2007], wyraźnie wskazują, że wykorzystywanie tylko jednego wzorca może doprowadzić do takiej sytuacji, że obiekt uznany za najlepszy według jednej metody, może okazać się najgorszy, według drugiej metody.

Warto podkreślić, że mierniki syntetyczne pozwalają zarówno uporządkować rozważane obiekty, jak również dokonać ich grupowania według zadanych przez użytkownika progów.

Dużą klasę mierników syntetycznych wykorzystywanych w badaniach naukowych i praktyce gospodarczej można wyrazić jako odległość od jednego lub dwu wzorców. Jeśli symbolem d oznaczmy miarę odległości między obiektami, symbolami \mathbf{u}_{min} , \mathbf{u}_{max} wektory opisujące, odpowiednio obiekt *najgorszy* i *najlepszy* {w naszym przypadku $\mathbf{u}_{min}=\mathbf{u}_0$, $\mathbf{u}_{max}=\mathbf{u}_{m+1}$ }, charakteryzowane za pomocą cech będących stymulantami, to dla $\mathbf{x} \in [\mathbf{u}_{min}, \mathbf{u}_{max}]$ znane z literatury są następujące mierniki syntetyczne [Hellwig 1968; Kukuła 2000; Binderman A. 2006, 2007; Hwang, Yoon 1981]:

$$M_1(\mathbf{u}) = \frac{d(\mathbf{u}_{min}, \mathbf{u})}{d(\mathbf{u}_{min}, \mathbf{u}_{max})},$$

$$M_2(\mathbf{u}) = 1 - \frac{d(\mathbf{u}_{max}, \mathbf{u})}{d(\mathbf{u}_{min}, \mathbf{u}_{max})},$$

$$M_3(\mathbf{u}) = \frac{M_1(\mathbf{u})+M_2(\mathbf{u})}{2} = \frac{1}{2} + \frac{d(\mathbf{u}_{min}, \mathbf{u})-d(\mathbf{u}_{max}, \mathbf{u})}{2d(\mathbf{u}_{min}, \mathbf{u}_{max})},$$

$$M_4(\mathbf{u}) = \sqrt{M_1(\mathbf{u})M_2(\mathbf{u})}$$

$$M_5(\mathbf{u}) = \frac{M_1(\mathbf{u})}{1+M_1(\mathbf{u})-M_2(\mathbf{u})} = \frac{d(\mathbf{u}_{min}, \mathbf{u})}{d(\mathbf{u}_{min}, \mathbf{u})+d(\mathbf{u}_{max}, \mathbf{u})}.$$

Mierniki M_1 i M_2 wykorzystują w zasadzie tylko jeden wzorec, natomiast mierniki M_3 , M_4 , M_5 , wykorzystują dwa wzorce. Mierniki te można potraktować jako narzędzia rozwiązywania wielokryterialnych problemów decyzyjnych. Mierniki M_1 i M_2 wykorzystują jedno kryterium natomiast M_3 , M_4 , M_5 wykorzystują dwa kryteria.

Podane wyżej mierniki są znormalizowane tj.

$$0 \leq M_k(\mathbf{u}) \leq 1 \text{ dla } \mathbf{u} \in \langle \mathbf{u}_0, \mathbf{u}_{m+1} \rangle \text{ dla } k=1,2,\dots,5.$$

Zauważmy ponadto, że:

$$M_k(\mathbf{u}_0) = 0, M_k(\mathbf{u}_{m+1}) = 1.$$

W pracy [Hellwiga 1968] podany miernik obiektów przy zastosowaniu standaryzacji cech, wykorzystuje obiekt najlepszy \mathbf{u}_{max} i miernik M_2 . Miernik M_5 związany jest z metodą TOPSIS [Hwang, Yoon 1981]. Metodę TOPSIS można zakwalifikować do metod rankingowych służących do podejmowania decyzji wielokryterialnych. Metodami tego typu są również znane metody SAW (*Simple Additive Weighting*) oraz AHP (*Analytical Hierarchy Process*).

W szeregu prac m.in. [Binderman A. 2006, 2007, 2011] podano teorie i zastosowania miernika M_3 (M_4) w ocenie regionalnego zróżnicowania rolnictwa w Polsce, pokazano, że stosowanie metod opartych tylko na jednym wzorcu, w wielu przypadkach prowadzi do otrzymania błędnych wyników, które nie spełniają warunków poprawności, według Jacksona, zobacz również [Kukuła, Luty 2017].

WYNIKI ANALIZY DANYCH

Aby zilustrować problem budowy stabilnego wskaźnika syntetycznego służącego do porządkowania obiektów opisanych za pomocą wielu wskaźników cząstkowych wykorzystaliśmy dane z lat 2004-2019 dotyczące efektywności gospodarstw w krajach UE specjalizujących się w uprawach polowych. Jako cel przyjęliśmy uporządkowanie krajów UE pod względem efektywności oraz zmiany poziomu efektywności w tej zbiorowości krajów w badanym okresie. Jako zestaw zmiennych do oceny efektywności gospodarowania w kraju UE przyjęto następujący (według autorskiego wyboru) zestaw:

- X1 – wartość produkcji roślinnej na 1 ha UR (SE136)),
- X2 – wartość dodana bruto w zł (SE410)),
- X3 – wartość dodana brutto na jednostkę produkcji roślinnej (SE410 / SE135),
- X4 – powierzchnia UR (SE026),
- X5 – czas pracy na jednostkę powierzchni (SE011)/ (SE026),
- X6 – udział pracy na pozostałą działalność zarobkową w ogółem nakładów pracy (%) (Share of OGA work /AWU (SE022),
- X7 – maszyny, urządzenia i środki transportu (SE455),
- X8 – nawozy w zł (SE295),
- X9 – środki ochrony roślin w zł (SE300),
- X10 – dopłaty do produkcji roślinnej (SE (610) + (611) +(612) + (613) +(631),
- X11– amortyzacja w zł (SE360),
- X12 – wielkość ekonomiczna (SE005).

Zmienne X5, X6 i X11 to destymulanty, a pozostałe zmienne to stymulanty.

Dane zostały znormalizowane dwoma najbardziej popularnymi sposobami, a mianowicie przy wykorzystaniu formuły unitaryzacji zerowanej oraz standaryzacji. Wykorzystaliśmy dwa zestawy wag v_1 oraz v_2 : czyli wagi jednakowe oraz wagi wykorzystujące macierz korelacji pomiędzy zmiennymi X1-X12 i miary rozproszenia tych zmiennych po unormowaniu. Jako wskaźniki syntetyczne przyjęto te najbardziej popularne: W1 i W2 oraz WS1 i WS2, to średnia lub średnia ważona w wierszu macierzy danych znormalizowanych odpowiednio przy wykorzystaniu unitaryzacji zerowanej i standaryzacji, W3 i W6 oraz WS3 i WS6, to wskaźniki wykorzystujące wzorzec negatywny oraz odpowiednio wagi v_1 i v_2 oraz normalizację (unitaryzację zerowaną i standaryzację). Analogicznie W4 i W7 oraz WS4 i WS7 to wskaźniki wykorzystujące wzorzec pozytywny. Natomiast W5 i W8 oraz WS5 i WS8 to wskaźniki wykorzystujące dwa wzorce (metoda TOPSIS). W ten sposób mamy 16 możliwych rankingów w każdym roku z badanego okresu. Można też dokonać podziału krajów na grupy pod względem efektywności tego sektora w poszczególnych krajach UE.

Wyniki przeprowadzonych obliczeń wskazują, że rankingi są nawet znacząco różniące się w zależności od użytego wskaźnika syntetycznego. W tabeli 2 przedstawiono jak wyglądają rankingi w roku 2019. Natomiast podobieństwo tych rankingów mierzone wartością wskaźnika podobieństwa dwu wektorów, który liczy się takim samym wzorem jak współczynnik korelacji zamieszczono w górnym trójkącie tabeli 1. Ponadto w tabeli 1 w dolnym trójkącie zamieszczono współczynniki korelacji pomiędzy wartościami 17 wskaźników syntetycznych w całym badanym okresie 2004-2019.

Tabela 1. Współczynniki korelacji pomiędzy wartościami wskaźników 17 wskaźników syntetycznych w okresie 2004-2019 (dolny lewy trójkąt tabeli) oraz współczynniki korelacji pomiędzy rangami wartości tych wskaźników w roku 2019 (górną trójkąt tabeli)

	W1	W2	W3	W4	W5	W6	W7	W8	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	W _{sr}
W1		0.973	0.792	0.984	0.985	0.752	0.979	0.984	0.931	0.859	0.689	0.932	0.913	0.832	0.851	0.907	0.974
W2	0.984		0.824	0.952	0.982	0.812	0.987	0.983	0.861	0.792	0.607	0.875	0.832	0.735	0.806	0.822	0.924
W3	0.927	0.930		0.739	0.859	0.978	0.778	0.86	0.698	0.538	0.725	0.645	0.722	0.585	0.518	0.707	0.759
W4	0.981	0.960	0.841		0.958	0.701	0.956	0.957	0.956	0.913	0.694	0.968	0.935	0.885	0.905	0.932	0.98
W5	0.995	0.986	0.939	0.974		0.823	0.973	0.999	0.88	0.792	0.657	0.878	0.87	0.772	0.788	0.863	0.939
W6	0.876	0.920	0.971	0.787	0.896		0.758	0.825	0.647	0.488	0.692	0.609	0.661	0.521	0.478	0.639	0.717
W7	0.971	0.975	0.839	0.990	0.969	0.813		0.974	0.874	0.813	0.604	0.895	0.845	0.747	0.837	0.836	0.932
W8	0.995	0.986	0.938	0.974	1.000	0.896	0.969		0.882	0.791	0.661	0.879	0.871	0.771	0.786	0.861	0.94
WS1	0.940	0.881	0.821	0.944	0.926	0.709	0.909	0.926		0.945	0.817	0.977	0.987	0.947	0.914	0.978	0.982
WS2	0.894	0.837	0.732	0.922	0.876	0.619	0.895	0.876	0.979		0.668	0.968	0.921	0.96	0.983	0.921	0.911
WS3	0.814	0.737	0.817	0.758	0.804	0.684	0.705	0.804	0.901	0.832		0.722	0.861	0.773	0.591	0.851	0.781
WS4	0.923	0.875	0.745	0.963	0.911	0.652	0.939	0.911	0.976	0.976	0.807		0.949	0.938	0.959	0.938	0.966
WS5	0.922	0.862	0.812	0.926	0.914	0.700	0.890	0.914	0.991	0.962	0.925	0.970			0.877	0.994	0.97
WS6	0.827	0.745	0.756	0.804	0.806	0.616	0.753	0.806	0.929	0.918	0.945	0.863	0.935		0.91	0.955	0.907
WS7	0.857	0.814	0.646	0.918	0.843	0.555	0.904	0.843	0.937	0.975	0.733	0.980	0.927	0.828		0.875	0.891
WS8	0.921	0.861	0.812	0.925	0.913	0.700	0.888	0.912	0.990	0.961	0.925	0.969	0.999	0.935	0.925		0.961
W _{sr}	0.986	0.957	0.893	0.979	0.981	0.820	0.964	0.981	0.979	0.947	0.863	0.964	0.970	0.885	0.914	0.969	

Źródło: opracowanie własne

Łatwo można sprawdzić (np. sumując wiersze macierzy korelacji powstałych z tych dwu trójkątów z tabeli 1), że zarówno w przypadku rankingów, w 2019 roku jak i wartości 17 wskaźników, w całym okresie najbardziej podobnym do innych jest wskaźnik W_{sr} . Wskaźnik utworzony, jako średnia arytmetyczna z wartości pozostałych 16-tu wskaźników, po ich unormowaniu przy użyciu techniki zwanej unitaryzacją zerowaną. W przypadku takiego zestawu nie przypadkowo nieznacznie gorszym reprezentantem (rozważając podobieństwo do innych) jest wskaźnik W1. Między innymi w literaturze przedmiotu w przypadku, gdy nie ma wskazań eksperckich dotyczących wag rekomenduje się zazwyczaj używanie wskaźników

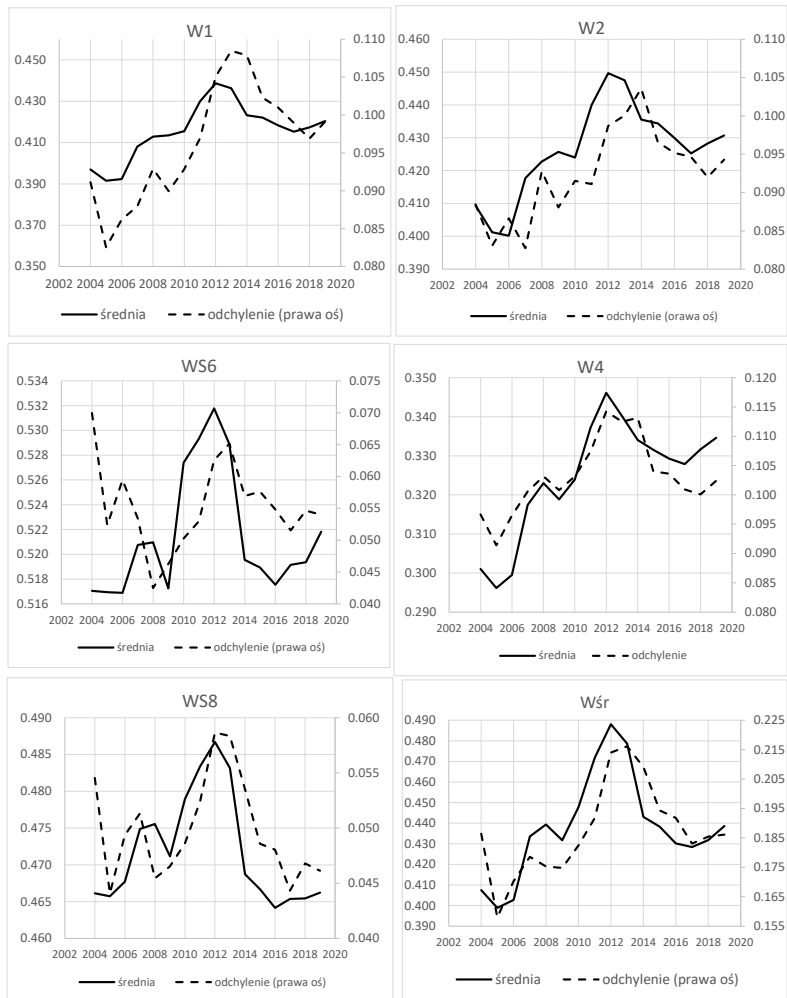
zbudowanych w oparciu o tzw. jednakowe wagi zmiennych. Mimo, że użyliśmy tu najbardziej popularnych wskaźników, to rankingi różnią się dość znacznie (por. tabela 2, ostatnia kolumna). Jednakże ogólna tendencja w wartościach wszystkich wskaźników, gdy opisujemy całą UE (por rysunek 1) jest podobna. Dlatego zazwyczaj jeśli ktoś wybiera konkretny, nawet bez uzasadnienia, to może otrzymać wyniki zadawalające. Jednakże warto zwrócić uwagę na jakie zmiany jest wrażliwy używany wskaźnik syntetyczny. Na rysunku 2 przedstawiono kształtowanie się wartości wskaźników zaprezentowanych na rysunku 1 w przypadku Polski. Jak widać zmiany wartości mają podobny kierunek, ale względna wielkość zmian różni się znacząco.

