

**Współczesne wyzwania
gospodarcze w analizach
uczestników programu
MŁODY NAUKOWIEC SGH**

REDAKCJA NAUKOWA **RADOSŁAW MALIK**

**Współczesne wyzwania
gospodarcze w analizach
uczestników programu
MŁODY NAUKOWIEC SGH**

Recenzje

Wawrzyniec Michalczyk

Ewa Oziewicz

Redakcja językowa polska

Anna Gaudy-Piątek

Redakcja językowa angielska

Mirosław Szymański

© Copyright by Szkoła Główna Handlowa w Warszawie, Warszawa 2022

Wszelkie prawa zastrzeżone. Kopiowanie, przedrukowywanie i rozpowszechnianie całości lub fragmentów niniejszej publikacji bez zgody wydawcy zabronione.

Wydanie I

ISBN 978-83-8030-550-2

Oficyna Wydawnicza SGH – Szkoła Główna Handlowa w Warszawie

02-554 Warszawa, al. Niepodległości 162

www.wydawnictwo.sgh.waw.pl

e-mail: wydawnictwo@sgh.waw.pl

Projekt i wykonanie okładki

Magdalena Limbach

Skład i łamanie

DM Quadro

Druk i oprawa

volumina.pl Daniel Krzanowski

ul. Ks. Witolda 7–9

71–063 Szczecin

tel. 91 812 09 08

e-mail: druk@volumina.pl

Zamówienie 63/V/22

Spis treści

Wprowadzenie	7
Radostław Malik	
Symulacyjne porównanie zdolności odtwarzania frakcji populacyjnych przez metody wielokrotnej imputacji dla danych nominalnych	11
Hubert Drążkowski	
Analiza wykorzystania rozwiązań opartych na chmurze obliczeniowej przez branżę finansową w 2020 roku na świecie	49
Szymon Dziubak	
Brand Loyalty to Popular Music Group Brands. The Role of Brand Attachment and Brand Community Relationships	71
Martyna Fira	
Kluczowe bariery regulacyjne i zagrożenia dla rozwoju morskiej energetyki wiatrowej (offshore) w Polsce	97
Bruno Jasic	
Weighting Schemes for Momentum Trading Strategies in the CEE Currency Universe	125
Radostław Jawor	
Jaka powinna być nowa polska partia? Maksymalizacja wyniku wyborczego partii politycznej za pomocą hybrydowego algorytmu Simulated Annealing – Multi-Start Tabu Search	147
Radostław Jawor, Szymon Dziubak	
Modelowanie procesów w kontekście projektów robotyzacji procesów biznesowych	167
Jakub Stelmaszczyk	
Behawioralne przyczyny zgonów związanych z pandemią COVID-19. Przestrzenna analiza ekonometryczna śmiertelności w Unii Europejskiej w 2020 roku	191
Mateusz Szysz	
Dokąd zmierzamy? Wartości w reklamach w XXI wieku	209
Anna Kozłowska, Agnieszka Wojtukiewicz	

Wprowadzenie

Radosław Malik

Niniejsza monografia naukowa jest zbiorem dziewięciu rozdziałów napisanych przez studentów Szkoły Głównej Handlowej w Warszawie, uczestników drugiej edycji programu „Młody Naukowiec SGH”. Celem tego programu jest rozwój naukowy utalentowanych młodych ludzi, którzy już na etapie studiów licencjackich i magisterskich pokazali swoje zainteresowania naukowe i chęć rozwoju kompetencji ważnych w pracy badawczej. Teksty opublikowane w tej monografii wskazują na szerokie horyzonty autorów oraz ciekawość w podejmowaniu nowatorskich tematów badawczych, co dowodzi sukcesu programu „Młody Naukowiec SGH”. Ambitna tematyka tekstów oraz zastosowanie innowacyjnych narzędzi badawczych przez autorów, dla których jest to często debiut publikacyjny, stanowi zapowiedź ich przyszłych sukcesów naukowych.

W rozdziale pierwszym monografii Hubert Drażkowski dokonał analizy porównawczej metod wielokrotnej imputacji dla wielowymiarowych danych nominalnych, wskazując na dominację jednego z badanych algorytmów. Wyniki pracy stanowią istotną wskazówkę dotyczącą właściwego doboru metod do odtwarzania wielowymiarowych zależności populacyjnych i mogą być wykorzystane w doskonaleniu przyszłych analiz ekonomicznych w wielu dziedzinach.

W drugim rozdziale Szymon Dziubak przeprowadził analizę zastosowania usług wykorzystujących chmury obliczeniowe przez branżę finansową w 2020 r., pokazując zróżnicowane zastosowanie tych rozwiązań zależnie od lokalizacji przedsiębiorstwa, jego profilu i dojrzałości biznesowej. Wyniki analizy wskazują również na nieoczywiste zależności między strategicznymi trendami technologicznymi a wykorzystaniem chmur obliczeniowych.

Trzeci rozdział monografii autorstwa Martyny Firy dostarcza interesujących rezultatów na temat lojalności wobec marki na przykładzie popularnych grup muzycznych. Unikalny zestaw danych zebranych w społecznościach fanów posłużył autorce do przetestowania hipotez mających uzasadnienie w literaturze dotyczącej teorii marki i wskazania na zależności między lojalnością wobec marki a identyfikacją ze społecznością, zaangażowaniem w społeczność oraz postrzeganiem podobieństwem między społecznością a marką.

W czwartym rozdziale monografii Bruno Jasic podjął istotne zagadnienie rozwoju morskiej energetyki wiatrowej w Polsce, wskazując na bariery regulacyjne oraz na inne zagrożenia dla rozwoju tej istotnej dla gospodarki i środowiska działalności. Wyniki analizy barier regulacyjnych zostały powiązane z długookresową strategią transformacji energetycznej, dostarczając ważnych wskazówek dla polityki gospodarczej państwa.

W piątym rozdziale monografii, napisanym przez Radosława Jawora, przedmiotem analizy jest wpływ różnych metod wazenia aktywów i sygnałów kupna/sprzedaży na wyniki strategii momentum na portfelach zbudowanych na czterech walutach krajów Europy Środkowej i Wschodniej. Wyniki analizy dostarczają ciekawych wniosków, wskazując najbardziej efektywną strategię, oraz pokazując, że założone metody wazenia aktywów nie powodują istotnej poprawy wyników w analizowanych koncepcjach.

W szóstym rozdziale monografii Radosław Jawor i Szymon Dziubak, wykorzystując zaawansowane metody analityczne, dokonali nowatorskiej analizy poglądów politycznych wyborców oraz partii politycznych w Polsce, w kontekście wyboru preferowanego ugrupowania w wyborach parlamentarnych. Wyniki przeprowadzonej symulacji wskazują, że zupełnie nowa partia polityczna o odpowiednio sformułowanych poglądach mogłaby zdobyć ponad 19% głosów i zostać trzecią pod względem liczby mandatów partią w Sejmie w kolejnych wyborach parlamentarnych.

Síódmy rozdział monografii autorstwa Jakuba Stelmaszczyka dotyczy aktualnego i ważnego biznesowo zagadnienia modelowania procesów w projektach wykorzystujących automatyzację procesów biznesowych. Wnioski z pracy sugerują konieczność podniesienia jakości i standaryzacji diagramów jako narzędzi analitycznych wykorzystywanych w projektach, oraz stanowią istotny postulat, który może być zastosowany w praktyce biznesowej i przyczynić się do podniesienia efektywności w realizacji projektów.

Ósmy tekst, będący częścią tej monografii, napisany przez Mateusza Szysza, dostarcza interesujących wniosków na temat behawioralnych przyczyn zgonów w pandemii COVID-19 w Unii Europejskiej w 2020 r., co jest ważnym i aktualnym tematem badawczym. Autor wskazał m.in. na zależność między wysokim zaufaniem do rządu i ekspertów medycznych a mniejszym przyrostem zgonów na COVID-19. Analiza potwierdza również, że regiony o niższych nakładach na służbę zdrowia i większej gęstości zaludnienia były bardziej podatne na negatywne skutki pandemii.

W ostatnim tekście zamieszczonym w monografii Agnieszka Wojtukiewicz i dr Anna Kozłowska podejmują tematykę pandemii w zupełnie innym ujęciu, analizując jej wpływ na przekazy reklamowe w obszarze prezentowanych wartości.

Analiza spotów reklamowych przeprowadzona przez autorki wskazuje na zmianę w przedstawianych wartościach, która jednak jest nierównomierna i w dużym stopniu zależy od rodzajów komunikatów reklamowych oraz kategorii reklamowanych produktów.

Różnorodność tematyki badawczej w tekstach zgromadzonych w niniejszej monografii, nowatorstwo podejmowanych analiz i wykorzystanie często zaawansowanych narzędzi badawczych wskazują na wysoki potencjał naukowy i duże ambicje autorów – uczestników programu „Młody Naukowiec SGH”, oraz stanowią uzasadnienie dla rozwoju tego programu w Szkole Głównej Handlowej w Warszawie.

Jako redaktor naukowy tej monografii chciałbym podziękować osobom, które wspierały uczestników drugiej edycji programu „Młody Naukowiec SGH”. Pani Prorektor ds. nauki dr hab. Agnieszce Chłoń-Domińczak za zainicjowanie tego programu i inspirację do podejmowania ambitnych wyzwań przez jego uczestników. Opiekunom merytorycznym uczestników programu za wsparcie w rozwoju naukowym i mentoring, jakiego udzielili studentom. Autorom warsztatów rozwijających kompetencje naukowe uczestników, dr. hab. Jakubowi Growcowi – Kierownikowi Katedry Ekonomii Ilościowej, prof. dr. hab. Wojciechowi Pachowi – Dziekanowi Szkoły Doktorskiej SGH, pani Lidii Tomaszewskiej – Kierownikowi Działu Obsługi Projektów SGH, dr. hab. Annie Visvizi – Kierownikowi Zakładu Międzynarodowej Polityki Ekonomicznej i dr. hab. Andrzejowi Zybale – Kierownikowi Katedry Polityki Publicznej, za dzielenie się wiedzą, pokazywanie nieoczywistych możliwości i inspiracje, jakie przekazali uczestnikom. Program nie przebiegałby tak sprawnie, gdyby nie troska i zaangażowanie pani Anny Ulatowskiej z Działu Nauki SGH, czuwającej nad wszystkimi sprawami administracyjnymi.

Opiekunowie merytoryczni II edycji programu „Młody Naukowiec SGH”:

- dr hab. Beata Czarnacka-Chrobot, prof. SGH,
- dr hab. Bogumił Kamiński, prof. SGH,
- dr Agnieszka Kłos,
- dr hab. Barbara Kowalczyk, prof. SGH,
- dr Anna Kozłowska,
- dr Marek Kwas,
- dr hab. Dorota Niedziółka,
- dr Przemysław Polak,
- dr hab. Jerzy Surma, prof. SGH,
- dr hab. Andrzej Torój, prof. SGH,
- dr Wojciech Trzebiński.

Symulacyjne porównanie zdolności odtwarzania frakcji populacyjnych przez metody wielokrotnej imputacji dla danych nominalnych

Hubert Drążkowski

Streszczenie

Niniejszy rozdział porównuje kilka obecnych w literaturze metod wielokrotnej imputacji dla wielowymiarowych danych nominalnych. Porównanie odbywa się na drodze symulacji, wielokrotnego próbkowania z rzeczywistej bazy danych i każdorazowo amputowania wartości według mechanizmu MCAR albo MAR. Następnie aplikowane są algorytmy wielokrotnej imputacji. Zbadane zostały dwie metody uczenia maszynowego, dwie głębokiego uczenia i siedem klasyfikowanych do klasycznej teorii statystyki. Celem badania była ocena i porównanie zdolności odtwarzania wielowymiarowych zależności populacyjnych. Jakość imputacji mierzona była poprzez pomiar trafności prognozy idiosynkratycznej brakującej wartości, obciążenie estymatorów i ich zmienność w oszacowaniu populacyjnych wielowymiarowych frakcji, oraz pokrycia przedziałów ufności uzyskane za pomocą reguł Rubina. Z analizy wynika, że algorytm MICE CART dominował nad pozostałymi metodami. Algorytmy MIDA, GAIN, nieparametryczny hot deck i MICE LDA nie uzyskiwały nominalnego pokrycia przedziałów ufności.

Słowa kluczowe: wielokrotna imputacja, symulacja, uczenie maszynowe, braki danych, dane nominalne, uczenie głębokie

1. Wstęp

Powszechnym problemem, z którym zmagają się badacze próbujący dokonać analiz statystycznych na pewnych zbiorach danych, są występujące w nich braki. Mogą one powstawać na skutek wielu różnych mechanizmów. Czasem brak obecności wartości jednej ze zmiennych dla danej obserwacji wynika z błędów operacyjnych

wprowadzania danych, umyślnego opuszczania pewnych pytań w kontekście badań ankietowych, łączenia różnych mniejszych baz danych w jeden duży zbiór lub czułości przyrządów pomiaru. To tylko kilka z możliwych, zgoła odmiennych wyjaśnień pochodzenia zjawiska, a zdaje się, że w tym względzie ogranicza nas jedynie wyobraźnia. Występowanie braków danych można zaobserwować w wielu dziedzinach, począwszy od badań społecznych [Stavseth, Clausen, Røislien, 2019], przez badania medyczne [Pedersen i in., 2017], a skończywszy na danych gospodarczych [Grochowina, 2015]. Często zjawisko to dotyczy danych ankietowych, na których bazują analizy z dziedziny mikroekonometrii, socjologii ekonomicznej, badań marketingowych itp. [Anand, Mamidi, 2020].

Obecność braków danych może powodować zniekształcenia rozkładów zmienionych obserwowanych w próbie, prowadzić do uzyskania obciążonych niektórych estymatorów bazujących na niepełnych danych oraz spadku mocy narzędzi statystycznych. Jak dowodzi mnogość publikacji na temat radzenia sobie z niepełnymi rekordami, analitycy nie są na szczęście skazani na bezradność i często ze struktury brakujących danych, w połączeniu z charakterystyką obserwowalnych zmiennych, można próbować luki uzupełniać, poprawiając tym samym jakość wnioskowania statystycznego [Lin, Tsai, 2020; van Buuren, 2018].

W literaturze wyróżnia się trzy podstawowe mechanizmy powstawania braków danych, wyszczególnione przez Rubina w artykule z 1976 r. [Rubin, 1976]. Są to:

- mechanizm całkowicie losowy MCAR (ang. *Missing Completely at Random*), zakładający, że prawdopodobieństwo występowania braku informacji nie zależy od wartości obserwowalnych ani nieobserwowalnych zmiennych;
- mechanizm losowy MAR (ang. *Missing at Random*), zakładający, że prawdopodobieństwo występowania braku danych dla danej zmiennej nie zależy od wartości nieobserwowalnych, ale może zależeć od wartości obserwowalnych zmiennych;
- mechanizm nielosowy MNAR (ang. *Missing Not at Random*), zakładający, że prawdopodobieństwo braku informacji o danej zmiennej zależy od nieobserwowalnych wartości.

Po rozpoznaniu i przyjęciu założenia o mechanizmie powstawania braków danych, następnym krokiem w analizie jest próba pozbycia się lub uwzględnienia i imputacji nieobserwowalnej części zbioru. Poprzez imputację rozumie się zespół metod, które uzupełniałyby obecne braki w próbie, tak aby stosowanie standardowych metod dla danych pełnych było możliwe.

Pierwszym historycznie pomysłem, który pojawił się w praktyce, było usuwanie tych obserwacji, które posiadają co najmniej jedną niezaobserwowaną wartość którejś ze zmiennych (ang. *listwise deletion*). Zabieg ten mógłby zostać wykonany w celu wykorzystania metod stworzonych dla danych pełnych. Wadą takiego podejścia jest

utrata sporej części dostępnych informacji, a ponadto nierzadko zmuszałoby to do znaczącego obniżenia wielkości próbki. Takie podejście jest dopuszczalne w sytuacji występowania relatywnie małej ilości luk i mechanizmu MCAR. W przypadku innych procesów generowania braków, zabieg ten mógłby prowadzić do poważnych zniekształceń wnioskowania [van Buuren, 2018].

Drugim z podejść wspomnianym już wcześniej jest próba uzupełniania braków danych pewnymi prognozowanymi wartościami na podstawie obserwowalnych informacji. W literaturze przedmiotu znajdują się metody jednokrotnej i wielokrotnej imputacji. Te pierwsze polegają na jednokrotnym imputowaniu brakującej wartości metodami deterministycznymi i wyciąganiu wniosków, estymacji parametrów na ich podstawie. Te drugie z wymienionych podejść bazują na wielokrotnym użyciu metod stochastycznych imputacji i wyciąganiu wniosków na podstawie kilku powstałych duplikatów zbioru danych z różnymi zaimputowanymi wartościami po uprzednim ich zagregowaniu [Rubin, 1987].

W przypadku jednokrotnej imputacji problem braków danych traktowany był głównie jako problem prognozy brakującej wartości. Jak zauważono wraz z rozwojem dziedziny, ważnym zagadnieniem w kontekście uzupełniania braków danych jest próba pomiaru niepewności związanej z uzupełnianiem luk, kwantyfikowania zmienności potencjalnych wyników uzyskanych za pomocą konkretnych metod. To właśnie wielokrotna imputacja daje narzędzie, które pozwalałoby mierzyć, jak zmienne mogą być estymatory ze względu na występującą nieobserwowalność części informacji. W matematycznym sensie metoda wprowadza poprawkę na wariancję estymatora.

Niniejsza praca bierze sobie za cel kompleksowe porównanie kilku metod wielokrotnej imputacji w przypadku, kiedy obecne są jedynie dane o wartościach nominalnych. W wielu artykułach autorzy zwracają uwagę na mnogość dostępnych metod, a przy tym na pilną potrzebę ich porównania [Lin, Tsai, 2020; Jadhav, Pramod, Ramanathan, 2019]. Zaznaczają relatywnie małą liczbę tego typu badań algorytmów dla danych jakościowych. Praca ma pomóc w przyszłości badaczom w bardziej świadomym i łatwiejszym wyborze narzędzi radzenia sobie z brakami danych. Poddana weryfikacji zostanie hipoteza, iż można wskazać najlepszą w niżej opisanym sensie, dominującą nad pozostałymi metodę wielokrotnej imputacji dla danych nominalnych.

W artykule obrano za cel zbadanie jakości imputacji danych nominalnych kilkunastu metod wielokrotnej imputacji, dwóch metod uczenia maszynowego, dwóch głębokiego uczenia i siedmiu klasyfikowanych do klasycznej teorii statystyki. Porównanie odbywa się na drodze symulacji, tj. wielokrotnego próbkowania z pełnej bazy danych i każdorazowo amputowania wartości według MCAR albo

MAR w przypadku różnych proporcji nieobserwowalnych wielkości (10%, 30%, 50% albo 33%*) [Akanke, Li, Reiter, 2017]. Następnie aplikowane są algorytmy imputowania danych.

Jakość imputacji mierzona jest na kilka sposobów. Po pierwsze, poprzez pomiar trafności prognozy idiosynkratycznej brakującej wartości, tak jak odbywa się to w przypadku badań imputacji jednokrotnej. Po drugie, za pomocą dostępnych narzędzi wielokrotnej imputacji i reguł Rubina. Uwzględniony jest pomiar błędu oszacowania frakcji populacyjnej dla jednej zmiennej oraz badania zdolności odtwarzania wielowymiarowych zależności poprzez oszacowanie frakcji dla kombinacji wartości dwóch zmiennych. Weryfikowane są pokrycia przedziałów ufności oraz ich szerokość. Badane są rozkłady miar jakości. Końcowy ranking metod konstruowany jest w oparciu o narzędzia nieparametrycznej statystyki. Zmienność metod mierzona jest w dwójnasób, zarówno dzięki narzędziom wielokrotnej imputacji, jak i zaprojektowanemu badaniu symulacyjnemu.

W pierwszym rozdziale zaprezentowano podstawowe koncepty oraz analizę literatury przedmiotu, ze szczególnym uwzględnieniem badań dotyczących wielowymiarowych danych nominalnych. W kolejnym przedstawiono użyte w badaniu dane, szczegółowy schemat symulacji wraz z wykorzystywanymi do ewaluacji metrykami oraz zwięzłe zestawienie zastosowanych algorytmów. W rozdziale trzecim zaprezentowano wyniki i dokonano kompleksowego porównania narzędzi służących do imputacji. Na końcu znalazły się wnioski z analiz oraz wymienione ograniczenia badania i możliwe kierunki rozbudowy poszukiwań.

2. Podstawy teoretyczne i przegląd literatury

2.1. Trzy mechanizmy powstawania braków

Ważnym w kontekście modelowania braków danych jest zrozumienie różnych typów mechanizmów powstawania braków danych, które zostały wymienione na wstępie. Są to MCAR, MAR i MNAR.

Mechanizm całkowicie losowy (MCAR) zakłada niezależność statystyczną pomiędzy brakami odpowiedzi a obserwacjami. Powstające braki nie są w żaden sposób związane ani z obserwowalną, ani z nieobserwowalną częścią danych. Luki pojawiają się więc całkowicie losowo, ale mogą się one koncentrować na wybranym zestawie zmiennych. Przykładem przyczyny powstawania takiego mechanizmu może być używanie przyrządu pomiarowego o wyczerpujących się bateriach, zagubienie dokumentów itp.

Mechanizm losowy (MAR) zakłada występowanie zależności pomiędzy obecnością braków danych a wartościami obserwowalnymi w pozostałych zmiennych dla danej obserwacji. Zachodzenie takiego mechanizmu byłoby korzystne dla badacza, ponieważ mógłby modelować nieobserwowalne wielkości za pomocą tych dostępnych. Badając, czy powinniśmy przyjąć założenie MCAR czy MAR, można porównywać dwie grupy obserwacji podzielone przy warunku występowania braków albo pełnej informacji. W przypadku występowania znaczących różnic w rozkładach pomiędzy tymi dwiema, wykrytych testami t-Studenta lub którymiś z nieparametrycznych testów, można przyjąć, że mamy do czynienia z MAR. Źródłem powstawania takiego mechanizmu mogą być np. różnice płci widoczne w odmiennym postrzeganiu drażliwych tematów w badaniach ankietowych itp.

Mechanizm nielosowy (MNAR) najbardziej zniekształca wnioskowanie statystyczne. Zakłada bowiem, iż braki danych występują warunkowo w zależności od przyjmowanej wartości danej zmiennej lub zależą od nieobserwowanych zmiennych. Niestety obecność tej formy procesu powstawania braków jest najbardziej niekorzystna, dlatego że jest praktycznie nierozróżnialna od mechanizmu MAR, a za to znacząco zaburza wygląd obserwowalnego rozkładu zmiennej nim dotkniętej. Rozwiązaniem powstałego problemu byłby dostęp do informacji o części nieobserwowalnej. Przyczyną obecności takiej struktury braków może być wpływ intensywności modelowanego zjawiska, jak np. choroby psychicznej lub nierejestrowania pewnych płatności o niskiej wartości itp.

2.2. Wielokrotna imputacja

Wśród metod wykorzystywanych do imputacji braków należy być świadomym podstawowego rozróżnienia. Algorytmy dzieli się na deterministyczne i stochastyczne. Metody deterministyczne, otrzymawszy ten sam zestaw danych na wejściu, zawsze zaimputują tę samą wartość – wykorzystywane są w przypadku imputacji jednokrotnej. Metody stochastyczne to te, które wprowadzają pewną losowość podczas zabiegu imputacji, potencjalnie pomimo tego samego zestawu danych, na którym operujemy, za każdym razem imputując trochę inną wartość. Pomysł, aby wykorzystać te ostatnie, umożliwił rozwój metod wielokrotnej imputacji.

Wielokrotna imputacja przebiega w kilku etapach. Po wybraniu metody imputacji danych należy zdecydować się na liczbę duplikatów tego samego zbioru danych, na których zostanie kilkakrotnie zastosowana ta sama stochastyczna metoda uzupełniania braków. Kolejnym krokiem jest zastosowanie wybranej metody, a ostatecznie zebranie wyników, przez co powstaje nowy punktowy estymator wraz z przedziałem ufności uzyskanym dzięki obecności kilku wersji imputacji.

Do tej pory mowa była wyłącznie o możliwości konstruowania przedziałów ufności w przypadku wielokrotnej imputacji na podstawie tzw. metod Rubin [Rubin, 1987]. W tym rozdziale zostaną wyjaśnione detale procedury. Poniższe rozważania zostały wyprowadzone za pomocą narzędzi analizy bayesowskiej. Detale wyprowadzenia można znaleźć w załączniku.

Oznaczmy jako Y ramkę danych, dostępną próbę, która składa się z obserwowalnych wartości Y_{obs} i nieobserwowalnych wartości Y_{mis} . Niech Q oznacza naszą zmienną oszacowania (np. wektor), która może być funkcją przyjmującą jako argument całą, w pełni obserwowalną populację lub parametrem modelu dopasowanym do takiej populacji (np. parametr strukturalny w regresji, frakcja, parametr położenia). Teraz wprowadźmy estymator bazujący na dostępnej próbce oznaczony przez $\hat{Q}(Y)$, który charakteryzowany jest także poprzez wariancję z próby $\hat{U}(Y)$ (ang. *within imputation variance*). Załóżmy, że $Y_{mis}^{(1)}, Y_{mis}^{(2)}, \dots, Y_{mis}^{(M)}$ będą dotyczyły M wersji zaimputowanych luk w wejściowej ramce danych dla wartości Y_{mis} . Wyżej wymieniony estymator można by też zapisać dla m -tej imputacji poprzez $\hat{Q}^{(m)} = \hat{Q}(Y_{obs}, Y_{mis}^{(m)})$ oraz analogicznie dla $\hat{U}^{(m)}$. Dalej wprowadźmy estymatory agregujące (ang. *pooling*) informacje z kolejnych wersji imputacji. Kolejno jest to średnia estymatora naszego obiektu celu:

$$\bar{Q} = \sum_{m=1}^M \frac{\hat{Q}^{(m)}}{M},$$

prosta średnia z wariancji po wszystkich wersjach imputacji:

$$\bar{U} = \sum_{m=1}^M \frac{\hat{U}^{(m)}}{M}$$

oraz wariancja estymatora pomiędzy imputacjami:

$$B_m = \sum_{m=1}^M \frac{(\hat{Q}^{(m)} - \bar{Q})^2}{M - 1}.$$

Całkowita wariancja (ang. *total variance*) może zostać ujęta poprzez wzór:

$$T_m = \bar{U}_M + \left(1 + \frac{1}{M}\right) B_M.$$

Ostatnie wyrażenie odpowiada za podobną do ANOVA dekompozycję wariancji. Obiekt oznaczony powyżej jako B_m estymuje nadwyżkę zmienności związaną z brakującą informacją (ang. *between imputation variance*). Czynnikiem wchodzący

multiplikatywnie do sumy przy B_m , $\left(1 + \frac{1}{M}\right)$ odpowiada za poprawkę na obciążenie

estymatora dla skończonej liczby wersji imputowanych zbiorów [Rubin, 1987].

Chcąc stworzyć przedział ufności, autor pomysłu odniósł się do założenia o asymptotycznej normalności różnicy zagregowanego estymatora i prawdziwej wartości parametru:

$$\lim_{M \rightarrow \infty} (\bar{Q}_M - Q) \sim N(0, T_\infty).$$

Oczywiście w praktyce całkowita wariancja nie jest znana *a priori*. Ten obiekt trzeba estymować. Z teorii wiadomo [Murray, 2018], że w takim przypadku, gdy do wyrażenia wstawiamy empiryczną wariancję, można skorzystać z rozkładu t-Studenta:

$$(\bar{Q}_M - Q) \sim t_{v,M}(0, T_M).$$

Odpowiednia liczba stopni swobody jest opisana w załączniku. Odpowiedzią na wyjściowe pytanie, jak wygląda przedział ufności, za pomocą którego możemy badać m.in. pokrycie, jest wyrażenie, które wyznacza przedział ufności na poziomie $100(1 - \alpha)\%$

$$\bar{Q}_M \mp t_{v, 1 - \frac{\alpha}{2}} \sqrt{T_M}.$$

2.3. Problem modelowania danych wielowymiarowych

Podczas implementowania wielokrotnej imputacji można przyjąć dwie strategie modelowania. Pierwsza z nich to JM (ang. *joint modeling*), a druga to FCS (ang. *fully conditional specification*) [van Buuren, Brand, Groothuis-Oudshoorn, Rubin, 2006]. W przypadku strategii JM należy określić wspólny rozkład wielowymiarowy dla wszystkich zmiennych w ramce danych (np. wielowymiarowy rozkład normalny). Z jednej strony taka specyfikacja jest dobrze ugruntowana w podstawach teoretycznych [van Buuren, 2018], ale problematyczne staje się precyzyjne stworzenie takiego rozkładu, który oddawałby wszystkie zależności pomiędzy zmiennymi, zwłaszcza w przypadku wysoce wielowymiarowych danych. Alternatywą jest modelowanie poszczególnych warunkowych jednowymiarowych rozkładów dla każdej zmiennej z osobna. Najpopularniejszym algorytmem aplikującym FCS jest MICE (ang. *multiple imputation by chained equations*), który używa uogólnionych modeli liniowych dla każdego brzegowego rozkładu rekursywnie [van Buuren, Groothuis-Oudshoorn, 2011]. Łatwo zauważyć, że to podejście cechuje prostota i niesamowita elastyczność, niestety brak jest w dziedzinie imputacji rozwiniętych solidnych teoretycznych podstaw matematycznych dla tej metody. Może się tak

zdarzyć, że modelowanie każdego z rozkładów z osobna będzie dawało źle określony wspólny wielowymiarowy model [Arnold, Press, 1989]. Pomimo teoretycznych zarzutów badania symulacyjne pokazywały, że w praktyce algorytm MICE bardzo dobrze modeluje braki [Raghunathan, Lepkowski, Hoewyk, Solenberger, 2000; van Buuren i in., 2006; van Buuren, 2007].