Tabela 2. Rankingi sektorów rolniczych państw UE w roku 20219 oparte o wartości 16 wskaźników syntetycznych oraz wskaźnik W_{sr}

kraj	W1	W2	W3	W4	W5	W6	W7	W8	WS1	WS2	WS3	WS4	WS5	WS6	WS7	WS8	W _{sr}	min	max	delta
BEL	9	8	10	10	9	8	8	8	9	11	12	8	10	11	10	14	8	8	14	6
BGR	11	11	8	11	10	11	11	10	10	13	8	12	9	13	13	9	10	8	13	5
CYP	23	21	14	24	22	13	23	22	25	26	20	25	25	25	26	25	23	13	26	13
CZE	4	4	4	4	4	4	5	4	4	5	4	5	4	5	7	4	4	4	7	3
DAN	10	10	20	9	11	23	10	11	12	6	26	10	13	8	5	12	11	5	26	21
DEU	5	5	7	5	5	7	4	5	5	4	9	4	5	7	3	5	5	3	9	6
ELL	24	23	21	25	24	20	22	24	24	24	22	24	24	24	22	24	24	20	25	5
ESP	15	13	9	15	13	9	15	13	15	17	10	17	14	15	18	13	15	9	18	9
EST	12	12	22	12	12	22	12	12	14	14	25	14	16	17	14	16	14	12	25	13
FRA	6	6	5	6	6	5	6	6	7	10	7	7	7	9	9	7	7	5	10	5
HRV	20	20	15	21	19	18	20	19	20	21	16	22	20	22	21	20	21	15	22	7
HUN	17	19	16	18	17	19	18	17	16	18	15	18	15	16	19	15	18	15	19	4
IRE	7	7	6	7	7	6	7	7	6	7	5	6	6	6	6	6	6	5	7	2
ITA	18	18	23	19	20	21	16	20	18	16	19	16	21	21	15	21	19	15	23	8
LTU	16	15	12	16	14	10	17	14	19	22	14	20	17	18	24	17	17	10	24	14
LUX	21	26	25	17	25	26	26	25	8	8	3	11	8	4	11	8	13	3	26	23
LVA	14	14	18	14	15	17	14	15	17	20	17	15	19	20	20	19	16	14	20	6
NED	2	2	2	3	3	2	1	3	2	2	2	3	2	2	2	2	2	1	3	2
OST	13	16	24	13	16	24	13	16	13	12	11	13	11	12	12	10	12	10	24	14
POL	25	24	17	26	23	16	25	23	26	27	24	27	26	26	27	26	25	16	27	11
POR	22	17	19	20	21	15	19	21	22	19	21	21	23	23	17	23	22	15	23	8
ROU	19	22	13	22	18	14	21	18	21	23	13	23	18	19	25	18	20	13	25	12
SUO	26	27	27	23	26	27	27	26	23	15	23	19	22	14	16	22	26	14	27	13
SVE	8	9	11	8	8	12	9	9	11	9	18	9	12	10	8	11	9	8	18	10
SVK	1	1	1	2	1	1	3	1	1	1	1	1	2	1	4	1	1	1	4	3
SVN	27	25	26	27	27	25	24	27	27	25	27	26	27	27	23	27	27	23	27	4
UKI	3	3	3	1	2	3	2	2	3	3	6	1	3	3	1	3	3	1	6	5

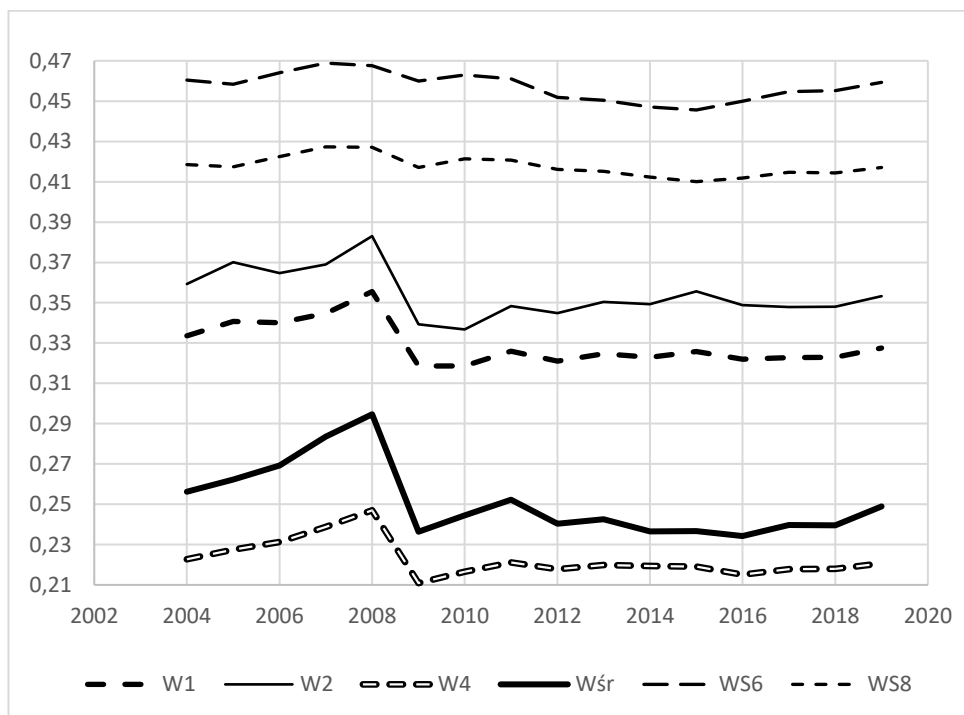
Źródło: opracowanie własne

Rysunek 1. Wartości wybranych wskaźników syntetycznych oceniających efektywność sektorów gospodarstw UE w okresie 2004-2019



Źródło: opracowanie własne

Rysunek 2. Wartości wybranych wskaźników syntetycznych oceniających efektywność sektora gospodarstw o specjalizacji uprawy polowe w Polsce w okresie 2004-2019



Źródło: opracowanie własne

UWAGI KOŃCOWE

W pracy przedstawiliśmy na praktycznym przykładzie sektora, który nie podlega szybkim zmianom, konsekwencje wyboru wskaźnika syntetycznego przeznaczonego do porządkowania obiektów opisanych wieloma cechami. Ze względu na rozmiar pracy ograniczyliśmy się zarówno do najczęściej stosowanych metod normalizacji zmiennych wskaźnikowych jak i do stosowanych technik (mechanizmów) tworzenia wskaźników syntetycznych je wykorzystujących. Nawet przy takim zestawie wskaźników syntetycznych otrzymuje się zazwyczaj różne uporządkowania obiektów oraz różne wyniki grupowania wykorzystujące wartości poszczególnych wskaźników. W naszym przykładzie dotyczącym efektywności przy podziale państw na 4 grupy według 3 rekomendowanych w literaturze progów: średnia - odchylenie, średnia, średnia + odchylenie zaobserwowaliśmy - co prawda nieliczne - zmiany o 2 grupy w zależności który wskaźnik użyliśmy.

Dlatego jeśli nie ma zdecydowanych wytycznych od strony praktyków, które determinują wagi, sposób normalizacji i technikę budowy wskaźnika, to według nas

wskazówka zacytowana we wstępie, a podana przez L. Breimana powinna być naszym zdaniem stosowana. Nasz przykład pokazuje, że najlepszym reprezentantem tych 17-tu rankingów jest W_{sr} , który powstał poprzez uśrednienie unormowanych wartości poszczególnych 16-tu wskaźników. W tym przypadku uznaliśmy wszystkie wskaźniki za jednakowo ważne. Oczywiście w wielu praktycznych sytuacjach można zastosować wagi przy tworzeniu W_{sr} w oparciu o wskazówki ekspertów z danej dziedziny.

Na zakończenie warto podkreślić, że taka metoda postępowania pozwala na stworzenie dość stabilnego wskaźnika syntetycznego. Jednak znacząco można polepszyć sposób tworzenia wskaźnika syntetycznego poprzez współpracę z praktykami przy wykorzystaniu więcej niż dwu wzorców (czyli wykorzystanie wzorców pośrednich stworzonych przez praktyków). Jak zbudować taki wskaźnik można przeczytać w pracy [Binderman, Borkowski, Szczesny 2020].

BIBLIOGRAFIA

- Binderman A. (2011) Wielokryterialne metody analizy zróżnicowania polskiego rolnictwa w 2009 roku. *Metody Ilościowe w Badaniach Ekonomicznych*, 12(2), 58-68.
- Binderman A. (2007) Wielowymiarowa analiza regionalnego zróżnicowania rolnictwa w Polsce. Praca doktorska, SGGW, Warszawa.
- Binderman A. (2006) Klasyfikacja obiektów oparta na dwóch wzorcach. *EiOGŻ, Zeszyty Naukowe SGGW*, nr 60, Warszawa, 25-37.
- Binderman Z. (2015) Zagadnienia poprawnie postawione w ekonomii i zarządzaniu (klasyfikacja i porządkowanie obiektów, pomiar koncentracji). *Розвиток національної економіки: теорія і практика, Матеріали Міжнародної науково-практичної конференції 3-4 квітня 2015 року Частина 2 Івано-Франківськ*, 318-321.
- Binderman Z. (2011) Matematyczne aspekty metod radarowych. *Metody Ilościowe w Badaniach Ekonomicznych*, 12(2), 69-79.
- Binderman Z. (2010) Zjawisko niedosytu w polu preferencji indukowanej przez miernik dwuwzorcowy. *Metody Ilościowe w Badaniach Ekonomicznych*, 11(2), 62-70.
- Binderman Z., Borkowski B., Szczesny W. (2020) Wykorzystanie wzorców pośrednich do budowy mierników syntetycznych, *Quantitative Methods in Economics*, 21(3), 157-167.
- Binderman Z., Borkowski B., Koszela G., Kozera R., Szczesny W. (2017) On the Choice of Synthetic Measures for Assessing Economic Effects. *Quantitative Methods in Economics*, 18(1), 7-17.
- Binderman Z., Borkowski B., Kozera R., Prokopenya A., Szczesny W. (2018) On Mathematical Modelling of Synthetic Measures. *Mathematical Modelling and Analysis*, 23(4), 699-711. <https://doi.org/10.3846/mma.2018.042>
- Binderman Z., Borkowski B., Szczesny W. (2008) O pewnej metodzie porządkowania obiektów na przykładzie regionalnego zróżnicowania rolnictwa. *Metody Ilościowe w Badaniach Ekonomicznych*, 9, 39-48.
- Binderman, Z., Borkowski B., Szczesny W. (2010) Radar measures of structures' conformability. *Quantitative Methods in Economics*, 11(1), 45-59.

- Binderman, Z., Borkowski B., Szczesny W. (2012) Radar Coefficient of Concentration. *Quantitative Methods in Economics*, 13(2), 7-21.
- Binderman Z., Borkowski B., Szczesny W. (2013a) Synthetic radar measures that factor in weights of features. *Quantitative Methods in Economics*, 14(1), 37-52.
- Binderman Z., Borkowski B., Szczesny W. (2013b) Zastosowanie metryki Minkowskiego do pomiaru zmian koncentracji. *Metody Ilościowe w Badaniach Ekonomicznych*, 14(3), 27-38.
- Binderman Z., Borkowski B., Szczesny W. (2015) Applications of Minkowski's Metric in Measuring Changes of Concentration of Value Added in Agriculture, Forestry, Fishing And Hunting Sectors. *Acta Scientiarum Polonorum, Oeconomia*, 14(1), 17-28.
- Borkowski B., Wiliński A., Szczesny W., Binderman Z. (2020). Mathematical analysis of synthetic measures based on radar charts. *Mathematical Modelling and Analysis*, 25(3), 473-489. <https://doi.org/10.3846/mma.2020.11223>
- Breiman L. (1998) Bias-variance, Regularization, Instability and Stabilization. [in:] C. M. Bishop (ed.), *Neural Networks and Machine Learning*, Springer-Verlag, 27-56.
- Breiman L., (1996) Arcing Classifiers. Technical Report 460, Department of Statistics, University of California, CA, USA, February.
- Breiman L. (1994), Bagging Predictors. Technical Report 420, Department of Statistics, University of California, CA, USA.
- Cieślak M. (1974) *Modele zapotrzebowania na kadry kwalifikowane*. PWN, Warszawa.
- Cieślak M. (1993) Ekonomiczne zastosowanie mierników syntetycznych ze zmiennym wzorcem. [w:] *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*, AE, Kraków.
- Gatnar E., Walesiak M. (2009) *Statystyczna analiza danych z wykorzystaniem programu R*. PWN, Warszawa.
- Jackson D. M. (1970) *The Stability of Classifications of Binary Attribute Data*. Cornell University.
- Jackson D. M. (1969). Comparison of classifications. [in:] A. J. Cole (ed.) *Numerical Taxonomy*, 91-111. New York: Academic Press.
- Jackson D. M. (1969a) *Automatic Classifications and Information Retrieval*. Cambridge.
- Hellwig Z. (1968) Zastosowanie metody taksonomicznej do typologicznego podziału krajów ze względu na poziom ich rozwoju oraz zasoby i strukturę kwalifikowanych kadr. *Przegląd Statystyczny*, 4.
- Hwang C. L., Yoon K. (1981) *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, New York.
- Kukuła K. (2000) *Metoda unitaryzacji zerowanej*. PWN, Warszawa.
- Kukuła K., Luty L. (2017) Jeszcze o procedurze wyboru metody porządkowania liniowego. *Przegląd Statystyczny*, 64(2), 163-176.
- Młodak A. (2006) *Analiza taksonomiczna w statystyce regionalnej*. DIFIN, Warszawa.
- Panek T. (2009) *Statystyczne metody wielowymiarowej analizy porównawczej*. Wydawnictwo SGH, Warszawa.
- Pluta W. (1976) *Wielowymiarowa analiza porównawcza w badaniach ekonomicznych*. PWE, Warszawa.
- Zeliaś A. (2000) *Taksonomiczna analiza przestrzennego zróżnicowania poziomu życia w Polsce w ujęciu dynamicznym*. Wydawnictwo AE Kraków.

ABOUT THE PROBLEMS OF APPLICABILITY OF SYNTHETIC MEASUREMENTS FOR ORDERING OBJECTS

Abstract: The problem of building a stable synthetic index used to organize the objects described with multiple partial indexes has been and is the subject of the authors' considerations. In our research, we aimed at the construction of a measure that would be insensitive to the input data, the results that would not depend on the method of normalization of the variables, the choice of the distance measure (similarity) and the selection of features. Our experience is consistent with the results of the work of the outstanding American statistician L. Breiman. A single classifier (synthetic measure) may be far from optimal, while combinations of many result in a classifier that is close to optimal and stable. Unfortunately, when "weak" classifiers are used, the combination may result in an even worse classifier. In the work, we presented the consequences of choosing a synthetic indicator for organizing objects described by many features on a practical example of farms in the EU countries covered by the FADN survey.

Keywords: synthetic measure, linear ordering of objects, classification, weights

JEL classification: C1, G2

WSKAŹNIKI DOBROBYTU O CHARAKTERZE ODSETKÓW – PRZEGLĄD

Dominik Sieradzki  <https://orcid.org/0000-0001-5843-3413>

Instytut Ekonomii i Finansów
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
e-mail: dominik_sieradzki@sggw.edu.pl

Streszczenie: Artykuł ma charakter przeglądowy. Celem badań było określenie wykorzystywanych wskaźników dobrobytu, które mają charakter odsetków. Na samym początku przedstawiono szkic historyczny ekonomii dobrobytu. Następnie przedstawiono opisane w literaturze domeny dobrobytu i zidentyfikowano odpowiadające im wskaźniki dobrobytu. W dalszej części przedstawiono klasyfikację wskaźników do poszczególnych podgrup - tradycyjnych mierników gospodarki, zmodyfikowanych rachunków narodowych, miar indeksowych oraz indyktorów środowiskowych. Na końcu omówiono źródła danych dla wskaźników dobrobytu, a następnie wskazano wskaźniki, które mają charakter odsetków.

Słowa kluczowe: ekonomia dobrobytu, pomiar dobrobytu, wskaźniki dobrobytu, odsetki

JEL classification: I31, I39

WSTĘP

Dyskusja na temat dobrobytu ma swój początek w XVIII wieku, kiedy to w 1776 roku Adam Smith opublikował swoje największe dzieło „Badania nad naturą i przyczynami bogactwa narodów”. Była to pierwsza systematyczna praca w dziedzinie ekonomii. Z pojęciem dobrobytu związane są takie terminy, jak: dobrobyt ekonomiczny, zamożność, poziom materialny czy jakość życia. Pierwsza koncepcja dobrobytu opiera się na doktrynie utylitaryzmu, której propagatorem w XVIII wieku był Jeremy Bentham [Bentham 1781]. Utylitaryzm zakłada, że szczęście grupy ludzi jest istotniejsze niż szczęście pojedynczych osób. Kontynuatorami rozważań na temat utylitaryzmu byli: współpracownik J. Benthama, ekonomista James Mill oraz jego syn John Stuart Mill, który

<https://doi.org/10.22630/MIBE.2020.21.4.19>

opowiadał się za utylitaryzmem pośrednim. Miał on być skupiony wokół zmian w działaniu instytucji, reguł postępowania czy ludzkich charakterów [Mill 1863]. Pierwszym zastosowaniem doktryny utylitaryzmu było słynne, wspomniane wyżej dzieło Adama Smitha [Smith 1954]. Adam Smith, David Ricardo oraz James Mill wzrost dobrobytu utożsamiali z reguły ze wzrostem wytworzonego produktu.

Wspomniana doktryna utylitaryzmu stała się podstawą ekonomii dobrobytu, która wyodrębniła się na przełomie wieku XIX i XX. Moment jej powstania datuje się umownie na rok 1912, kiedy Arthur Cecil Pigou opublikował swoją pracę [Pigou 1912], w której podał definicję ekonomii dobrobytu. Była to pierwsza teoretyczna praca poświęcona ekonomii dobrobytu. Ta gałąź ekonomii wywodzi się z neoklasycznej ekonomii Alfreda Marshalla i od wspomnianego wcześniej A. C. Pigou oraz z matematycznej szkoły lozańskiej reprezentowanej przez Vilfredo Pareto i Enrico Barone'a. Głównym zadaniem ekonomii dobrobytu jest określenie poziomu i struktury dobrobytu ekonomicznego. Popularnym wskaźnikiem do pomiaru jego poziomu jest Produkt Krajowy Brutto (PKB) lub PKB per capita, jako wartość rynkowa dóbr wytworzonych w danym kraju w ciągu roku [Czarny 2011] lub wartość dóbr w przeliczeniu na jednego mieszkańca. Rozważania teoretyczne w zakresie ekonomii dobrobytu nierozzerwalnie wiążą się z kwestią pomiaru dobrobytu. Najbardziej znanym wskaźnikiem pomiaru dobrobytu jest wspomniany syntetyczny wskaźnik Produktu Krajowego Brutto, który jest wyliczany według określonej metodologii [GUS 2020]. Duża część wskaźników dobrobytu jest uzyskiwana w wyniku przeprowadzania badań reprezentacyjnych.

MIERZENIE DOBROBYTU

W drugiej połowie XX wieku nad kwestią dobrobytu zaczęły intensywnie skupiać się inne nauki społeczne, w szczególności socjologia i psychologia, przez co zagadnienie dobrobytu poszerzano o elementy inne niż te dochodowe i materialne. Rozumienie dobrobytu wyszło poza poziom dochodu pieniężnego czy konsumpcję i zaczął być on traktowany jako dobrostan ogólny, poziom życia czy jakość życia.

W badaniach na temat pomiaru dobrobytu można wyodrębnić trzy nurty badawcze: pomiar dobrobytu ekonomicznego, społeczno-ekonomicznego oraz psychicznego [Diener, Suh 1997]. W każdym z tych nurtów część wykorzystywanych wskaźników jest możliwa do otrzymania po przeprowadzeniu badania reprezentacyjnego. Przykładem są badania sondażowe, które zawierają pytania pozwalające na subiektywną ocenę użyteczności dochodów respondentów [Kot 2000] czy poziomu ich zadowolenia [Blanchflower i in. 2000]. Prócz surowych wskaźników dobrobytu konstruuje się miary indeksowe, które uwzględniają różne aspekty dobrobytu. Są to przykładowo: Indeks Warunków Życia (Living Conditions Index), Indeks Jakości Życia QOL (Quality of Life Index) czy Indeks Rozwoju Społecznego HDI (Human Development Index).

W swojej budowie indeksy te uwzględniają m. in. takie aspekty jak warunki mieszkaniowe, kwestie żywnościowe, zdrowie oraz opiekę zdrowotną, aktywność sportową i kulturalną, bezpieczeństwo (określane przez stopę morderstw i gwałtów), stan rodzinny (stopa rozwodów), ale również aspekty ekonomiczne - warunki pracy, migracje, PKB na mieszkańca, bezrobocie, nierówność dochodową, ubóstwo, oczekiwaną długość życia czy wskaźnik skolaryzacji i alfabetyzmu. Przy konstrukcji poszczególnych indeksów należy określić wartość wszystkich jego składowych. W niektórych przypadkach istnieje możliwość ustalenia wartości prawdziwej, np. PKB na jednego mieszkańca. Część danych jest publikowana przez Główny Urząd Statystyczny jako wynik badania reprezentacyjnego.

W dotychczasowych badaniach naukowych nie wypracowano jednoznacznej definicji dobrobytu, której brak spowodowany jest jego złożonością. Rozpatruje się dwa określenia dobrobytu: w odniesieniu do osobistego życia poszczególnych jednostek i ich dobrostanu (ang. well-being) oraz w odniesieniu do wpływu gospodarki danego państwa na ogólny dobrostan społeczeństwa (ang. economic welfare). Obecnie dobrobyt danej jednostki jest rozumiany wielowymiarowo jako poziom realnego majątku, poczucia bezpieczeństwa ekonomiczno-społecznego (np. bezpieczeństwo zatrudnienia, stabilność cen, czy dostateczna ochrona zdrowia), posiadane umiejętności i wiedza oraz prowadzony styl życia [Kasprzyk 2013]. W publikacjach o tematyce dobrobytu używa się zamiennie takich określeń, jak poziom zamożności, poziom życia czy jakość życia [Kot i in. 2004]. Poziom zamożności dotyczy wartości materialnych, np. wielkości posiadanego majątku. Kategoria jakości czy poziomu odnosi się do szczęścia, zasobów i potrzeb. Jakość życia to pojęcie złożone, które obejmuje wiele dziedzin. Są to kategorie takie, jak konsumpcja, niemierzalne stany indywidualnego zadowolenia, satysfakcji, szczęście płynące z konsumpcji [Wronowska 2016].

Zgodnie z pierwszą koncepcją, dobrobyt można mierzyć przy pomocy tradycyjnych mierników potencjału gospodarki, które opierają się na Systemie Rachunków Narodowych. Należą do nich m.in. Produkt Krajowy Brutto (PKB), Produkt Narodowy Netto (PNN), Produkt Narodowy Brutto (PNB), Produkt Krajowy Netto (PKN) czy Produkt Krajowy Brutto na jednego mieszkańca (PKB per capita). W zamyśle, wartości powyższych miar miały stanowić pewne przybliżenie rozwoju gospodarki i dobrobytu danego społeczeństwa. Istnieje jednak wiele poglądów [Zienkowski 1979] zwracających uwagę na niedoskonałości tych wskaźników jako mierników dobrobytu. Najważniejszy z nich mówi, że miary te nie uwzględniają szeroko rozumianej jakości życia reprezentowanej np. przez stan zdrowia czy edukację. Miary te nie biorą również pod uwagę aspektu nierównomiernego podziału dochodu w społeczeństwie [Mankiw, Taylor 2016]. Z powodu tych wad [Talberth i in. 2007] podjęto próbę modyfikacji tradycyjnych mierników potencjału gospodarki poprzez uwzględnienie jakościowych aspektów życia [Cieślik 2008]. Miary takie określa się jako zmodyfikowane rachunki narodowe.

Jedną z nich jest Miernik Dobrobytu Ekonomicznego (Measure of Economic Welfare – MEW) [Nordhaus, Tobin 1972]. W swojej konstrukcji MEW bierze pod uwagę m.in. konsumpcję, wartość czasu wolnego, nakłady na ochronę środowiska czy straty z powodu zniszczenia środowiska i urbanizacji. Niedoskonałość stanowi subiektywny dobór jego składowych, brak uwzględniania kwestii rozdziału bogactwa w społeczeństwie oraz to, że szacunki poszczególnych składowych MEW są oparte na słabej jakości danych.

Drugą opracowaną propozycją w tej grupie jest Wskaźnik Ekonomicznych Aspektów Dobrobytu (Index of the Economic Aspects Welfare - EAW) [Zolotas 1981]. Miara ta opiera się na zestawie jakościowych zmian dobrobytu [Cieślik 2008]. Wskaźnik ten bierze pod uwagę wydatki na konsumpcję prywatną, które są korygowane o wydatki na dobra trwałego użytku, reklamy, koszty wyczerpywania się zasobów naturalnych, koszty społeczne związane z zanieczyszczeniem środowiska, koszty dojazdów do miejsca pracy oraz wydatki sektora prywatnego na zdrowie i edukację. Autor tej propozycji w swoich badaniach weryfikował hipotezę o malejącej funkcji wzrostu gospodarczego w rozwiniętych społeczeństwach. Hipoteza ta została potwierdzona w badaniach empirycznych, w których tempo wzrostu wskaźnika EAW dla zaawansowanych społeczeństw jest niższe od dynamiki wzrostu Produktu Narodowego Brutto w jednym okresie. Spowodowane jest to tym, że składowe odjęte od prywatnej konsumpcji rosną szybciej niż sam PNB [Zolotas 1981].

Mniej popularną propozycją pomiaru dobrobytu jest wykorzystanie „zazielenionej” wersji PKB. Jest to suma konsumpcji, oszczędności i usług środowiskowych [Żylicz 2007]. Miarę tę traktuje się jako liniowe przybliżenie funkcji dobrobytu. Sformułowanie „zazieleniony” spowodowane jest uwzględnianiem w budowie tego wskaźnika spożycia usług środowiskowych i inwestycji w zasoby naturalne.

Innym wskaźnikiem uwzględniającym środowisko jest Indeks Zrównoważonego Dobrobytu Ekonomicznego (Index of Sustainable Economic Welfare - ISEW) [Cobb, Daly 1989]. Jest on oparty na wydatkach konsumpcyjnych, pomniejszonych o straty wynikające z nierównomiernego rozkładu dochodów, koszty eksploatacji zasobów i niszczenie środowiska naturalnego (hałas, zanieczyszczenie wód i powietrza). Wskaźnik ten bierze pod uwagę aspekty naturalne, lecz nie uwzględnia ważnych elementów dobrobytu, jak na przykład oszczędności ludności [Noga 1996].

Innym podejściem do pomiaru dobrobytu jest zastosowanie miar indeksowych. Określają one jakość i poziom życia mieszkańców kraju, biorąc pod uwagę wiele wymiarów dobrobytu.

W latach 60. XX wieku zaproponowany został Wskaźnik Dobrobytu Ekonomicznego Netto (Net Economic Welfare - NEW) [Beckerman, Bacon 1966], co dało początek opracowaniu indykatów bazujących na metodach taksonomicznych. Składowe tego wskaźnika to m.in. realna prywatna konsumpcja, roczna produkcja cementu czy liczba pojazdów drogowych. W związku z tym, że

jest to miernik taksonomiczny wymagana jest konieczność określenia punktu odniesienia, czyli wzorca. Zazwyczaj dla tego wskaźnika za wzorzec przyjmuje się Stany Zjednoczone.

Następnym przykładem podejścia taksonomicznego jest tzw. metoda genewska, nazywana również odległością genewską [Drewnowski, Scott 1966]. Metoda ta opiera się na założeniu, że potrzeby ludzkie można podzielić na siedem kategorii: wyżywienie, mieszkanie, zdrowie, wykształcenie, rekreacja, bezpieczeństwo społeczne i materialne. Stopień zaspokojenia każdej z kategorii mierzy się stosując indyktor ilościowy oraz dwa indykatory jakościowe. Podejście takie charakteryzuje się wielowymiarowością, lecz napotyka ono na problem ustalenia wzorca. Występuje również trudność ujednoczenia tej metody na wszystkie państwa [Cieślik 2008].

Problematyczna w stosowaniu metod taksonomicznych jest konieczność określenia wzorca. Wady tej nie ma ważny syntetyczny wskaźnik rozwoju – Indeks Rozwoju Społecznego (Human Development Index – HDI).

Został on opracowany wspólnie przez Mahbubul Haq i Amartya Sena i opublikowany w raporcie Programu Narodów Zjednoczonych ds. Rozwoju [UNDP 1990]. Wskaźnik HDI bierze pod uwagę trzy wymiary rozwoju: długość życia, wiedzę (mierzoną wskaźnikami opisującymi piśmiennictwo społeczeństwa oraz długością trwania nauki) i standard życia reprezentowany przez PKB per capita, skorygowany parytetem siły nabywczej. Indeks ten przyjmuje wartość z przedziału od 0 do 1. Dla krajów wysoko rozwiniętych wskaźnik przyjmuje wartość powyżej 0,8, kraje średnio rozwinięte otrzymują HDI od 0,5 do 0,8, zaś kraje o niskim poziomie rozwoju poniżej 0,5. Powstała również modyfikacja tego wskaźnika - Wskaźnik Zróżnicowania Rozwoju Społecznego ze Względu na Płeć (Gender Development Index - GDI). Wskaźnik GDI bierze pod uwagę dodatkowo nierównomierny udział każdej płci w sferach, które są uwzględniane w HDI. Wyzwaniem przy obliczaniu obu tych miar jest dostępność wiarygodnych danych statystycznych.