2.4. Przegląd literatury

Wyczerpującą wysokopoziomową analizę artykułów opublikowanych w latach 2006–2017 można znaleźć w przeglądzie literatury autorstwa Lin i Tsai *Missing value imputation: a review and analysis of the literature* [2020]. Autorzy zwracają uwagę na częstą dychotomię wśród porównań wyłącznie w ramach metod wywodzących się z klasycznej statystyki po jednej stronie, zaś po drugiej tylko na porównywanie metod uczenia maszynowego między sobą. Kolejną obserwacją, którą czynią, jest częste wykorzystywanie tylko jednej metryki do oceny jakości metod imputacji. Ponadto gros publikacji bada jedynie sytuację, w której zakłada się obecność mechanizmu MCAR. Niniejszy tekst ma się przyczynić do przełamania tego trendu poprzez uwzględnienie zarówno mechanizmów MCAR, jak i MAR oraz porównanie za pomocą wielu metryk.

Temat złotego standardu analizy narzędzi imputacji danych był już często poruszany [Vink, 2016; Salfrán, Jordan, Spiess, 2016; Wang, Akande, Poulos, Li, 2022]. Obecny konsensus przyjmuje, że badanie powinno rozpoczynać się od pełnego zbioru danych (rzeczywistego lub syntetycznego/wygenerowanego), a następnie amputacji pewnej części zbioru zgodnie z przyjętym mechanizmem powstawania braków [Schouten, Lugtig, Vink, 2018]. Kolejnym krokiem powinno być wykorzystanie wybranego zbioru algorytmów, które uzupełniałyby luki. Oprócz zdolności do prognozowania brakujących wartości (ang. *accuracy*), powinno się mierzyć niepewność związaną z użyciem metody w konkretnych warunkach. Na przykład można to uczynić, wykorzystując narzędzia wielokrotnej imputacji i wnioskowania na temat pewnych parametrów, tym samym uzyskując przedziały ufności. O teorii związanej z wielokrotną imputacją powstało kilka syntetycznych opisów [Rubin, 1996; Harel, Zhou, 2007; Murray, 2018]. Dodatkowy poziom zmienności wyników może być oceniony na drodze wielokrotnego pobierania próbek z danego zbioru i próby odtworzenia parametrów populacyjnych [Audigier, Husson, Josse, 2017; Akande i in., 2017; Wang i in., 2022] relatywnie do zdolności próby pełnej, dostępnej przed amputacją. Optymalne byłoby badanie zdolności odtwarzania własności wielowymiarowych rozkładów oryginalnych danych [Little, Rubin, 1987]. Ponadto sugeruje się analizę własności algorytmów, ich zbieżności i czasu trwa-

nia wykonywanych procedur [Lin, Tsai, 2020]. Wśród pozostałych pomysłów, jak porównywać pomiędzy sobą narzędzia imputacji, można znaleźć badanie wpływu braków danych na jakość klasyfikacji lub grupowania danych [Farhangfar, Kurgan, Dy, 2008; Luengo, García, Herrera, 2012; Audigier, Niang, 2020].

Syntetycznego podsumowania podstaw teorii doczekała się także polska literatura. O matematycznych podstawach teorii można poczytać w pozycji Wesołowskiego i Tarczyńskiego [2016], zaś w publikacji Pokropka [2018] znajduje się wysokopoziomowy, historyczny przegląd dziedziny.

Stosunkowo często w literaturze pojawiają się podsumowania i analiza kilku metod dla danych ciągłych [Miształ, 2019; Jadhav i in., 2019; Hong, Lynn, 2020]. Coraz częściej można też spotkać artykuły traktujące o problematyce uzupełniania danych mieszanych [Luengo i in., 2012; Salfrán, 2018; Bertsimas, Pawłowski, Zhuo, 2018; Wang i in., 2022]. Stosunkowo rzadko jednak weryfikowana była jakość imputacji dla danych jakościowych nominalnych.

2.5. Przegląd literatury dla danych nominalnych

Badanie *A comparison of incomplete data methods for categorical data* [2016] stworzone przez van der Palma, van der Arka i Vermunta porównuje ze sobą cztery metody imputacji. Są to: wielokrotna imputacja z użyciem zmiennych ukrytych (ang. *multiple imputation using latent class analysis*), procedura największej wiarygodności dla niepełnych danych (ang. *maximum likelihood for incomplete data*), wielokrotna imputacja z wykorzystaniem log-liniowych modeli (ang. *multiple imputation using a log-linear model*) oraz PMM MICE (ang. *multivariate imputation using chained equations PMM*). Autorzy tworzą syntetyczne dane zgodnie z modelem logitowym z interakcjami (7 zmiennych) w dwóch wielkościach próbek (200 i 1000 obserwacji), i na nich wprowadzają mechanizm MAR w 10%, 20% i 40%. Wykonują 1000 symulacji. Do porównania wykorzystują oszacowania parametrów w modelu logitowym, z którego generowali dane. Badają obciążenie, odchylenie standardowe oszacowania parametru i obciążenie odchylenia standardowego parametru [van der Palm, van der Ark, Vermunt, 2016]

W artykule *An Empirical Comparison of Multiple Imputation Methods for Categorical Data* [2017] autorzy Akande, Li oraz Reiter skupiają się na ewaluacji trzech metod wielokrotnej imputacji. Wykorzystują obszerny zbiór danych (17 zmiennych) zebranych podczas cenzusu USA. Do badanych metod należy DPMPM (ang. *bayesian joint distribution Dirichlet process mixture model*), wielokrotna imputacja z wykorzystaniem regresji logitowej (ang. *multiple imputation logistic regression*), wielokrotna imputacja z wykorzystaniem algorytmu drzewa decyzyjnego CART

(ang. *multiple imputation CART trees*). Na drodze symulacji pobierają 200 próbek o wielkościach 1000 lub 10 000 obserwacji, a następnie amputują zmienne mechanizmem MAR albo MCAR w 30% i 40%. Do porównania wyników wykorzystują kwantyle rozkładu Rel.MSE (opis w następnym rozdziale) dla spełniających założenia centralnego twierdzenia granicznego frakcji pojedynczych, par i potrójnych rozkładów brzegowych. Na dodatek obliczają odchylenia standardowe metodami Monte Carlo i pokrycie frakcji zawartych w przedziałach ufności uzyskanych za pomocą wielokrotnej imputacji. Na końcu twórcy rozszerzają swoje badanie, korzystając z analizy wrażliwości wyników [Akande i in., 2017].

Publikacja *MIMCA: Multiple imputation for categorical variables with multiple correspondence analysis* [2017] autorstwa Audigiera i in. w głównej mierze skupia się na zaprezentowaniu nowatorskiego podejścia do imputacji zmiennych jakościowych z wykorzystaniem wielokrotnej analizy korespondencji (ang. *multiple correspondence analysis*). MCA konfrontowane jest z DPMPM (ang. *Dirichlet process mixture of products of multinomial distributions model*), wielokrotną imputacją z wykorzystaniem lasów losowych (ang. *multiple imputation by random forests*), wielokrotną imputacją z wykorzystaniem log-liniowych modeli (ang. *multiple imputation using the log-linear model*), wielokrotną imputacją z wykorzystaniem regresji logitowej (ang. *multiple imputation by logistic regressions*). Zaproponowaną przez siebie metodę porównują na prawdziwych sześciu zbiorach danych. Zbiór danych traktują jako populację i na niej dopasowują modele regresji logistycznej; parametr, który oszacowali, traktują jako prawdziwy, referencyjny. Następnie losują ze zwracaniem pomiędzy losowaniami 200 próbek różnej wielkości (zależnie od wielkości zbioru danych) i amputują MCAR 20%. Badają obciążenie estymatorów parametrów strukturalnych modelu oraz pokrycie prawdziwych wielkości parametrów, używając przedziałów ufności dla wielokrotnej imputacji [Audigier i in., 2017].

Badanie Ghorbaniego i Desmaraisa *Performance Comparison of Recent Imputation Methods for Classification Tasks over Binary Data* [2017] skupia się na wpływie imputacji na zmianę błędu predykcji dla różnych algorytmów klasyfikacji. Są to naiwny klasyfikator Bayesa oraz jego wersja z drzewiastym rozszerzeniem (ang. *tree augmented naive Bayes*), regresja logistyczna i SVM z jądrem radialnym. Uczenia pracują z dwiema metodami pojedynczej imputacji, tj. missForest i hot deck oraz dwiema wielokrotnej imputacji, tj. algorytmu EM oraz stochastycznej regresji logistycznej. Aplikują mechanizm MCAR z różnymi proporcjami braków danych (5, 10, 20, 30, 40, 50%) na 14 zbiorach danych z dużą liczbą zmiennych (od 20 do 100 zmiennych) binarnych [Ghorbani, Desmarais, 2017].

Publikacja *Missing Data Imputation For Supervised Learning* [2018] Poulosa i Valle bada wpływ imputacji danych w przypadku występowania zmiennych jako-

ściowych na zmianę błędu klasyfikacji na algorytmy drzewa decyzyjnego, lasów losowych, sieci neuronowych. W badaniu biorą udział dwa zbiory, wykorzystane są zastane braki danych, dodatkowo dane treningowe są zaburzone MCAR o różnych proporcjach braków danych (od 10 do 40 co 10%). Badacze porównują imputację k – najbliższych sąsiadów, SVM, regresję logistyczną, lasy losowe z użyciem indykatorów (ang. *one hot encoding*) kategorii pod tytułem „brak informacji” jako punktem odniesienia [Poulos, Valle, 2018].

Praca *How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data* Stavseth i in. [2019] zajmuje się porównaniem sześciu metod wielokrotnej imputacji. W badaniu obecne są metody wielokrotnej imputacji z wykorzystaniem algorytmu maksymalizującego wartość oczekiwaną z bootstrapem (ang. *multiple imputation using expectation – maximization with bootstrapping*), wielokrotnej imputacji z wykorzystaniem wielokrotnej analizy korespondencji (ang. *multiple imputation using multiple correspondence analysis*), wielokrotnej imputacji z wykorzystaniem zmiennych ukrytych (ang. *multiple imputation using latent class analysis*), wielokrotnej imputacji przy użyciu hot deck (ang. *multiple hot deck imputation*), *multivariate imputation by chained equations logistic regression*, *multivariate imputation by chained equations random forests*. Autorzy korzystają z danych zebranych w kwestionariuszu dotyczącym uzależnienia od opiatów (cztery zmienne). Zmniejszając zbiór jedynie do w pełni obserwowalnych rekordów, pobierają próbki o dwóch wielkościach (200 i 1000 obserwacji) za pomocą losowania ze zwracaniem. Tworzą 400 pełnych wersji zbiorów, a następnie amputują 5%, 10%, 20%, 40%, zgodnie z mechanizmem MAR opartym na zależnościach, które zaobserwowali z powstających w kwestionariuszu braków danych. Jakość imputacji analizują na drodze porównania obciążenia parametrów regresji logitowej dla pełnych danych, zarówno tych po amputacji, jak i imputacji. Do tego liczą odchylenie standardowe oszacowania parametrów strukturalnych i pokrycie przedziału ufności prawdziwego parametru [Stavseth i in., 2019].

W odniesieniu do powyższych, niniejsza praca ma na celu porównać szerokie spektrum algorytmów imputacji, dotąd niespotkanych w tak dużym gronie w kontekście zmiennych nominalnych i w takim zestawieniu. Wyróżnia je zastosowanie zarówno metod uczenia maszynowego, głębokiego uczenia i klasycznych metod statystycznych. Wykorzystana jest rzeczywista baza danych z siatką różnych wielkości luk w danych amputowanych za pomocą MCAR i MAR. Porównanie jakości imputacji odbywa się na wielu poziomach. Oprócz badania zdolności predykcyjnych prognozowania idiosynkratycznych wartości, za pomocą zasad Rubina wielokrotnej imputacji badane jest odtwarzanie brzegowych frakcji dla wielu kategorii.

Pozwala to tym samym wnioskować na temat oddawania wielowymiarowych rozkładów przez narzędzia imputacji. Sprawdzane jest obciążenie, pokrycie przedziałów ufności, szerokość przedziałów dla estymatorów stworzonych na drodze wielokrotnej imputacji. Ponadto mierzony jest czas wykonywania się obliczeń i ewentualne pojawiające się problemy ze zbieżnością algorytmów.

3. Opis metod badania

3.1. Opis danych

Badanie zostanie przeprowadzone w oparciu o wykorzystanie ogólnodostępnej bazy danych pochodzącej z pakietu R *kernlab*¹. W bazie danych zostały usunięte obserwacje, tak aby do samego badania były użyte wyłącznie w pełni obserwowalne rekordy.

Zbiór *income* pochodzi z badania ankietowego firmy marketingowej przeprowadzonego w galerii handlowej w San Francisco. Posiada 6876 w pełni obserwowalnych obserwacji, 13 atrybutów o liczności kategorii od 2 do 9. Jest to średniej wielkości zbiór. Dokładny opis atrybutów obecnych w ramce danych znajduje się w załączniku.

Oryginalne pełne zbiory danych będą stanowiły odtąd w badaniu populację. Z populacji będą pobierane próbki o wielkości 1000 obserwacji 100 razy dla danego schematu amputacji. Za każdym razem na danej próbie będzie wykonana amputacja według mechanizmu określonego na danym etapie badania (MCAR albo MAR) z różnymi proporcjami braków (10, 30, 50% albo 33%), a następnie wykonana wielokrotna imputacja.

Amputacja mechanizmem MCAR zostanie wykonana jednostajnie na wszystkich zmiennych, tj. w przypadku proporcji braków q , każda z obserwacji ma prawdopodobieństwo obserwowalności $1-q$. W przypadku wystąpienia obserwacji o brakach danych dla wszystkich cech, uzupełniana była losowo dla tej obserwacji jedna wartość z oryginalnej ramki danych.

W celu przetestowania algorytmów z wykorzystaniem indukowania braków mechanizmem MAR zmienne zostały podzielone na w pełni obserwowalne i będące obiektami amputacji. O czterech zmiennych w próbie dostępne są wszystkie informacje. Pozostałe siedem cech warunkowo w zależności od wartości zmiennych w pełni obserwowanych, poddawanych jest różnym proporcjom braków, które

¹ <https://cran.r-project.org/web/packages/kernlab/kernlab.pdf> (dostęp: 25.07.2021).

dają średnio 30% brakujących informacji o wartościach tych zmiennych, a 17% w stosunku do całej realizacji próby. Szczegóły dotyczące mechanizmu znajdują się w załączniku. Dla tak przygotowanej ramki danych analizy miar oceny jakości odbywają się z wyłączeniem obserwowalnych zmiennych.

Praca w dwójnasób bada zmienność wyników i mierzy niepewność indukowaną poprzez obecność braków. Po pierwsze dokonuje się wielokrotne próbkowanie z oryginalnej bazy, po drugie każda taka instancja poddawana jest procedurze wielokrotnej imputacji z $M = 10$. Taki zabieg był już promowany w kilku nowszych pracach w dziedzinie [Audigier i in., 2017; Akande i in., 2017; Akande, Wongkamthong, 2021; Wang i in., 2022]. Pozostało rozstrzygnąć, jaki obiekt wybrać na zmienną oszacowania. W przypadku zmiennych nominalnych dobrym kandydatem zdaje się być wybór frakcji „populacyjnych” prawdopodobieństw *a priori*. Skupienie się na frakcjach dla par warunkowych prawdopodobieństw (ang. *marginal probabilities*) zapewniłoby weryfikację zdolności odtwarzania wielowymiarowych zależności narzędzia imputującego. Ważnym przy tym elementem analizy jest wykluczenie frakcji, które nie zapewniałyby zachodzenia centralnego twierdzenia granicznego. Dlatego, przy oznaczeniu wielkości próby jako n oraz wielkości frakcji danej wartości dla zmiennej w próbie jako p , z badania wykluczone są te warunkowe prawdopodobieństwa, które nie spełniają $np > 10$ lub $n(1 - p) > 10$.

3.2. Opis miar jakości

Wychodząc naprzeciw zarzutom stawianym w kierunku dotychczas wydanych prac, przywołanych w publikacji Lin i Tsai [2020], jakoby zbyt często używano w nich jedynie pojedynczych metryk ewaluacji, jakość imputacji będzie mierzona za pomocą kilku. W wielu pracach, w których algorytmy były przedstawiane po raz pierwszy, ich porównanie odbywało się relatywnie w stosunku do usuwania wierszy (ang. *listwise deletion*). Zostało dobrze ugruntowane, że w tym kontekście metody imputacji poprawiają szeroko pojęte wnioski. Dlatego w niniejszej pracy nie jest stosowany ten punkt odniesienia, a badane są zdolności odtwarzania frakcji populacyjnych przez metody wielokrotnej imputacji dla danych nominalnych. Wyniki agregowane są dla wszystkich frakcji.

Trafność (ang. *accuracy*): Ze względu na historyczne zażyłości, bliżej związane z pojedynczą imputacją (ang. *single imputation*), mierzona będzie zdolność predykcyjna poszczególnych algorytmów do odtwarzania oryginalnych wartości.

$$\text{Trafność} = \frac{\sum_{i=1}^n \sum_{j \in S} M_{ij} \mathbb{I}(\hat{Y}_{ij} = Y_{ij})}{\sum_{i=1}^n \sum_{j \in S} M_{ij}},$$

gdzie n – wielkość próby, \hat{Y}_{ij} – zaimputowana wartość, Y_{ij} – oryginalna wartość, S – zbiór indeksów zmiennych obecnych w ramce danych, M_{ij} – indykator odpowiadający za występowanie amputowanej wartości.

Rel.MSE: Kwadratowa funkcja straty, błąd średniokwadratowy uwzględnia zarówno obciążenie, jak i wariancję estymatora. W przypadku tej metryki badana jest relatywna zdolność odtwarzania parametru populacyjnego dla próbki po imputacji w stosunku do próbki pełnej przed amputacją.

$$Rel.MSE = \frac{\sum_{h=1}^H (\bar{Q}_h - Q)^2}{\sum_{h=1}^H (\bar{q}_h - Q)^2},$$

gdzie H – liczba próbek w symulacji, \bar{Q}_h – zagregowany estymator z wielokrotnej imputacji dla symulacji h , Q – wartość estymowana populacyjna, \bar{q}_h – estymator wartości populacyjnej z pełnej próby. Przeanalizowany będzie rozkład *Rel.MSE* w celu uzyskania bardziej informatywnych wniosków.

Absolutny błąd wystandaryzowany (ang. *absolute standardized bias*): Mierzy obciążenie estymatora wielokrotnej imputacji. Standaryzacja dokonywana jest w celu wzięcia pod uwagę bliskich zera oszacowań frakcji.

$$ASB = \sum_{h=1}^H \frac{|\bar{Q}^{(h)} - Q|}{HQ},$$

gdzie oznaczenia jak powyżej. Również jak w przypadku miary 2, porównanie będzie wykonywane w odniesieniu do rozkładu *ASB*.

Pokrycie (ang. *coverage*): Jak zostało to wcześniej umotywowane, ważnym aspektem ewaluacji jest badanie pokrycia przedziałów ufności wielkości estymowanej, wyznaczonych na poziomie istotności α oznaczonych CI_h^α , posumowanych po wszystkich symulacjach:

$$Coverage = \sum_{h=1}^H \frac{\mathbb{I}(Q \in CI_h^\alpha)}{H}.$$

Symulacja przeprowadzona jest $H=100$ razy, za każdym razem wykonywana jest procedura badania opisana poniżej.

Z oryginalnej, w pełni obserwowalnej bazy danych, która traktowana jest jako populacja, pobierana jest próbka w schemacie bez zwracania.

Na próbce dokonywana jest estymacja populacyjnych frakcji spełniających założenia konieczne dla CTG w próbce.

Na próbce dokonywana jest amputacja według wybranego mechanizmu powstawania braków (MCAR albo MAR w proporcjach 10%, 30%, 50% albo 33%*).

Na amputowanym zbiorze wykonywane są procedury wielokrotnej imputacji z $M=10$ za pomocą różnych algorytmów. Konstruowany jest punktowy i przedziałowy estymator zgodnie z regułami Rubina.

Dla każdej wersji zbioru danych zapisywana jest trafność.

Dla zagregowanych wyników mierzone jest pokrycie przedziału ufności i jego szerokość oraz miary $Rel.MSE$, ASB .

Na samym końcu, aby podjąć próbę agregacji wniosków, wyniki zostaną przedstawione graficznie oraz zostanie stworzony ranking dla każdej z miar. Wyniki zagregowane będą przeanalizowane za pomocą narzędzi statystyki nieparametrycznej [Gheyas, Smith, 2010; Jadhav i in., 2019]. Zostaną wykorzystane testy Friedmana i Page'a.

3.3. Opis użytych metod wielokrotnej imputacji

Najpopularniejszym środowiskiem do obliczeń statystycznych używanym w praktyce imputacji jest R. Aktualny przegląd dostępnych pakietów, które mogą się badaczowi przydać, można znaleźć na stronie <https://cran.r-project.org/web/views/MissingData.html>. Dziewięć z wykorzystanych metod było zaimplementowanych w R. Dwie metody głębokiego uczenia zostały napisane w Pythonie, wraz z całym kodem obsługującym symulację.

Detale dotyczące hiperparametrów dla poszczególnych algorytmów znajdują się w załączniku. Zostały zastosowane ustawienia domyślne algorytmów lub te proponowane w artykułach, które wprowadzały daną metodę do dziedziny.

Wśród używanych metod możemy dokonać pewnego podziału taksonomicznego. Wyróżniamy metody uczenia maszynowego, metody głębokiego uczenia i te, które zaliczane są do klasycznych narzędzi teorii statystyki [Lin, Tsai, 2020]. Szczegółowy opis algorytmów oraz analiza teoretycznych własności nie są celem niniejszego badania. Wszelkie detale na ich temat można znaleźć w artykułach, które przytoczono podczas wymieniania każdej z metod.

1) Metody statystyczne

- DPMPM: wykorzystuje metody opierające się o modele ze zmienną ukrytą (ang. *latent class model*). Implementacja DPMPM (ang. *Dirichlet process mixture of products of multinomial distributions*) znajduje się w pakiecie R, `NPBayesImputeCat`.
- AMELIA: metoda jest oparta o strategię modelowania JM poprzez zastosowanie wielowymiarowego rozkładu normalnego i algorytmu EM. Jej implementację można znaleźć w pakiecie R `AMELIA`, a opis u Honakera, Kinga

i Blackwella [2011]. Jest on reprezentantem metod opartych na największej wiarygodności.

- POLYREG: algorytm jest oparty o strategię modelowania FCS MICE. Regresję multinomialną (ang. *polytomous regression*) opisano w pracy Branda [1999], a zaimplementowano w pakiecie R *mice*. Jest on reprezentantem uogólnionych modeli liniowych.
- PMM: PMM (ang. *predictive mean matching*) opiera się o algorytm MICE wykorzystujący FCS, jego opis można znaleźć w dziele van Buurena [2018], a implementację w pakiecie R *mice*. Metoda należy do klasy hot deck, ale wykorzystuje regresję do wyboru dawców (ang. *donors*).
- MIMCA: MCA (ang. *multiple correspondence analysis*) wykorzystuje strategię JM. Jej opis znajduje się w artykule Audigiera i in. [2017], a ich implementacja w pakiecie R *missMDA*. Jest to algorytm korzystający z teorii redukcji wymiarowości.
- HD: opis nieparametrycznej metody hot deck znajduje się w pracy Cranmera i Gilla [2013, a jej implementacja w „hot.deck R package”.
- LDA: opis LDA znaleźć można w dziele Branda [1999], a implementację w *mice*. Inspiracją do powstania algorytmu były klasyfikatory budowane w oparciu o regułę Bayesa.

2) Metody uczenia maszynowego

- CART: drzewa decyzyjne (ang. *decision trees*) CART korzystają z FCS, opisane są w publikacji Burgette, Reitera [2010] lub van Buurena [2018], a zaimplementowane w *mice*.
- RF: lasy losowe (ang. *random forests*) korzystają z FCS, opisane są w pracy Doove, van Buurena i Dusseldorp [2014], a zaimplementowane w *mice*.

3) Metody głębokiego uczenia

- GAIN: GAIN jest opisany w pracy Yoon, Jordon, Schaar [2018] i modyfikacja dla wielokrotnej imputacji w Wang i in. [2022]. Została wykorzystana implementacja Zhenhua Wang². Algorytm wykorzystuje sieci neuronowe typu GAN (ang. *generative adversarial networks*).
- MIDA: MIDA jest opisany w Gondara, Wang [2018] i modyfikacja dla wielokrotnej imputacji w Wang i in. [2022]. Została wykorzystana implementacja Zhenhua Wang. Algorytm opiera się w swojej konstrukcji na architekturze autokoderów odszumiających (ang. *denoising autoencoders*).

² https://github.com/zhenhua-wang/MissingData_DL (dostęp: 30.06.2021).

4. Wyniki badania

Wyniki badania zostały zebrane dla każdego z mechanizmów braków dla każdej frakcji każdej ze zmiennych w bazie danych (47) i dla kombinacji par (1412). Każdą z metod umieszczono w tabeli z wynikami Rel.MSE, ASB. Pokrycia zostały wyrysowane na boxplotach. Na końcu znajduje się opis czasu obliczeń algorytmów. Do załącznika została przeniesiona tabela z rozkładem trafności w konkretnej imputacji dla prób (10×100 oszacowań) oraz relatywna szerokość przedziału ufności pomiędzy pełną próbą a amputowaną.

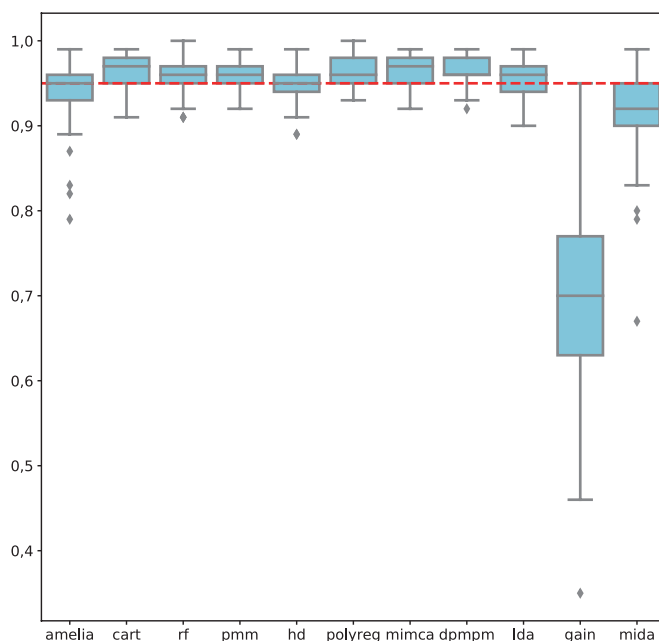
4.1. Symulacja MCAR

Zgodnie z intuicją wraz ze wzrostem proporcji występowania braków można zauważyć przesuwanie się wykresów skrzynkowych (ang. *boxplots*) pokrycia frakcji populacyjnych przez przedziały ufności Rubina w dół. Zwiększa się także wariancja wśród pokryć. Dla niektórych frakcji wartości kombinacji dwóch zmiennych pokrycia są na niskich poziomach, co jest zapewne spowodowane niskimi wartościami tych frakcji w populacji. Dla kombinacji dla danej proporcji mechanizmu braków zazwyczaj można zaobserwować wyższą wariancję pokrycia. Średnia szerokość przedziałów ufności jest podobna pomiędzy metodami, waha się od 0,02 do 0,1 w skrajnych przypadkach dla frakcji dla kombinacji par. Dla algorytmu GAIN można dostrzec przesunięcie rozkładu średniej szerokości przedziału ufności o mniej więcej 0,01 w kierunku węższych oszacowań. Kolejna metoda, która w mniejszym stopniu, ale wciąż wyróżnia się na tle pozostałych algorytmów w tym względzie, to HD; uzyskuje także węższe przedziały ufności oraz w mniejszym stopniu LDA. Wynik metody hot deck wydaje się intuicyjny, biorąc pod uwagę jej proces uzupełnienia braków obecnymi obserwowalnymi wartościami.

Potwierdziły się wyniki z pracy Wang i in. [2022], że metody głębokiego uczenia, tj. MIDA i GAIN osiągają niższą medianę pokryć od pozostałych metod. Przy czym GAIN radzi sobie zauważalnie gorzej niż MIDA. Mediana pokrycia dla metod PMM, DPMPM, MIMCA na wszystkich analizowanych wykresach osiąga co najmniej nominalny poziom 0,95. Dla tych dwóch ostatnich potwierdza to wyniki z publikacji Audigiera i in. [2017] oraz artykułu Akande i in. [2017]. Pośród algorytmów MICE, najgorsze wyniki w kontekście pokrycia przeciętnie osiąga LDA. Zaskakujące są wyniki gorsze dla lasów losowych w porównaniu do zwykłego drzewa CART; te wyniki także potwierdzają wnioski z analizy z pracy Wang i in. [2022]. Wśród metod typu hot deck parametryczna metoda PMM przewyższa nieparametryczny HD. Potwierdził się wynik z artykułu Akande i in. [2017],

że metody oparte na uogólnionych modelach liniowych, tj. POLYREG, radzą sobie gorzej od MIMCA, CART i DPMPM, natomiast w kontekście wszystkich analizowanych metod nie są to duże różnice. Algorytmy AMELIA, GAIN i MIDA najbardziej straciły na zwiększeniu proporcji brakujących danych, patrząc po średniej zmianie mediany pokrycia. Przyglądając się minimum z pokryć dla frakcji, można ocenić zdolność oddawania najmniejszych z nich. Dokładne wyniki można zaobserwować na rysunkach od 1 do 6.