Następną propozycją z tej grupy jest Wskaźnik Biedy Społecznej (Human Poverty Index - HPI) [Lawn 2003]. Uwzględnia on takie aspekty jak m. in.: konsumpcja indywidualna, usługi związane z użytkowaniem dróg i autostrad, przestępczość, niepełne zatrudnienie czy emisja dwutlenku węgla.

Metodyka konstrukcji HDI stanowiła bazę do budowy szeregu innych propozycji, które uwzględniają środowisko naturalne. Można wyróżnić takie wskaźniki jak: Indeks Trwałego Rozwoju Społecznego (Sustainable Human Development Index – SHDI) [Costantini, Monni 2005], Indeks Rozwoju Społecznego Uwzględniający Stopień Zanieczyszczenia Środowiska (Pollution Sensitive Human Development Index – HDPI) [Florczak 2008], Indeks Zagrożenia Środowiska Naturalnego (Environment Endangerment Index – EEI) [Neumayer 2007]. Wskaźniki te uwzględniają dodatkowo takie aspekty jak np. stopień zalesienia, emisja gazów cieplarnianych czy jakość środowiska naturalnego.

Ważną kwestię w ekonomii dobrobytu odgrywa kwestia sprawiedliwego podziału dobrobytu [Dalton 1920]. W praktyce badawczej ma to odzwierciedlenie w pomiarze równomierności rozkładu dochodów w społeczeństwie. Aspekt ten jest brany pod uwagę w konstrukcji niektórych wspomnianych wskaźników dobrobytu. Najczęściej wykorzystywaną miarą określającą równomierność rozkładu dochodów jest współczynnik Giniego [Gini 1912]. Przyjmuje on wartość w przedziale od 0 do 1. Przyjmuje się, że dla krajów rozwiniętych współczynnik ten wynosi około 0,3, z kolei w państwach rozwijających się jego wartość jest wyższa. Zbliżonym narzędziem badawczym jest indeks Theila [Peacock i in. 1988]. Jest on rzadziej stosowany, ale on również bada rozkład dochodu, uwzględniając dodatkowo grupy państw. Biorąc pod uwagę poziom i dystrybucję dochodów na jednego mieszkańca Sen opracował na tej podstawie pewną miarę dobrobytu [Sen 1992]. Według jego koncepcji wysoki poziom dochodu na mieszkańca nie jest tożsamy z wysokim poziomem dobrobytu. Należy uwzględnić również rozkład dochodu. Mianowicie kraje o niskich dochodach, ale mające równomierny ich podział w społeczeństwie oraz kraje o wysokim poziomie dochodu na mieszkańca oraz nierównomiernym ich rozkładzie, mogą zostać uznane za kraje o tym samym poziomie dobrobytu. Innym stosowanym wskaźnikiem pomiaru nierówności dochodowych jest indeks Zengi [Zenga 1990], przyjmujący również wartości z przedziału [0,1]. Określa on przeciętną koncentrację dochodów oraz poziom ich nierówności. Wartości wyższe świadczą o wysokiej nierówności w danej zbiorowości. Z kolei w opracowaniach Eurostatu do pomiaru zróżnicowania dochodowego wykorzystuje się miary oparte na kwantylach tzw. wskaźniki zróżnicowania decylogowego i kwintylogowego [Foster i in. 1984; Panek 2011; Jędrzejczak, Pekasiewicz 2020]. Określają one ile razy dochody osób najbogatszych są wyższe od dochodów osób najbiedniejszych. Wysokie wartości wskaźników kwintylogowego i decylogowego świadczą o dużych nierównościach dochodowych.

Ostatnią grupą miar dobrobytu są tak zwane indykatory środowiskowe. Metodologie ich konstrukcji koordynują międzynarodowe instytucje: ONZ, Bank Światowy oraz OECD. Przykładowymi wskaźnikami z tej grupy są: wskaźnik stopnia degradacji atmosfery (uwzględnia m. in. zniszczenie warstwy ozonowej), wskaźniki stanu gleby (uwzględnia m. in. zużycie nawozów czy powierzchnię pustynnienia) oraz wskaźniki zanieczyszczenia wód (uwzględnia m. in. zużycie wód powierzchniowych czy występowanie bakterii coli w wodzie pitnej). Indykatory środowiskowe stają się coraz bardziej popularne i w połączeniu z innymi wskaźnikami dobrobytu określają dość precyzyjnie poziom dobrobytu w danym państwie.

Powyższe podejścia do pomiaru dobrobytu uwzględniają różne jego aspekty, tzw. domeny dobrobytu. Osobnym obszarem badawczym jest pomiar dobrobytu z perspektywy przedsiębiorstw. Stosowanym tu narzędziem jest tak zwany test koniunktury, który składa się z odpowiedniego kwestionariusza, utworzonego według ścisłej metodologii [GUS 2013]. Analizę aplikacyjną i teoretyczną tego

narzędzia badawczego można znaleźć w innych licznych pracach z tej tematyki [Męczarski 1998; Silgoner 2007; Kalinowski 2015].

Przystępując do pomiaru dobrobytu staje się przed dylematem wyboru dziedzin życia, które należy wziąć pod uwagę. Dotychczas nie został opracowany jeden powszechnie stosowany zestaw dziedzin życia. W badaniach, te dziedziny życia określa się zazwyczaj domenami życia. Dotychczas opracowano różne podejścia do identyfikacji tych domen. Rozwiązania te zostały zaproponowane przez różne ośrodki i zespoły badawcze. Część z nich ma charakter międzynarodowy, zaś inne są stosowane w krajowych badaniach. W dalszej części zostaną przedstawione polskie koncepcje określające domeny dobrobytu.

POLSKIE PODEJŚCIA DO BADANIA JAKOŚCI ŻYCIA

W Polsce również prowadzone są badania na temat dobrobytu i jakości życia. Opierają się one na identyfikacji grup potrzeb, które mogą być traktowane jako domeny jakości życia. Potrzebę można zdefiniować jako odczuwany przez jednostkę stan braku czegoś [Lisowski 2002]. W poniższej tabeli 1 przedstawiono wyróżnione grupy potrzeb w badaniach przeprowadzonych na przestrzeni lat.

Tabela 1. Przegląd polskich badań dotyczących jakości życia

Lata badań	Wyróżnione grupy potrzeb	Źródło
1968-1971	wyżywienie, mieszkanie, ochrona zdrowia, wykształcenie, rekreacja, zabezpieczenie społeczne, zagospodarowanie materialne	Luszniewicz [1982]
1988-1989	wyżywienie, mieszkanie, zdrowie, oświata, rekreacja, zabezpieczenie społeczne, zabezpieczenie materialne, środowisko naturalne	Krawczak i in. [1990]
1992	dochody osobiste ludności, degradacja środowiska naturalnego, poziom zdrowotności, warunki pracy, stosunki społeczne i bezpieczeństwo publiczne	Śmiłowska [1995]
1994	stan biologiczny (wyżywienie, mieszkanie, zdrowie, środowisko naturalne, wypoczynek), stan zawodowy (posiadanie pracy, warunki pracy, czas pracy, płace), stan materialny (oszczędności, ceny, dobra trwałe), stan edukacyjny (oświata dzieci i młodzieży, oświata dorosłych, kultura i sztuka), stan społeczny (bezpieczeństwo socjalne, egalitaryzm dochodowy, patologia społeczna, więzi rodzinne i społeczne, polityka)	Słaby [1994]
2004	dochody, wydatki gospodarstwa domowego, spożycie artykułów żywnościowych, warunki mieszkaniowe, wyposażenie gospodarstw domowych w przedmioty trwałego użytkowania, ochrona zdrowia i opieka społeczna, edukacja, kultura i wypoczynek	GUS [2004]

Lata badań	Wyróżnione grupy potrzeb	Źródło
2005	dochody i sposób ich gospodarowania, wyżywienie, zasobność materialna, warunki mieszkaniowe, edukacja, kultura i wypoczynek, opieka zdrowotna, stan obywatelstwa społecznego, rynek pracy, niepełnosprawność, dyskryminacja	Czapiński i Panek [2006]

Źródło: opracowanie własne na podstawie [Panek 2016]

Grupy potrzeb, które zostały wyróżnione w polskich badaniach są zbieżne z domenami życia w koncepcjach zagranicznych. W ostatnim przeprowadzonym badaniu w ramach projektu „Diagnoza społeczna” wyróżniono następujące domeny jakości życia [Czapiński, Panek 2015]: kapitał społeczny, dobrostan psychiczny, dobrostan fizyczny, poziom cywilizacyjny, dobrobyt materialny, stres życiowy, patologie. Z kolei w badaniach jakości życia przeprowadzanych przez Główny Urząd Statystyczny przyjęto domeny życia sformułowane przez Eurostat [Szukielójć-Bieńkuńska 2015]. Poniżej przedstawiono listę wskaźników dobrobytu według GUS w podziale na domeny jakości życia [Szukielójć-Bieńkuńska 2015].

1. Materialne warunki życia

- przeciętny miesięczny dochód rozporządzalny,
- wskaźnik zróżnicowania kwintylowego dochodów rozporządzalnych,
- wskaźnik zagrożenia ubóstwem skrajnym,
- wskaźnik zagrożenia ubóstwem wielowymiarowym,
- wskaźnik dobrej samooceny budżetu domowego,
- udział wydatków na żywność i napoje bezalkoholowe w wydatkach ogółem,
- wskaźnik złych warunków sanitarnych,
- gospodarstwa domowe mieszkające w zbyt dużych/małych mieszkaniach,
- brak pieniędzy na żywność,
- brak pieniędzy na przynajmniej jeden tydzień wakacji raz w roku,
- brak samochodu z przyczyn finansowych.

2. Główny rodzaj aktywności, praca

- wskaźnik zatrudnienia,
- stopa bezrobocia,
- stopa bezrobocia długotrwałego,
- młodzież o wydłużonym czasie pracy,
- pracujący na podstawie umowy na czas określony,
- mediana wynagrodzeń,
- mediana świadczeń emerytalnych,
- wskaźnik ogólnego zadowolenia z pracy.

3. Zdrowie

- wskaźnik dobrej/złej samooceny zdrowia,
- osoby w wieku 15 lat i więcej z nadwagą i otyłością,
- osoby palące tytoń codziennie,
- rezygnacja z wizyty u lekarza z powodów finansowych/listy oczekujących.

4. Edukacja

- osoby posługujące się językami obcymi,
- osoby powyżej 25 roku życia z wykształceniem co najwyżej gimnazjalnym,
- osoby w wieku 15-64 lat uczestniczące w kształceniu i szkoleniu,
- odsetek osób w wieku 16-74 lat korzystających z komputera/internetu.

5. Czas wolny i relacje społeczne

- satysfakcja z ilości czasu wolnego,
- zaufanie do innych osób.

6. Bezpieczeństwo ekonomiczne i fizyczny

- zaległości w opłatach,
- poczucie bezpieczeństwa w miejscu zamieszkania,
- brak możliwości pokrycia nieoczekiwanego wydatku.

7. Państwo i podstawowe prawa, aktywność obywatelska

- zaufanie do systemu politycznego/prawnego,
- poczucie zagrożenia dyskryminacją,
- frekwencja w wyborach do sejmiku RP.

8. Jakość środowiska w miejscu zamieszkania

- narażenie na nadmierny hałas,
- narażenie na zanieczyszczenie,
- zadowolenie z terenów rekreacyjnych.

9. Subiektywny dobrostan

- ogólne zadowolenie z życia,
- złożony wskaźnik złego samopoczucia,
- poczucie sensu życia.

Drugi polski system wskaźników był stosowany w ramach badania „Diagnoza społeczna” (badanie realizowane cyklicznie w latach 1989-2015 przez Radę Monitoringu Społecznego). Poniżej przedstawiono listę wskaźników dobrobytu stosowanych w tym badaniu w podziale na domeny jakości życia (Czapiński i Panek 2015).

1. Kapitał społeczny
 - aktywność na rzecz środowiska lokalnego,
 - udział w wyborach,
 - stosunek do demokracji.
2. Dobrostan psychiczny
 - poczucie szczęścia,
 - ocena minionego roku.
3. Dobrostan fizyczny
 - poważna choroba w minionym roku,
 - niepełnosprawność.
4. Dobrostan społeczny
 - brak poczucia osamotnienia,
 - liczba przyjaciół.
5. Poziom cywilizacyjny
 - poziom wykształcenia,
 - posiadanie nowoczesnych urządzeń komunikacyjnych,
 - posiadanie prawa jazdy,
 - czynna znajomość języków obcych.
6. Dobrobyt materialny
 - dochód gospodarstwa domowego na osobę.
7. Stres życiowy
8. Patologie
 - nadużywanie alkoholu i narkotyków,
 - wizyty u psychiatry lub psychologa,
 - bycie sprawcą lub ofiarą łamania prawa.

Oba polskie systemy wskaźników charakteryzują się wielowymiarowością. Ponadto uwzględniają one zarówno obiektywne, jak i subiektywne aspekty jakości życia.

WSKAŹNIKI DOBROBYTU O CHARAKTERZE ODSETKÓW

Część wskaźników dobrobytu (jakości życia) dotyczy zjawisk dychotomicznych, czyli takich, w których obserwacje są zerojedynkowe. Celem badania jest wnioskowanie o odsetku obiektów w całej populacji, które mają pewną własność. Odsetek ten wyraża się często w procentach. W zależności od dziedziny odsetek ten określa się mianem wskaźnika struktury, proporcji, frakcji

czy częstości względnej. Przykładowa populacja $U=\{u_1, \dots, u_N\}$ składa się z N elementów i przedmiotem badania jest informacja o pewnej cesze \mathbf{Y} . Wartości cechy \mathbf{Y} dla każdego obiektu to N -wymiarowy wektor \mathbf{Y} :

$$\mathbf{Y} = (Y_1, \dots, Y_N)^T. \quad (1)$$

W przypadku wskaźnika struktury poszczególne współrzędne wektora \mathbf{Y} przyjmują wartości 1 w przypadku, gdy i -ty obiekt charakteryzuje się badaną własnością, lub 0 w przeciwnym przypadku. Wskaźnik struktury oznacza się z reguły jako θ i definiuje się jako

$$\theta = \bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i = \frac{M}{N}, \quad (2)$$

gdzie M oznacza liczbę elementów w populacji z pewną własnością. Z przytoczonych wskaźników jakości życia część z nich ma charakter frakcji. Poniżej dokonano ich identyfikacji w podziale na badania reprezentacyjne, które dotyczą dobrobytu [GUS 2017].

1. Badanie Budżetów Gospodarstw Domowych

- Zasięg ubóstwa skrajnego - odsetek gospodarstw domowych, w których poziom wydatków (obejmujących również wartość artykułów otrzymanych bezpłatnie oraz wartość spożycia naturalnego) był niższy od przyjętej granicy ubóstwa skrajnego (minimum egzystencji).
- Wskaźnik dobrej samooceny budżetu domowego - odsetek gospodarstw domowych, którym według ich własnej oceny pieniędzy starcza na wiele bez specjalnego oszczędzania lub które mogą pozwolić sobie na pewien luksus.
- Wskaźnik złej samooceny budżetu domowego - odsetek gospodarstw domowych, które na co dzień bardzo oszczędnie gospodarują pieniędzmi lub którym według ich własnej oceny pieniędzy nie starcza na zaspokojenie podstawowych potrzeb.