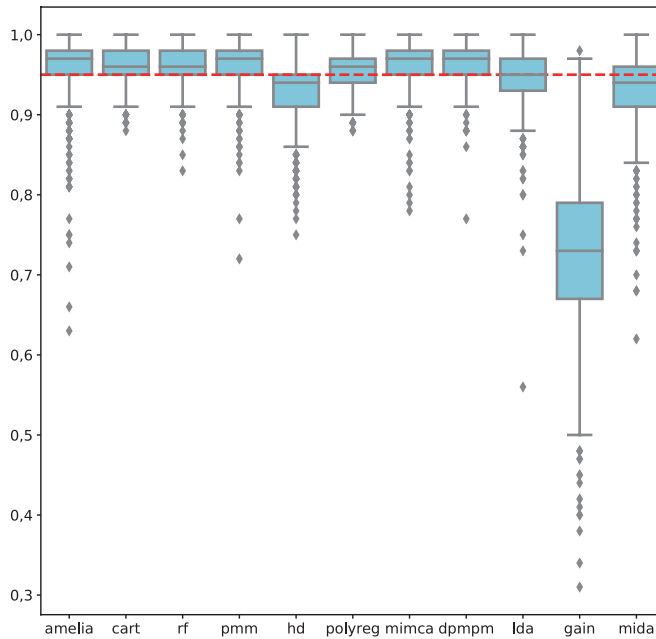
Rysunek 1. Pokrycie dla frakcji jednej zmiennej MCAR 0,1



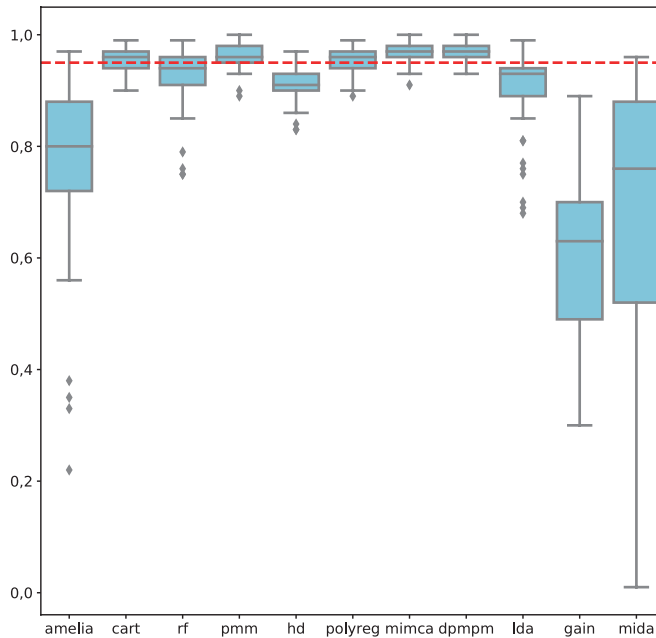
Źródło: opracowanie własne.

Fakt, że Rel.MSE jest wyższe dla frakcji pojedynczych zmiennych wynika z tego, że sam błąd średniokwadratowy dla pełnej próby, który znajduje się w mianowniku wyrażenia, jest mniejszy niż dla frakcji kombinacji wartości dla par. Wraz ze wzrostem proporcji braków różnice wśród algorytmów uwidaczniają się.

Dla algorytmów AMELIA, GAIN i MIDA można dostrzec bardzo grube ogony rozkładów Rel.MSE. To sugeruje, że metody mają dużą wariację wśród oszacowań pomiędzy różnymi próbami i zmiennymi. Wyżej wymienione zdają się radzić sobie najgorzej na tle pozostałych metod. Stabilne wyniki zawsze uzyskują algorytmy drzewiaste CART i RF ze wskazaniem na ścisłą dominację wśród wyników tego pierwszego.

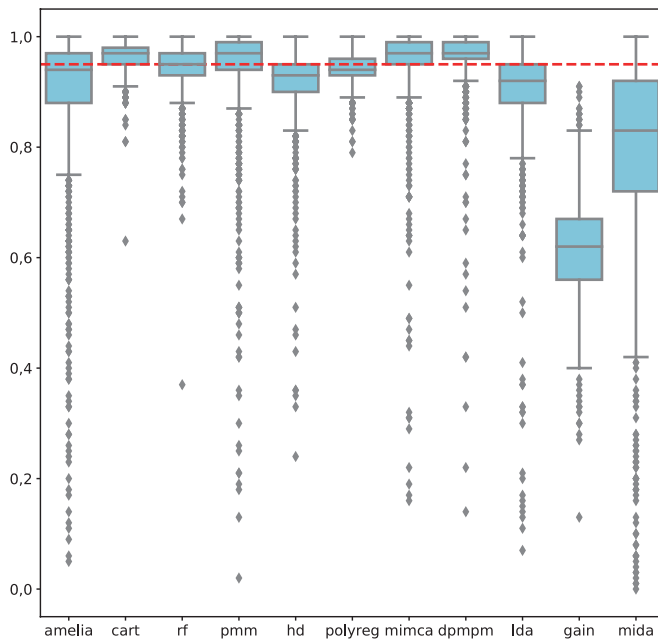
Rysunek 2. Pokrycie dla frakcji kombinacji dwóch zmiennych MCAR 0,1

Źródło: opracowanie własne.

Rysunek 3. Pokrycie dla frakcji jednej zmiennej MCAR 0,3

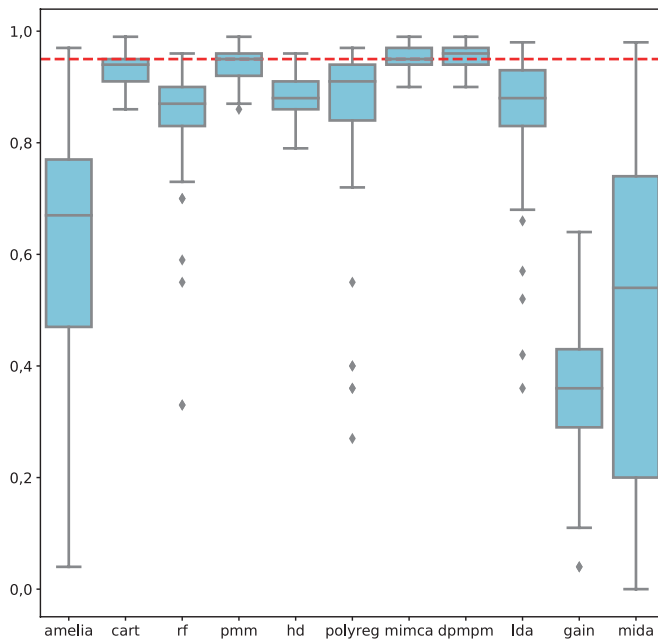
Źródło: opracowanie własne.

Rysunek 4. Pokrycie dla frakcji kombinacji dwóch zmiennych MCAR 0,3

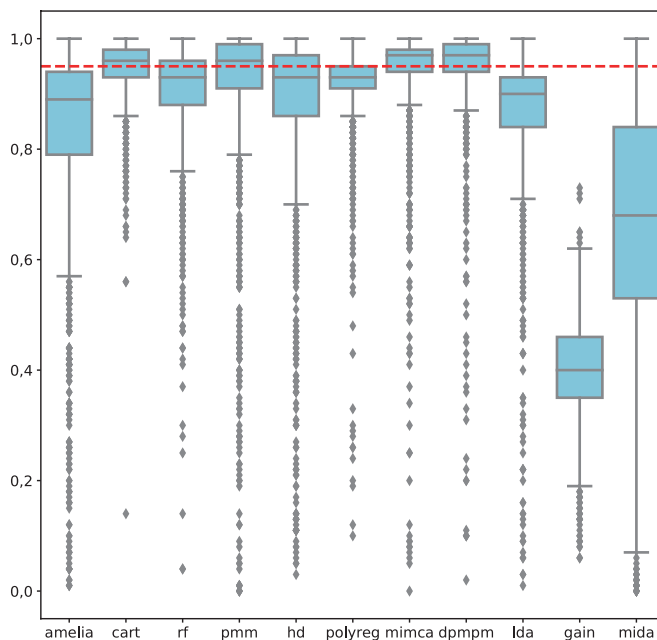


Źródło: opracowanie własne.

Rysunek 5. Pokrycie dla frakcji jednej zmiennej MCAR 0,5



Źródło: opracowanie własne.

Rysunek 6. Pokrycie dla frakcji kombinacji dwóch zmiennych MCAR 0,5

Źródło: opracowanie własne.

Algorytmy LDA, HD i POLYREG często radzą sobie gorzej w odtwarzaniu frakcji populacyjnych od MIMCA, DPMPM, PMM, CART, RF, ale nie jest to jednostajna reguła. Patrząc na różnice w rozkładach, widać sensowność porównywania ich nie względem mediany, a całościowo, biorąc pod uwagę kilka kwantyli.

Wykonanie dla każdej z 12 tabel testu Friedmana z poprawką Benjaminiego-Hochberga na wielokrotne testowanie daje wniosek, że metody nie radzą sobie jednakowo pod względem miar ASB i Rel.MSE, biorąc pod uwagę ich wyniki dla kwantyli podanych w tabeli. Można wśród nich wskazać pewien porządek. Test Page'a z poprawką Benjaminiego-Hochberga potwierdził występowanie porządku wśród metod dla wszystkich tabel oprócz tabeli dla rozkładu ASB przy brakach 50% MCAR³.

Wszystkie metody uzyskują znacznie wyższe ASB dla frakcji dla par, co potwierdza intuicję, że oddawanie charakterystyk wielowymiarowych jest trudniejszą sztuką. Algorytmy MIDA i AMELIA uzyskują najgorsze wyniki. Podobnie jak w przypadku Rel.MSE, najlepsze wyniki zdaje się uzyskiwać metoda CART. Algorytm PMM zdecydowanie gorzej radzi sobie w oddawaniu wielowymiarowych zależności w stosun-

³ Szczegóły odnośnie do wskazanego trendu z hipotezy alternatywnej w załączniku.

ku do jego relatywnych wyników dla pojedynczych frakcji. Zarówno dla ASB, jaki i Rel.MSE algorytm POLYREG traci na wzroście ilości braków relatywnie bardziej niż najlepsi konkurenci, spadając w rankingach.

Tabela 1. Kwantyle Rel.MSE dla 10% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,13	1,02	1,04	1,03	1,06	1,00	1,03	1,04	1,01	1,14	1,10
25%	1,20	1,06	1,07	1,07	1,11	1,05	1,08	1,07	1,07	1,20	1,25
50%	1,38	1,10	1,10	1,14	1,20	1,11	1,14	1,11	1,10	1,34	1,42
75%	1,63	1,17	1,18	1,17	1,27	1,17	1,17	1,18	1,22	1,51	1,57
90%	1,97	1,20	1,22	1,21	1,34	1,20	1,22	1,20	1,28	1,82	2,00
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,03	0,98	0,98	0,95	1,08	1,04	0,94	0,96	1,03	1,16	1,07
25%	1,10	1,03	1,04	1,01	1,15	1,10	1,01	1,01	1,10	1,26	1,15
50%	1,20	1,09	1,10	1,08	1,23	1,17	1,08	1,08	1,18	1,38	1,24
75%	1,34	1,15	1,16	1,16	1,32	1,25	1,17	1,15	1,27	1,54	1,37
90%	1,58	1,22	1,23	1,27	1,41	1,31	1,28	1,22	1,38	1,73	1,60

Źródło: opracowanie własne.

Tabela 2. Kwantyle Rel.MSE dla 30% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	2,01	1,20	1,29	1,25	1,59	1,23	1,27	1,19	1,31	1,92	2,01
25%	2,91	1,29	1,38	1,29	1,69	1,33	1,35	1,31	1,42	2,72	2,78
50%	3,91	1,40	1,61	1,45	1,87	1,50	1,45	1,46	1,67	3,39	3,94
75%	6,17	1,50	1,78	1,61	2,06	1,71	1,58	1,58	2,02	3,90	6,28
90%	9,98	1,58	2,17	1,73	2,22	2,06	1,75	1,71	3,03	8,48	9,2
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,46	1,06	1,11	0,98	1,32	1,39	1,03	1,02	1,39	2,12	1,67
25%	1,76	1,20	1,25	1,17	1,48	1,57	1,21	1,17	1,58	2,58	1,99
50%	2,32	1,34	1,43	1,40	1,72	1,77	1,40	1,37	1,85	3,16	2,48
75%	3,37	1,52	1,66	1,82	2,03	2,06	1,71	1,57	2,32	3,93	3,22
90%	5,39	1,78	2,00	2,93	2,49	2,34	2,29	1,88	3,00	5,11	4,88

Źródło: opracowanie własne.

Tabela 3. Kwantyle Rel.MSE dla 50% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	4,37	1,62	1,90	1,75	2,22	1,85	1,67	1,62	1,88	5,43	3,96
25%	6,41	1,82	2,24	2,01	2,40	2,18	1,93	1,89	2,27	8,14	5,46
50%	9,94	2,15	2,66	2,37	2,80	3,00	2,36	2,31	2,78	10,23	10,61
75%	14,23	2,40	3,44	2,67	3,09	4,64	2,58	2,51	3,87	15,60	15,52
90%	28,72	2,62	4,56	3,09	3,74	9,47	2,95	2,90	6,12	29,60	31,85
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	2,56	1,15	1,21	1,11	1,20	2,24	1,30	1,16	1,92	4,02	2,52
25%	3,44	1,48	1,63	1,51	1,62	2,77	1,64	1,49	2,44	5,19	3,41
50%	5,10	1,88	2,07	2,12	2,24	3,54	2,12	1,93	3,23	7,08	4,54
75%	8,86	2,46	2,73	3,26	3,36	4,61	2,79	2,48	4,58	10,14	6,52
90%	15,07	3,46	3,87	7,30	6,54	5,90	4,18	3,39	6,80	16,20	10,42

Źródło: opracowanie własne.

Tabela 4. Kwantyle ASB dla 10% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,51	0,12	0,04	0,11	0,08	0,12	0,09	0,11	0,13	0,29	0,55
25%	0,82	0,22	0,28	0,33	0,29	0,23	0,28	0,30	0,38	0,68	1,52
50%	2,50	0,48	0,69	0,61	0,80	0,59	0,44	0,61	0,82	1,25	3,03
75%	4,67	0,82	1,34	0,89	1,29	1,28	0,96	1,26	1,43	2,27	4,83
90%	8,21	1,18	2,31	1,25	1,86	1,56	1,53	1,70	2,74	3,80	8,14
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,72	0,18	0,23	0,27	0,22	0,19	0,27	0,24	0,31	0,36	0,81
25%	1,62	0,45	0,58	0,75	0,58	0,50	0,73	0,61	0,69	0,94	2,01
50%	3,45	1,03	1,40	1,72	1,33	1,07	1,79	1,34	1,58	2,00	4,15
75%	6,35	2,08	2,66	3,59	2,65	1,94	3,57	2,72	3,29	3,74	7,01
90%	9,59	3,48	4,41	6,32	3,97	3,25	6,27	4,50	5,73	6,07	9,87

Źródło: opracowanie własne.

Tabela 5. Kwantyle ASB dla 30% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,36	0,19	0,51	0,17	0,18	0,61	0,15	0,34	0,18	0,57	2,07
25%	3,49	0,29	1,02	0,34	0,41	0,88	0,32	0,55	0,90	1,94	5,63
50%	6,84	0,44	2,14	0,90	0,92	2,05	0,72	1,03	1,97	3,85	9,02
75%	12,07	1,00	3,70	1,46	1,70	3,31	1,68	1,91	4,79	7,57	14,78
90%	24,77	1,32	6,13	1,88	2,90	7,69	2,42	4,69	9,80	11,85	23,26
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,91	0,33	0,73	0,74	0,62	0,48	0,73	0,55	0,75	1,21	2,34
25%	4,68	0,92	1,77	1,82	1,58	1,25	1,71	1,36	1,84	2,98	6,74
50%	9,75	2,37	3,68	4,64	3,57	2,58	4,15	3,06	4,34	6,67	14,15
75%	17,83	4,64	7,15	10,69	7,17	4,97	8,16	6,23	9,00	11,06	21,59
90%	29,15	8,05	11,89	18,43	12,82	8,67	15,96	11,21	15,06	15,08	27,96

Źródło: opracowanie własne.

Tabela 6. Kwantyle ASB dla 50% wartości MCAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	2,68	0,05	0,68	0,15	0,28	1,16	0,27	0,42	0,63	1,75	5,51
25%	5,76	0,17	1,68	0,63	0,78	3,09	0,88	1,08	1,64	5,34	10,17
50%	11,91	0,55	3,28	1,47	1,44	6,19	1,36	2,08	4,05	10,91	16,60
75%	21,65	1,36	8,38	2,86	2,58	9,63	3,28	3,90	7,87	16,05	24,51
90%	42,90	2,11	9,75	4,16	4,20	21,91	5,21	6,72	13,54	25,51	36,74
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	3,82	0,84	1,44	1,23	1,19	1,46	1,05	1,09	1,34	3,60	5,27
25%	9,00	2,10	3,87	3,56	3,53	3,85	2,98	2,78	3,65	8,23	13,01
50%	19,53	5,45	8,03	9,62	9,04	8,68	6,82	6,22	7,89	16,44	27,77
75%	42,65	11,30	15,79	20,66	18,58	19,16	14,44	13,20	15,01	26,27	38,31
90%	73,21	20,90	26,10	39,19	32,91	32,47	27,05	25,95	25,84	33,86	46,89

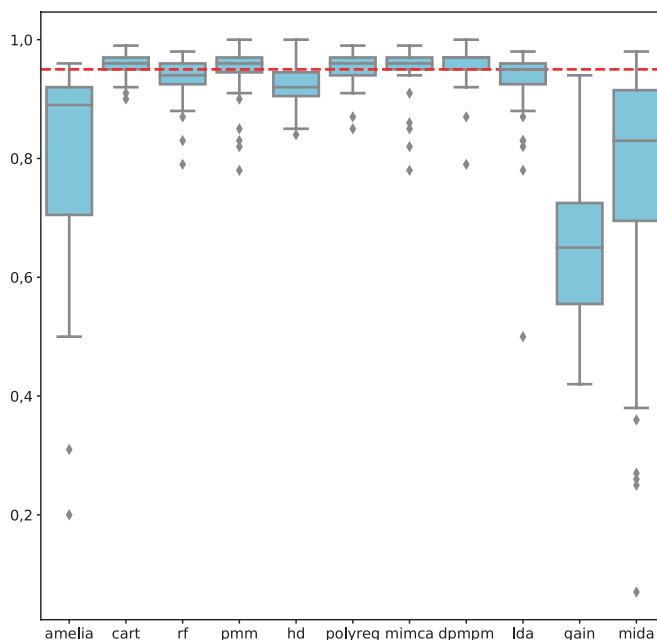
Źródło: opracowanie własne.

4.2. MAR

Wyniki: nominalne wartości miar nie mogą być porównywane wprost pomiędzy schematem z brakami typu MCAR i MAR, ponieważ zestaw zmiennych podanych mechanizmowi powstawania braków nie jest identyczny.

Metody AMELIA, GAIN, MIDA i HD nie uzyskują nominalnego poziomu pokrycia prawdziwych wartości przez przedziały ufności Rubin na poziomie istotności 0,05. W przypadku metod GAIN i HD może to być związane z węższymi przedziałami ufności, tj. także małą zmiennością pomiędzy poszczególnymi wersjami imputowanych zbiorów w ramach danej próbki. Wyniki zebrane są na rysunkach 7 i 8.

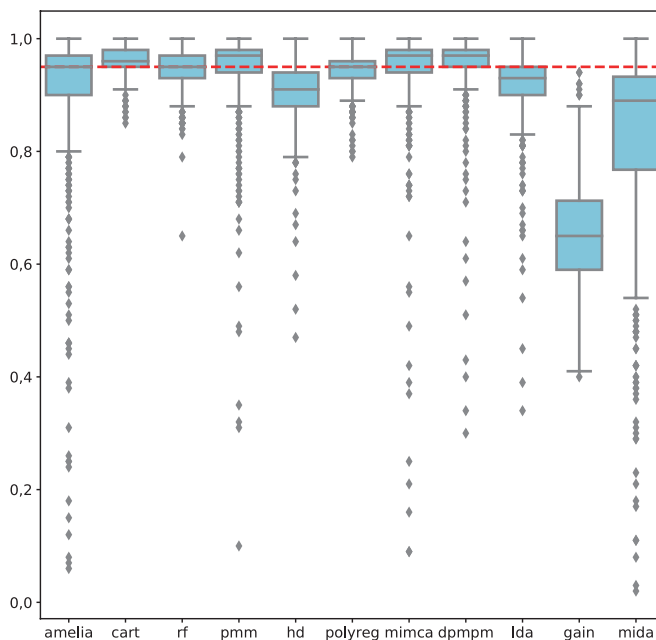
Rysunek 7. Pokrycie dla frakcji jednej zmiennej MAR



Źródło: opracowanie własne.

Analizując rozkład miary Rel.MSE, ponownie można dostrzec grube ogony wśród metod AMELIA, GAIN, MIDA, uwidocznione w tabeli 7. Ponownie, jak w przypadku MCAR, najkorzystniej zdaje sobie radzić metoda CART. W kontekście tejże miary, dobre wyniki dla problemu oddawania wielowymiarowych zależności osiągają CART, RF, PMM, MIMCA, DPMPM. GAIN osiąga lepsze wyniki niż MIDA i AMELIA.

Rysunek 8. Pokrycie dla frakcji kombinacji dwóch zmiennych MAR



Źródło: opracowanie własne.

Tabela 7. Kwantyle Rel.MSE dla MAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,55	1,10	1,23	1,11	1,21	1,11	1,13	1,12	1,15	1,36	1,24
25%	2,17	1,19	1,29	1,26	1,30	1,17	1,22	1,30	1,22	1,63	1,86
50%	3,10	1,33	1,46	1,41	1,46	1,37	1,41	1,43	1,30	2,21	2,38
75%	5,54	1,54	1,70	1,87	1,86	1,62	1,73	1,89	1,77	3,80	5,46
90%	8,64	1,78	2,02	2,30	2,03	2,22	2,21	2,11	2,69	4,56	9,66
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,22	1,00	1,05	1,00	1,23	1,19	0,99	0,99	1,19	1,44	1,21
25%	1,48	1,12	1,15	1,13	1,35	1,34	1,11	1,11	1,33	1,80	1,44
50%	1,99	1,24	1,33	1,33	1,57	1,55	1,30	1,29	1,55	2,35	1,87
75%	3,09	1,48	1,55	1,73	1,87	1,87	1,63	1,58	1,94	3,20	2,69
90%	5,20	1,74	1,85	2,70	2,20	2,29	2,35	2,01	2,76	4,33	5,37

Źródło: opracowanie własne.

Zaskakujący w kontekście wyżej wymienionych rezultatów jest bardzo korzystny wynik metody HD w przypadku miary ASB, widoczny w tabeli 8. Może to oznaczać w odniesieniu do wyników innych miar trafne przewidywanie rzadkich frakcji tegoż algorytmu. Grube ogony rozkładu ASB można dostrzec dla metod AMELIA i MIDA zarówno w przypadku frakcji pojedynczych, jak i frakcji dla par. Ponownie najlepsze wyniki uzyskuje CART. Analizując tabelę, można zwrócić uwagę na fakt, że rozkłady mają różne kształty i dla wielu metod kolejność wyników zmienia się podczas przechodzenia pomiędzy różnymi kwantylami. Test Friedmana wskazał na wpływ zastosowanej metody na ranking jakości imputacji.

Tabela 8. Kwantyle ASB dla MAR

Brzegowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,65	0,10	0,62	0,22	0,19	0,37	0,19	0,14	0,28	0,85	2,56
25%	3,59	0,35	1,04	0,57	0,40	0,76	0,53	0,65	0,59	1,43	4,39
50%	6,04	0,61	1,61	1,22	0,81	1,56	1,26	1,51	1,48	2,71	7,01
75%	13,06	1,63	3,24	4,35	1,42	3,18	2,97	3,85	2,89	4,50	10,75
90%	19,32	2,72	5,68	7,69	2,34	5,21	7,42	7,68	5,61	10,11	15,44
Dwuwymiarowe											
	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	1,71	0,40	0,64	0,54	0,29	0,49	0,59	0,62	0,54	0,69	2,26
25%	4,85	0,99	1,61	1,60	0,82	0,97	1,64	1,39	1,36	1,87	4,82
50%	9,94	2,27	3,18	4,82	2,10	2,56	4,44	3,27	3,23	4,73	9,73
75%	16,73	4,40	6,05	9,11	4,07	4,80	9,16	7,39	6,95	8,70	18,00
90%	25,84	7,38	10,66	15,30	7,13	8,31	14,79	11,93	12,24	12,66	25,09

Źródło: opracowanie własne.

4.3. Czasy wykonywania zadania

Obliczenia były wykonywane na Intel® Core™ i7-8750H CPU 2.20GHz, 16GB RAM z kartą graficzną NVIDIA GeForce GTX 1050 Ti. Modele głębokiego uczenia są dostosowane do korzystania z mocy obliczeniowej karty graficznej, dzięki czemu są szybsze niż RF czy CART. Niemniej jednak lasy losowe także mogą korzystać ze zrównoleglenia tworzenia się drzew. Warto zwrócić uwagę, że niektóre algorytmy są wrażliwe na proporcję brakujących wartości. Dla mechanizmu MCAR i proporcji braków 50% metoda AMELIA dla niektórych próbek doświadczała problemów numerycznych i algorytm nie osiągał zbieżności. Własności te i różnice w czasie

wykonywania się obliczeń mogą zyskiwać na znaczeniu przy rozważaniu bardzo dużych zbiorów danych, co pokazano w tabeli 9.

Tabela 9. Czas wykonywania się algorytmów

Czas dla 100 próbek i wielokrotnej imputacji	Nazwa algorytmu (proporcja braków, jeżeli zależne)
Krótki (0; 1h)	AMELIA (0,1), GAIN CART (0,1), LDA DPMPM, MIMCA, HD, PMM, MIDA
Średni (1h; 2h)	RF (0,1), CART (0,3, 0,5)
Długi (2h; 3h)	RF (0,3)
Bardzo długi (3h; +)	AMELIA (0,3, 0,5), RF (0,5), POLYREG

Źródło: opracowanie własne.

5. Wnioski

Dziedzina radzenia sobie z brakującymi danymi rozwija się prężnie. W wielu badaniach zostało udowodnione, że imputacja poprawia szeroko pojętą poprawność wnioskowania [van Buuren, 2018; Poulos, Valle, 2018]. W ostatnich latach prym wiodą metody wielokrotnej imputacji z powodu zdolności pomiaru niepewności związanej z występującymi brakami [Murray, 2018]. Coraz częściej w literaturze pojawiają się zdania o potrzebie dokonywania porównań na drodze symulacji pomiędzy dostępnymi metodami [van Buuren 2018; Jadhav i in., 2019], a tych dla danych nominalnych jest jeszcze stosunkowo mało [Akande i in., 2017]. Niniejsza praca analizowała zdolności odtwarzania frakcji populacyjnych pojedynczych i kombinacji dwóch zmiennych przez metody wielokrotnej imputacji dla danych nominalnych. Analizowane było pokrycie przedziałów ufności, Rel.MSE oraz ASB.

Stosunkową nowością związaną ze zwiększaniem się mocy obliczeniowej komputerów i ostatnią popularnością sieci neuronowych jest wykorzystanie metod głębokiego uczenia, w tym MIDA i GAIN [Yoon i in., 2018; Gondara, Wang, 2018]. Wyniki uzyskane w niniejszym badaniu potwierdzają te z pracy Wang i in. [2022], że wymienione algorytmy gorzej radzą sobie w oddawaniu frakcji populacyjnych dla danych nominalnych od sporej części pozostałych, starszych algorytmów. Dzieje się tak m.in. dlatego, że są to metody niezaadaptowane jeszcze do wielokrotnej imputacji. Inną kwestią jest problem strojenia hiperparametrów dla metod głębokich. Nie został do tej pory ustalony schemat, dzięki któremu w problemie braków danych wypracowano by sprawną metodę ich wyboru.

Za pomocą testu Page'a został wyłoniony pewien porządek wśród metod. Najlepiej radzi sobie algorytm drzewa decyzyjnego, a najgorzej metoda oparta o wielo-

wymiarowy rozkład normalny, GAIN oraz MIDA. Lasy losowe radziły sobie gorzej od algorytmu CART. Wyniki wskazują na dominację algorytmu CART nad innymi metodami. Metoda AMELIA uzyskuje gorsze wyniki na tle metod uczenia maszynowego i tych zaliczanych do klasycznej teorii statystyki. Strategie modelowania braków: GAIN, AMELIA, MIDA, HD i LDA uzyskują medianę pokrycia niższą niż nominalne 0,95. W wielu przypadkach ranking metod zmienia się w zależności od analizowania frakcji pojedynczych i kombinacji dwóch zmiennych oraz w zależności od miary jakości. Metoda POLYREG zauważalnie traci na zwiększaniu proporcji braków danych, natomiast PMM na przejściu z jednego do dwóch wymiarów porównania. Rozkłady miar Rel.MSE i ASB różnią się wśród metod, w tym kurtozą.

Praca ma kilka oczywistych ograniczeń. Niektóre z użytych metod wymagają strojenia hiperparametrów. W rozdziale zostały użyte zalecenia z dostępnej literatury lub ustawienia dlań domyślne. Warto byłoby w kolejnych pracach zbadać wrażliwość wyników poszczególnych algorytmów w tej kwestii. Podobną analizę można by rozszerzyć na większą liczbę baz danych lub na zbiór o większej liczbie obserwacji i atrybutów. Innym typem badania własności algorytmów imputacji, które warto zaznaczyć, byłoby przeanalizowanie ich wpływu na różne algorytmy klasyfikujące.

Załączniki

A1. Liczba stopni swobody

Liczba stopni swobody dla rozkładu t dla wyprowadzonych reguł Rubina to

$(M - 1) \left(1 + \frac{1}{r_M} \right)^2$, gdzie M to liczba zbiorów rozważanych w wielokrotnej imputacji,

a $r_M = \frac{\left(1 + \frac{1}{M} \right) B_M}{\bar{U}_M}$. Wyrażenie r_M mierzy relatywny wzrost wariancji spowodowanej nieobserwowalnością.