2. Europejskie Badanie Dochodów i Warunków Życia

- Wskaźnik złych warunków sanitarnych - odsetek gospodarstw domowych, które nie posiadają w mieszkaniu ciepłej wody bieżącej lub ustępu spłukiwanego wodą bieżącą lub łazienki z wanną/prysznicem.
- Gospodarstwa domowe mieszkające w zbyt dużych/małych mieszkaniach - odsetek gospodarstw domowych, które określają swoje mieszkanie jako zbyt duże/małe w stosunku do swoich potrzeb.
- Brak samochodu z przyczyn finansowych - odsetek gospodarstw domowych nieposiadających przynajmniej jednego samochodu prywatnego z powodów finansowych.
- Osoby o ograniczonej zdolności wykonywania codziennych czynności - odsetek osób w wieku 16 lat i więcej wskazujących na ograniczoną zdolność wykonywania podstawowych czynności, trwającą 6 miesięcy lub dłużej.

- Wskaźnik dobrej/złej samooceny stanu zdrowia - odsetek osób w wieku 16 lat i więcej, które oceniły ogólny stan swojego zdrowia jako dobry/zły lub bardzo dobry/zły.
 - Rezygnacja z wizyty u lekarza specjalisty z powodu listy oczekujących, braku skierowania - odsetek osób w wieku 16 lat i więcej, które nie skorzystały z konsultacji z lekarzem rodzinnym, pierwszego kontaktu lub specjalistą z powodu listy oczekujących lub braku skierowania wśród osób, które deklarowały potrzebę takiej konsultacji.
 - Satysfakcja z ilości czasu wolnego - odsetek osób w wieku 16 lat i więcej deklarujących, że są zadowolone lub bardzo zadowolone z ilości czasu wolnego, jakim dysponują.
 - Zaufanie do innych osób - odsetek osób w wieku 16 lat i więcej deklarujących, że zdecydowanie mają zaufanie lub raczej mają zaufanie do ludzi w ogólnym rozumieniu tego pojęcia.
 - Brak możliwości pokrycia nieoczekiwanego wydatku - odsetek gospodarstw domowych deklarujących brak możliwości pokrycia z własnych środków nieoczekiwanego wydatku w wysokości 1200 zł.
 - Zaległości w opłatach - odsetek gospodarstw domowych, które z powodów finansowych, w okresie 12 miesięcy przed badaniem, dwa lub więcej razy nie
 - były w stanie terminowo uiszczać takich opłat jak: rachunki za gaz, elektryczność, wodę, wywóz nieczystości stałych lub płynnych, podatków od nieruchomości i innych stałych opłat związanych z użytkowaniem mieszkania.
 - Poczucie bezpieczeństwa w miejscu zamieszkania - odsetek osób w wieku 16 lat i więcej deklarujących, że chodząc samotnie po zmroku w okolicy swojego miejsca zamieszkania czują się bardzo bezpiecznie lub raczej bezpiecznie.
3. Badanie Aktywności Ekonomicznej Ludności
- Wskaźnik ogólnego zadowolenia z pracy - odsetek osób pracujących w wieku 16 lat i więcej deklarujących, że są zadowolone lub bardzo zadowolone ze swojej obecnej sytuacji zawodowej.
 - Osoby powyżej 25. roku życia z wykształceniem wyższym - odsetek osób w wieku 25 lat i więcej posiadających wyższe wykształcenie
 - Osoby w wieku 15-64 lata uczestniczące w kształceniu lub szkoleniu - odsetek osób w wieku 15-64 lata uczących się lub doksztalających się wśród ludności w tej samej grupie wiekowej (w okresie czterech tygodni przed badaniem).

4. Europejskie Badanie Ankietowe Zdrowia

- Osoby w wieku 15 lat i więcej z nadwagą lub otyłością - odsetek osób z nadwagą lub otyłością wśród osób w wieku 15 lat i więcej. Nadwaga i otyłość jest klasyfikowana indeksem masy ciała (BMI), gdzie relacja masy ciała (w kg) do wysokości (podniesionej do kwadratu w metrach) przyjmuje wartość co najmniej 30 (otyłość) lub 25 (nadwaga).
- Odsetek osób w wieku 15 lat i więcej deklarujących codzienne palenie tytoniu - odsetek osób w wieku 15 lat i więcej, które deklarują palenie tytoniu każdego dnia.

5. Badanie Spójności Społecznej

- Osoby posługujące się językami obcymi - odsetek osób w wieku 16 lat i więcej, które deklarują posługiwanie się przynajmniej jednym językiem obcym w stopniu co najmniej komunikatywnym (umiejętność posługiwania się podstawowymi wyrażeniami oraz używania języka dla określenia znanych rzeczy i sytuacji).
- Zaufanie do innych osób.

6. Badanie Wykorzystywania Technologii Informacyjno-Telekomunikacyjnych (ICT) w gospodarstwach domowych

- Osoby w wieku 16-74 lata korzystające regularnie z komputera - odsetek osób w wieku 16-74 lata korzystających regularnie co najmniej raz w tygodniu z komputera w ciągu ostatnich 3 miesięcy.
- Osoby w wieku 16-74 lata korzystające regularnie z Internetu - odsetek osób w wieku 16-74 lat korzystających regularnie co najmniej raz w tygodniu z Internetu w ciągu ostatnich 3 miesięcy.

PODSUMOWANIE

W opracowaniu przedstawiono stosowane wskaźniki dobrobytu z uwzględnieniem podziału na różne domeny według polskiego podejścia. Ponadto przedstawiono klasyfikację stosowanych wskaźników i zidentyfikowano dla nich źródła danych. W ostatnim kroku dokonano identyfikacji wskaźników dobrobytu, które mają charakter odsetków. Wykorzystanym narzędziem badawczym była analiza literatury przedmiotu, w tym sprawozdań technicznych. Dalszy kierunek badań autora będzie dotyczył metodologii pobierania próby w badaniach dobrobytu.

BIBLIOGRAFIA

- Beckerman W., Bacon R (1966) International Comparisons of Income Levels: A Suggested New Measure. *The Economic Journal* 76(303), 519-536.
- Bentham J. (1780). *An Introduction to the Principles of Morals and Legislation*. London: T. Payne i Sons.

- Blanchflower D. G., Oswald A. J. (2000) Well-Being Over Time in Britain and the USA. Working Paper 7487. National Bureau of Economic Research.
- Cieślak E. (2008). Wybrane alternatywne sposoby mierzenia poziomu rozwoju gospodarczego. *Equilibrium*, 1(1-2), 145-160.
- Costantini V., Monni S. (2005) Sustainable Human Development for European Countries. *Journal of Human Development and Capabilities*, 6(3), 329-351.
- Czapiński J., Panek T. (2006) Diagnoza społeczna 2005. Warunki i jakość życia Polaków.
- Czapiński J., Panek T. (2015) Diagnoza społeczna 2015. Warunki i jakość życia Polaków.
- Czarny B. (2011) Podstawy ekonomii. Polskie Wydawnictwo Ekonomiczne.
- Dalton H. (1920) The measurement of the inequality of incomes. *The Economic Journal* 30(119), 348-361.
- Daly H. E., Cobb Jr. J. B. (1989) For the common good: redirecting the economy toward community, the environment, and a sustainable future. Beacon Press.
- Diener E., Suh E. (1997) Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research*, 40(1), 189-216.
- Drewnowski J., Scott W. (1966) The level of living index. Geneva, United Nations Research Institute for Social Development.
- Florczak W. (2008) Wskaźniki zrównoważonego rozwoju. *Wiadomości Statystyczne*, 3, 14-34.
- Foster J., Greer J., Thorbecke E. (1984) A Class of Decomposable Poverty Measures. *Econometrica*, 52(3), 761-766.
- Gini C. (1912) Variability and Mutability. C. Cuppini, Bologna.
- GUS (2004) Warunki życia ludności w 2004 r. Spraw. tech. GUS.
- GUS (2013) Badanie koniunktury gospodarczej. Zeszyt metodologiczny zaopiniowany przez Komisję Metodologiczną GUS.
- GUS (2017) Jakość życia w Polsce, edycja 2017. Spraw. tech. Główny Urząd Statystyczny.
- GUS (2020) Wstępny szacunek produktu krajowego brutto w II kwartale 2020 r.
- Jędrzejczak A., Pekasiewicz D. (2020) Changes in Income Distribution for Different Family Types in Poland. *International Advances in Economic Research*, 26(2), 135-146.
- Kalinowski S. (2015) Czynniki wpływające na opinie przedsiębiorców w kwestionariuszowych badaniach koniunktury. *Bank i Kredyt*, 46(4), 393-410.
- Kasprzyk B. (2013) Wybrane aspekty oceny dobrobytu ekonomicznego i jakości życia (ujęcie regionalne - Podkarpacie). Wydawnictwo Uniwersytetu Rzeszowskiego.
- Kot S. M. (2000) Ekonometryczne modele dobrobytu. Wydawnictwo Naukowe PWN, Warszawa.
- Kot S. M., Małowski A., Węgrzecki A. (red.) (2004) Dobrobyt społeczny, nierówności i sprawiedliwość dystrybucyjna. Wydawnictwo Akademii Ekonomicznej w Krakowie.
- Krawczak I. i in. (1990) Poziom życia ludności Polski w latach 1988-1989. SGPiS, Warszawa.
- Lawn P. (2003) A Theoretical Foundation to Support the Index of Sustainable Economic Welfare (ISEW), Genuine Progress Indicator (GPI), and Other Related Indexes. *Ecological Economics*, 44(1), 105-118.
- Lisowski A. (2002) Potrzeby społeczne i ich diagnozowanie. Red. A. Kurzynowski. Warszawa: SGH.

- Luszniewicz A. (1982) Statystyka społeczna. Warszawa, Polskie Wydawnictwo Ekonomiczne.
- Mankiw G. N., Taylor M. P. (2016) Makroekonomia. Polskie Wydawnictwo Ekonomiczne.
- Męczarski M. (1998) Zróżnicowanie koniunktury. Metody wnioskowania statystycznego. Prace i Materiały Instytutu Rozwoju Gospodarczego SGH.
- Mill J. S. (1863) Utilitarianism. Parker, Son, and Bourn, London.
- Neumayer E. (2007) Sustainability and Well-being Indicators. [in:] M. McGillivray (ed.) Human Well-Being, 193-213.
- Noga M. (1996) Pomiar dobrobytu społecznego uwzględniający stan i jakość środowiska przyrodniczego człowieka. Ekonomia i Środowisko, 19.
- Nordhaus W., James Tobin J. (1972) Is Growth Obsolete? [in:] Economic Research: Retrospect and Prospect, Economic Growth. 5, 1-80.
- Panek T. (2011) Ubóstwo, wykluczenie społeczne i nierówności. Teoria i praktyka pomiaru. Oficyna Wydawnicza SGH.
- Panek T. (red.) (2016) Jakość życia od koncepcji do pomiaru. Warszawa: Oficyna Wydawnicza Szkoła Główna Handlowa w Warszawie.
- Peacock W. G., Hoover G. A., Killian C.D. (1988) Divergence and Convergence in International Development: A Decomposition Analysis of Inequality in the World System. American Sociological Review, 53(6), 838-852.
- Pigou A. C. (1912) Wealth and Welfare. Macmillan i Company, limited.
- Sen A. (1992) Inequality Reexamined. Harvard University Press.
- Silgoner M. A. (2007) The Economic Sentiment Indicator: Leading Indicator Properties in Old and New EU Member States. Journal of Business Cycle Measurement and Analysis, 3(2), 199-215.
- Słaby T. (1994) Systemy wskaźników społecznych w polskich warunkach transformacji rynkowej. Monografie i Opracowania, Szkoła Główna Handlowa.
- Smith A. (1954) Badania nad naturą i przyczynami bogactwa narodów. Tom I. PWN.
- Szukielój-Bieńkuńska A. (2015) Pomiar jakości życia w statystyce publicznej. Wiadomości Statystyczne, 7, 19-32.
- Śmiłowska T. (1995) Zróżnicowanie poziomu i jakości życia ludności w przekroju terytorialnym. [w:] Z Prac Zakładu Badań Statystyczno-Ekonomicznych, GUS.
- Talberth J., Cobb C., Slattery N. (2007) The Genuine Progress Indicator 2006. Redefining Progress, Oakland.
- UNDP (1990) Human Development Report 1990. Concept and Measurement of Human Development. New York. <http://www.hdr.undp.org/en/reports/global/hdr1990>
- Wronowska G. (2016) Dobrobyt – ujęcie teoretyczne i pomiar. Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie, 12(948), 5-16.
- Zenga M. (1990) Concentration Curves And Concentration Indexes Derived From Them. [In:] Dagum C., Zenga M. (eds) Income and Wealth Distribution, Inequality and Poverty. Studies in Contemporary Economics. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-84250-4_7
- Zienkowski L. (1979) Poziom życia: metody mierzenia i oceny. PWE, Warszawa.
- Zolotas X. (1981) Economic Growth and Declining Social Welfare / Xenophon Zolotas. English. Bank of Greece Athens, xiii, 199 p.

Żylicz T. (2007) Trwały rozwój w teorii ekonomii [w:] B. Poskrobko (red.) Obszary badań nad trwałym i zrównoważonym rozwojem. Wydawnictwo Ekonomia i Środowisko, Białystok.

WELFARE INDICATORS IN THE FORM OF PROPORTIONS – REVIEW

Abstract: This article is revision-based. The aim of the research was to define the welfare measures used, which are in the form of proportion. At the beginning, a historical sketch of the welfare economy was presented. Then, the welfare domains described in the literature are presented and the corresponding welfare indicators were identified. In the next section of the article indicators classified into particular subgroups were presented. The subgroups are: traditional measures of the economy, modified national accounts, index measures and environmental indicators. The article ends with a discussion of the data sources for the welfare indicators and based on that the indicators in the form of proportion were indicated.

Keywords: welfare economics, welfare measure, welfare indicators, proportions

JEL classification: I31, I39

WIZUALIZACJA JAKO NARZĘDZIE WSPOMAGAJĄCE ANALIZĘ DANYCH W PROCESIE DYDAKTYCZNYM

Monika Zielńska-Sitkiewicz  <https://orcid.org/0000-0003-4829-3239>

Mariola Chrzanowska  <https://orcid.org/0000-0002-8743-7437>

Instytut Ekonomii i Finansów

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

e-mail: monika_zielinska_sitkiewicz@sggw.edu.pl

mariola_chrzanowska@sggw.edu.pl

Streszczenie Prezentacja informacji w postaci graficznej jest jedną z podstawowych form wizualizacji danych. Taka forma przedstawienia informacji jest dużym wsparciem zarówno podczas wstępnej, jak i dalszej analizy. Jednakże niewłaściwa forma graficzna może prowadzić do mistyfikacji, a w konsekwencji do błędnych wniosków. W niniejszej pracy zaprezentowano wybrane przykłady graficznej prezentacji danych zaczerpnięte z praktyki dydaktycznej autorek. Ponadto omówiono przypadki poprawnej oraz błędnej prezentacji danych.