A2. Opis zmiennych w bazie danych

Baza danych użyta w badaniu składała się z 13 zmiennych nominalnych:

- płeć (2 kategorie),
- wiek (7 kategorii),
- status związku (4 kategorie),

- poziom edukacji (5 kategorii),
- zawód (9 kategorii),
- język dominujący w domu (2 kategorie),
- klasa etniczna (5 kategorii),
- typ domu (4 kategorie),
- status zamieszkania (3 kategorie),
- ludzie poniżej 18 r.ż. w domu (4 kategorie),
- liczba osób zamieszkałych w domu (5 kategorii),
- rodzaj budżetu domowego (3 kategorie),
- czas zamieszkania w rejonie San Francisco/Oakland/San Jose (3 kategorie).

A3. Trafność

Prezentowane są kwantyle rozkładu trafności. Należy zwrócić uwagę na fakt, iż problem imputacji danych nie jest tożsamy z problemem prognozowania idiosynkratycznych wartości.

Tabela 10. Kwantyle trafności dla 10% wartości MCAR

	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,46	0,57	0,57	0,49	0,56	0,58	0,47	0,52	0,61	0,59	0,60
25%	0,46	0,58	0,58	0,50	0,57	0,59	0,47	0,53	0,62	0,60	0,61
50%	0,48	0,59	0,59	0,51	0,58	0,60	0,48	0,54	0,63	0,61	0,61
75%	0,49	0,60	0,60	0,52	0,59	0,60	0,49	0,55	0,64	0,62	0,63
90%	0,49	0,61	0,61	0,53	0,60	0,61	0,50	0,55	0,64	0,63	0,63

Źródło: opracowanie własne.

Tabela 11. Kwantyle trafności dla 30% wartości MCAR

	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,44	0,54	0,55	0,47	0,49	0,54	0,47	0,50	0,58	0,57	0,58
25%	0,45	0,54	0,55	0,47	0,50	0,55	0,47	0,51	0,58	0,58	0,58
50%	0,46	0,55	0,56	0,48	0,50	0,56	0,48	0,51	0,59	0,58	0,59
75%	0,46	0,56	0,56	0,48	0,51	0,56	0,48	0,52	0,59	0,59	0,59
90%	0,47	0,56	0,57	0,49	0,52	0,57	0,49	0,52	0,60	0,59	0,60

Źródło: opracowanie własne.

Tabela 12. Kwantyle trafności dla 50% wartości MCAR

	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,42	0,50	0,51	0,44	0,44	0,49	0,46	0,47	0,53	0,54	0,55
25%	0,43	0,50	0,51	0,44	0,44	0,49	0,46	0,48	0,53	0,55	0,55
50%	0,43	0,50	0,52	0,45	0,45	0,50	0,47	0,48	0,54	0,56	0,56
75%	0,44	0,51	0,52	0,45	0,45	0,50	0,47	0,49	0,54	0,56	0,56
90%	0,45	0,51	0,53	0,46	0,46	0,51	0,48	0,49	0,55	0,57	0,57

Źródło: opracowanie własne.

Analogiczne tabele z wynikami dla schematu amputacji danych według MAR znajdują się poniżej.

Tabela 13. Kwantyle trafności dla MAR

	AMELIA	CART	RF	PMM	HD	POLYREG	MIMCA	DPMPM	LDA	GAIN	MIDA
10%	0,46	0,57	0,58	0,51	0,56	0,57	0,48	0,51	0,62	0,60	0,61
25%	0,46	0,58	0,59	0,52	0,57	0,58	0,48	0,52	0,63	0,61	0,62
50%	0,47	0,59	0,59	0,53	0,58	0,59	0,49	0,53	0,63	0,62	0,63
75%	0,48	0,60	0,60	0,54	0,59	0,59	0,50	0,54	0,64	0,63	0,64
90%	0,49	0,60	0,61	0,54	0,60	0,60	0,51	0,55	0,65	0,64	0,64

Źródło: opracowanie własne.

Wyniki trafności w porównaniu do ASB wskazują, że GAIN i MIDA przewidyują częstsze klasy, ale nie oddają klas rzadkich.

A4. Hipotezy alternatywne w teście Page'a

Detale dotyczące porządku w hipotezie alternatywnej dla testu informującego o istnieniu porządku znajdują się w tabelach 14 i 15.

Tabela 14. Poszczególne alternatywne hipotezy w teście Page'a

Brzegowe
Rel.MSE 10% MCAR: MIDA \leq AMELIA \leq GAIN \leq HD \leq MIMCA \leq RF \leq DPMPM \leq LDA \leq PMM \leq CART \leq POLYREG
Rel.MSE 30% MCAR: MIDA \leq AMELIA \leq GAIN \leq HD \leq LDA \leq RF \leq POLYREG \leq MIMCA \leq PMM \leq DPMPM \leq CART
Rel.MSE 50% MCAR: GAIN \leq MIDA \leq AMELIA \leq POLYREG \leq LDA \leq HD \leq RF \leq PMM \leq MIMCA \leq DPMPM \leq CART

ASB 10% MCAR: MIDA ≤AMELIA ≤GAIN ≤LDA ≤HD ≤RF ≤DPMPM ≤PMM ≤POLYREG ≤MIMCA ≤CART
ASB 30% MCAR: MIDA ≤AMELIA ≤GAIN ≤RF ≤POLYREG ≤LDA ≤DPMPM ≤HD ≤PMM ≤MIMCA ≤CART
ASB 50% MCAR: MIDA ≤AMELIA ≤GAIN ≤RF ≤POLYREG ≤LDA ≤DPMPM ≤MIMCA ≤HD ≤PMM ≤CART
Dwuwymiarowe
Rel.MSE 10% MCAR: GAIN ≤MIDA ≤HD ≤AMELIA ≤LDA ≤POLYREG ≤RF ≤CART ≤MIMCA ≤PMM ≤DPMPM
Rel.MSE 30% MCAR: GAIN ≤AMELIA ≤MIDA ≤LDA ≤POLYREG ≤HD ≤RF ≤MIMCA ≤PMM ≤CART ≤DPMPM
Rel.MSE 50% MCAR: GAIN ≤AMELIA ≤MIDA ≤POLYREG ≤LDA ≤HD ≤MIMCA ≤PMM ≤RF ≤DPMPM ≤CART
ASB 10% MCAR: MIDA ≤AMELIA ≤GAIN ≤PMM ≤MIMCA ≤LDA ≤DPMPM ≤RF ≤HD ≤POLYREG ≤CART
ASB 30% MCAR: MIDA ≤AMELIA ≤GAIN ≤PMM ≤LDA ≤MIMCA ≤RF ≤HD ≤DPMPM ≤POLYREG ≤CART
ASB 50% MCAR: MIDA ≤AMELIA ≤GAIN ≤PMM ≤POLYREG ≤RF ≤HD ≤LDA ≤MIMCA ≤DPMPM ≤CART

Źródło: opracowanie własne.

Tabela 15. Poszczególne alternatywne hipotezy w teście Page'a

Brzegowe
Rel.MSE MAR: AMELIA ≤MIDA ≤GAIN ≤HD ≤DPMPM ≤RF ≤PMM ≤LDA ≤MIMCA ≤POLYREG ≤CART
ASB MAR: MIDA ≤AMELIA ≤GAIN ≤RF ≤POLYREG ≤PMM ≤DPMM ≤LDA ≤MIMCA ≤HD ≤CART
Dwuwymiarowe
Rel.MSE MAR: GAIN ≤AMELIA ≤MIDA ≤HD ≤LDA ≤POLYREG ≤PMM ≤RF ≤MIMCA ≤DPMPM ≤CART
ASB MAR: AMELIA ≤MIDA ≤GAIN ≤MIMCA ≤PMM ≤DPMPM ≤RF ≤LDA ≤POLYREG ≤CART ≤HD

A5. Hiperparametry

Hiperparametry dla wybranych metod⁴:

CART:

minbucket = 5

cp = 1*e-04

RF:

ntree = 10

⁴ <https://cran.r-project.org/web/packages/NPBayesImputeCat/NPBayesImputeCat.pdf> (dostęp: 30.06.2021).
<https://cran.r-project.org/web/packages/AMELIA/AMELIA.pdf> (dostęp: 30.06.2021).
<https://cran.r-project.org/web/packages/mice/mice.pdf> (dostęp: 30.06.2021).
<https://cran.r-project.org/web/packages/hot.deck/hot.deck.pdf> (dostęp: 30.06.2021).
https://github.com/zhenhua-wang/MissingData_DL (dostęp: 30.06.2021).

PMM:

donors = 5 ridge = $1 \cdot 10^{-5}$

AMELIA:

domyślne parametry

DPMPM

nrun ~ 10000

burn ~ 2000

k = 35

thin = 50

aalpha = 0,25

balpha = 0,25

HD

optimStep = 0,1

optimStop = 5

cutoff = 10

sdCutoff = 1

method = "best.cell"

LDA

MIMCA

ncp = 3

POLYREG

domyślne argumenty

GAIN

batch_size = 512

hint_rate = 0,13

alpha = 100

iterations = 200

MIDA

learning_rate = 0,001

hint_rate = 0,13

alpha = 100

iterations = 200

A6. Schemat MAR

Został zaprojektowany schemat braków MAR, poniżej w tabelach 16 i 17 detale dotyczące prawdopodobieństwa braków.

Tabela 16. Schemat braków według mechanizmu MAR, warunkowe prawdopodobieństwa nieobserwowalności dla kombinacji wieku i statusu związku

Status małżeński/wiek	14–24 lat	25–44 lat	45–64 lat	64+ lat
W małżeństwie/wdowa	0,15	0,5	0,3	0,6
W związku	0,15	0,3	0,5	0,6
Samotny	0,15	0,4	0,4	0,6

Źródło: opracowanie własne.

Tabela 17. Schemat braków według mechanizmu MAR, warunkowe prawdopodobieństwa nieobserwowalności dla kombinacji płci i wykształcenia

Płeć/edukacja	Przed trzecim stopniem	W trakcie lub po trzecim stopniu
Mężczyzna	0,3	0
Kobieta	0	0,3

Źródło: opracowanie własne.

Bibliografia

- Akande, O., Li, F., Reiter, J. (2017). An empirical comparison of multiple imputation methods for categorical data, *The American Statistician*, 71(2), s. 162–170.
- Akande, O., Wongkamthong, C. (2021). *A Comparative Study of Imputation Methods for Multivariate Ordinal Data*, arXiv:2010.10471.
- Anand, V., Mamidi, V. (2020). *Multiple Imputation of Missing Data in Marketing*, 2020 International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI), s. 1–6.
- Arnold, B.C., Press, S.J. (1989). Compatible Conditional Distributions, *Journal of the American Statistical Association*, 84(405), s. 152–156.
- Audigier, V., Husson, F., Josse, J. (2017). MIMCA: multiple imputation for categorical variables with multiple correspondence analysis, *Statistics and Computing*, 27(2), s. 501–518.
- Audigier, V., Niang, N. (2020). *Clustering with missing data: which equivalent for Rubin's rules?*, arxiv:2011.13694.
- Bertsimas D., Pawlowski C., Zhuo D. (2018). From Predictive Methods to Missing Data Imputation: An Optimization Approach, *The Journal of Machine Learning Research*, 18(1), s. 7133–7171.
- Brand, J. (1999). *Development, Implementation and Evaluation of Multiple Imputation Strategies for the Statistical Analysis of Incomplete Data Sets* (doctoral thesis). TNO Prevention and Health/Erasmus University Rotterdam.

- Burgette, L.F., Reiter, J.P. (2010). Multiple imputation for missing data via sequential regression trees, *American Journal of Epidemiology*, 172(9).
- van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification, *Statistical Methods in Medical Research*, 16(3), s. 219–242.
- van Buuren, S. (2018). *Flexible Imputation of Missing Data*. Boca Raton, FL: CRC Press.
- van Buuren, S., Brand, J.P.L., Groothuis-Oudshoorn, C.G.M., Rubin, D.B. (2006). Fully conditional specification in multivariate imputation, *Journal of Statistical Computation and Simulation*, 76(12), s. 1049–1064.
- van Buuren, S., Groothuis-Oudshoorn, C.G.M. (1999). *Flexible Multivariate Imputation by MICE*. PG/VGZ/99.054. Leiden: TNO Prevention and Health.
- van Buuren, S., Groothuis-Oudshoorn, K. (2011). MICE: multivariate imputation by chained equations in R, *Journal of Statistical Software*, 45(3), s. 1–67.
- CRAN, <https://cran.r-project.org/web/packages/AMELIA/AMELIA.pdf> (dostęp: 30.06.2021).
- CRAN, <https://cran.r-project.org/web/packages/hot.deck/hot.deck.pdf> (dostęp: 30.06.2021).
- CRAN, <https://cran.r-project.org/web/packages/mice/mice.pdf> (dostęp: 30.06.2021).
- CRAN, <https://cran.r-project.org/web/packages/missMDA/missMDA.pdf> (dostęp: 30.06.2021).
- CRAN, <https://cran.r-project.org/web/packages/NPBayesImputeCat/NPBayesImputeCat.pdf> (dostęp: 30.06.2021).
- Cranmer, S.J., Gill, J.M. (2013). We have to be discrete about this: a non-parametric imputation technique for missing categorical data, *British Journal of Political Science*, 43(2), s. 425–449.
- Doove, L.L., van Buuren, S., Dusseldorp, E. (2014). Recursive partitioning for missing data imputation in the presence of interaction effects, *Computational Statistics & Data Analysis*, (72), s. 92–104.
- Farhangfar, A., Kurgan, L., Dy, J. (2008). Impact of imputation of missing values on classification error for discrete data, *Pattern Recognition*, 41, s. 3692–3705.
- Gheyas, I.A., Smith, L.S. (2010). A neural network-based framework for the reconstruction of incomplete data sets, *Neurocomputing: An International Journal*, 73, s. 16–28.
- Ghorbani, S., Desmarais, M. (2017). Performance comparison of recent imputation methods for classification tasks over binary data, *Applied Artificial Intelligence*, 31, s. 1–22.
- Gondara, L., Wang, K. (2018). MIDA: Multiple Imputation Using Denoising Autoencoders. W: *Advances in Knowledge Discovery and Data Mining*, s. 260–272.
- Grochowina, D. (2015). Wpływ metod imputacji danych na skuteczność klasyfikacyjną modelu logitowego zastosowanego do prognozowania upadłości przedsiębiorstw, *Acta Universitatis Nicolai Copernici Oeconomia*, 45(2), s. 187–205.
- Harel, O., Zhou, X.H. (2007). Multiple imputation: review of theory, implementation and software, *Statistics in medicine*, 26, s. 3057–3077.
- Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data, *Journal of Statistical Software*, 45(7), 1–47.
- Hong, S., Lynn, H.S. (2020). Accuracy of random-forest-based imputation of missing data in the presence of non-normality, non-linearity, and interaction, *BMC Med Res Methodology*, 20(1), s. 199.

- Jadhav, A., Pramod, D., Ramanathan, K. (2019). Comparison of Performance of Data Imputation Methods for Numeric Dataset, *Applied Artificial Intelligence*, 33, s. 1–21.
- Lin, W.C., Tsai, C.F. (2020). Missing value imputation: a review and analysis of the literature (2006–2017), *Artificial Intelligence Review*, 53, s. 1487–1509.
- Little, R.J.A., Rubin, D.B. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley & Sons Inc.
- Luengo, J., García, S., Herrera, F. (2012). On the choice of the best imputation methods for missing values considering three groups of classification methods, *Knowledge and Information Systems*, 32, s. 77–108.
- Misztal, M. (2019). Comparison of selected multiple imputation methods for continuous variables-preliminary simulation study results, *Acta Universitatis Lodziensis. Folia Oeconomica*, 6, s. 73–98.
- Murray, J. (2018). Multiple imputation: a review of practical and theoretical findings, *Statistical Science*, 33, s. 73–96.
- van der Palm, D.W., van der Ark L.A., Vermunt J.K. (2016). A comparison of incomplete-data methods for categorical data, *Statistical Methods in Medical Research*, 25(2), s. 754–774.
- Pedersen, A.B., Mikkelsen, E.M., Cronin-Fenton, D., Kristensen, N.R., Pham, T.M., Pedersen, L., Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research, *Clinical Epidemiology*, 9, s. 157–166.
- Pokropek, A. (2018). Wybrane statystyczne metody radzenia sobie z brakami danych, *Polskie Forum Psychologiczne*, XXIII (2), s. 291–310.
- Poulos, J., Valle, R. (2018). Missing data imputation for supervised learning, *Applied Artificial Intelligence*, 32, s. 186–196.
- Raghunathan, T., Lepkowski, J., Hoewyk, J., Solenberger, P. (2000). A multivariate technique for multiply imputing missing values using a sequence of regression models, *Survey Methodology*, 27(1). s. 85–95.
- Rubin, D. (1976). Inference and missing data, *Biometrika*, 63(3), s. 581–592.
- Rubin, D. (1987). *Multiple imputation for Nonresponse in Surveys*. New York: John Wiley & Sons Inc.
- Rubin, D. (1996). Multiple imputation after 18 years, *Journal of the American Statistical Association*, 91(434), s. 473–489.
- Salfrán, D. (2018). *Multiple Imputation for Complex Data Sets* (doctoral thesis). Universitat Hamburg.
- Salfrán, D., Jordan, P., Spiess, M. (2016). *Missing data: On criteria to evaluate imputation methods*, 4, University of Hamburg.
- Schouten, R.M., Lugtig, P., Vink G. (2018). Generating missing values for simulation purposes: a multivariate amputation procedure, *Journal of Statistical Computation and Simulation*, 88(15), s. 2909–2930.
- Stavseth, M.R., Clausen, T., Røislien, J. (2019). How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data, *SAGE Open Medicine*, 7, s. 34–52.

- Vidotto, D., Vermunt, J.K., Kaptein, M.C. (2015). Multiple imputation of missing categorical data using latent class models: state of art, *Psychological Test and Assessment Modeling*, 57(4), s. 542–576.
- Vink, G. (2016). *Towards a standardized evaluation of multiple imputation routine*, <https://www.gerkovink.com/docs/Standardized%20evaluation.pdf> (dostęp: 30.06.2021).
- Wang, Z., Akande, O., Poulos, J., Li, F. (2022). *Are Deep learning models superior for missing data imputation in large surveys? Evidence from an empirical comparison*, arXiv:2103.09316.
- Wesołowski, J., Tarczyński, J. (2016). Podstawy matematyczne technik imputacyjnych, *Wiadomości Statystyczne*, 9, s. 7–54.
- Yoon, J., Jordon, J., Schaar, M. (2018). GAIN: Missing Data Imputation using Generative Adversarial Nets Proceedings of the 35th International Conference on Machine Learning, 80, s. 5689–5698.

Simulation Comparison of the Ability to Recreate Population Fractions by Multiple Imputation Methods for Nominal Data

Abstract

This article compares several multiple imputation methods for multivariate nominal data found in the literature. The comparison is carried out by simulation, multiple sampling from the actual database and, each time, by amputating the value according to the MCAR or MAR mechanism. The multiple imputation algorithms are applied then. Two methods of machine learning, two deep learning methods and seven methods classified under the classical statistical theory were examined. The aim of the paper was to study the ability to recreate multidimensional population relationships between variables. The quality of imputation is measured by accuracy of the prediction of the idiosyncratic missing value, bias of the estimators and their variability in the estimation of multivariate population fractions as well as the coverage of the confidence intervals obtained with Rubin's rules. The analysis shows that the MICE CART algorithm dominated over other methods. The MIDA, GAIN, non-parametric hot deck, and MICE LDA algorithms did not achieve the nominal confidence interval coverage.

Keywords: multiple imputation, simulation, machine learning, missing data, nominal data, deep learning

Analiza wykorzystania rozwiązań opartych na chmurze obliczeniowej przez branżę finansową w 2020 roku na świecie

Szymon Dziubak

Streszczenie

Rozwój nowych technologii opartych na chmurze obliczeniowej daje przedsiębiorstwom możliwości korzystania z zasobów IT na żądanie. Bazując na 214 studiach przypadków, opublikowanych przez głównych dostawców chmur obliczeniowych w 2020 r., sprawdzono jakie przedsiębiorstwa z branży finansowej i w jaki sposób używają usług i produktów opartych na chmurze obliczeniowej. Uwzględniono banki, firmy ubezpieczeniowe, firmy zapewniające zdywersyfikowane usługi finansowe i firmy technologiczne oferujące technologie finansowe. Wyniki przedstawiono z perspektywy klientów wdrażających chmurę obliczeniową w organizacjach, pięciu głównych dostawców chmur obliczeniowych oraz produktów i usług w chmurze. Sprawdzono, czy dostawcy chmur obliczeniowych oferują produkty i usługi uznane jako strategiczne trendy technologiczne oraz czy przedsiębiorstwa z branży finansowej korzystają z tego rodzaju produktów i usług. Wyniki badania wskazują, że najczęściej z oferty dostawców chmur obliczeniowych korzystają względnie młode banki, firmy technologiczne i firmy zapewniające zdywersyfikowane usługi finansowe z Europy lub Ameryki Północnej. Najczęściej wybierają Microsoft Azure, korzystając przede wszystkim z aplikacji biznesowych, mocy obliczeniowych i baz danych w chmurze. W niewielkim stopniu korzystają z produktów i usług w chmurze uznanych jako strategiczne trendy technologiczne.

Słowa kluczowe: chmura obliczeniowa, branża finansowa, trendy technologiczne

1. Wstęp

Jedna z uznanych definicji chmury obliczeniowej sformułowana w ramach National Institute of Standards and Technology definiuje ją jako model świadczenia usług, który umożliwia wygodny, wszechobecny dostęp do puli dzielonych zasobów obliczeniowych na żądanie. Zasoby mogą być szybko dodawane i odejmowane w zależności od potrzeb, przy marginalnym wysiłku dotyczącym zarządzania zasobami i przy minimalnej interakcji z dostawcą usług chmurowych [Mell, Grance, 2011, s. 2]. Bai [2019] definiuje chmurę jako zbiór usług informatycznych, które cechują się wysoką dostępnością, skalowalnością, wysokim poziomem bezpieczeństwa i zwinnością w tworzeniu i wdrażaniu nowych pomysłów. Największy dostawca chmury obliczeniowej typu infrastruktura jako usługa (ang. *Infrastructure as a Service* – IaaS) i platforma jako usługa (ang. *Platform as a Service* – PaaS) według firmy badawczej Gartner [Bala, Gill, Smith, Wright, Kevin, 2020], Amazon Web Services, definiuje chmurę obliczeniową jako dostarczanie zasobów IT na żądanie za pomocą Internetu z modelem płatności *pay-as-you-go*. Zamiast kupowania i utrzymywania centrów danych i serwerów, chmura obliczeniowa umożliwia dostęp do technologii takich jak moc obliczeniowa, bazy danych czy dyski zgodnie z aktualnymi potrzebami [Amazon Web Services, 2019].

Zainteresowanie chmurami obliczeniowymi stale rośnie. Na podstawie danych Google Trends liczba osób, które wyszukiwały hasło „Cloud Computing” na całym świecie od 2006 r., czyli od daty komercjalizacji usługi EC2 przez Amazon Web Services (AWS), uznawanej za pierwsze udostępnienie nowoczesnych usług chmurowych [Mateos, Rosenberg, 2011, s. 64], do 2020 r. wzrosła piętnastokrotnie. Co ciekawe, trend jest jeszcze bardziej widoczny w Polsce. Tutaj wzrost jest niemal dwudziestopięciokrotny. Analizie poddane zostały średnie arytmetyczne za okres 12 miesięcy (okresy od stycznia 2006 do grudnia 2006 i od stycznia 2020 do grudnia 2020) w celu wyeliminowania wahań okresowych. W analizie miesięcznych wartości da się zauważyć maksimum w kwietniu 2020. Fakt ten wyjaśnia wybuch pandemii COVID-19. Organizacje, w szczególności małe i średnie przedsiębiorstwa, aby stawić czoła restrykcjom nakładanym na działalność gospodarczą, rozpoczęły korzystanie z nowych technologii [Papadopoulos, Baltas, Balta, 2020, s. 1]. Analiza wykorzystania technologii opartych na chmurze obliczeniowej w 2020 r. w branży finansowej na świecie jest więc szczególnie ciekawa, ponieważ od początku 2020 r. gospodarka światowa mierzyła się z ograniczoną mobilnością i koniecznością szybkiej adopcji technologii informatycznych. To z kolei oznacza, że firmy nie będą miały czasu na zbudowanie własnej infrastruktury centrów danych i racjonalne będzie użycie chmury obliczeniowej oferującej usługi na żądanie.

Wyższe zainteresowanie usługami chmurowymi przekłada się na wzrost ich użycia. Microsoft zanotował wzrost przychodów z usług chmurowych o 30%, w tym wzrost przychodów z usług chmurowych Azure o 64% w 2019 r. w porównaniu do 2018 [Microsoft Corp., 2020]. Wzrost przychodów AWS wyniósł 36,5% r/r w analogicznym okresie [Amazon, 2020]. Biorąc pod uwagę podobny okres analizy, wzrost przychodów zanotowały także: Google Cloud o 53% [Alphabet, 2020], Oracle Cloud o 3% [Oracle Corp., 2020], IBM Cloud o 4,5% [IBM, 2020] i Alibaba Cloud o 56,4% [Alibaba Group, 2020].

Jak wskazuje badanie firmy IDC [IDC, 2021], globalne wydatki na technologie informacyjno-komunikacyjne (ang. *Information and Communication Technologies* – ICT) spadły w 2020 r. o 0,3% w porównaniu do 2019 r. Firma IDC upatruje tutaj szansę dla chmur obliczeniowych i innych usług informatycznych w formie subskrypcyjnym. Ma to związek ze spadkiem nakładów inwestycyjnych na ICT, więc przedsiębiorstwa są zmuszone zrezygnować z tworzenia własnych rozwiązań i budowy własnych centrów danych na rzecz rozwiązań, w których płaci się tylko za faktyczne użycie zasobów. IDC [IDC, 2022] twierdzi, że w 2021 r. wydatki na rozwiązania IaaS wzrosły o 27%. Co ciekawe, szacuje wzrost wydatków na aplikacje w chmurze typu SaaS (ang. *Software as a Service*) o 39% w 2021 r.

W literaturze występuje wiele pozycji dotyczących teoretycznych i praktycznych aspektów nowych technologii i chmury obliczeniowej [por. np. Sandholm, Lee, 2014]. Opisane są również zastosowania produktów i usług w chmurze w celu rozwiązania konkretnych problemów [por. np. Rashid, Zeshan, Noor, Asif, Kumlesh, 2020]. Brakuje jednak analizy faktycznych przypadków użycia chmury obliczeniowej.

Celem pracy jest zbadanie stanu rewolucji cyfrowej związanej z chmurami obliczeniowymi w branży finansowej w 2020 r.

2. Zastosowanie chmur obliczeniowych – przegląd literatury

Zagadnienie chmur obliczeniowych nie jest nowym tematem w literaturze przedmiotu. Rosenberg [Mateos, Rosenberg, 2011, s. 26] przekonuje, że chmura obliczeniowa jest bardzo obiecującą technologią, która może zostać wykorzystana przez firmy o różnej wielkości. Termin wydania publikacji może sugerować przedawnienie tej tezy w świetle dynamicznie rozwijających się nowych technologii, jednak nowsze prace wskazują na aktualność tego twierdzenia. Bai [2019, s. 26] wysuwa odważniejszą tezę, mówiąc, że chmura zapewnia wiele możliwości, a zakres jej implementacji nie powinien być ograniczany, ponieważ nie wiadomo, z jakimi potrzebami przyjdzie się mierzyć firmie i z jakich rozwiązań finalnie

skorzysta. W dodatku chmura obliczeniowa i wszystkie usługi, jakie oferuje, zmieniają się w czasie [Rountree, Castrillo, 2014, s. 2]. Analizowana technologia ma jednak swoje korzenie w tradycyjnej infrastrukturze informatycznej. Erl [Erl, Mahmood, Puttini, 2013, s. 92] podkreśla, że technologie, które umożliwiły stworzenie chmur obliczeniowych i systemów na nich opartych, były znane przed pojawieniem się jej. Wymienia tutaj sieć i architekturę Internetu, technologię centrów danych, wirtualizację, technologię webową, konteneryzację i możliwość używania jednej usługi przez wielu użytkowników (ang. *multitenant application*). Joshua [Joshua, Ogwueleka, 2013, s. 45] dodaje do tego terminy takie jak architektura zorientowana na usługi (ang. *Service Oriented Architecture – SOA*) czy inne modele obliczeń, takie jak obliczenia sieciowe (ang. *grid computing*) czy obliczenia o wysokiej wydajności (ang. *High Performance Computing – HPC*).

Jednym z największych wyzwań, z jakimi przychodzi się mierzyć firmom chcącym zaadoptować chmurę, jest bezpieczeństwo [Bhowmik, 2017, s. 11], w tym w branży finansowej [Citigroup, 2020, s. 49; Elnagdy, Qiu, Gai, 2016, s. 1]. Dostawcy chmur obliczeniowych tworzą wiele rozwiązań mających chronić oferowane przez nich usługi. Niestety w środowisku bezpieczeństwa chmurowego istnieją jeszcze liczne obszary, które wymagają poprawy [Singh, Chatterjee, 2017, s. 112]. Coppolino [Coppolino, Mazzeo, Salvatore, Romano 2016, s. 127–129] podkreśla, że chmura jest narażona na wiele rodzajów ryzyka i określa sieć (ang. *network*), sprzęt (ang. *hardware*) i hiperwizory (ang. *hypervisors*) jako elementy najbardziej narażone na ataki. Vacca [2017] opisuje główne zabezpieczenia w chmurze, wskazując jednocześnie na ich słabe punkty. Dotson [2019] proponuje natomiast obszerny przegląd zabezpieczeń oferowanych przez dostawców chmur. Problem bezpieczeństwa związanego z chmurą obliczeniową będzie utrzymywał się wraz z rozwojem usług chmurowych, jednak dostawcy chmur obliczeniowych czynią znaczne wydatki na stworzenie nowych i poprawę istniejących zabezpieczeń oferowanych rozwiązań. Firma Gartner [Bala i in., 2020] przewiduje, że globalne wydatki na bezpieczeństwo chmurowe wzrosną o 33,3% w 2020 r. w porównaniu do 2019 r. Wartość ta jest najwyższa spośród wszystkich kategorii inwestycji w bezpieczeństwo, których łączny wzrost ma wynieść 2,4% w analogicznym okresie. W czasie przygotowywania artykułu nie były dostępne dane pozwalające skonfrontować przewidywania z rzeczywistymi wydatkami na bezpieczeństwo chmurowe.

Jauhainen [Huttunen i in., 2019, s. 22] określa cztery typowe obszary użycia Big Data w finansach i księgowości: modelowanie warunków skrajnych, predykcja fraudów, predykcja cen akcji i modelowanie ilościowe. Jak zauważa, wiele z zastosowań Big Data w finansach opiera się na chmurze obliczeniowej, ponieważ oferuje ona inteligentne narzędzia, pozwalające na szukanie przyczyn danego zdarzenia, a nie

jedynie stwierdzanie faktu jego zajścia. Podobną tezę wysuwa Yang [Yang, Huang, Li, Liu, Hu, 2016, s. 15], który uważa, że chmura obliczeniowa pozwala rozwiązać wyzwania, jakie niesie Big Data. Zwraca szczególną uwagę na obliczenia o wysokiej wydajności (ang. *High Performance Computing* – HPC). Do tematu HPC w finansach, dokładnie w bankowości inwestycyjnej, nawiązuje Dempster [Dempster, Kanninen, Keane, Vynckier, 2018]. Opisuje matematyczne problemy napotykane w finansach i ich implementacje za pomocą obliczeń o wysokiej wydajności. Na tworzenie klastrów HPC, będących połączeniem wielu usług w chmurze, pozwalają analizowani dostawcy chmur obliczeniowych. Proponują oni również szereg sposobów na integrację usług chmurowych w celu otrzymania klastra HPC. Jedną z komercyjnych implementacji HPC jest implementacja Apache Spark stworzona przez Databricks. Jest ona udostępniana przez Azure i AWS i stanowi alternatywę dla połączeń wielu usług chmurowych w celu utworzenia klastra HPC proponowanych przez głównych dostawców chmur obliczeniowych [Databricks, 2020].

Istota usług chmurowych w branży finansowej nie sprowadza się jedynie do zagadnień związanych z przechowywaniem i analizą dużych zbiorów danych. Dostawcy chmury obliczeniowej oferują również technologię blockchain. Raport firmy IDC [IDC, 2020] przewiduje, że skumulowany roczny wskaźnik wzrostu dla tej technologii w latach 2018–2023 wyniesie 57,1% na całym świecie. Obliczenia dokonane zostały z uwzględnieniem korekty związanej z pandemią COVID-19. Technologia blockchain jest oferowana w chmurze przez głównych dostawców chmur obliczeniowych w modelu Blockchain-as-a-Service [Onik, Miraz, 2019, s. 5]. O wysokiej wadze technologii blockchain w finansach przekonują A. Tapscott i D. Tapscott [2017] czy Guo i Chen [2016].

Raport firmy Deloitte [2019, s. 3–4] podkreśla również istotę sztucznej inteligencji oraz robotyzacji w branży finansowej. Tadapaneni [2020, s. 2–4] wymienia szereg zastosowań sztucznej inteligencji w finansach. Wysuwa istotną konkluzję stanowiącą o wykładniczym wzroście implementacji sztucznej inteligencji w branży finansowej. Lista zastosowań dyskutowana jest również przez Beilinga i Qiao [2016, s. 1–2], którzy zwracają szczególną uwagę na rynki finansowe. Raport firmy McKinsey [Balasubramanian, Libarikian, McElhaney, 2020] opisuje szerokie spektrum zmian, jakie może przynieść zastosowanie sztucznej inteligencji i jej pochodnych w ubezpieczeniach. Na zastosowania sztucznej inteligencji i innych technologii w ubezpieczeniach uwagę zwraca także Lehmann [2018, s. 6–7]. Bai [2019, s. 123] podkreśla, że wzrost zainteresowania sztuczną inteligencją ma swoje podstawy w rozwoju chmury obliczeniowej. Jest to spowodowane dużymi zasobami obliczeniowymi oferowanymi przez chmurę, niezbędnymi do budowy modeli AI (ang. *Artificial Intelligence*) oraz możliwością skalowania i tworzenia shardów.

Robotyzację z kolei należy odróżnić od tradycyjnej automatyzacji [Sobczak, 2019, s. 2; van der Aalst, Bichler, Heinzl, 2018, s. 2]. Tradycyjna automatyzacja znajduje swoje zastosowania w łatwych, powtarzalnych procesach skupionych w ramach jednego systemu informatycznego. Robotyzacja procesów biznesowych (ang. *Robotic Process Automation* – RPA) jest stosowana przy mniej powtarzalnych czynnościach łączących różne systemy. Systemy RPA mogą być implementowane i używane w chmurze (jak produkty i usługi firm takich jak UiPatch czy Blue Prism). Raport firmy Citi [Cole, Shah, 2020, s. 4] podkreśla, że automatyzacja zajmuje drugie miejsce na liście priorytetów dla departamentów zarządzania skarbcem. Obszar zarządzania skarbcem może być jednym z modułów systemu ERP (ang. *Enterprise Resource Planning*) lub oddzielnym systemem. Jedną z korzyści, które niesie za sobą implementacja systemu klasy ERP jest automatyzacja [Shang, Seddon, 2000, s. 8]. Szybką i tanią, w porównaniu do tradycyjnego modelu wdrożenia, metodą uzyskania systemu ERP jest implementacja systemu w chmurze obliczeniowej [Weng, Hing, 2014, s. 3–4].

Niniejszy przegląd literatury nie wyczerpuje wszystkich możliwości wykorzystania chmury obliczeniowej i usług opartych na chmurze w finansach, jednak stanowi przegląd technologii zyskujących na popularności.

3. Analiza chmury obliczeniowej w kontekście strategicznych trendów technologicznych

Dostawcy chmur obliczeniowych oferują duży zakres usług, które bazują na analizowanej technologii. Warto jednak rozpatrzyć, czy analizowani dostawcy produktów są w stanie udostępniać produkty i usługi wchodzące w skład strategicznych trendów technologicznych według firmy Gartner. Każdego roku Gartner publikuje listę 10 kluczowych technologii. Analiza dokonana została poprzez porównanie trendów z lat 2019 i 2020 z usługami, które są obecnie oferowane przez dostawców chmur obliczeniowych, uznanych za liderów, jeżeli chodzi o wizję rozwoju, tj.: Amazon Web Services, Microsoft Azure i Google Cloud Platform [Bala i in., 2020]. Tematem analizy są usługi chmurowe, więc pominięte zostały trendy: demokratyzacja, wzmocnienie człowieka, przejrzystość i identyfikowalność, rozwój oparty na sztucznej inteligencji i bezpieczeństwo sztucznej inteligencji, które nie dają się opisać w kategoriach konkretnych usług, a są pewną zmianą w ich funkcjonowaniu.

Na podstawie porównania strategicznych trendów technologicznych z produktami chmurowymi (por. tabela 1) można przyjąć, że chmura obliczeniowa oferuje nowoczesne rozwiązania informatyczne. Teza ta pozwala stwierdzić, że dostawcy

chmur obliczeniowych są w stanie dostarczyć przedsiębiorstwom z sektora finansowego szeroki zakres produktów zaspokajający większość ich potrzeb.

Tabela 1. Porównanie strategicznych trendów technologicznych z produktami chmurowymi

Nazwa trendu	AWS	Azure	GCP
Hiperautomatyzacja	AppRobotic RPA	integracja Azure Cognitive Services z UiPath Enterprise RPA Platform w chmurze Azure, Project Bonsai	integracja usług ML oferowanych przez GCP z UiPath Enterprise RPA Platform w Google Cloud
Multidoświadczanie	Amazon Sumerian	Spatial Anchors	–
Przetwarzanie brzegowe	kombinacja usług w ramach AWS	Azure IoT Edge	Google Mobile Edge Cloud, Anthos at the Edge
Rozproszona chmura	integracja usług z wielu lokalizacji za pomocą rozwiązań odnoszących się do sieci i dostarczania treści (ang. <i>Networking and Content Delivery</i>)		
Autonomiczne rzeczy	AWS RoboMaker	kombinacja usług w ramach Azure	kombinacja usług w ramach GCP
Blockchain	Amazon Managed Blockchain	Azure Blockchain Service	–
Rozszerzona analityka	–	–	–
Cyfrowe bliźniaki	kombinacja usług AWS IoT Core i Amazon Sumerian	Azure Digital Twins	kombinacja usług Cloud IoT Core, Cloud Pub/Sub i Cloud Firestore
Inteligentne przestrzenie	Amazon WorkSpaces	Windows Virtual Desktop	Google Workspaces
Obliczenia kwantowe	Amazon Braket, Amazon Quantum Solutions Lab	Azure Quantum	–

Źródło: opracowanie własne na podstawie: Garfinkel, 2018; Panetta, 2019; produktów i usług w chmurze będących w ofercie AWS, Azure i GCP.

4. Metoda badawcza

Problem badawczy skupia się na sprawdzeniu jakie przedsiębiorstwa z branży finansowej i w jaki sposób korzystają z produktów i usług opartych na chmurze obliczeniowej. Podział przedsiębiorstw z branży finansowej oparty został na The Global Industry Classification Standard [MSCI, 2018]. Wyróżnia się: banki, firmy ubezpieczeniowe i firmy zapewniające zdedyfikowane usługi finansowe (nazywane również innymi instytucjami finansowymi). Do podziału dodano firmy technologiczne zapewniające technologie finansowe (nazywane firmami technologicznymi), które definiuje się jako przedsiębiorstwa opracowujące innowacyjne rozwiązania

finansowe i ubezpieczeniowe oparte na technologiach informatycznych [Puschmann, 2016, s. 2]. Zwiększenie zakresu podziału ma na celu przeprowadzenie dokładniejszej analizy. Badanie wykonano z zamiarem poszerzenia obecnej wiedzy na temat użycia produktów opartych na chmurze obliczeniowej w branży finansowej. Wiedza ta może być użyta zarówno przez przedsiębiorstwa z branży finansowej szukające możliwości wykorzystania chmury obliczeniowej, jak i przedsiębiorstwa oferujące rozwiązania oparte na chmurze obliczeniowej do identyfikowania trendów rynkowych. Okres analizy obejmuje 2020 r. Wybór jednorocznego okresu analizy ma dostarczyć wiedzy o aktualnych zjawiskach i trendach występujących na rynku, a także ma na celu ułatwienie porównywania wyników między okresami.

Badanie polega na analizie przypadków użycia produktów opartych na chmurze obliczeniowej opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, IBM Cloud, Oracle Cloud i Alibaba Cloud. Przeanalizowane zostały przypadki użycia opublikowane w 2020 r. Dane zbierane były od października 2020 do stycznia 2021.

Zebrane dane dotyczące przedsiębiorstw rozszerzone zostały o klasyfikację przedsiębiorstwa ze względu na jego działalność (czy badane przedsiębiorstwo jest bankiem, firmą ubezpieczeniową, inną instytucją finansową czy firmą technologiczną zapewniającą technologie finansowe), kraj pochodzenia przedsiębiorstwa, obszar działalności, rok założenia firmy i liczbę zatrudnionych pracowników. Natomiast odpowiadające im produkty i usługi sklasyfikowane zostały według następujących kategorii:

- analityka (ang. *analytics*),
- blockchain,
- moce obliczeniowe (ang. *compute*),
- usługi doradcze (ang. *consulting*),
- konteneryzacja (ang. *containers*),
- bazy danych (ang. *databases*),
- narzędzia programistyczne/developerskie (ang. *developer tools*),
- DevOps,
- moce obliczeniowe dla użytkowników końcowych (ang. *end user computing*),
- Internet rzeczy (ang. *Internet of things*),
- uczenie maszynowe (ang. *machine learning*),
- zarządzanie zasobami chmurowymi (ang. *management & governance*),
- sieci i dostarczanie treści (ang. *networking & content delivery*),
- aplikacje biznesowe (ang. *business applications*),
- bezpieczeństwo, tożsamość i zgodność (ang. *security, identity and compliance*),
- zasoby dyskowe (ang. *storage*).

W celu zachowania zgodności z nazwami publikowanymi przez dostawców chmur obliczeniowych w późniejszej części pracy wykorzystane zostaną nazwy angielskojęzyczne.

5. Wyniki badań i analiza

5.1. Analiza klientów

Ze względu na zróżnicowanie przedsiębiorstw w analizowanej branży, istnieje potrzeba sprawdzenia, jakie organizacje zostały uwzględnione w badaniu i jak wygląda adopcja chmury obliczeniowej ze względu na wybrane cechy przedsiębiorstw. Wyniki zostały pokazane w tabeli 2.

Tabela 2. Przeanalizowane studia przypadków w podziale na geograficzny obszar działalności i klasyfikację przedsiębiorstwa oraz liczbę zatrudnionych pracowników

Geograficzny obszar działalności i klasyfikacja przedsiębiorstwa	Liczba przedsiębiorstw wg liczby zatrudnionych pracowników				Suma
	1-49	50-999	1000-9999	10 000+	
Krajowy	26	55	34	23	138
Bank	1	14	13	11	39
Firma technologiczna	17	23	3	0	43
Firma ubezpieczeniowa	1	2	10	5	18
Inna instytucja finansowa	7	16	8	7	38
Transnarodowy	6	20	24	23	73
Bank	0	2	4	10	16
Firma technologiczna	6	11	8	1	26
Firma ubezpieczeniowa	0	1	3	4	8
Inna instytucja finansowa	0	6	9	8	23
Wielonarodowy	0	0	0	3	3
Bank	0	0	0	2	2
Firma ubezpieczeniowa	0	0	0	1	1
Suma	32	75	58	49	214

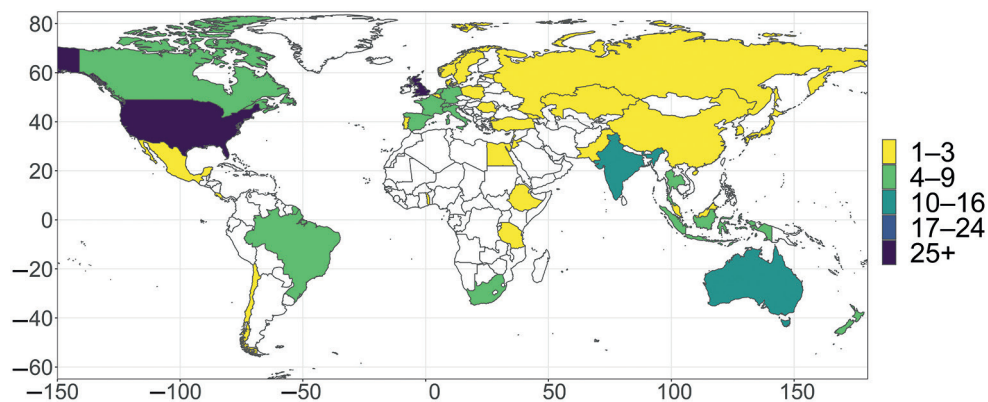
Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Jednym z wniosków płynących z analizy jest niewielka liczba studiów przypadków odnoszących się do firm ubezpieczeniowych w stosunku do pozostałych kategorii

klasyfikacji przedsiębiorstw. Do podobnych wniosków dochodzi Mueller [Mueller i in., 2015, s. 2], stwierdzając, że firmy ubezpieczeniowe wolniej adoptują narzędzia cyfrowe w porównaniu do m.in. banków. Upatruje jednak szansy na szybszą cyfryzację firm ubezpieczeniowych dzięki organizacjom typu startup. W analizowanym zbiorze danych są jedynie cztery przypadki użycia chmury obliczeniowej dotyczące organizacji typu InsurTech, gdy organizacji typu FinTech jest 51. Przytoczone liczby mogą wskazywać na niewielką cyfryzację firm ubezpieczeniowych, przejawiającą się niewielką liczbą organizacji, które skorzystały z produktów opartych na chmurze obliczeniowej. W krótkim okresie nie widać także perspektyw na zmianę tego trendu, ponieważ niewiele przedsiębiorstw typu InsurTech, opracowujących innowacyjne produkty i narzędzia dla firm ubezpieczeniowych, skorzystało z innowacyjnych produktów chmurowych.

Kraj, w którym przedsiębiorstwo stanowiące podmiot analizy ma siedzibę, może odgrywać znaczącą rolę w decyzji o użyciu publicznej chmury obliczeniowej. Rysunek 1 przedstawia mapę, która została stworzona za pomocą pogrupowania liczby analizowanych przedsiębiorstw ze względu na kraj, w którym dane przedsiębiorstwo ma siedzibę.

Rysunek 1. Podział przeanalizowanych studiów przypadków ze względu na kraj siedziby przedsiębiorstwa



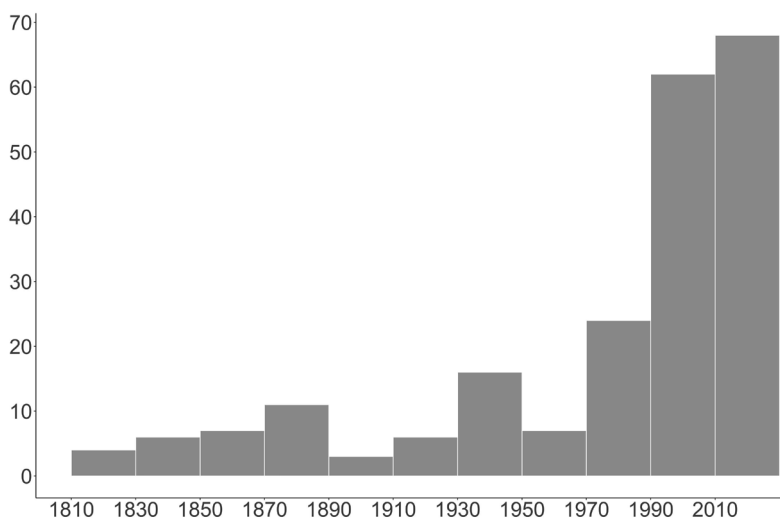
Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Większość przedsiębiorstw z branży finansowej, które wdrożyły rozwiązania oparte na chmurze obliczeniowej, ma swoją siedzibę w Europie (38%), nieco mniej w Azji (26%) i Ameryce Północnej (20%). W Australii i Oceanii liczba analizowanych przedsiębiorstw stanowi 7% całości, w Afryce 6%, a w Ameryce Południowej jedynie

3%. Wewnątrz kontynentów zróżnicowanie jest również bardzo wysokie. W Afryce aż 46% przedsiębiorstw jest z Republiki Południowej Afryki, w Ameryce Północnej Stany Zjednoczone i Kanada odpowiadają za 93% analizowanych organizacji, w Ameryce Południowej 5 na 6 przypadków użycia pochodzi z Brazylii. W Azji najbardziej innowacyjne okazały się przedsiębiorstwa z Singapuru (11 przedsiębiorstw), Indii (10 przedsiębiorstw), Indonezji (7 przedsiębiorstw), Tajlandii (7 przedsiębiorstw) i Hongkongu (6 przedsiębiorstw). W Europie liderem jest Wielka Brytania (28 przedsiębiorstw), następnie Francja (9 przedsiębiorstw) i Włochy (8 przedsiębiorstw).

Rok założenia organizacji odgrywa znaczną rolę w procesie adopcji chmury obliczeniowej. Młodsze przedsiębiorstwa są bardziej skłonne do jej adopcji (por. rysunek 2). Ma to związek z dużą liczbą firm technologicznych dostarczających technologie finansowe, spośród których 60 zostało założonych po 2000 r. Przedsiębiorstwa te, jak i pozostałe przedsiębiorstwa założone w czasie od 2000 r., mogą mieć mniejszy kapitał i poszukiwać technologii w modelu płatności za faktycznie wykorzystane usługi (jak chmura obliczeniowa), zamiast budować własną infrastrukturę. Dodatkowo ogólna liczba przedsiębiorstw założonych i istniejących w tym okresie jest większa niż liczba firm założonych w poszczególnych wcześniejszych okresach.

Rysunek 2. Rozkład analizowanych przedsiębiorstw ze względu na rok założenia



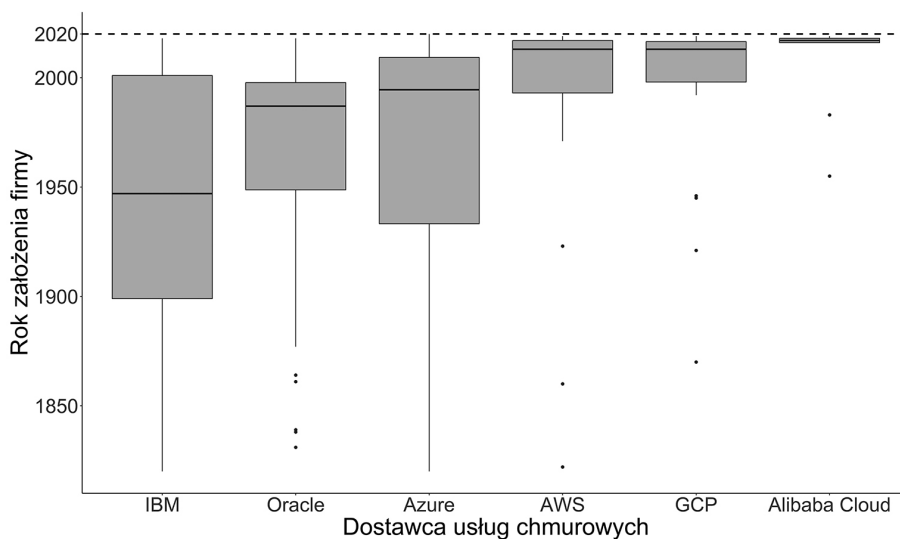
Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Udział analizowanych przedsiębiorstw w próbie założonych po 1990 r. to 62,1%, natomiast założonych w XXI w. – 46,8% organizacji.

5.2. Analiza dostawców chmur

Rysunek 3 obrazuje standardowe wykresy pudełkowe roku założenia analizowanych przedsiębiorstw z branży finansowej w zależności od dostawcy chmury obliczeniowej. Analiza wykresu pozwala stwierdzić, że rok założenia przedsiębiorstwa wpływa na decyzję o wyborze dostawcy chmury obliczeniowej.

Rysunek 3. Rok powstania firmy a dostawca usług chmurowych



Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Wyjaśnieniem tego faktu jest analiza udziału firm technologicznych w łącznej liczbie przedsiębiorstw po grupowaniu według dostawcy usług chmurowych. Firmy technologiczne zapewniające usługi finansowe to najczęściej organizacje typu Fin-Tech. Popularność tego typu organizacji zaczęła gwałtownie rosnąć dopiero w ciągu ostatnich kilku lat [Goldstein, Karolyi, Jiang, 2019, s. 2]. Dla Alibaba Cloud i Google Cloud Platform wskaźnik ten wynosi ponad 50%, gdy dla IBM, Oracle i Microsoft Azure – poniżej 25%.

Liczbę przedsiębiorstw korzystających z produktów i usług poszczególnych dostawców chmury obliczeniowej ze względu na klasyfikację przedsiębiorstwa przedstawia tabela 3. Przedsiębiorstwa z branży finansowej najczęściej korzystały z produktów i usług chmurowych firmy Microsoft Azure (68), Amazon Web Services (45) i Oracle (44). Z drugiej strony rzadziej wybierane były Google Cloud Platform (27), IBM (21) i Alibaba Cloud (9). Wszystkich analizowanych klientów Alibaba

posiada w Azji. AWS ma najwięcej swoich klientów w Azji (40%), natomiast rynek europejski okazał się najbardziej atrakcyjny dla pozostałych dostawców chmur, tj. Azure (46%), GCP (48%), IBM (71%) i Oracle (30%). Przedsiębiorstwa z Europy i Ameryki Północnej najczęściej korzystały z produktów i usług chmurowych Microsoft Azure (odpowiednio 38% i 49%), natomiast w Azji najpopularniejszy okazał się Amazon Web Services (39%).

Tabela 3. Klasyfikacja przedsiębiorstwa a dostawca usług chmurowych

Dostawca usług chmurowych	Bank	Firma technologiczna	Firma ubezpieczeniowa	Inna instytucja finansowa
Alibaba Cloud	–	7	1	1
AWS	10	20	1	14
Azure	19	14	12	23
GCP	3	15	3	6
IBM	8	5	3	5
Oracle	17	8	7	12

Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Liczbę przedsiębiorstw w zależności od geograficznego obszaru działalności i zatrudnienia w podziale według dostawców przedstawia tabela 4. Na jej podstawie zaproponowany został wskaźnik liczby przedsiębiorstw transnarodowych i wielonarodowych do ogólnej liczby przedsiębiorstw dla danego dostawcy w celu sprawdzenia z czyich usług i produktów korzystają przedsiębiorstwa finansowe prowadzące działalność zarówno krajową, jak i zagraniczną. Przyjmuje on najwyższe wartości dla Amazon Web Services (42,2%), Oracle (38,6%) i Microsoft Azure (36,8%). Oznacza to, że udział klientów będących organizacjami z branży finansowej prowadzącymi jednocześnie działalność krajową i zagraniczną jest najwyższy wśród wymienionych dostawców chmury.

Udział klientów, którzy zatrudniają powyżej 1000 pracowników, jest najwyższy dla IBM (71%), a następnie Oracle (68%) i Microsoft Azure (54%). Dla Google Cloud Platform jest to 37%, dla Amazon Web Services 31%, a dla Alibaba Cloud jedynie 11%. Należy zwrócić uwagę na podobieństwo tego rankingu z rankingiem pokazującym zależność między rokiem powstania firmy a dostawcą usług chmurowych. Klienci firmy IBM i Oracle zatrudniają najwięcej pracowników i najdłużej działają na rynku. Natomiast klienci Alibaba Cloud, GCP i AWS są organizacjami najmłodszymi i zatrudniającymi najmniej pracowników.

Tabela 4. Przeanalizowane studia przypadków w podziale na geograficzny obszar działalności i liczbę zatrudnionych pracowników a dostawca chmury obliczeniowej

Geograficzny obszar działalności i liczba zatrudnionych pracowników	Dostawca usług chmurowych					
	Alibaba Cloud	AWS	Azure	GCP	IBM	Oracle
Krajowy	7	26	43	19	16	27
1-49	1	5	6	7	3	4
50-999	5	17	18	6	3	6
1000-9999	1	3	12	3	6	9
10 000+	0	1	7	3	4	8
Transnarodowy	2	19	24	7	4	17
1-49	1	4	0	1	0	0
50-999	1	5	7	3	0	4
1000-9999	0	4	7	3	1	9
10 000+	0	6	10	0	3	4
Wielonarodowy	0	0	1	1	1	0
10 000+	0	0	1	1	1	0

Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

5.3. Analiza produktów i usług w chmurze

W kontekście strategicznych trendów technologicznych jedno z przedsiębiorstw z branży finansowej skorzystało z technologii blockchain. 10 przedsiębiorstw skorzystało z rozwiązań wchodzących w skład inteligentnych przestrzeni (miejsc pracy), w tym: pięć firm z produktu AWS, cztery z produktu GCP i jedna z produktu Azure; ponadto jedno przedsiębiorstwo skorzystało z produktu umożliwiającego stworzenie rozproszonej chmury dostarczonego przez IBM – Gateway Services (por. tabela 5). Zatem jedynie 12 spośród 577 przypadków użycia dotyczy strategicznych trendów technologicznych. Wskazuje to na ogromne wyzwania, jakie stoją przed przedsiębiorstwami z branży finansowej w transformacji cyfrowej.

Tabela 5 przedstawia ranking najczęściej używanych kategorii produktów ze wskazaniem liczebności dla poszczególnych dostawców. Liderami w dostarczaniu rozwiązań z kategorii *Business Applications* jest Microsoft Azure i Oracle. Najpopularniejszą usługą w tej kategorii jest Microsoft Teams (20 przypadków użycia), Oracle Cloud HCM (8) i Microsoft Intune (7). W kategorii *Compute* liderem jest AWS. W tej dziedzinie znalazło się 14 przypadków użycia przetwarzania bezserwerowego (ang. *serverless computing*). Liderem w tej podkategorii jest Google

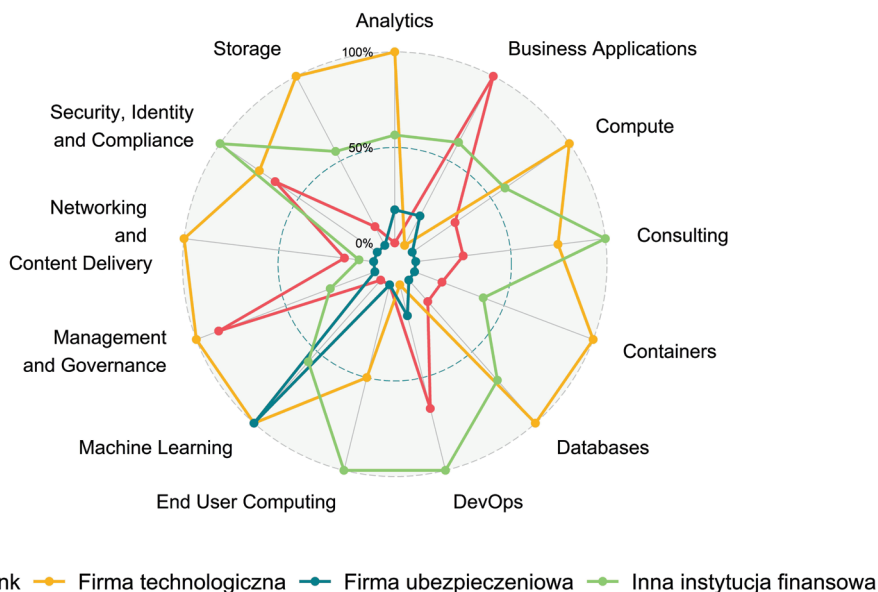
Cloud Platform z siedmioma przypadkami użycia. Liderem w kategorii *Databases* jest Microsoft Azure. Z nierelacyjnych baz danych zaklasyfikowanych pod kategorią *Databases* skorzystało 11 firm: sześć z produktów Azure, cztery – GCP i jedna – IBM. Najczęściej wybieranym dostawcą dla *Analytics* jest Google Cloud Platform, które to rozwiązanie jednocześnie oferuje najpopularniejszą hurtownię danych w chmurze – BigQuery (11 przypadków użycia). Azure prowadzi w kategorii *Security, Identity and Compliance* przede wszystkim dzięki oferowanej usłudze logowania wieloskładnikowego (ang. *Single Sign-On* – SSO) – Azure Active Directory (17 przypadków użycia). Najpopularniejszym dostawcą przestrzeni dyskowych w chmurze jest AWS z usługą Amazon Simple Storage Service, dla której zanotowano 13 przypadków użycia. Przedsiębiorstwa z branży finansowej najczęściej zdecydowały się na budowanie kontenerów w chmurze AWS i GCP, używając najczęściej technologii Kubernetes. Liderem w kategorii *Machine Learning* jest Azure, które oferuje popularną usługę HPC: Azure Databricks (AI) – pięć przypadków użycia. Widoczną pozycję lidera w kategorii *DevOps* ma Microsoft Azure.