Słowa kluczowe: prezentacja graficzna, szeregi czasowe, praktyka dydaktyczna

JEL classification: C19, Y10

WPROWADZENIE

Graficzna prezentacja danych i wyników analiz pełni bardzo ważną rolę w procesie dydaktycznym. Jest ona istotna zarówno na wczesnym etapie, czyli podczas poznawania i interpretowania zależności opisywanych przez dane, jak i później, czyli podczas prezentacji tych zależności innym osobom.

Jak dowodzi P. Biecek, dobra grafika statystyczna powinna pokazywać informację zawartą w danych liczbowych. Powinna to robić w taki sposób, by łatwo było odczytać i zrozumieć związek pomiędzy informacją a danymi. Obrazować, jak duże są pewne wielkości, jak ryzykowne są pewne rozwiązania, jak wyglądają zależności pomiędzy zjawiskami [Biecek 2014].

Dość trudnym zadaniem dydaktycznym w procesie kształcenia studentów jest przekazanie im wiedzy na temat badania i oceny różnego rodzaju danych

<https://doi.org/10.22630/MIBE.2020.21.4.20>

ilościowych. Jednym ze sposobów prezentacji analizy jest jej opracowanie w formie graficznej. Pozwala to w przystępny sposób zaprezentować zależności zachodzące pomiędzy mierzonymi bądź obserwowanymi wielkościami [Lenik i in. 2007].

Dla człowieka najbardziej naturalnym i najlepiej rozwiniętym źródłem informacji o obserwowanych obiektach jest zmysł wzroku. Według E. Dale'a¹, który opracował w 1946 r. „Cone of Experience” (Stożek Doświadczeń), ludzie uczący się mogą znacznie poprawić zdolność zapamiętywania, przyswajając wiedzę w oparciu o formy audio-wizualne oraz korzystając z doświadczenia [Thalheimer 2006].

Ponadto liczne badania potwierdziły, że jeśli proces dydaktyczny zostanie uzupełniony metodami interaktywnymi, z odpowiednio opracowaną grafiką, to korzyści poznawcze mogą być dużo wyższe. Jedno z takich badań przeprowadzone przez G. L. Adamsa wykazało następujące zyski:

- skuteczność nauczania może być większa o 56%;
- zrozumienie tematu może wzrosnąć od 56% do 60%;
- oszczędność czasu może wynieść od 38% do 70%;
- zakres przyswojonej wiedzy może być o 25% – 50% szerszy (por. [Adams 1992]).

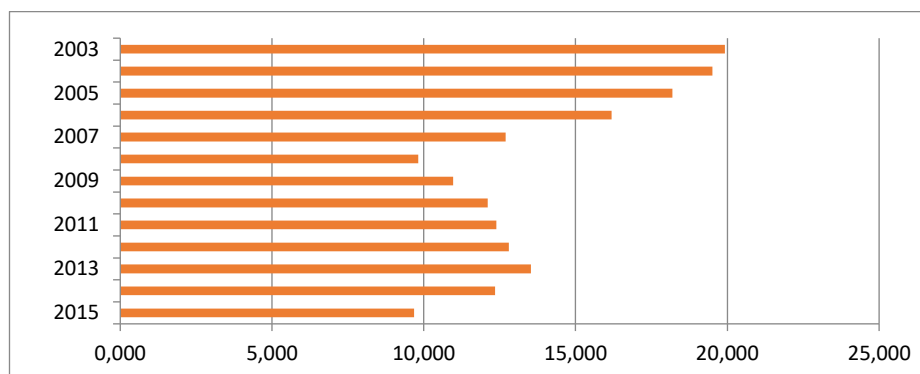
Zatem zwizualizowanie prezentowanego zagadnienia przyczynia się do lepszego zapamiętywania oraz budowania systemów skojarzeń, czyli instalowania tzw. haków pamięciowych.

Głównym celem pracy jest podkreślenie pomocniczej roli wizualizacji danych w kolejnych etapach analizy danych ze szczególnym uwzględnieniem danych czasowych. Cel został zrealizowany za pomocą kilku przykładów, które pozwolą bezpośrednio uzasadnić stosowanie metod wizualizacji podczas przeprowadzonych analiz. Każdy z przedstawionych przykładów jest propozycją wykorzystania metod wizualizacji danych na jednym z etapów eksploracji szeregu czasowego. Stanowią one również pewien zbiór dydaktycznych doświadczeń przedyskutowanych przez autorki ze studentami na różnych etapach ich edukacji.

Przykład z projektu studenckiego realizowanego na 1 roku studiów na kierunku Informatyka i Ekonometria (por. rysunek 1) obrazuje, jak bardzo potrzebne jest kształcenie w kierunku prawidłowego tworzenia i interpretacji grafiki danych czasowych. Uproszczona automatyzacja wielu narzędzi do tworzenia grafiki w popularnych programach komputerowych (np. MS Excel) powoduje, że autorzy nie przywiązują należytej uwagi do kontroli uzyskanych efektów wizualnych i nie dostrzegają błędów merytorycznych.

¹ Dale E. (1946, 1954, 1969). Audio-visual methods in teaching. New York: Dryden.

Rysunek 1. Grafika obrazująca stopę bezrobocia, wykonana przez studenta 1 roku, kierunku Informatyka i Ekonometria



Źródło: projekt studenta. Stopa bezrobocia (dane GUS)

W trakcie procesu dydaktycznego należy zatem zwracać szczególną uwagę na prawidłową i przemyślaną formę prezentacji danych.

WYKRYWANIE DANYCH ODSTAJĄCYCH

Analiza pierwotnego materiału statystycznego pozwala wyodrębnić jednostki, dla których wartości zmiennej znacznie odbiegają od pozostałych, czyli tzw. jednostki (obserwacje) odstające. Nietypowość informacji spowodowana jest na ogół niejednorodnością zbiorowości statystycznej, z której została pobrana próba, czy też nieoczekiwanymi zmianami, jakie zaszły w badanym zjawisku. Przyczyną jej powstania może być również błąd popełniony podczas pomiaru lub zapisywania wyników. Jak podaje P. Dittman, obserwacjami nietypowymi (odstającymi, odizolowanymi) nazywane są obserwacje, które znacząco różnią się od całego zbioru danych. Z tego powodu mogą one wywierać większy wpływ na oszacowania parametrów modelu niż pozostałe obserwacje² [Dittman 2000].

Obserwacje odstające wyraźnie różnią się wartością od pozostałych. Może to być efekt występujących w badanej zmiennej silnych wahań o charakterze losowym lub błędnego zapisu. Najprostszym sposobem wykrywania danych nietypowych dla danych jednowymiarowych jest ich wizualizacja na wykresie liniowym (rysunki 2 i 3) oraz wykresie pudełkowym ramka-wąsy (rysunek 4).

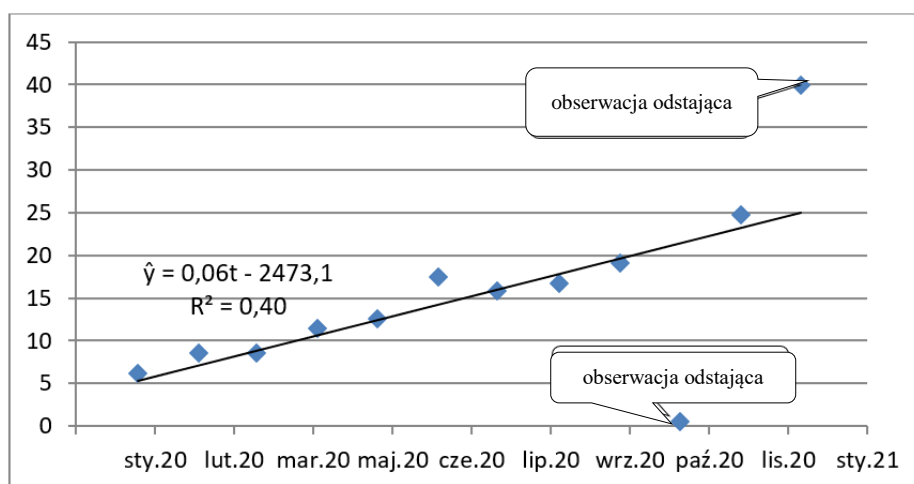
² W analizie zjawisk ekonomicznych informacja o pojawieniu się obserwacji nietypowych może wskazywać na zmianę tendencji analizowanego zjawiska, a w przypadku wartości ekstremalnych na występowanie tzw. punktów zwrotnych w analizie badanego zjawiska.

Tabela 1. Miesięczna sprzedaż (w tys. sztuk) zabawek typu Auto w 2020 roku

sty-20	lut-20	mar-20	kwi-20	maj-20	cze-20
6,17	8,57	8,61	11,47	12,54	17,49
lip-20	sie-20	wrz-20	paź-20	lis-20	gru-20
15,90	16,78	19,07	0,50	24,71	40,00

Źródło: dane umowne

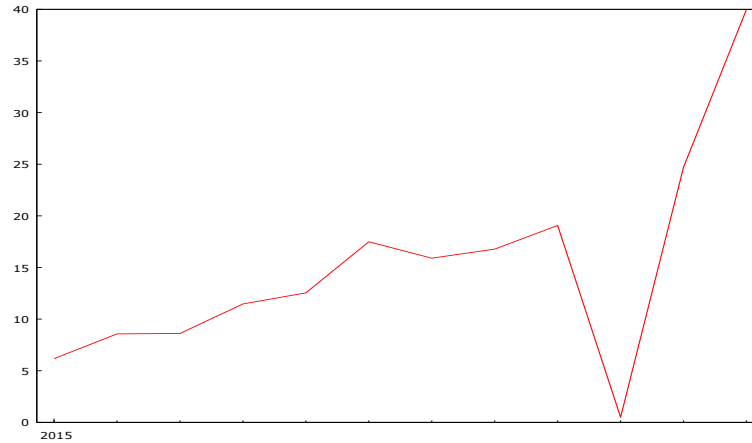
Rysunek 2. Prezentacja miesięcznej sprzedaży zabawek typu Auto w 2020 roku (tys. szt.)



Źródło: dane umowne (ilustracja MS Excel)

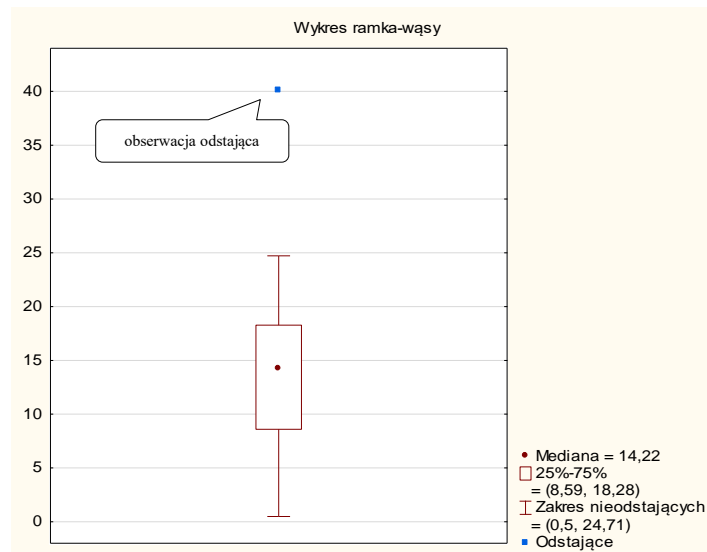
Rysunki 2 i 3 zawierają prezentacje miesięcznej sprzedaży zabawek typu Auto. Oba wykresy zostały zbudowane w dwóch programach powszechnie wykorzystywanych podczas zajęć praktycznych. Automatyzacja procedur tworzenia grafiki może jednak prowadzić do błędnych interpretacji ze względu na przypadkowe zniekształcenie percepcji danych, co obrazuje rysunek 3 – komunikat graficzny nie sugeruje w sposób czytelny występowania obserwacji nietypowych.

Rysunek 3. Prezentacja miesięcznej sprzedaży zabawek typu Auto w 2020 roku (tys. szt.)



Źródło: dane umowne (ilustracja pakiet R)

Rysunek 4. Wykres pudełkowy dla miesięcznej sprzedaży zabawek typu Auto w 2020 roku (tys. szt.)



Źródło: dane umowne (ilustracja pakiet Statistica).

Z kolei zaproponowany przez J. Tukeya [Tukey 1977] wykres pudełkowy ramka-wąsy ma kształt prostokąta z dołączonymi po bokach „wąsami”. Skala przedstawiająca zakres danej zmiennej jest umieszczona równoległe do boku prostokąta i „wąsów”. Obserwacje odstające to takie, których odległość przekracza

1,5 wysokości pudełka³. Obserwacje, których odległość od mediany jest wyższa niż trzykrotna długość prostokąta nazywane są obserwacjami ekstremalnymi. Na rysunku 4 przedstawiono klasyczny wykres pudełkowy dla miesięcznej sprzedaży zabawek typu Auto.

Zauważmy, że student może zwizualizować dane na kilka sposobów w zależności od oprogramowania używanego na zajęciach. Należy jednak podkreślić za P. Bieckiem, że nie jest ważne, czy na wykresie przedstawione są dobre wielkości lub zależności, lecz to, czy zostaną one prawidłowo z tego wykresu odczytane (por. P. Biecek [2014] s. 110). W omawianym przykładzie wykres pudełkowy ramka-wąsy jest najbardziej czytelny i spójny, stanowi bowiem dobry kompromis pomiędzy ilością informacji a ich zwieżłością. Ważnym celem nauczania wydaje się zatem zachęcenie słuchacza, by korzystał z wielu udostępnianych mu narzędzi służących do graficznej prezentacji danych. Istotne jest, by potrafił wybrać prawidłową formę ich ilustracji, z której wyciągnie poprawne wnioski.

PROSTE SPOSOBY IDENTYFIKACJI POSTACI FUNKCYJNEJ MODELU

Popularnym miernikiem sprawdzania jakości modelu zbudowanego Klasyczną Metodą Najmniejszych Kwadratów jest współczynnik determinacji R^2 . Określa on poziom dopasowania zmiennej objaśnianej Y do danych empirycznych. Wiara studentów w wiarygodność tego miernika bywa bezwarunkowa i często jest to podstawowy sposób weryfikacji zbudowanego modelu.

Przykładem, który może podważyć taką opinię studentów jest tzw. kwartet Anscombe'a [Anscombe 1973]. Są to cztery zestawy danych specjalnie dobranych przez angielskiego statystyka F. Anscombe'a w taki sposób, aby w każdym z nich występowały identyczne mierniki statystyczne, takie jak średnia arytmetyczna, wariancja, współczynnik korelacji, równanie regresji liniowej czy współczynnik determinacji. Ten zbiór danych (por. tabela 2), opublikowany już w 1973 r., wskazuje, jak bardzo istotna jest prezentacja graficzna w procesie analizy statystycznej. Dla wszystkich par zmiennych $((X1, Y1), (X2, Y2), (X3, Y3)$ oraz $(X4, Y4))$ oszacowany za pomocą KMNK model ma postać: $\hat{y}_i = 0,5 + 3,0x_i$, a współczynnik determinacji wynosi $R^2 = 0,67$.