Tabela 5. Kategoria produktu a dostawca chmury

Kategoria produktu	Dostawca chmury						Suma
	Alibaba Cloud	AWS	Azure	GCP	IBM	Oracle	
Business applications	1	0	42	0	2	32	77
Compute	10	20	10	13	9	11	73
Databases	8	12	19	15	1	10	65
Analytics	0	16	14	20	4	6	60
Security, Identity and Compliance	3	11	27	8	5	2	56
Storage	4	21	5	6	10	0	46
Containers	0	17	7	17	3	0	44
Machine Learning	0	9	22	7	4	0	42
Management & Governance	2	11	10	3	5	2	33
Networking & Content Delivery	8	8	8	5	2	1	32
Consulting	0	4	6	0	6	0	16
DevOps	0	0	14	1	0	0	15
End User Computing	0	5	1	4	0	0	10
Developer Tools	0	0	2	1	1	1	5
IoT	0	0	2	0	0	0	2
Blockchain	0	1	0	0	0	0	1

Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Rysunek 4. Porównanie wykorzystania produktów i usług danej kategorii ze względu na klasyfikację przedsiębiorstwa



Źródło: opracowanie własne na podstawie studiów przypadków opublikowanych przez Amazon Web Services, Microsoft Azure, Google Cloud Platform, Oracle Cloud, IBM Cloud i Alibaba Cloud w 2020 r.

Rysunek 4 pokazuje, które typy przedsiębiorstw z branży finansowej są bardziej skłonne do adopcji poszczególnych produktów i usług chmurowych. Przy analizie wykresu należy uwzględnić liczebność danych typów przedsiębiorstw. Jednym z najbardziej widocznych wniosków jest wysoki wskaźnik użycia produktów i usług z kategorii *Machine Learning* przez firmy ubezpieczeniowe, których jest najmniej w analizowanej grupie (27). Na uwagę zasługuje również znaczna liczba zaimplementowanych gotowych aplikacji biznesowych przez banki. Firmy technologiczne zapewniające usługi finansowe mają wyraźną przewagę w użyciu produktów z kategorii *Networking and Content Delivery* nad pozostałymi rodzajami przedsiębiorstw. Natomiast przedsiębiorstwa zapewniające zdywersyfikowane usługi finansowe mają największą przewagę nad pozostałymi w kategorii *End User Computing*, w skład której wchodzi produkty i usługi będące częścią strategicznego trendu technologicznego „inteligentne przestrzenie”.

Dla banków jest tylko jeden przypadek użycia produktu lub usługi wchodzącej w skład strategicznych trendów technologicznych, dla firm technologicznych – dwa, dla firm ubezpieczeniowych – jeden, a dla innych instytucji finansowych – pięć.

6. Podsumowanie

Problem badawczy w niniejszym tekście dotyczył zdiagnozowania przedsiębiorstw z branży finansowej korzystających z produktów i usług opartych na chmurze obliczeniowej oraz sposobu korzystania z nich przez te przedsiębiorstwa. W tym celu przeanalizowano studia przypadków opublikowane przez głównych dostawców chmur obliczeniowych w 2020 r. Wyniki analiz przedstawione zostały z trzech perspektyw: klienta, dostawcy chmury obliczeniowej oraz produktów i usług w chmurze.

Przeprowadzona analiza doprowadziła do sformułowania następujących wniosków:

- 1) Z usług dostawców publicznej chmury obliczeniowej korzystają przede wszystkim względnie młode banki, firmy technologiczne i firmy zapewniające zdywersyfikowane usługi finansowe prowadzące działalność krajową i transgraniczną, mające swoją siedzibę w Europie lub Ameryce Północnej.
- 2) Najpopularniejszym dostawcą chmury wśród analizowanych przedsiębiorstw jest Microsoft Azure.
- 3) Najczęściej wybierano aplikacje biznesowe, moce obliczeniowe i bazy danych w chmurze.

Konfrontacja analiz ze strategicznymi trendami technologicznymi pozwala stwierdzić, że branża finansowa nie jest pionierem w adopcji nowych technologii opartych na chmurze obliczeniowej. Zasadniczą konkluzją płynącą z przedstawionej powyżej analizy jest wniosek, iż adopcja technologii informatycznych, w szczególności chmury obliczeniowej ze wszystkimi produktami, jakie oferuje, ciągle stanowi dla przedsiębiorstw z branży finansowej duże wyzwanie. W rozdziale zawarta jest analiza wykorzystania chmur obliczeniowych w branży finansowej w 2020 r. Wybranie konkretnego roku pozwala na porównywanie analiz z innych okresów z analizami zawartymi w badaniu.

Kwestiami do przyszłych badań jest podobna analiza dla innych okresów, porównanie zmian i trendów. Ważne wydaje się również skonfrontowanie badania z inną branżą. Interesująca może okazać się także analiza jakościowa problemu, która może dostarczyć odpowiedzi na następujące pytania: 1) Dlaczego firmy branży finansowej decydują się na adopcję chmury obliczeniowej i jakie mają z tego korzyści?; 2) Jak chmura obliczeniowa przyczynia się do zmiany procesów w firmie z branży finansowej?; 3) Od czego zależy duże zróżnicowanie przeanalizowanych przedsiębiorstw w omawianym obszarze?; 4) Dlaczego firmy ubezpieczeniowe rzadziej korzystają z produktów i usług chmurowych niż pozostałe przedsiębiorstwa z branży finansowej? Wyniki analizy w niniejszym rozdziale mogą stanowić wartościowy punkt wyjścia dla wskazanych powyżej kierunków przyszłych badań.

Bibliografia

- van der Aalst, W.M.P., Bichler, M., Heinzl, A. (2018). Robotic Process Automation, *Business & Information Systems Engineering*, 60(4), s. 269–272.
- Alibaba Group (2020). *Alibaba Group Announces December Quarter 2019 Results*, <https://www.alibabagroup.com/en/ir/presentations/pre200213.pdf> (dostęp: 30.11.2020).
- Alphabet (2020). *Alphabet Announces Fourth Quarter and Fiscal Year 2019 Results*, https://abc.xyz/investor/static/pdf/2019Q4_alphabet_earnings_release.pdf (dostęp: 30.11.2020).
- Amazon (2020). *Amazon.com Announces Fourth Quarter Sales up 21% to \$87.4 Billion*, <https://press.aboutamazon.com/news-releases/news-release-details/amazoncom-announces-fourth-quarter-sales-21-874-billion/> (dostęp: 30.11.2020).
- Amazon Web Services (2019). *What is cloud computing?*, <https://aws.amazon.com/what-is-cloud-computing/> (dostęp: 4.12.2020).
- Bai, H. (2019). *Zen of Cloud: Learning Cloud Computing by Examples* (ed. 2). Boca Raton: CRC Press.
- Bala, R., Gill, B., Smith, D., Wright, D., Kevin, J. (2020). *Gartner Magic Quadrant for Cloud Infrastructure and Platform Services*.
- Balasubramanian, R., Libarikian, A., McElhaney, D. (2020). *Insurance 2030 – The Impact of AI on the Future of Insurance*. New York: Digital McKinsey & Company.
- Beling, P., Qiao, O. (2016). Decision analytics and machine learning in economic and financial systems, *Environment Systems and Decisions*, 2016(36), 109–113.
- Bhowmik S. (2017). *Cloud Computing*. Cambridge: Cambridge University Press.
- Citigroup (2020). *Annual Report 2019*, <https://www.citigroup.com/citi/investor/quarterly/2021/annual-report/> (dostęp: 4.12.2020).
- Cole, D., Shah, A. (2020). *Treasury Digitalization: Market Perspectives*.
- Coppolino, L., Mazzeo, G., Salvatore, D., Romano, L. (2016). Cloud Security: Emerging Threats and Current Solutions, *Computers & Electrical Engineering*, 59, s. 126–140.
- Databricks (2020). *Azure Databricks*, <https://databricks.com/product/azure> (dostęp: 6.12.2020).
- Databricks (2020). *Databricks Unified Data Analytics on AWS*, <https://databricks.com/aws> (dostęp: 6.12.2020).
- Deloitte (2019). *Disruptive digital technologies in the financial services industry*, <https://www2.deloitte.com/us/en/pages/financial-services/articles/disruptive-digital-technologies-in-the-financial-services-industry.html> (dostęp: 6.12.2020).
- Dempster, M.A.H., Kannianen, J., Keane, J., Vynckier, E. (2018). *High Performance Computing in Finance: Problems, Methods and Solutions*. Boca Raton: CRC Press.
- Dotson, C. (2019). *Practical Cloud Security: A Guide for Secure Design and Deployment*. O'Reilly.
- Elnagdy, S.A., Qiu, M., Gai K. (Eds.). (2016). *Understanding Taxonomy of Cyber Risks for Cyber-security Insurance of Financial Industry in Cloud Computing*. IEEE.
- Erl, T., Mahmood, Z., Puttini, R. (2013). *Cloud Computing: Concepts, Technology & Architecture*. Prentice Hall.

- Garfinkel, J. (2018). *Gartner Top 10 Strategic Technology Trends for 2019*. Orlando, Fla, <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/> (dostęp: 14.12.2020).
- Goldstein, I., Karolyi, A., Jiang, W. (2019). To FinTech and beyond, *The Review of Financial Studies*, 32, s. 1647–1661.
- Guo, Y., Chen, L. (2016). Blockchain application and outlook in the banking industry, *Financial Innovation*, 2(1).
- Huttunen, J., Jauhiainen, J., Lehti, L., Nylund, A., Martikainen M., Lehner O.M. (2019). Big data, Cloud Computing and data science applications in finance and accounting, *ACRN Journal of Finance and Risk*, 8, s. 16–30.
- IBM (2020). *IBM Annual Report 2019*, https://www.ibm.com/investor/att/pdf/IBM_Annual_Report_2020.pdf (dostęp: 30.11.2020).
- IDC (2020). *Worldwide Forecast for Share of Cloud Spending by Category 2021*, <https://www.idc.com/promo/customerinsights?tab=topic-one&modal=tile-PublicCloud> (dostęp: 4.12.2020).
- IDC (2021). *IDC – Global ICT Spending*, <https://www.idc.com/promo/global-ict-spending/forecast> (dostęp: 22.03.2022).
- IDC (2022). *IDC Reports Worldwide Blockchain Spending to Slow Down to US\$ 4.3 Billion in 2020*, <https://www.idc.com/getdoc.jsp?containerId=prAP46625520> (dostęp: 22.03.2022).
- Joshua, A., Ogwueleka, F. (2013). Cloud Computing with Related Enabling Technologies, *International Journal of Cloud Computing and Services Science*, s. 40–49.
- Lehmann, M. (2018). The Impact of Digitalization on the Insurance Value Chain and the Insurability of Risks, *Geneva Papers on Risk and Insurance – Issues and Practice*, 43, s. 359–396.
- Mateos, A., Rosenberg, J. (2011). *The Cloud at Your Service*. Manning Publications Co.
- Mell, P., Grance, T. (2011). *The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology*, National Institute of Standards and Technology Special Publication.
- Microsoft Corp. (2020). *Microsoft Commercial Strength Powers Second Quarter Results*, <https://news.microsoft.com/2020/01/29/microsoft-commercial-strength-powers-second-quarter-results-2/> (dostęp: 30.11.2020).
- MSCI (2018). *The Global Industry Classification Standard*, <https://www.msci.com/gics> (dostęp: 14.12.2020).
- Mueller, F., Naujoks, H., Singh, H., Schwarz, G., Schwedel, A., Thomson, K. (2015). *Global Digital Insurance Benchmarking Report 2015: Pathways to success in a digital world*.
- Onik, M.M., Miraz, M. (Eds.). 2019. *Performance Analytical Comparison of Blockchain-as-a-Service (BaaS) Platforms*. Springer.
- Oracle Corp. (2020). *Q4 Fiscal 2019 Earnings*, <https://investor.oracle.com/investor-news/news-details/2020/Oracle-Announces-Fiscal-2020-Fourth-Quarter-and-Fiscal-Full-Year-Financial-Results/default.aspx> (dostęp: 30.11.2020).
- Panetta, K. (2019). *Gartner Top 10 Strategic Technology Trends for 2020*, <https://www.gartner.com/smarterwithgartner/gartner-top-10-strategic-technology-trends-for-2020/> (dostęp: 14.12.2020).

- Papadopoulos, T., Baltas, K., Balta, M. (2020). The use of digital technologies by small and medium enterprises during COVID-19: Implications for theory and practice, *International Journal of Information Management*, 55.
- Puschmann, T. (2016). Fintech, *Business & Information Systems Engineering*, 59, s. 69–76.
- Rashid, N., Zeshan, A., Noor, S., Asif, L., Kumlesh, K. (2020). Cloud Computing Applications: A Review, *EAI Endorsed Transactions on Cloud Systems*, 6(17).
- Rountree, D., Castrillo, I. (2014). *The Basics of Cloud Computing: Understanding the Fundamentals of Cloud Computing in Theory and Practice*. Waltham: Elsevier.
- Sandholm, T., Lee, D. (2014). Notes on Cloud computing principles. *Journal of Cloud Computing*, 3(21).
- Shang, S., Seddon, P. (2000). *A Comprehensive Framework for Classifying the Benefits of ERP Systems*.
- Singh, A., Chatterjee, K. (2017). Cloud security issues and challenges: A survey. *Journal of Network and Computer Applications*, 79, s. 88–115.
- Sobczak, A. (2019). Developing a robotic process automation management model, *Business Informatics*, 2(52), s. 85–100.
- Tadapaneni, N. (2020). Artificial Intelligence in Finance and Investments, *International Journal of Innovative Research in Science, Engineering and Technology*, 5, 2792–2795.
- Tapscott, A., Tapscott, D. (2017). How Blockchain Is Changing Finance, *Harvard Business Review*, 2017.
- Vacca, J. (2017). *Cloud Computing Security: Foundations and Challenges*. CRC Press.
- Weng, F., Hing, M.-C. (2014). Competition and Challenge on Adopting Cloud ERP, *International Journal of Innovation, Management and Technology*, 5, 309–313.
- Yang, C., Huang, Q., Li, Z., Liu, K., Hu, F. (2016). Big Data and cloud computing: innovation opportunities and challenges, *International Journal of Digital Earth*, s. 13–53.

Analysis of Solutions Based on Cloud Computing as Used by the Financial Sector in the World in 2020

Abstract

The development of new cloud-based technologies enables companies to make use of IT resources on demand. Based on 214 case studies published by major cloud computing providers in 2020, the present study shows which financial sector companies make use of cloud computing products and services and in which way. It accounts for banks, insurance companies, diversified financials and financial technology companies. The results are presented from the perspective of clients, cloud providers and cloud-based products and services. The article checks whether cloud providers offer products and services recognised as strategic technology trends and whether financial sector companies use this kind of products and services. The research results indicate

that relatively new banks, financial technology companies and diversified financials most often use cloud, all of them based in Europe and North America. They most often choose Microsoft Azure, primarily using business applications, computing services and databases in the cloud. They are reluctant to use cloud-based products and services recognised as strategic technology trends.

Keywords: cloud computing, financial sector, technology trends

Brand Loyalty to Popular Music Group Brands. The Role of Brand Attachment and Brand Community Relationships

Martyna Fira

Abstract

This study extends the concept of brand attachment to the context of cultural brands, specifically brands of popular music groups. The author collates the idea of music group brand with branding theories and investigates the impact of attachment to a music group brand on music group brand loyalty. Brand community identification, brand community commitment and perceived community-brand similarity, interpreted in the context of music fan communities, were considered as factors of music brand attachment. The data collected in a survey among members of international fan communities of three popular music groups were used to test the formulated hypotheses. The results show that attachment to a music group brand is positively related to brand loyalty and fully mediates the relationship between brand community identification and brand loyalty, as well as the relationship between the perceived community-brand similarity and brand loyalty. The brand community identification is also shown to positively affect brand community commitment among music fan communities. The outcomes of the study were discussed in relation to their implications for theory and practice.

Keywords: brand loyalty, brand attachment, brand community, artist brand, music marketing.

1. Introduction

Brands are an inherent part of contemporary economies and societies. Looking from a consumer's perspective, a strong brand can facilitate the choice between a variety of goods and services, reduce the perceived risk by acting as a proof of quality and help consumers create their desired identity [Voeth, Herbst, 2008, p. 74]. Branding is applied not only to consumer goods or services and corporate entities but also to places [Hankinson, 2004], events [e.g. Leenders, 2010; Čáslavová, Petrácková,

2011; Walsh, Clavio, Lovell, Blaszk, 2013], people [Thomson, 2006]; non-profit organisations [Stride, Lee, 2007] and public institutions [Gromark, Melin, 2013] as well as artists [Schroeder, 2005] and cultural products [O'Reilly, 2005; O'Reilly, Kerrigan, 2013; Kohli, Melewar, Yen, 2014; Baumgarth, O'Reilly, 2014].

Cultural products, such as books, films, TV series, podcasts, magazines, newspapers, video games, fashion, crafts, architecture, performing arts, photography and music [UNESCO, 2009, p. 24] are categorised as experiential goods [Larce-neux, 2001, p. 3]. This means that symbolic, hedonic and aesthetic aspects prevail over functional aspects of these goods in the act of consumption, in line with Holbrook and Hirschman's [1982] framework. The definition of cultural products by UNESCO [2009, p. 87] also highlights their symbolic nature, stating that such products "convey ideas, symbols and ways of life".

The brand attachment construct, defined by Park, MacInnis, Priester, Eisingerich and Iacobucci [2010, p. 4] as "the strength of the bond connecting the brand with the self", is particularly relevant for experiential goods, as both brand attachment and experiential consumption convey emotional aspects of interactions with brands. Although this construct was examined with regard to hotel brands [Li, Lu, Bogicevic, Bujisic, 2019; Peng, Chen, 2019; Liu, Kou, Guan, Hu, Pu, 2020], restaurants [Bahri-Ammari, Van Niekerk, Khelil, Chtioui, 2016], luxury fashion [Kaufmann, Petrovici, Gonçalves Filho, Ayres, 2016] and human brands of individual celebrities [Thomson, 2006; Huang, Lin, Phau, 2015], there is scarce research on brand attachment, its factors and impact on brand loyalty in the context of cultural products, specifically music. Such research concerning music brands has a solid justification, as the global recorded music market has been growing consecutively for six years [IFPI, 2021]. Eighty-nine percent of Internet users aged 16–64 across the world use music streaming services and fifty-four percent admit that they love music [IFPI, 2019, p. 7]. The role of interactions with brand communities (which – in case of brands of music artists – can be described as fan communities and can largely affect the attitude to an artist) in shaping attachment to music brands, as brands of music groups are referred to by Obiegbu, Larsen, Ellis and O'Reilly [2019], is still understudied. Previous studies setting music artists in the theoretical marketing framework examined the general marketing principles for popular music and live music [O'Reilly, 2007; O'Reilly, Larsen, Kubacki, 2016], experiential aspects of brand loyalty in a digital music fandom [Obiegbu et al., 2019], criticism towards a music brand in an online fan community [Obiegbu, Larsen, Ellis, 2019], human branding [Huang et al., 2015], effect of social media interactions with a music artist brand on music sales [Saboo, Kumar, Ramani, 2016], brand community concept applied to music fan communities [Schau, Muñiz, 2007] and nostalgia, brand

trust and brand attachment in live music performance sector [Hidayati, Isabela, Yuwananda, Gunadi, 2021]. However, the brand attachment concept has not been studied yet in the context of brands of music groups (not solo artists); there are also few quantitative studies connecting music brand attachment and loyalty to music consumers' relationships with brand (fan) communities. This study addresses the gap by examining the relationship between brand loyalty, brand attachment and brand community-related constructs – namely brand community commitment, brand community identification and perceived community-brand similarity – with regard to music group brands.

This paper expands upon the branding aspects of music artists and reviews the literature on brand attachment, brand communities and possible relations between them in the context of music group brands. The developed hypotheses are tested in a survey conducted in May – June 2021 on a sample of active members of international music fan communities (participants of online forums and social media groups dedicated to three different popular music groups: U2, Coldplay and Florence + The Machine). These particular music groups have been chosen as they are well-known artists who have engaged and diverse fan communities, gathering people of different ages, professions and regions of the world. The paper is concluded with theoretical and managerial implications of the study as well as possible directions for future research.

2. Theoretical background and hypothesis development

2.1. Brand role in the music industry: bands as brands

The music market has changed enormously due to easier access to the Internet and the expansion of streaming services. However, the very idea remains unchanged: being a commercially successful music artist is not only about the talent and inspiration, but also about the hard work of many professionals. These professionals include the artists themselves, songwriters, lyricists, music producers, designers, photographers, A&R (artist & repertoire) departments, salespeople and, finally, marketers [IFPI, 2016, pp. 2–3]. Record companies work in the same way as other businesses, aiming at generating profit. In order to make released music profitable, record labels do not hesitate to invest in marketing. According to IFPI [2016, p. 6], marketing and promotion are the biggest item of spending on them. On the other hand, the Internet offers music artists more independence from record companies [Davis, Delaney, Kettler, 2011, p. 108], giving them the opportunity to present and

promote their work via social media and streaming platforms. Yet, if a music artist – either a world-known star with a record deal or an independent young band – would like to reach a certain audience and enjoy a commercial success in addition to achieving artistic goals, branding can be a useful marketing technique. In the end, “having a band is great, but it’s the brand that pays the bills,” as stated by Strauss [2005, p. 87]. Popular music, which is also a commodity to be marketed or consumed, creates value and is subject to market forces [Frith, 2005, p. 15], and musicians, like any other artists, can be seen as brand managers [Schroeder, 2005].

Let us consider the famous logo of The Rolling Stones, rated as the most iconic design of all time [The Independent, 2018] or the characteristic logos of AC/DC, Iron Maiden and Metallica, which have also earned an important place in the pop culture iconography. They fit the basic definition of a brand: “a name, term, design, symbol or any other feature that identifies one seller’s goods or service as distinct from those of other sellers” [American Marketing Association, 2021]. In the case of brands of music artists, “seller” means simply the artists themselves, and the phrase “goods or service” refers to recorded music. When consumers see the tongue and lips logo on a music album cover, they immediately know that this is The Rolling Stones’ recording. A recognisable brand is definitely an asset that distinguishes one music artist from others, but its role is more compound. The brand concept has evolved over the years, and the application of more complex brand definitions in the context of music group brands provides a much richer perspective.

The focus in academic discussion on brand management has shifted recently from the positivist to constructivist paradigm [Heding, Knudtzen, Bjerre, 2016, p. 21–22]. Brand management approaches evolved from company or brand owner-oriented, to consumer-oriented, to culture- and context-oriented. The latter approach supports the notion that brands are not shaped solely by their owners, they are co-constructed by multiple “authors”: companies (brand owners themselves), popular culture, customers and influencers [Holt, 2003], as well as brand community members [Muñiz, O’Guinn, 2001] and other stakeholders [Preece, Kerrigan, 2015]. In other words, brands are subjects of social construction, apart from managerial construction [O’Reilly, 2005], through meaning-making processes [McCracken, 1986]. Such conclusions lay ground for the concept of cultural branding [Holt, 2004; O’Reilly, 2005; Schroeder, Salzer-Mörling, 2006], focused on the importance of brands in the construction of consumers’ identities and defining brands as “cultural artefacts moving through history” [Holt, 2004, p. 215]. This conceptualisation corresponds with the nature of brands of music artists: they are undoubtedly rooted in culture, their value is co-created by multiple stakeholders (artists, music consumers, fan communities, record companies, critics, media) and

they play a role in the creation of music consumers' identities through their creative work [Frith, 1996; DeNora, 1999; Tarrant, North, Hargreaves, 2002].

Stern [2006] argues on the basis of brand literature review that brands are multi-dimensional constructs that can be assessed in terms of four categories: nature, function, locus and valence. Each category can be described with opposing terms: the brand nature can be either literal or metaphoric, function may be an entity or a process, locus may refer to the mind or world, and valence may be positive or negative. Music group brands, seen as symbols of certain lifestyles or social groups [e.g. Lewis, 1992], consumed in a symbolic way [Larsen, Lawson, Todd, 2012], co-created over time by a wide audience that gives them particular meanings, creating multiple associations in music consumers' minds and having generally positive connotations due to being a part of culture and responding to consumers' needs for aesthetics, tap into the metaphoric-process-mind-positive category of brands.

2.2. Brand attachment, brand communities and their relation to brand loyalty

The marketing theory indicates that consumers are able to develop emotional attachment to brands in a similar way as they develop relationships with people [e.g. Fournier, 1998; Thomson, MacInnis, Park, 2005]. Brand attachment is defined as the strength of the bond connecting the consumer with the brand [Park et al., 2010]. This bond is represented by a rich and accessible memory network that involves thoughts and feelings about the brand in relation to the self. The authors describe brand attachment as a two-dimensional construct, comprising brand-self connection (cognitive and emotional connection between the brand and its user) and brand prominence (ease and frequency of evoking brand-related thoughts and feelings). The constructs resemble Belk's [1988] concept of self-extension, particularly in terms of brands which are parts of consumers' identities; however, brand attachment expands upon the idea of an extended self by adding the dimension of brand prominence. The brand attachment theory can be viewed in relation to the psychology-based attachment theory [Ainsworth, Bowlby, 1991], which suggests that strong emotional attachment to somebody results in higher commitment to them and higher willingness to make sacrifices for them, as cited by Thomson et al. [2005]. Likewise, strong brand attachment may have a positive effect on the constructs that indicate commitment to a brand or willingness to devote resources (i.e. money and time) to a brand, e.g. brand loyalty [Thomson et al., 2005; Park et al., 2010]. Brand loyalty, described by Oliver [1999, p. 34] as a deeply rooted commitment to rebuy or repatronise a preferred brand consistently over time, despite situational factors

and competitive marketing efforts that may cause switching behaviours, involves two phases: (1) attitudinal loyalty, which leads to (2) action (or behavioural) loyalty, understood as an act of purchase. However, behavioural loyalty might not be a fully appropriate measure of loyalty in the context of brands of music groups, as purchasing a music album of a band or a ticket to their concert is not necessarily an indicator of loyalty – it might be as well a purchase made out of curiosity or as a status symbol [Obiegbu, Larsen, Ellis, 2020, p. 3]. Hence, the focus in this study was put on attitudinal loyalty, consisting of a cognitive component (preference of a brand over its alternatives), affective component (cognition and affect for a brand) and conative component (behavioural intention to repurchase a brand, which, in the case of experiential goods, could be described as an intention to repeat interactions and experiences with a brand) [Oliver, 1999; Back, Parks, 2003]. Attitudinal brand loyalty indeed highlights the aspect of commitment and willingness for resource engagement, which can be potentially predicted by brand attachment. Thus, the following hypothesis was formulated:

H1. Attachment to a brand of a music group has a positive effect on attitudinal loyalty to this brand.

One of the first studies which contributed to the constitution of the brand attachment phenomenon were studies on brand communities and subcultures of consumption [Schouten, McAlexander, 1995]. The idea of brand communities was expanded by Muñoz and O’Guinn [2001, p. 412], who defined them as “specialized, non-geographically bound community, based on a structured set of social relations among admirers of a brand”. Members of brand communities feel connected not only to the brand but also to each other; they distinguish “true” members of the community from the others and define themselves in relation to brands that they do not consume, as opposed to their favourite brand. They have certain community rituals and traditions, celebrate the brand history, share brand stories among themselves, feel a sense of duty to the community and assist other community members in the use of the brand. This description fits the characteristics of music group brand communities or fandoms, as such communities are generally called in the cultural sectors [O’Reilly, Larsen, Kubacki, 2013]. Their members often call themselves “a family”, have their own definition of a “true fan” (e.g. a person who has spent much time as a fan of a music group or experienced some historical moments in a group’s history [Obiegbu et al., 2020]), often show their preference for one group in opposition to other music artists, have their own traditions, such as theme parties with their favourite band music, share facts from the band history and their personal experiences with the band on online forums and social media

groups. They also assist each other in the consumption of a brand, e.g. by sharing new recordings and information about the band plans for concert tours with community members or helping fellow fans get tickets for a concert. Such strong brand communities can shape brand loyalty [Muñiz, O'Guinn, 2001; Jang, Olfman, Ko, Koh, Kim, 2008; Fournier, Lee, 2009]. However, brand communities are different and operate on different levels [McAlexander, Schouten, Koenig, 2002] so the relationship between brand community participation and brand loyalty can be mediated by various mechanisms.

In terms of specific brand community-related concepts, available research indicates a positive effect of brand community identification and commitment on brand loyalty [Jang et al., 2008; Fournier, Lee, 2009]. Brand community identification, based on the social identity theory [Tajfel, Turner, 2004], according to which people, besides their self-identity, define themselves by their group membership [Mael, Ashforth, 1992, p. 104], has been proven to have a positive impact on brand community commitment among Chinese car club members [Zhou, Zhang, Su, Zhou, 2012]. It was also evidenced earlier that members who had a high level of identification with a brand community were more likely to maintain the relationship with this community by engagement in community-based discussions and interaction with other community members [Muñiz, Schau, 2005; Algesheimer, Dholakia, Herrmann, 2005]. After all, those who do not identify with a group would not be generally interested in committing their time and effort to this group. As communities gathered around music group brands are characterised by high member participation and involvement, just like communities gathered around automotive brands [Algesheimer et al., 2005, p. 20], a similar effect can be expected in their case:

H2. Brand community identification is positively related to commitment to brand community of a music group.

Brand community commitment is expressed as a feeling that the continuous relationship between an individual and the brand community they belong to is valuable [Jang et al., 2008]. Community commitment can also be demonstrated as a feeling of loss at the thought of the community no longer existing [Wasko, Faraj, 2005, p. 48] or as a concern for the fate of the community [Mathwick, Wiertz, De Ruyter, 2008, p. 839]. An active, committed participation in brand communities, including music group brand communities, may bring the brand psychologically closer to the consumer-community member [Malär, Krohmer, Hoyer, Nyffenegger, 2011, p. 44] and, therefore, a consumer (or a fan) might develop a stronger bond with the brand, making it a steady part of everyday life that often is on top of the consumer's mind. In recent studies, community commitment among online brand

communities had a positive indirect effect on brand loyalty [Munnukka, Karjaluoto, Tikkanen, 2015] and a positive effect on brand attitude [Wang, Cao, Park, 2019], which is a construct related to brand attachment that can also serve as a predictor of loyalty behaviours, according to Park et al. [2010, pp. 7–11]. Moreover, positive correlation between brand community commitment and brand attachment was confirmed in a study by Zhou et al. [2012]. Hence, it is expected that

H3. Brand community commitment is positively related to attachment to a brand of a music group.

Consumers tend to have a more positive attitude towards the brands whose personalities are similar to the consumers' personality – either their actual or ideal selves. In other words, consumers often buy certain brands or interact with them because these brands allow them to express their identities, which is known in the literature as self-congruence or self-congruity [Aaker, 1997, p. 353]. The self-congruence effect can evolve not only from a perceived similarity between the personality of a consumer and the personality of a brand itself but also from the similarity between a consumer and a typical brand user [Helgeson, Supphellen, 2004], e.g. a fellow brand community member. A study by Malär et al. [2011] indicated a positive effect of actual self-congruence on brand attachment. The similarity between a brand and its community might explain an emotional bond between a consumer-community member and the brand: if brand community members consider a brand similar to the community they belong to, they also consider the brand similar to themselves in a certain sense. Community-brand similarity might then act similarly to actual self-congruence in the studies of Malär et al. [2011] and Rabbane, Roy and Spence [2020], being positively related to brand attachment. It would also support recent findings on the so-called “self-congruity effect of music” [Greenberg, Matz, Schwartz, Fricke, 2020], suggesting that music artist whose public personas are similar to the listener's personality are preferred by the listener. The above considerations result in the following hypothesis:

H4. Perceived community-brand similarity is positively related to attachment to a music group brand.

3. Research method and results summary

3.1. Method

Two-hundred seven participants from international music fan communities of three music groups: U2, Coldplay and Florence + The Machine (49.3% females, $M_{\text{age}} = 33$, $SD = 12.3$, $n_{\text{U2}} = 121$, $n_{\text{Coldplay}} = 32$, $n_{\text{FATM}} = 54$), aged 14–67, took part in an online survey in May – June 2021. These particular music groups were chosen for the study because of their worldwide popularity, as well as diverse and vibrant fan communities. Irish rock band U2 began their career in the early 1980s. They have released fourteen studio albums so far, having experienced ups and downs all throughout their long career. In addition, the band is known even among light music listeners due to their famous leader Bono and his philanthropic activities. British band Coldplay is one of the most iconic music groups of the 2000 s with discography covering different genres, from melancholic soft rock songs to pop hits. Florence + The Machine is a British indie band active in the music industry for over ten years, recognisable for a charismatic female leader with a distinctive style.

The data were collected among two groups of respondents: international fan communities and Polish fan communities. The survey was published in two language versions – English and Polish – on both international and local Polish online forums and Facebook groups dedicated to U2, Coldplay and Florence + The Machine. The response rate was higher among Polish fandoms; hence, the majority of respondents (49.8%) were Polish, followed by the Americans (17.4%) and the British (7.7%).

Ninety-six percent of respondents were studying (learning) or/and working. There was one missing entry regarding age, which was substituted by a median of age in the total sample. All demographic characteristics of the sample are presented in Table 1.

Table 1. Sample demographics

Country of origin	Number of respondents	% of total sample
Australia	3	1.4
Austria	1	0.5
Belgium	2	1.0
Brazil	1	0.5
Canada	6	2.9
China	1	0.5
Croatia	1	0.5

cont. Table 1

Country of origin	Number of respondents	% of total sample
Czech Republic	1	0.5
Denmark	1	0.5
Dominican Republic	1	0.5
Germany	9	4.3
Greece	1	0.5
Ireland	3	1.4
Italy	4	1.9
Kenya	1	0.5
Korea	1	0.5
Mexico	1	0.5
Poland	103	49.8
Slovakia	2	1.0
Slovenia	1	0.5
Spain	3	1.4
The Netherlands	7	3.4
UK	16	7.7
Uruguay	1	0.5
USA	36	17.4
Age		
14–19	12	5.8
20–29	74	35.7
30–39	41	19.8
40–49	52	25.1
50–59	22	10.6
60–67	5	2.4
No data	1	0.5
Occupation		
Studying/learning	54	26.1
Working	134	64.7
Studying/learning and working	11	5.3
None of the above	8	3.9
Gender		
Female	102	49.3
Male	102	49.3
Other	2	1.0
I do not wish to disclose	1	0.5

Source: Author's own data.

All respondents who completed the English version of the questionnaire had at least basic knowledge of English; forty-one percent of them had advanced or fluent knowledge and fifty-one percent were native English speakers. The survey questions were adjusted to each music group brand included in the study, i.e. survey published on U2 forums included questions formulated in direct relation to U2, e.g. “U2 is my first choice when I feel like listening to music.” Participants were informed about the scientific purpose of the study and reassured about the anonymity of their responses. Some of the questions were asked in a projective way to reduce a social desirability bias, e.g. “Some people may feel close to fan communities of their favourite music artists. How is it about you?”

At the beginning of the survey, brand loyalty was measured, using the scale adopted from Back and Parks [2003]; $\alpha = .86$. Then, brand attachment was measured with the scale adopted from Park et al. [2010]; $\alpha = .89$, followed by a five-item scale measuring a perceived community-brand similarity [Zhou et al. 2012]; $\alpha = .88$. Afterwards, scales inspired by Zhou et al. [2012] were used to measure brand community commitment ($\alpha = .92$) and brand community identification ($\alpha = .87$). Each measurement used a seven-point Likert scale coded from 1 to 7. All items used in the study are featured in Table 2.

Table 2. Measurement scales used in the study

Construct	Items used	References	Reliability
Attitudinal brand loyalty*	Even if another music artist that I like offered lower prices for a live show ticket than [Band X] *** and I could only afford one concert per year, I would still go to [Band X] concert. If [Band X] releases a new music album, I intend to buy it. [Band X] is my first choice when I feel like listening to music. I love listening to [Band X]. I feel better when I listen to [Band X]’s music. I like [Band X] more than other music artists. [Band X] records music of a higher quality as compared to any other music artist. No other music artist performs better in live concerts than [Band X]. I believe [Band X]’s music is more valuable than the music from other artists.	Inspired by Back & Parks [2003]	$\alpha = .86$
Brand attachment**	Is [Band X] part of you and who you are? Do you feel personally connected to [Band X]? Are your thoughts and feelings toward [Band X] often automatic, coming to mind seemingly on their own? Do your thoughts and feelings toward [Band X] come to your mind naturally and instantly?	Adopted from Park et al. [2010]	$\alpha = .89$

cont. Table 2

Construct	Items used	References	Reliability
Perceived community-brand similarity*	If [Band X] brand (their visual identification, image and reputation, differentiation, values and ways of communication) was imagined as a person, that person would have a personality similar to personalities of [Band X] fans. There is a similarity between values of [Band X] fans and values represented by [Band X]. There is a similarity between the style of [Band X] fans and the style of [Band X]. There is a similarity between what I feel when I think about [Band X] fans and what I feel when I think about [Band X]. There is nothing similar between [Band X] fans and [Band X]. (R)	Inspired by Zhou et al. [2012]	$\alpha = .88$
Brand community commitment*	I would feel a loss if social media groups or forums dedicated to [Band X] were no longer available. I care about the future of [Band X] fan community. I feel a great deal of loyalty to [Band X] fan community. I intend to maintain the relationship I have with [Band X] fan community indefinitely. The relationships I have with other fans of [Band X] are important to me.	Adopted from Zhou et al. [2012]	$\alpha = .92$
Brand community identification*	When I think about [Band X] fan community: I usually say "we" rather than "they" I see myself as a part of this community I enjoy the successes of [Band X] fan community almost as much as I enjoy my successes If someone praises [Band X] fan community, it feels to me like a personal compliment If someone criticises [Band X] fan community, it feels to me like a personal insult I am interested in what others think about [Band X] fan community	Adopted from Zhou et al [2012]	$\alpha = .87$

* Anchored by (1 = strongly disagree and 7 = strongly agree).

** Anchored by (1 = not at all and 7 = very much).

*** Band X = U2/Coldplay/Florence + The Machine.

Source: Author's own elaboration based on Back & Parks [2003], Park et al. [2010] and Zhou et al. [2012].

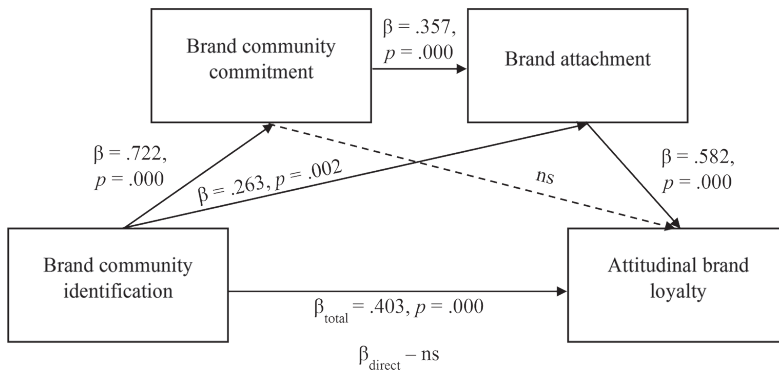
3.2. Results

Attitudinal brand loyalty served as a dependent variable in both tested mediation models. The first tested model used brand community identification as an independent variable; brand community commitment and brand attachment served as sequential mediators. The second model included brand attachment as the only mediator. Taking into account the overrepresentation of Polish respondents in the sample, the moderation effect of being a Polish citizen was examined with regard to the considered theoretical models (PROCESS, Hayes model 85 and 59). However, the interaction effect with all the considered variables was insignificant in both models (CI95%_{index of moderated mediation_Model 1} [-.021, .026; CI95%_{index of moderated media-}

tion_Model 2 [-.166,.203], meaning that the overrepresentation of Polish respondents does not affect the results.

As expected in H2, there was a positive effect of brand community identification on community commitment in the first model (PROCESS, Hayes model 6, Figure 1; $\beta = .722, p = .000$). Community identification was also positively related to brand attachment ($\beta = .263, p = .002$). Next, brand community commitment had a positive impact on brand attachment ($\beta = .357, p = .000$), providing support for H3. Brand attachment demonstrated a positive effect on attitudinal brand loyalty ($\beta = .582, p = .000$), supporting H1. The total and indirect effect of mediation were positive and statistically significant (total effect: $\beta = .403, p = .000$; indirect effect: CI95% [.207,.470]), while the direct effect of brand community identification on attitudinal brand loyalty was not significant ($p > .4$). Such results suggest a full mediation through brand community commitment and brand attachment. However, the indirect effect was also significant for brand attachment as the only mediator, with the exclusion of brand community commitment (CI95% [.043, .272]). Apart from that, the positive direct effect of brand community identification on brand attachment suggests a partial mediation through brand community commitment.

Figure 1. Brand community commitment and brand attachment as mediators between brand community identification and attitudinal loyalty to a music group brand

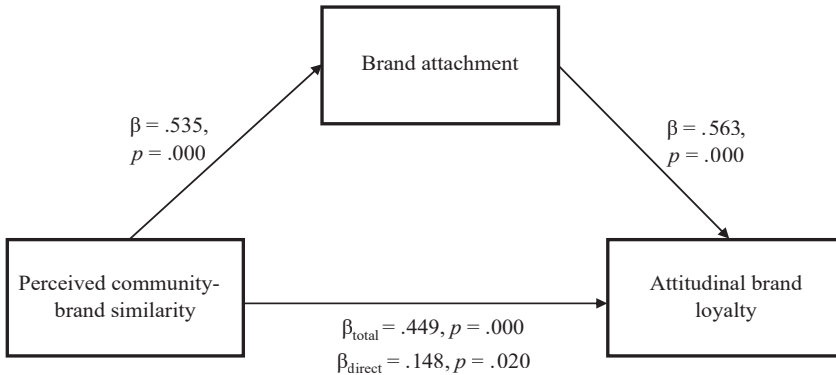


Source: Author's own data.

The second model (Figure 2) included perceived community-brand similarity as an independent variable. Brand attachment served as a mediator. Perceived community brand-similarity was positively related to brand attachment, in support of H4 ($\beta = .535, p = .000$), and brand attachment had a positive effect on attitudinal brand loyalty ($\beta = .563, p = .000$). Both total and direct effect of brand-community similarity on attitudinal brand loyalty were positive and statistically significant

(total effect: $\beta = .449, p = .000$; direct effect: $\beta = .148, p = .020$), indicating a partial mediation through brand attachment.

Figure 2. Brand attachment as a mediator between perceived community-brand similarity and attitudinal loyalty to a music group brand

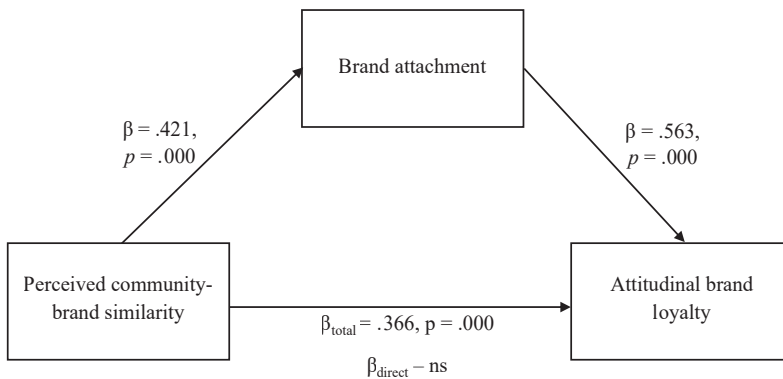
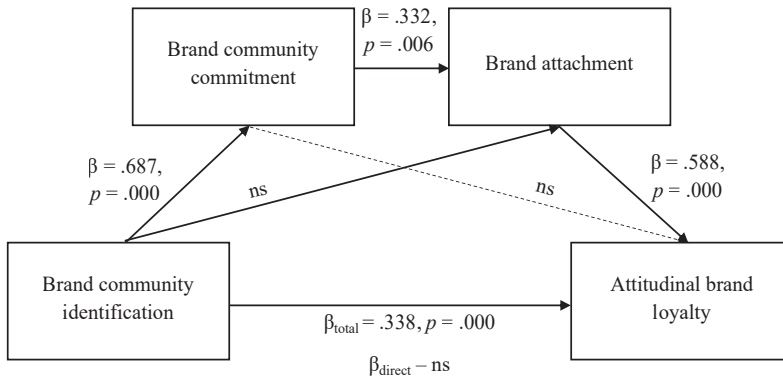


Source: Author's own data.

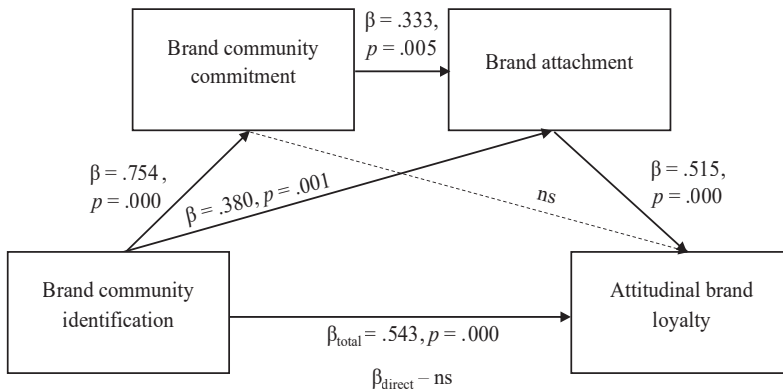
In order to control for potentially distorting effects of personal variables such as age or nationality, the above models were tested in four subgroups: respondents aged below 33 or precisely 33 (median of age in total sample) vs. respondents aged above 33 and respondents from developed countries vs. respondents from developing countries, grouped according to International Monetary Fund classification (Figure 3). The mechanisms confirmed in the total sample were also present in each analysed subgroup. However, contrary to the total sample, there was no direct effect of brand community identification on brand attachment in the sequential mediation model among younger respondents, which suggests a full mediation between community identification and brand attachment through community commitment. The indirect effect of community identification on attitudinal brand loyalty through brand attachment only, with the exclusion of community commitment, was not confirmed either. There also occurred a full mediation between perceived community-brand similarity and attitudinal brand loyalty through brand attachment in this subgroup. In the subgroup of respondents from developing countries, the mediation between brand community identification and attitudinal brand loyalty through brand community commitment and brand attachment was only partial, suggesting the role of some other factors in shaping this relationship. Unlike in the total sample, there was also a full mediation between community-brand similarity and attitudinal brand loyalty through brand attachment among respondents from developing countries.

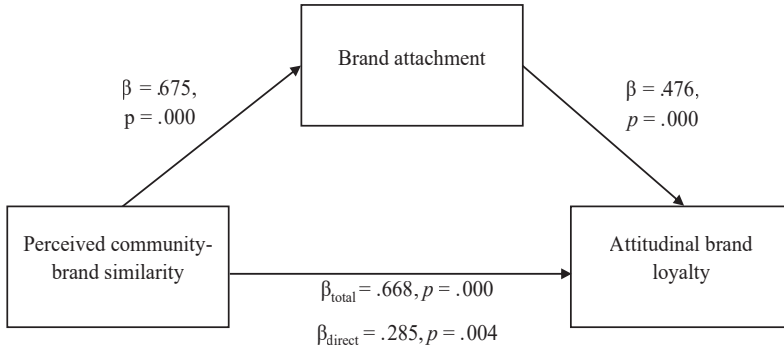
Figure 3. Intermediate mechanisms between brand community identification, perceived community-brand similarity and attitudinal loyalty to a music group brand among age groups (age below median vs. age above median) and nationality groups (developed vs. developing countries)

Age below or equal to median

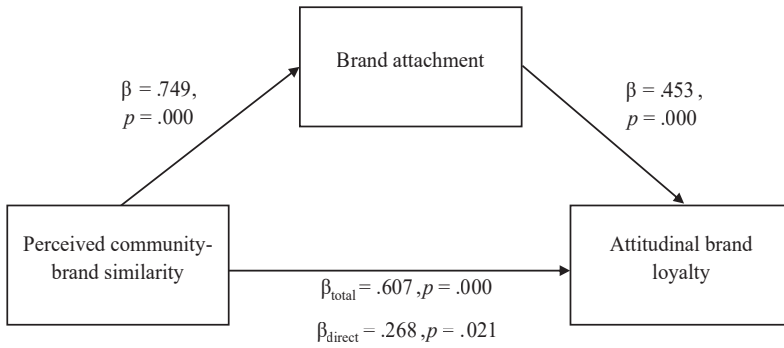
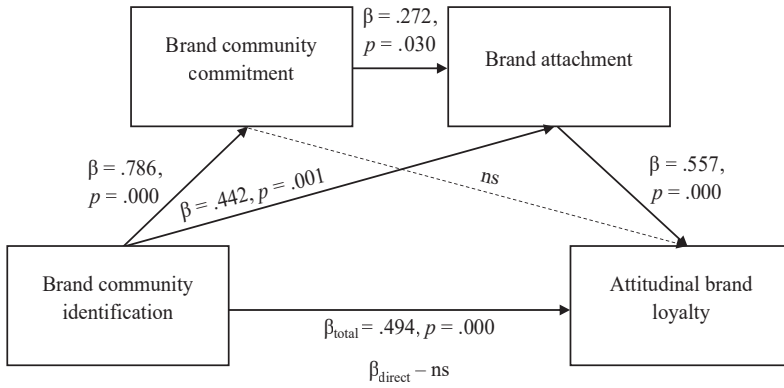


Age above median

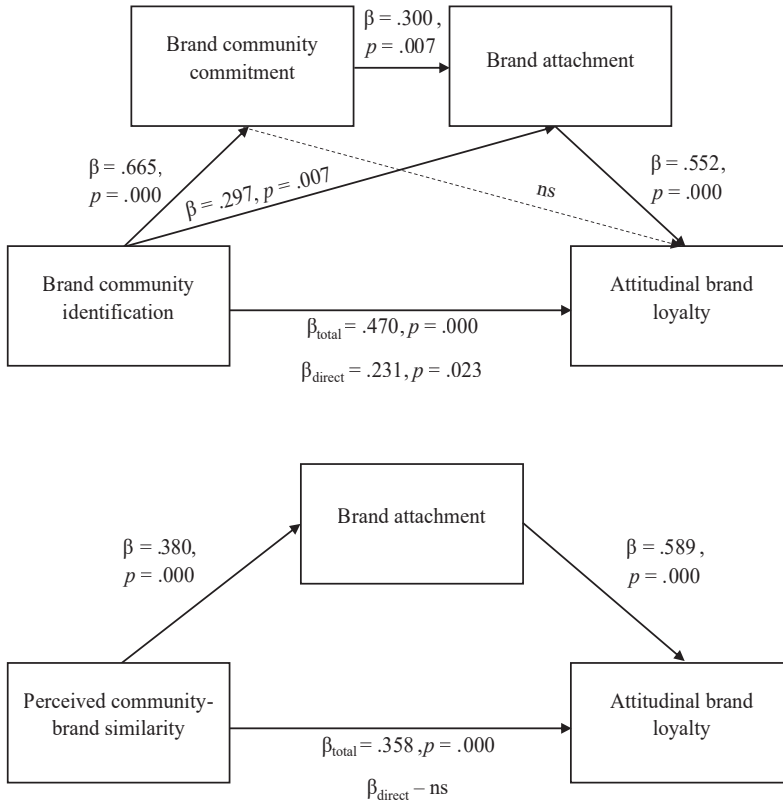




Developed countries*



Developing countries**



* Developed countries: USA, UK, Canada, Australia, Germany, The Netherlands, Belgium, Austria, Denmark, Greece, Ireland, Italy, Spain, Slovenia, Slovakia, Czech Republic, Korea.

** Developing countries: Poland, Croatia, Brazil, Mexico, Uruguay, China, Kenya, Dominican Republic.

Source: Author's own data.

4. Discussion, implications and limitations

This paper contributes to the branding theory by assessing for the first time the intermediate mechanisms between brand community-related constructs, brand attachment and brand loyalty with regard to brands of music groups in an international context. The conducted study confirms mechanisms evidenced in previous studies on examples of other categories of brands (Jang et al., 2008 – mobile phone and automotive brands; Algesheimer et al., 2005, Zhou et al., 2012 – car brands; Malär et al. 2011 – FMCGs, durable consumer goods, services and retailing brands; Munnukka et al., 2015 – home décor and footwear brands; Rabbanee et al., 2020 – sportswear brands) in a new context of experiential, symbolic brands. This

study acknowledges that brands of music groups are also subjects to already known mechanisms shaping brand loyalty among brand communities. The results show that the relation between identification with a community formed around a music group brand and attitudinal loyalty to this group is fully mediated by attachment to this group. As there is a partial mediation through brand community commitment between brand community identification and brand attachment, commitment to the community can be considered as an additional mediating variable between community identification and attitudinal brand loyalty. However, community commitment is not necessary for the mediating effect to occur among older music fans. A person who identifies with the community around the preferred music group brand is more likely to develop emotional attachment to the brand and, as a result, display a higher level of loyalty to it, even without being committed to the fan community. On the other hand, young fans have stronger need to maintain relationships with fandoms built around their idols, as it helps them build their individual as well as collective identity as fans [Lacasa, de la Fuente, García-Pernía, Cortés, 2017, p. 52]. Younger fans are also more avid Internet users, which gives them more opportunities to interact with online fan communities and, consequently, build commitment to them; as a result, brand community commitment becomes a necessary variable, along with brand attachment, in the mediation mechanism between brand community identification and brand loyalty. A partial mediation between community identification and brand attachment among older members of fan communities suggests that there might be additional mediating factors, other than community commitment. These might be particularly factors concerning older fans, such as familiarity with the community, positive image of the community or time spent as a community member. Those fans of music groups who have identified with the fan community for a long time, perceive it in a positive way, know the community well, have met numerous fellow fans and have even become friends with some of them; they might be more inclined to feel attached to the community and to the music group itself, as the group is the thing that puts community members together. With regard to the fans from developing countries, the mediation between brand community identification and brand loyalty is partial, which indicates that there might be other mediating factors, on top of community commitment and brand attachment. Perhaps, these factors may involve temporal and spatial distance. All bands included in the study are Western European, they play concerts more often in the developed than developing countries; as a consequence, fans in the developed countries have more opportunities for personal interactions with them.

Perceived community-brand similarity serves in the context of music group brands as a factor with a positive effect on brand attachment. If a member of

a music group fan community finds the community and the group alike, they may be attached and loyal to this music group. However, partial mediation between perceived community-brand similarity and attitudinal brand loyalty, particularly among older fans and fans from developed countries, suggests that there might be more factors mediating the relationship, e.g. brand attitude (music groups which have much in common with their fans might be assessed more positively, which may translate into brand loyalty [Alhaddad, 2015]) or brand trust (people may be inclined to trust brands that have a personality similar to them [Pentina, Zhang, Basmanova, 2013], and brand trust demonstrates positive impact on brand loyalty [Zehir, Şahin, Kitapçı, Özşahin, 2011; Laroche, Habibi, Richard, Sankaranarayanan, 2012; Chinomona, 2016]).

From the managerial standpoint, identification with a fan community of a music group is not sufficient to build loyalty to the music group brand. Attachment to the brand of a music group is essential to shape loyalty to the group among fan community members. One of the factors that have a positive effect on brand attachment in the context of music fan communities, especially among younger community members, is community commitment. Even though fan communities are usually member-initiated communities, managers of music bands who would like to maintain a loyal fanbase could take some actions that engage fans and make them want to remain within the community. For instance, they may give their fan community a distinct name related to the band work which would give fans a sense of unity, e.g. Lady Gaga named her fan community “*Little Monsters*,” which refers back to the name of her second album [VICE 2017]. The name even appears on the official merchandise. Community commitment might be also enhanced by music groups expressing interest in their fan communities and appreciation for them, e.g. by sharing fans’ experiences related not only to the artist and their music but also to the fan community. Furthermore, as similarity between a music group and its fan community also induces attachment and loyalty, music bands may appeal to the characteristics shared with their fans. For example, the artist-initiated contest for the best styling for Florence + The Machine fans who follow Florence Welch’s style at concerts and festivals, wearing wreaths of flowers and glitter make-up, would probably positively affect their attachment to the band.

This study is not free of limitations. Although the sample included fans of three various music groups, all of them represent popular rock-pop-indie music. The mechanisms considered in the study may differ between music groups representing different genres, including more niche genres, such as heavy metal, hip-hop, electronic music or jazz. A similar study on a larger sample could also prove the obtained results. Furthermore, the research on the same mechanisms extended

to other categories of cultural goods, e.g. book series, films or TV series, would provide a valuable perspective. Apart from that, scholars may take into consideration other constructs applicable to cultural brands, apart from brand attachment and brand loyalty; for instance, brand love, brand religiosity and brand trust could be examined in the context of cultural brands and fan communities.

References

- Ainsworth, M.S., Bowlby, J. (1991). An ethological approach to personality development, *American Psychologist*, 46(4), pp. 333–341.
- Algesheimer, R., Dholakia, U.M., Herrmann, A. (2005). The social influence of brand community: Evidence from European car clubs, *Journal of Marketing*, 69(3), pp. 19–34.
- Alhaddad, A. (2015). Perceived quality, brand image and brand trust as determinants of brand loyalty, *Journal of Research in Business and Management*, 3(4), pp. 1–8.
- American Marketing Association (2021). <https://www.ama.org/topics/branding> (accessed: 31.05.2021).
- Back, K.J., Parks, S.C. (2003). A brand loyalty model involving cognitive, affective, and conative brand loyalty and customer satisfaction, *Journal of Hospitality & Tourism Research*, 27(4), pp. 419–435.
- Bahri-Ammari, N., Van Niekerk, M., Khelil, H.B., Chtioui, J. (2016). The effects of brand attachment on behavioral loyalty in the luxury restaurant sector, *International Journal of Contemporary Hospitality Management*, 28(3), pp. 559–585.
- Baumgarth, C., O'Reilly, D. (2014). Brands in the arts and culture sector, *Arts Marketing: An International Journal*, 4 (1/2), pp. 2–9.
- Belk, R.W. (1988). Possessions and the extended self, *Journal of Consumer Research*, 15(2), pp. 139–168.
- Cáslavová, E., Petrácková, J. (2011). The brand personality of large sport events, *Kinesiology*, 1, pp. 91–106.
- Chinomona, R. (2016). Brand communication, brand image and brand trust as antecedents of brand loyalty in Gauteng Province of South Africa, *African Journal of Economic and Management Studies*, 7(1), pp. 124–139.
- Davis, D.C., Delaney, B.P., Kettler, H.M. (2011). Fandom of the Internet: Musician Communication with Fans. In: *Rock Brands: Selling Sound in a Media Saturated Culture* (pp. 101–116), E.B. Christian (Ed.). Plymouth: Lexington Books.
- DeNora, T. (1999). Music as a technology of the self, *Poetics*, 27(1), pp. 31–56.
- Fournier, S. (1998). Consumers and their brands: Developing relationship theory in consumer research, *Journal of Consumer Research*, 24(4), pp. 343–373.
- Fournier, S., Lee, L. (2009). Getting Brand Communities Right, *Harvard Business Review*, 87(4), pp. 105–111.