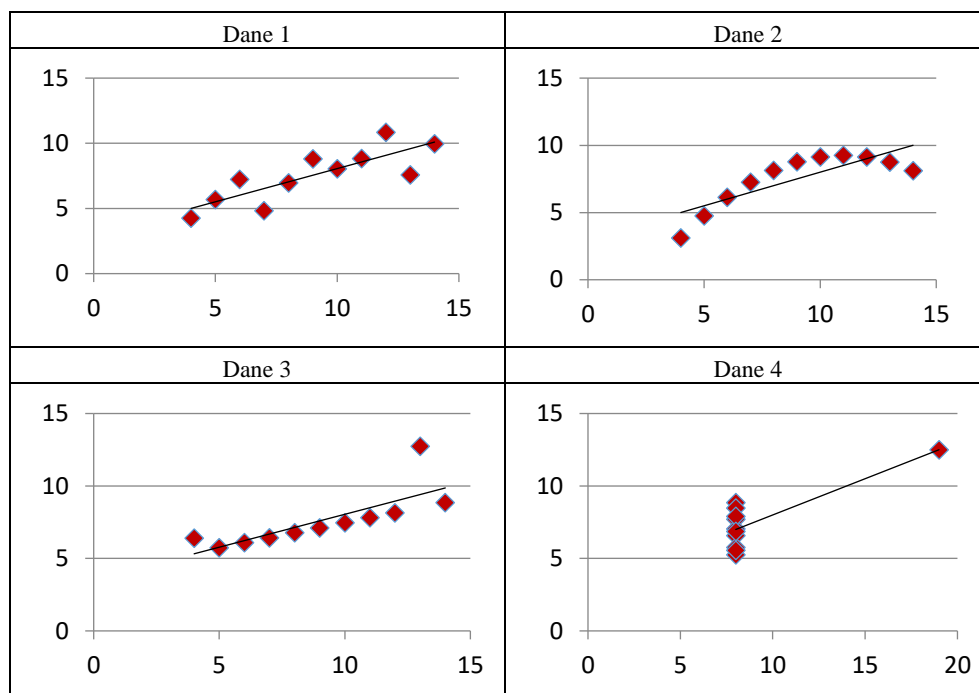
³ Ta odległość jest zalecana w literaturze, w praktyce decyzja należy do badacza.

Tabela 2. Zbiór danych zaproponowany przez F. Anscombe'a

Dane 1	$X1$	10,00	8,00	13,00	9,00	11,00	14,00	6,00	4,00	12,00	7,00	5,00
	$Y1$	8,04	6,95	7,58	8,81	8,83	9,96	7,24	4,26	10,84	4,82	5,68
Dane 2	$X2$	10,00	8,00	13,00	9,00	11,00	14,00	6,00	4,00	12,00	7,00	5,00
	$Y2$	9,14	8,14	8,74	8,77	9,26	8,10	6,13	3,10	9,13	7,26	4,74
Dane 3	$X3$	10,00	8,00	13,00	9,00	11,00	14,00	6,00	4,00	12,00	7,00	5,00
	$Y3$	7,46	6,77	12,74	7,11	7,81	8,84	6,08	6,39	8,15	6,42	5,73
Dane 4	$X4$	8,00	8,00	8,00	8,00	8,00	8,00	8,00	19,00	8,00	8,00	8,00
	$Y4$	6,58	5,76	7,71	8,84	8,47	7,04	5,25	12,50	5,56	7,91	6,86

Źródło: Anscombe F. J. [1973] s. 17–21.

Rysunek 5. Graficzna prezentacja zbioru Anscombe'a



Źródło: opracowanie własne

W tym przypadku to komunikat graficzny pozwala szybko zauważyć kardynalne błędy, jakie można popełnić, stosując regresję liniową dla danych numer 2 i 4 (rysunek 5).

Należy podkreślić, że sprawność studentów w odczytywaniu i interpretowaniu danych z szeregów i tablic wzrasta wraz z rosnącą w trakcie studiów częstotliwością ich pracy na różnych zbiorach danych. W przedstawionym


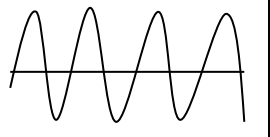
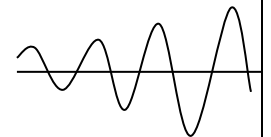
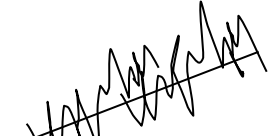
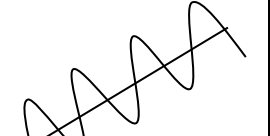

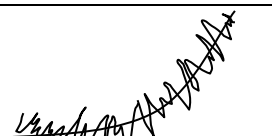
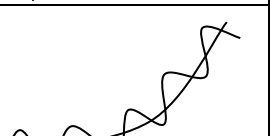
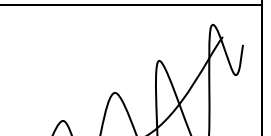
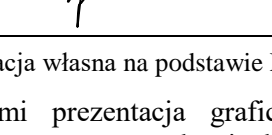
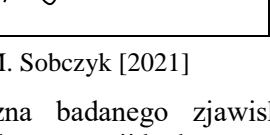
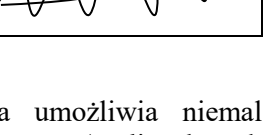
przykładzie „siła” przekazu prezentacji graficznej nie budzi wątpliwości i przypadek ten warto wprowadzać do zajęć praktycznych ze względu na eksponowane pułapki interpretacyjne.

DEKOMPOZYCJA SZEREGU CZASOWEGO

Celem analizy szeregów czasowych jest skonstruowanie modelu pewnego zjawiska w oparciu o obserwowane w czasie zmiany pewnych mierzalnych wielkości opisujących ten proces. Zgodnie z ogólnym założeniem, analizowany przebieg składa się z części systematycznej (trend, składowa stała, wahania sezonowe i cykliczne), w oparciu, o którą buduje się model, oraz części przypadkowej (szumu, wahań przypadkowych).

W analizie szeregów dąży się do wyodrębnienia i pomiaru czynników determinujących rozważane zjawisko, dokonując dekompozycji szeregu czasowego. Wykorzystując zbudowany model, można dokonywać predykcji (eksploracji) przebiegu szeregu lub jego składowych.

Tabela 3. Elementy składowej systematycznej

Sezonowość	Brak sezonowości	Sezonowość addytywna	Sezonowość multiplikatywna
Trend			
Brak trendu			
Trend liniowy			
Trend nieliniowy			

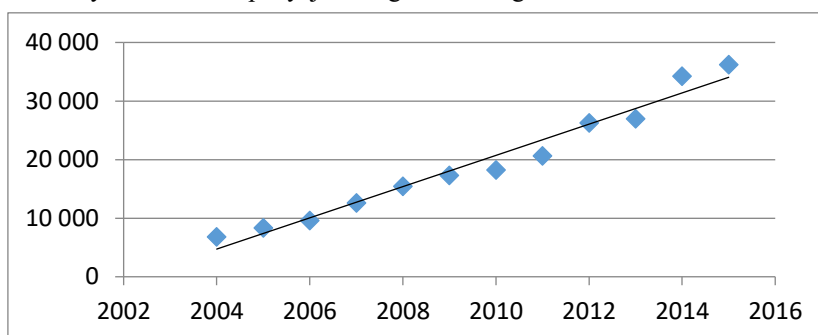
Źródło: ilustracja własna na podstawie M. Sobczyk [2021]

Czasami prezentacja graficzna badanego zjawiska umożliwia niemal natychmiastowe przeprowadzenie dekompozycji badanego szeregu. Analiza danych przedstawionych na rysunku 6 pozwala wyodrębnić stałą systematyczną (trend rosnący) oraz niewielkie wahania przypadkowe.

W praktyce jednak czasami trudno jednoznacznie stwierdzić, jakiego typu wahania występują przy eksploracji strumienia danych zawierających czas (rysunki 7 i 8).

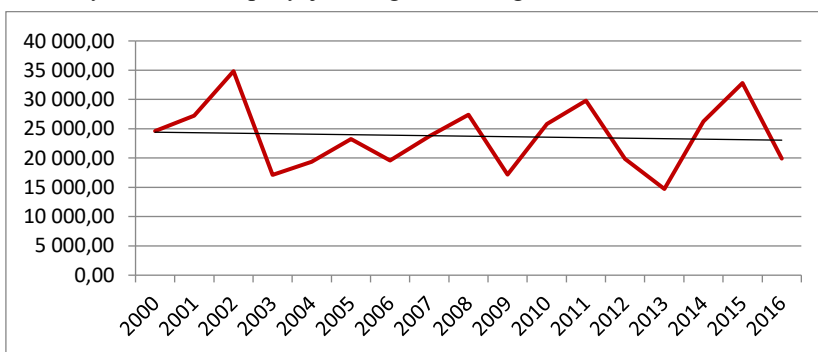
W szeregu prezentowanym na rysunku 7 nakreślona linia trendu pozwala stwierdzić występowanie nieznacznej tendencji spadkowej. W tym przypadku można również zaobserwować wyraźne wahania przypadkowe. Połączenie linią ciągłą surowych danych czasowych pozwala poprawić czytelność interpretacji wizualizacji. Ten komunikat graficzny był prawidłowo analizowany przez większość studentów w trakcie zajęć praktycznych.

Rysunek 6. Przykład 1 dekompozycji szeregu czasowego



Źródło: opracowanie własne na danych umownych

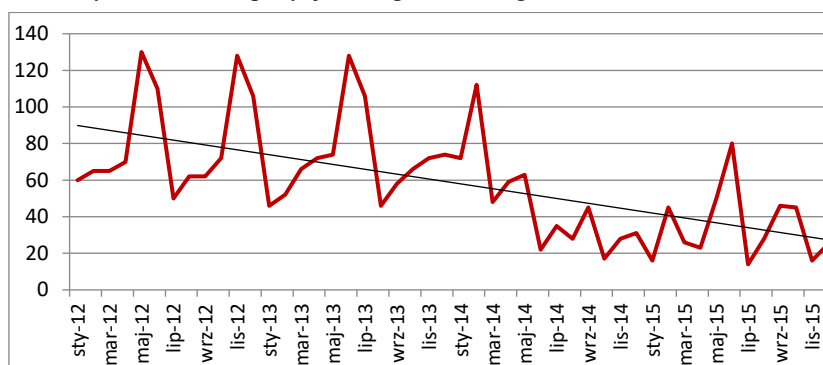
Rysunek 7. Przykład 2 dekompozycji szeregu czasowego



Źródło: opracowanie własne na danych umownych

Z kolei analizując wykres ilustrujący przykładowe dane przedstawione na rysunku 8, można zauważyć niestacjonarność wariancji tego szeregu. W okresie od stycznia 2012 do kwietnia 2014 występuje trend malejący oraz wahania sezonowe. W kolejnych miesiącach badane zjawisko nadal charakteryzuje tendencja spadkowa, ale bez wyraźnych wahań sezonowych, za to z wahaniami przypadkowymi.

Rysunek 8. Przykład 3 dekompozycji szeregu czasowego



Źródło: opracowanie własne na danych umownych

Eksploatacja danych z przykładu 3 zestawionych jedynie w formie tabelarycznej sprawiała studentom trudność i skutkowała błędnymi wnioskami. W prawidłowo przeprowadzonej dekompozycji tego szeregu należy bowiem uwzględnić dwa różne okresy badawcze, co wyraźnie sugeruje komunikat graficzny.

WNIOSKI

Zaprezentowane przykłady z praktyki dydaktycznej pozwalają na sformułowanie wniosku o konieczności stosowania metod wizualizacji danych na każdym z etapów eksploracji. Ilustracja graficzna pozwala bowiem na wizualizację tego, czego nie można odczytać z tabeli zawierającej dane.

Należy uczyć studentów, że komunikat graficzny powinien być spójny i jak najbardziej czytelny dla odbiorcy oraz przedstawiać to, co istotne. Ponadto wykresy danych mogą stanowić również szczególnie użyteczne narzędzie podczas wyboru metod badawczych. Powinno się jednak pamiętać o krytycznej weryfikacji prezentacji graficznych pod kątem tego, co dokładnie na nich widać, a co może być zakłócone tym co chciałby zobaczyć badacz.

BIBLIOGRAFIA

- Adams G. L. (1992) Why Interactive? Multimedia & Videodisc Monitor. Falls Church, Va, March 1992.
- Anscombe F. J. (1973) Graphs in Statistical Analysis. American Statistician, s. 17-21.
- Biecek P. (2014) Odkrywać! Ujawniać! Objasnić! Zbiór esejów o sztuce prezentowania danych. Fundacja Naukowa SmarterPoland, Warszawa.
- Dittmann P. (2000) Metody prognozowania sprzedaży w przedsiębiorstwie. Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław.

- Lenik K., Dziedzic K., Czerkawska A. (2007) Wykorzystanie wybranych form graficznych do przedstawiania wyników badań doświadczalnych procesu tarcia i zużycia. *Postępy Nauki i Techniki*, 1, 25-32.
- Tukey J. (1977) *Exploratory Data Analysis*. Addison-Wesley Publishing Company.
- Thalheimer W. (2006) People remember 10%, 20%...Oh Really?, www.willatworklearning.com (blog Willa'a Thalheimer'a Phd, Columbia University in the City of New York).
- Sobczyk M. (2021) *Statystyka*. PWN, Warszawa.

VISUALIZATION AS AN INSTRUMENT OF SUPPORTING THE TEACHING PROCESS IN DATA ANALYSIS

Abstract Presentation of information in a graphical form is one of the basic forms of data presentation. It is a great support during both the preliminary and further analysis. However, an incorrect graphical form can lead to misinterpretation and, in consequence, to erroneous conclusions. This paper presents some examples of graphical data visualisation that come from authors' teaching experience. The article includes cases of both correct and incorrect data presentation.

Keywords: graphical data visualisation, time series, teaching experience

JEL classification: C19, Y10

SIMULATION OF PARTICLE SWARM OPTIMIZATION FOR INVESTMENTS ON STOCK MARKET

Maciej Janowicz  <https://orcid.org/0000-0002-1584-2089>

Institute of Information Technology
Warsaw University of Life Sciences – SGGW
e-mail: maciej_janowicz@sggw.edu.pl

Andrzej Zembruski  <https://orcid.org/0000-0001-5943-3968>

Institute of Information Technology
Warsaw University of Life Sciences – SGGW
e-mail: andrzej_zembruski@sggw.edu.pl

Abstract: This work reports simulations performed using Particle Swarm Optimization (PSO) as applied to investments on the stock market. About 480 stocks belonging to *the S&P500* index have been taken into account. A naive approach has been developed in which one simulation step corresponded to one trading period. As a second ingredient of the investment strategy, the relative strength of an asset has been employed. The results are analyzed with respect to the parameters of PSO.

Keywords: particle swarm optimization; stock market

JEL classification: C53, C63

INTRODUCTION

There exists quite a widespread opinion that, with the advent of large-scale application of machine learning, especially deep learning, to the stock market analysis and investment, discretionary trading is a matter of the past. The argument goes by making a comparison with, e.g., the recent spectacular triumphs of the intelligent machine players over humans in the *Go* game [Silver et al., 2016, Silver et al., 2018]. This conception is, however, misleading: human traders, either with their Excel charts or with their primitive, homemade random forests never have to fight against the full, enormous power of the Big-Corporate deep-learning

<https://doi.org/10.22630/MIBE.2020.21.4.21>

algorithms. It is discretionary traders' skills to move across the stock markets and win relatively small profits using the play of corporate giants for their advantage.

The author's personal motivation to undertake simulations has been related to the most elementary notions and problems of the reinforcement learning associated with the *N-armed bandits* [Robbins 1952]. One can easily draw a comparison between a group of "one-armed bandit" machines in a casino and the stocks on the stock market. Modified *N-armed bandits* have been applied for the construction of stock-market portfolios already in the past decade. Now, one way to proceed further would be to include a multi-agent version of the *N-armed bandit* problem. One of the possible solutions has been to employ one of many swarm algorithms. We have decided to use particle swarm optimization chiefly because of its simplicity and obvious connection with many-particle dynamical systems known from physics. Needless to say, apart from the profit itself (or related quantities like Sharpe ratio) there is nothing really to optimize, that is, there is no well-defined target function. But the way to adjust PSO to the task of finding a sub-optimal investment policy is, theoretically, rather simple. One should "only" find a suitable way to define "best particle positions" in the evolving environment of the market, from which the "best swarm position" can already be easily selected.

This work is indebted to the papers by Blackwell, Branke, and coworkers, please see, e.g., [Blackwell et al. 2002; Blackwell et al. 2004; Blackwell et al. 2006; Blackwell et al. 2008]; their work contains thorough discussions of the problems associated with dynamic optimization problems. Still, a very elementary approach has been employed here: one will not know whether and which modifications of that approach are needed in a new kind of application without having it extensively tested.

The main body of the work is organized as follows. In Section II, we describe how the simulations have been conducted. Section III contains the main results illustrated in several figures. Section IV comprises a summary and some concluding remarks.

METHODS

Instead of using a discrete version of PSO, the interval $[0,1]$ has been divided into L subintervals $\{S_k\}$ ($k = 0,1,2,\dots,L-1$), each corresponding to a stock. The Python module *yfinance* has been used to download the stock data. In the simulations, a kind of "day trading" approach has been applied. That is, it has been understood that if a particle landed in the subinterval $\{S_k\}$ corresponding to the k -th stock, the agent associated with that particle bought shares of that stock for all available cash with opening price and then sold all the shares with the closing prices. It has not been attempted to take into account spread, possible slippages or taxes.

Standard PSO equations of the form:

$$x(t + 1) - x(t) = v(t)$$

$$v(t + 1) - v(t) = a(t)$$

have been solved; the acceleration $a(t)$ has been given as:

$$a(t) = -\gamma v(t) + c_1 r_1 (p_{pb}(t) - x(t)) + c_2 r_2 (p_{sb}(t) - x(t)),$$

where r_1 and r_2 are random numbers drawn from the uniform distribution on the interval $[0,1]$, γ , c_1 and c_2 are matrix constants (in fact, they have been simply scalars in the present simulations. The vector p_{pb} is the "best position" of an individual particle, and p_{sb} is the "best position" of the whole swarm.

As has been stressed in the introduction, while we use PSO algorithm, we do not really solve an optimization problem. Rather, we use PSO in the spirit of N -armed bandit and attempt to find an efficient policy of buying and selling stocks. The (time-dependent) target functions is determined only implicitly via the specification of a method to obtain p_{pb} and p_{sb} .

Needless to say, PSO does not provide a trading strategy on its own. The determination of p_{pb} and p_{sb} has served as a necessary second ingredient. It has been assumed that $p_{pb}(t)$ is a vector of positions of particles with the largest relative strength of the corresponding stocks visited by the particle. That relative strength has been measured by a fictitious profit that would be achieved if a "trader" associated with the particle bought the asset at the time $t - t_0$ to sell it at the time $t - 1$. The quantity p_{sb} has been scalar because the pseudo-dynamics of the swarm has been one-dimensional. Profits have been calculated as a sum of all the profits accumulated by the "best" particle while visiting various stocks.

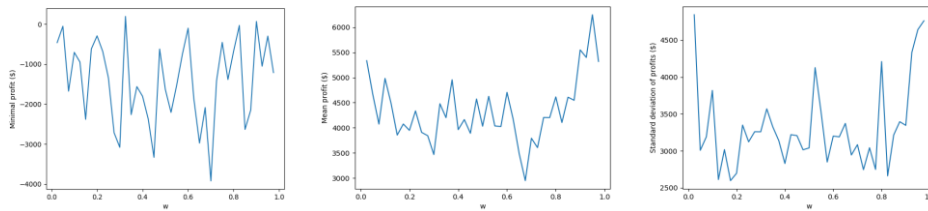
Two pairs of constants c_1, c_2 have been applied: $c_1 = c_2 = 0.5$ and $c_1 = c_2 = 1.49445$ (the latter choice is canonical). To the author's surprise, no appreciable change has been achieved when switching from the first to the second pair.

1383 trading days starting with 2015.01.01 and ending with 2020.06.30 have been taken for simulations. This has been enough to take into account both the bearish market associated with the lockdown due to the pandemia and subsequent periods of bullish developments. During that time, the value of S&P500 index has increased from 2104.50 to 3500.31, that is, by 66.325%.

RESULTS

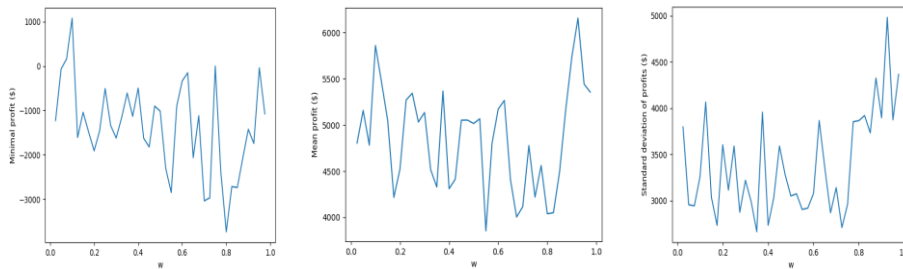
Swarms containing from 4 to 120 particles have been employed. 100 runs have been performed for every number of particles and delay parameters to obtain some reasonable statistics. This has allows us to find the mean profits and their standard deviation on averaging over the runs. In the figures below, the profits for the worst performing (that is, least profitable) runs have also been shown. In Figure 1-2 profits obtained as functions of the “inertia weight” $w = 1 - \gamma$ have been shown for the number of particles equal to 4 and 10.

Figure 1. Minimal profit (left), mean profit (center), and the standard deviation (right) of profits as functions of the “inertia weight” w for the number of particles $N = 4$, $t_0 = 2$, and $c_1 = c_2 = 0.1$



Source: own simulations

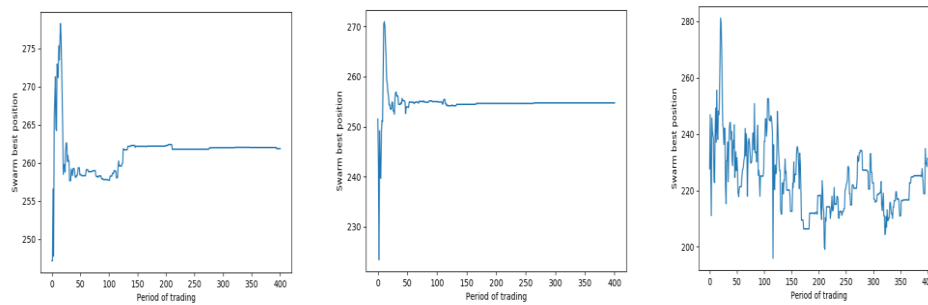
Figure 2. Minimal profit (left), mean profit (center), and the standard deviation (right) of profits as functions of the “inertia weight” w for the number of particles $N = 10$, $t_0 = 2$, and $c_1 = c_2 = 0.1$



Source: own simulations

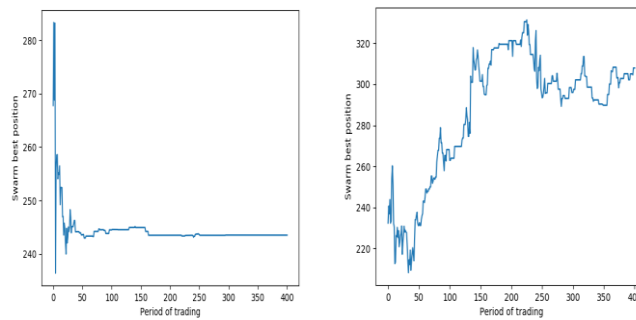
In Figures 3-4, examples of the trajectories of the best swarm positions (averaged over 100 runs) have been displayed for various number of particles and several inertial weights w .

Figure 3. Examples of trajectories of the best swarm position averagef over 100 runs for the number of particles $N = 4$, $t_0 = 2$, and $c_1 = c_2 = 0.1$. The for the “inertial weight” w has been equal to 0.1 (left), 0.5 (center) and 0.9 (right).



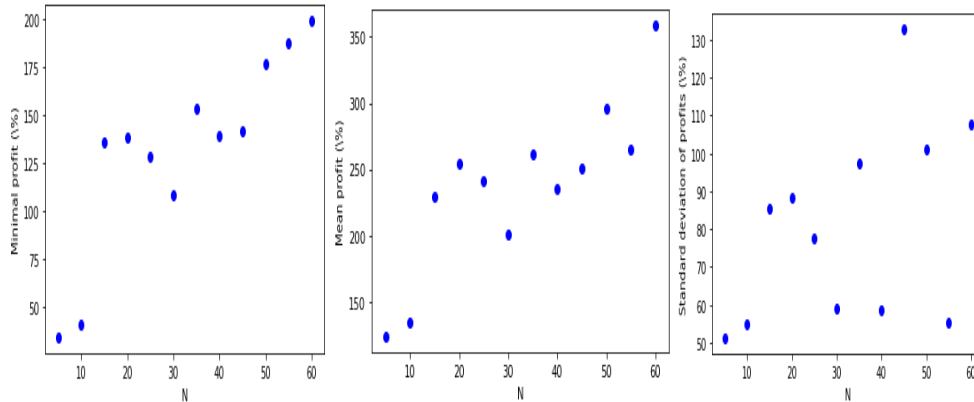
Source: own simulation

Figure 4. Examples of trajectories of the best swarm position averagef over 100 runs for the number of particles $N = 10$, $t_0 = 2$, and $c_1 = c_2 = 0.1$. The for the “inertial weight” w has been equal to 0.1 (left), 0.5 (center) and 0.9 (right).



Source: own simulation

Figure 5. The dependence of profits (in percent) on the number of particle in the swarm for $t_0 = 130$, $w = 0.5$, and $c_1 = c_2 = 2.0$: profit for the worst run (left), mean profit (center), and the standard deviations (right)



Source: own simulation

CONCLUDING REMARKS

To conclude, simulations of the behavior of particle swarms in an environment made by the stocks belonging to the *S&P500* index have been performed. It appears that while the method does not bring spectacular results, it does generate profits and it can serve as a valuable addition to the arsenal of traders, especially, as the author believes, individual ones. The necessary condition for such efficiency is operating with a sufficiently large number of particles in the swarm.

It appears that, if the number of particles is sufficiently large, the swarm can indeed quite efficiently find its best position near the stocks with high relative strength. As can be seen from Figure 5, even the least profitable runs clearly outperform the benchmark if we accept it to be the value of the *S&P* index itself.

The results obtained here must, of course, be understood as preliminary. We have not attempted to perform any optimization of the (hyper-)parameters. Therefore, no time-dependent cross-validation has been performed. It is possible that the best parameters can be strongly market-dependent.

REFERENCES

- Blackwell T. M., Bentley P. J. (2002) Dynamic Search with Charged Swarms. [in:] W. B. Langdon et al. (ed.) Proc. of the 2002 Genetic and Evolutionary Computation Conference, 19-26.

- Blackwell T. M., Branke J. (2004) Multi-swarm Optimization in Dynamic Environments. [in:] G. R. Raidl (ed.) Applications of Evolutionary Computing, Lecture Notes in Computer Science, 3005, 489-500. Springer, Berlin, Germany.
- Blackwell T. M., Branke J. (2006) Multi-swarms, Exclusion and Anti-convergence in Dynamic Environments. IEEE Transactions on Evolutionary Computation, 10(4), 459-472.
- Blackwell T. M., Branke J., Li X. (2008) Particle Swarms for Dynamic Optimization Problems. [in:] C. Blum and D. Merkle (eds.) Swarm Intelligence, 193-217 (Springer).
- Burtini G., Loeppky J., Lawrence R. (2015) Improving Online Marketing Experiments with Drifting Multi-armed Bandits. Proceedings of the 17th International Conference on Enterprise Information Systems, 2, 630-636. DOI: 10.5220/0005458706300636.
- Hoffman M.D., Brochu E., de Freitas N. (2011) Portfolio Allocation for Bayesian Optimization. [in:] The Conference on Uncertainty in Artificial Intelligence.
- Kennedy J., Eberhart R. C., Shi Y. (2001) Swarm Intelligence. Morgan Kaufmann Publishers, San Francisco.
- Robbins H. (1952) Some Aspects of the Sequential Design of Experiments. Bulletin of the American Mathematical Society, 58(5), 527-535, doi:10.1090/S0002-9904-1952-09620-8.
- Shen W., Wang J., Jiang Y.-G., Zha H. (2015) Portfolio Choices with Orthogonal Bandit Learning. [in:] Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (IJCAI 2015).
- Silver D., Huang A., Maddison C. et al. (2016) Mastering the Game of Go with Deep Neural Networks and Tree Search. Nature 529, 484-489. <https://doi.org/10.1038/nature16961>
- Silver D., Schrittwieser J., Simonyan K. et al. (2017) Mastering the Game of Go without Human Knowledge. Nature 550, 354-359. <https://doi.org/10.1038/nature24270>

REVIEWERS COOPERATING WITH THE JOURNAL IN 2020

WARSAW UNIVERSITY OF LIFE SCIENCES – SGGW

Bolesław Borkowski	Rafik Nafkha
Hanna Dudek	Luiza Ochnio
Konrad Furmańczyk	Łukasz Pietrych
Mariusz Hamulczuk	Marcin Ratajczak
Stanisław Jaworski	Wiesław Szczesny
Andrzej Karpio	Monika Zielińska-Sitkiewicz
Joanna Kisielińska	Dorota Żebrowska-Suchodolska
Joanna Landmesser-Rusek	

OTHER UNIVERSITIES

Agata Boratyńska (Warsaw School of Economics - SGH)
Joanna Dębicka (Wrocław University of Economics and Business)
Romana Głowicka-Wołoszyn (Poznań University of Life Sciences)
Joanna Górka (Nicolaus Copernicus University in Toruń)
Zbigniew Korzeb (Białystok University of Technology)
Agnieszka Kozera (Poznań University of Life Sciences)
Karol Kukuła (University of Agriculture in Cracow)
Iwona Markowicz (University of Szczecin)
Kesra Nermend (University of Szczecin)
Anna Sączewska-Piotrowska (University of Economics in Katowice)
Ewa M. Syczewska (Warsaw School of Economics - SGH)