- Frith, S. (1996). Music and Identity. In: *Questions of Cultural Identity* (pp. 108–127), S. Hall, P. Du Gay (Eds.). SAGE Publications, Inc.
- Frith, S. (2005). Popular Music and the Local State. In: *Rock and Popular Music: Politics, Policies, Institutions* (pp. 15–24), T. Bennett, S. Frith, L. Grossberg, J. Shepherd, G. Turner (Eds.). New York: Routledge.
- Greenberg, D.M., Matz, S.C., Schwartz, H.A., Fricke, K.R. (2020). The self-congruity effect of music, *Journal of Personality and Social Psychology*. DOI: 10.1037/pspp0000293.
- Gromark, J., Melin, F. (2013). From market orientation to brand orientation in the public sector, *Journal of Marketing Management*, 29 (9–10), pp. 1099–1123.
- Hankinson, G. (2004). Relational network brands: towards a conceptual model of place brands, *Journal of Vacation Marketing*, 10(2), pp. 109–21.
- Heding, T., Knudtzen, C.F., Bjerre, M. (2016). *Brand Management: Research, Theory and Practice*. 2nd Ed. Abingdon: Routledge.
- Helgeson, J.G., Supphellen, M. (2004). A conceptual and measurement comparison of self-congruity and brand personality: The impact of socially desirable responding, *International Journal of Market Research*, 46(2), pp. 205–233.
- Hidayati, A.N., Isabela B.A., Yuwananda Y.S., Gunadi W. (2021). The Effect of Nostalgia Emotion to Brand Trust and Brand Attachment towards Repurchase Intention, *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(3), pp. 4492–4503.
- Holbrook, M.B., Hirschman, E.C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun, *Journal of Consumer Research*, 9(2), pp. 132–140.
- Holt, D.B. (2003). Brands and branding, *Harvard Business Review*, March, pp. 1–12.
- Holt, D.B. (2004). *How brands became icons: The principles of cultural branding*. Cambridge, MA: Harvard Business Press.
- Huang, Y.A., Lin, C., Phau, I. (2015). Idol attachment and human brand loyalty, *European Journal of Marketing*, 49 (7/8), pp. 1234–1255.
- IFPI (2016). *Investing in Music. The Value of Record Companies*, <https://www.riaa.com/wp-content/uploads/2017/01/ifpi-iim-report-2016.pdf> (accessed: 10.06.2021).
- IFPI (2019). *Music Listening 2019*, <https://www.ifpi.org/wp-content/uploads/2020/07/Music-Listening-2019-1.pdf> (accessed: 24.06.2021).
- IFPI (2021). *Global Music Report 2020*, <https://gmr2021.ifpi.org/report> (accessed: 08.06.2021).
- Jang, H., Olfman, L., Ko, I., Koh, J., Kim, K. (2008). The influence of on-line brand community characteristics on community commitment and brand loyalty, *International Journal of Electronic Commerce*, 12(3), pp. 57–80.
- Kaufmann, H.R., Petrovici, D.A., Gonçalves Filho, C., Ayres, A. (2016). Identifying moderators of brand attachment for driving customer purchase intention of original vs counterfeits of luxury brands, *Journal of Business Research*, 69(12), pp. 5735–5747.
- Kohli, G.S., Melewar, T.C., Yen, D. (2014). Investigating the strategic marketing significance of brand love in developing and nurturing consumer-brand relationships via film branding: A brandscape perspective, *The Marketing Review*, 14(4), pp. 382–404.
- Lacasa, P., de la Fuente, J., Garcia-Pernia, M., Cortés, S. (2017). Teenagers, fandom and identity, *Persona Studies*, 3(2), pp. 51–65.

- Larceneux, F. (2001). Critical opinion as a tool in the marketing of cultural products: the experiential label, *International Journal of Arts Management*, 3(2), pp. 60–70.
- Laroche, M., Habibi, M.R., Richard, M.O., Sankaranarayanan, R. (2012). The effects of social media based brand communities on brand community markers, value creation practices, brand trust and brand loyalty, *Computers in Human Behavior*, 28(5), pp. 1755–1767.
- Larsen, G., Lawson, R., Todd, S. (2012). The symbolic consumption of music. In: *New Horizons in Arts, Heritage, Nonprofit and Social Marketing* (pp. 79–94), R. Bennett, F. Kerrigan, D. O'Reilly (Eds.). New York: Routledge.
- Leenders, M.A. (2010). The relative importance of the brand of music festivals: a customer equity perspective, *Journal of Strategic Marketing*, 18(4), pp. 291–301.
- Lewis, G.H. (1992). Who do you love? The dimensions of musical taste. In: *Popular music and communication* (pp. 134–151), J. Lull (Ed.). SAGE Publications, Inc.
- Li, Y., Lu, C., Bogicevic, V., Bujisic, M. (2019). The effect of nostalgia on hotel brand attachment, *International Journal of Contemporary Hospitality Management*, 31(2), pp. 691–717.
- Liu, Y., Kou, Y., Guan, Z., Hu, J., Pu, B. (2020). Exploring hotel brand attachment: The mediating role of sentimental value, *Journal of Retailing and Consumer Services*, 55. DOI: 10.1016/j.jretconser.2020.102143.
- Mael, F., Ashforth, B.E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification, *Journal of Organizational Behavior*, 13(2), pp. 103–123.
- Malär, L., Krohmer, H., Hoyer, W.D., Nyffenegger, B. (2011). Emotional brand attachment and brand personality: The relative importance of the actual and the ideal self, *Journal of Marketing*, 75(4), pp. 35–52.
- Mathwick, C., Wiertz, C., De Ruyter, K. (2008). Social capital production in a virtual P3 community, *Journal of Consumer Research*, 34(6), pp. 832–849.
- McAlexander, J.H., Schouten, J.W., Koenig, H.F. (2002). Building brand community, *Journal of Marketing*, 66(1), pp. 38–54.
- McCracken, G. (1986). Culture and consumption: A theoretical account of the structure and movement of the cultural meaning of consumer goods, *Journal of Consumer Research*, 13(1), pp. 71–84.
- Muñiz, A.M., O'Guinn, T.C. (2001). Brand Community, *Journal of Consumer Research*, 27(4), pp. 412–432.
- Muñiz, A.M., Schau, H.J. (2005). Religiosity in the abandoned Apple Newton brand community, *Journal of Consumer Research*, 31(4), pp. 737–747.
- Munnukka, J., Karjaluo, H., Tikkanen, A. (2015). Are Facebook brand community members truly loyal to the brand? *Computers in Human Behavior*, 51, pp. 429–439.
- O'Reilly, D. (2005). Cultural brands/branding cultures, *Journal of Marketing Management*, 21(5–6), pp. 573–588.
- O'Reilly, D. (2007). The marketing of popular music. In: *Arts Marketing* (pp. 16–35), F. Kerrigan, P. Fraser, M. Ozbilgin (Eds.). London: Routledge.
- O'Reilly, D., Kerrigan, F. (2013). A view to a brand: introducing the film brandscape, *European Journal of Marketing*, 47 (5/6), pp. 769–789.

- O'Reilly, D., Larsen, G., Kubacki, K. (2013). Music, Fans and Fandom. In: *Music, Markets and Consumption* (pp. 162–180), D. O'Reilly, G. Larsen, K. Kubacki (Eds.). Oxford: Goodfellow Publishers.
- O'Reilly, D., Larsen, G., Kubacki, K. (2016). Marketing live music. In: *Coughing and clapping: Investigating audience experience* (pp. 29–42), K. Burland, S. Pitts (eds). London: Routledge.
- Obiegbu, C.J., Larsen, G., Ellis, N. (2019). The critical music fan: the role of criticality in collective constructions of brand loyalty, *Arts and the Market*, 9(1), pp. 65–80.
- Obiegbu, C.J., Larsen, G., Ellis, N. (2020). Experiential brand loyalty: Towards an extended conceptualisation of consumer allegiance to brands, *Marketing Theory*, 20(3), pp. 251–271.
- Obiegbu, C.J., Larsen, G., Ellis, N., O'Reilly, D. (2019). Co-constructing loyalty in an era of digital music fandom: An experiential-discursive perspective, *European Journal of Marketing*, 53(3), pp. 463–482.
- Oliver, R.L. (1999). Whence Consumer Loyalty? *Journal of Marketing*, 63 (4_suppl1), pp. 33–44.
- Park, C.W., MacInnis, D.J., Priester, J., Eisingerich, A.B., Iacobucci, D. (2010). Brand attachment and brand attitude strength: Conceptual and empirical differentiation of two critical brand equity drivers, *Journal of Marketing*, 74(6), pp. 1–17.
- Peng, N., Chen, A. (2019). Examining consumers' luxury hotel stay repurchase intentions-incorporating a luxury hotel brand attachment variable into a luxury consumption value model, *International Journal of Contemporary Hospitality Management*, 31(3), pp. 1348–1366.
- Pentina, I., Zhang, L., Basmanova, O. (2013). Antecedents and consequences of trust in a social media brand: A cross-cultural study of Twitter, *Computers in Human Behavior*, 29(4), pp. 1546–1555.
- Preece, C., Kerrigan, F. (2015). Multi-stakeholder brand narratives: An analysis of the construction of artistic brands, *Journal of Marketing Management*, 31 (11–12), pp. 1207–1230.
- Rabbanee, F.K., Roy, R., Spence, M.T. (2020). Factors affecting consumer engagement on online social networks: self-congruity, brand attachment, and self-extension tendency, *European Journal of Marketing*, 54(6), pp. 1407–1431.
- Saboo, A.R., Kumar, V., Ramani, G. (2016). Evaluating the impact of social media activities on human brand sales, *International Journal of Research in Marketing*, 33(3), pp. 524–541.
- Schau, H.J., Muñiz, A.M. (2007). Temperance and religiosity in a non-marginal, non-stigmatized brand community. In: *Consumer Tribes* (pp. 144–162), B. Cova, R.V. Kozinets, A. Shankar (Eds.). New York: Routledge.
- Schouten, J.W., McAlexander, J.H. (1995). Subcultures of consumption: An ethnography of the new bikers, *Journal of Consumer Research*, 22(1), pp. 43–61.
- Schroeder, J.E. (2005). The artist and the brand, *European Journal of Marketing*, 39 (11/12), pp. 1291–1305.
- Schroeder, J.E., Salzer-Mörling, M. (Eds.). (2006). *Brand Culture*. Abingdon: Routledge.
- Stern, B.B. (2006). What does brand mean? Historical-analysis method and construct definition, *Journal of the Academy of Marketing Science*, 34(2), pp. 216–223.
- Strauss, S. (2005). *The Small Business Bible: Everything You Need to Know to Succeed in Your Small Business*. Hoboken, New Jersey: John Wiley & Sons, Inc.

- Stride, H., Lee, S. (2007). No Logo? No Way. Branding in the Non-Profit Sector, *Journal of Marketing Management*, 23 (1–2), pp. 107–122.
- Tajfel, H., Turner, J. (2004). An Integrative Theory of Intergroup Conflict. In: *Organizational Identity: A Reader* (pp. 56–65), M.J. Hatch, M. Schultz (Eds.). New York: Oxford University Press.
- Tarrant, M., North, A.C., Hargreaves, D.J. (2002). Youth Identity and Music. In: *Musical identities* (pp. 134–150), R.A.R. MacDonald, D.J. Hargreaves, D. Miell (Eds.). New York: Oxford University Press.
- The Independent (2018). *Rolling Stones Logo Is Most Iconic Design of All Time, Poll Claims*, <https://www.independent.co.uk/life-style/rolling-stones-logo-tshirt-lips-design-mick-jagger-keith-richards-ronnie-wood-a8597306.html> (accessed: 12.06.2021).
- Thomson, M. (2006). Human brands: Investigating antecedents to consumers' strong attachments to celebrities, *Journal of Marketing*, 70(3), pp. 104–119.
- Thomson, M., MacInnis, D.J., Park, C.W. (2005). The ties that bind: Measuring the strength of consumers' emotional attachments to brands, *Journal of Consumer Psychology*, 15(1), pp. 77–91.
- UNESCO (2009). *The 2009 UNESCO Framework for Cultural Statistics (FCS)*, <https://unstats.un.org/unsd/statcom/doc10/BG-FCS-E.pdf> (accessed: 08.06.2021).
- VICE (2017). *Lady Gaga Pioneered Online Fandom Culture as We Know It*, <https://www.vice.com/en/article/pakq59/lady-gaga-online-fandom-culture-little-monsters> (accessed: 25.06.2021).
- Voeth, M., Herbst, U. (2008). The concept of brand personality as an instrument for advanced non-profit branding – an empirical analysis, *Journal of Nonprofit & Public Sector Marketing*, 19(1), pp. 71–97.
- Walsh, P., Clavio, G., Lovell, M.D., Blaszk, M. (2013). Differences in Event Brand Personality Between Social Media Users and Non-Users, *Sport Marketing Quarterly*, 22(4), pp. 214–223.
- Wang, X.W., Cao, Y.M., Park, C. (2019). The relationships among community experience, community commitment, brand attitude, and purchase intention in social media, *International Journal of Information Management*, 49, pp. 475–488.
- Wasko, M.M., Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice, *MIS Quarterly*, 29(1), pp. 35–57.
- Zehir, C., Şahin, A., Kitapçı, H., Özşahin, M. (2011). The effects of brand communication and service quality in building brand loyalty through brand trust; the empirical research on global brands, *Procedia-Social and Behavioral Sciences*, 24, pp. 1218–1231.
- Zhou, Z., Zhang, Q., Su, C., Zhou, N. (2012). How do brand communities generate brand relationships? Intermediate mechanisms, *Journal of Business Research*, 65(7), pp. 890–895.

Lojalność wobec marek popularnych grup muzycznych: rola przywiązania do marki i relacji ze społecznością wokół marki

Streszczenie

W rozdziale poddano analizie koncepcję przywiązania do marki w kontekście sektora kultury, posługując się przykładem marek zespołów muzycznych. Zestawiono ideę marki grupy muzycznej z teoretycznymi koncepcjami marki oraz zbadano wpływ przywiązania do marki zespołu muzycznego na lojalność wobec niej. Identyfikacja ze społecznością wokół marki, zaangażowanie w społeczność oraz postrzegane podobieństwo pomiędzy społecznością a marką, interpretowane w kontekście społeczności fanów, rozpatrywano jako czynniki kształtujące przywiązanie do marki. Dane zebrane w ankiecie przeprowadzonej wśród członków międzynarodowych społeczności fanów trzech popularnych zespołów posłużyły do testowania postawionych hipotez. Wyniki potwierdzają pozytywną zależność pomiędzy przywiązaniem do marki zespołu a lojalnością wobec marki i mediacyjną rolę przywiązania do marki w relacjach między identyfikacją ze społecznością wokół marki i podobieństwem społeczności do marki a lojalnością wobec marki. Potwierdzono także identyfikację ze społecznością wokół marki zespołu jako czynnik kształtujący zaangażowanie w społeczność. Omówiono ponadto znaczenie wyników badania dla teorii i praktyki.

Słowa kluczowe: lojalność wobec marki, przywiązanie do marki, społeczność wokół marki, marka artysty, marketing muzyki

Kluczowe bariery regulacyjne i zagrożenia dla rozwoju morskiej energetyki wiatrowej (offshore) w Polsce

Bruno Jasic

Streszczenie

Celem tekstu jest ocena barier regulacyjnych dla rozwoju morskiej energetyki wiatrowej (offshore) w Polsce z punktu widzenia strategii transformacji energetycznej oraz wyzwań w procesie rozwoju źródeł niskoemisyjnych. W pracy objęto analizą rozwój technologii offshore w Europie i jej rosnącą rolę w ramach krajowych mocy wytwórczych, wykorzystując do tego badania porównawcze oparte o aktualne wyniki badań i raporty sektorowe. Ryzyko wystąpienia barier regulacyjnych oceniono jako wysokie z uwagi na brak infrastruktury przesyłowej zdolnej do przyjęcia dodatkowych mocy wytwórczych, scentralizowany model zarządzania infrastrukturą energetyczną, konkurencyjność ze strony konwencjonalnych mocy wytwórczych oraz niewystarczający rozwój sieci transgranicznych. Wskazano, że należy dążyć do określenia długofalowej strategii transformacji energetycznej, z zachowaniem świadomości nierównego potencjału ekonomicznego państw biorących udział w procesie transformacji. W celu realizacji przyjętego modelu transformacji energetycznej niezbędne będzie powiązanie krajowych rynków energii (regionalizacja) oraz usunięcie barier regulacyjnych blokujących rozwój energetyki odnawialnej i niskoemisyjnej.

Słowa kluczowe: bezpieczeństwo energetyczne, odnawialne źródła energii (OZE), polityka energetyczna, regionalizacja rynków energii, transformacja energetyczna

1. Wprowadzenie

Transformacja energetyczna w kierunku rozwoju energetyki wykorzystującej energię pochodzącą ze źródeł odnawialnych jest w Polsce procesem o zmiennym tempie i intensywności. Pomimo wypracowanego konsensusu naukowego

odnoszącego się do przyczyn nadchodzącego – i już trwającego – kryzysu klimatycznego, przez ostatnich kilkanaście lat niewiele zmieniło się w kwestii wykorzystania konwencjonalnych mocy wytwórczych w Polsce, które nadal odpowiadają za zdecydowaną większość krajowej produkcji energii elektrycznej. Wbrew oczekiwaniom środowiska międzynarodowego, wyrażanym w szczególności w postaci realizowanej przez Unię Europejską polityki wspierania technologii odnawialnych źródeł energii, polska energetyka opiera się na wytwarzaniu energii elektrycznej w procesie spalania węgla kamiennego lub brunatnego. Choć udział energii ze źródeł odnawialnych wzrósł w ostatnich latach i obecnie wynosi ok. 20% produkcji [Urząd Regulacji Energetyki (URE), 2020, s. 10], to tempo wprowadzania zmian w strukturze krajowego mixu energetycznego budzi wątpliwości ze strony partnerów unijnych, i może przyczynić się do pogłębiania nieefektywności i niekonkurencyjności polskiej gospodarki [Wojtkowska-Łodej, 2014, s. 279], w szczególności tej o profilu energochłonnym.

Zachowawczy kierunek obrany przez polską politykę energetyczną znajduje odzwierciedlenie w postaci barier regulacyjnych i zagrożeń o charakterze prawnym, ekonomicznym, społecznym oraz politycznym, które negatywnie wpływają nie tylko na tempo dostosowania krajowego sektora energetycznego do wymagań rynkowych oraz oczekiwań wyrażanych przez partnerów na arenie międzynarodowej, ale również na samą możliwość przeprowadzenia transformacji energetycznej w kierunku pożądanym przez procesy związane z XXI-wiecznym paradygmatem prowadzenia polityki klimatycznej. Problemów z zainicjowaniem transformacji energetycznej w skali mogącej przynieść realną poprawę stanu środowiska naturalnego doświadcza właściwie cała Unia Europejska. Opóźnienia w realizacji krajowych planów modernizacyjnych i zaniechania na szczeblu programowym sprawiają, że polski przypadek stanowić może egzemplifikację przyczyn spowolnienia tempa przemian, które mają odzwierciedlenie w wielu państwach regionu Europy Środkowo-Wschodniej [Księżopolski, Kotlewski, Maśloch, 2020, s. 142].

Precyzyjne rozpoznanie źródeł powstania barier regulacyjnych oraz instytucjonalnych zagrożeń dla oczekiwanego tempa rozwoju energetyki ze źródeł odnawialnych oraz niskoemisyjnych jest kluczowe w kontekście dążenia do ich eliminacji jako przeszkód na drodze do osiągnięcia zarówno krajowych, jak i unijnych celów w zakresie redukcji emisji gazów cieplarnianych. Plan rozwoju gospodarki niskoemisyjnej ma na celu jednak nie tylko przeciwdziałanie zmianom klimatycznym. Być może nawet bardziej istotnym wyzwaniem w tym aspekcie jest oczekiwanie na skuteczną odpowiedź państwa wobec następujących zjawisk: wzrostu cen energii elektrycznej dla odbiorców przemysłowych oraz konsumentów, zagrożeń dla efektywności i konkurencyjności krajowej gospodarki w starciu z państwami dys-

ponującymi tańszą energią elektryczną oraz konieczności dostosowania krajowego systemu energetycznego do nasilających się trendów w zakresie prosumenckiej produkcji energii, mikrogeneracji oraz tworzenia lokalnych klastrów energii z wykorzystaniem technologii magazynowania energii elektrycznej [Maćkowiak-Pandera, Rubczyński, 2020, s. 16–18].

Z uwagi na fundamentalną rolę energii elektrycznej w rozwoju współczesnej gospodarki, problem zapewnienia bezpiecznych, stabilnych i niezależnych od nieprzewidywalnych zjawisk atmosferycznych źródeł energii stanowi największe wyzwanie nadchodzącego stulecia. Co więcej, prognozuje się, że rozwój nowoczesnych technologii cyfrowych, upowszechnienie pracy w systemie zdalnym lub mieszanym oraz gwałtownie rosnące zapotrzebowanie na energię elektryczną w transporcie (e-mobilność) i w gospodarstwach domowych (ogrzewanie mieszkań zimą oraz chłodzenie latem) przyczyni się do wzrostu zapotrzebowania na energię elektryczną, której nie będzie w stanie zaspokoić wzrost efektywności energetycznej oraz transformacja polegająca wyłącznie na zastąpieniu emisyjnych, konwencjonalnych źródeł wytwórczych energią ze źródeł odnawialnych – niezbędne będą zmiany o charakterze strukturalnym. Należy również zadać pytanie o znaczenie restrykcyjnej polityki klimatycznej w sektorze energetycznym prowadzonej przez Unię Europejską, która odpowiada łącznie za jedynie ok. 10% światowych emisji dwutlenku węgla [Szpak, 2020, s. 45].

W tak zarysowanym kontekście zdecydowanie wyróżnia się wiodąca rola potencjału morskich farm wiatrowych (offshore), uznawanych obecnie za najbardziej obiecującą technologię z punktu widzenia realizacji ścieżki dążenia do zeroemisyjności i neutralności klimatycznej [Dennison, Loss, Soderstrom, 2021]. Korzystne warunki atmosferyczne na Morzu Bałtyckim, efektywne systemy wsparcia z zakresu pomocy publicznej oraz warunki dostosowujące rynek energii elektrycznej do preferencyjnego traktowania energii ze źródeł odnawialnych sprawiają, że rozsądne jest oparcie polskiej transformacji energetycznej w dominujący sposób o wytwarzanie energii elektrycznej przez morskie farmy wiatrowe. Z dynamicznym wzrostem potencjału wytwórczego energetyki offshore wiązać się mogą jednakże problemy o charakterze instytucjonalno-ekonomicznym oraz psychologiczno-społecznym, które to natomiast związane są zarówno z kwestiami technologicznymi, finansowymi oraz politycznymi, jak i nawet nastawieniem lokalnych społeczności oraz sprzeciwem wobec intensywnej industrializacji regionów kraju najbardziej atrakcyjnych turystycznie [Mrozowska, 2020].

Niniejszy tekst przedstawia kluczowe wyzwania stojące przed rozwojem energetyki wiatrowej typu offshore w Polsce, proponując rozwiązania w zakresie kompletności krajowego systemu przesyłowego w ramach regionalnych rynków

energii, tworzenia efektu synergii poprzez współpracę morskich farm wiatrowych z elektrowniami atomowymi jako źródłami stabilnych, niskoemisyjnych mocy wytwórczych. Celem rozdziału jest również ocena barier regulacyjnych dla rozwoju morskiej energetyki wiatrowej w Polsce z punktu widzenia długofalowej strategii transformacji energetycznej oraz kluczowych wyzwań w procesie rozwoju współpracy innych źródeł niskoemisyjnych i odnawialnych na rynku energii. Ostatnim z problemów jest kwestia umiejętnego zarządzania siecią na poziomie regionalnym (międzynarodowym) w ten sposób, by systemy przesyłowe mogły błyskawicznie reagować na nagłe fluktuacje w sferze podaży bez ryzyka zduszenia popytu na energię elektryczną w danej chwili [van der Burgt, 2017, s. 9–15], na co odpowiedzią może być rozwój technologii elektrolizy wodoru, magazynów energii, elektromobilności lub właśnie wspieranie rozwoju energetyki atomowej czy hydroelektrowni. Należy jednocześnie podkreślić, że pomysł spalania gazu ziemnego jako *paliwa przejściowego* ma charakter dyskwalifikujący z uwagi na jego znikome znaczenie dla poprawy wskaźników redukcji emisji gazów cieplarnianych [US Environmental Protection Agency, 2021].

2. Perspektywy dalszego rozwoju energetyki wiatrowej w Polsce

Dynamiczny rozwój technologii OZE w pierwszych dwóch dekadach XXI w. zaczyna wyhamowywać, co postępuje pomimo konsekwentnego wsparcia finansowego z zakresu pomocy publicznej oferowanej ze strony państw członkowskich oraz Unii Europejskiej [Dennison i in., 2021, s. 3–4]. Należy ocenić, że obniżenie dynamiki przyrostu mocy zainstalowanej w sektorze OZE na szczeblu unijnym wskazuje na wystąpienie czynników hamujących dotychczasowy trend, które nie są czynnikami o charakterze czysto finansowym [Książkowski i in., 2020, s. 131]. Również w przypadku Polski dynamiczny rozwój OZE nie wpłynął znacząco na strukturę krajowego miksu energetycznego, w którym podobnie istotną rolę, co energia wiatrowa lub słoneczna, zaczyna odgrywać import energii elektrycznej [URE, 2020, s. 31–32].

Polska jest państwem o przeciętnym potencjale teoretycznym energii pochodzącej ze źródeł odnawialnych – efektywność promieniowania słonecznego w ciągu całego roku kalendarzowego doznaje istotnego uszczerbku przez okresowe zachmurzenia oraz krótkie dni w miesiącach zimowych, natomiast potencjał energii pochodzącej z wiatru realizuje się głównie na obszarze Morza Bałtyckiego, gdzie elektrownie wiatrowe budowane są kosztowną technologią offshore [Mirowski, 2017, s. 7–8]. Z uwagi na zdecydowanie najwyższy potencjał teoretyczny wynikający ze stałego charakteru wiatrów na Morzu Bałtyckim, to właśnie energetyka wia-

trowa offshore stanowić może o dalszym wzroście udziału OZE w polskim miksie energetycznym [Ministerstwo Klimatu i Środowiska (MKiŚ), 2021, s. 62]. Nie bez znaczenia pozostaje kwestia transferu technologii poprzez koordynowany międzynarodowy łańcuch dostaw w zakresie materiałów, usług oraz *know-how* towarzyszącego energetyce wiatrowej typu offshore, co zagwarantować ma uchwalona w 2020 r. specustawa wiatrowa [Teraz Środowisko, 2021b, s. 16].

Zdecydowanie mniej obiecująco rozwijać się będzie energetyka odnawialna wykorzystująca energię słoneczną z uwagi na granice potencjału do wzrostu. Stanie się tak, mimo że sprzyjające otoczenie legislacyjne, preferencyjne traktowanie prosumentów, dostępność cenowa technologii fotowoltaicznej i brak istotnych ograniczeń z zakresu ochrony środowiska (mogących spełniać analogiczną funkcję co *reguła 10H*) sprawiają, że w ostatnich latach można mówić o dynamicznej ekspansji farm fotowoltaicznych, które jednak powinny w przeciągu najbliższych dziesięciu lat osiągnąć szczytowe nasycenie sieci wolumenem produkcji. Uważa się, że w perspektywie do 2050 r. energia słoneczna odpowiadać będzie za jedynie ok. kilka procent krajowej produkcji energii. Wynika to bowiem z naturalnych poziomów średniego nasłonecznienia powierzchni Polski. Energetyka słoneczna może natomiast odegrać znaczącą rolę w procesie rozwoju elektromobilności w strefach podmiejskich, zwiększenia efektywności energetycznej przedsiębiorstw oraz obniżeniu kosztów ogrzewania lub chłodzenia gospodarstw domowych. Wykorzystanie korzystnego profilu elektrowni słonecznych, które najwięcej energii produkują w okresie dużego nasłonecznienia pokrywającego się ze szczytem dobowym zapotrzebowania na energię elektryczną, sprzyjać będzie redukcji obciążenia krajowych sieci przesyłowych oraz zrównoważeniu strat przesyłowych przy ekstremalnie wysokich temperaturach. Co więcej, elektrownie słoneczne doświadczają w Polsce mniejszych problemów związanych z brakiem dostępnych punktów przyłączeniowych do sieci z uwagi na znaczące rozproszenie skali instalowanych mocy wytwórczych [Najwyższa Izba Kontroli (NIK), 2014, s. 13–28].

W przypadku energetyki opartej na pozyskiwaniu energii elektrycznej z wiatru sytuacja jest bardziej skomplikowana i wymaga odmiennej, w większym stopniu długofalowej strategii w celu zapewnienia jej dynamicznego rozwoju. Niewątpliwie jednak przed energetyką wiatrową otwierają się szerokie możliwości wynikające z przygotowania polskiej gospodarki do wspierania rozwoju energetyki wiatrowej w modelu offshore, co stanowić będzie odwrócenie dotychczasowych trendów. Rozwój otoczenia instytucjonalnego wokół energetyki wiatrowej w Polsce można bowiem podzielić na trzy diametralnie odmienne etapy, z których każdy wyróżnia się innym podejściem w zakresie intensywności wsparcia. Pierwszy z nich, w okresie niedojrzałości technologicznej i inwestycyjnej po stronie krajowych

wytwórców, realizowany był w latach 2001–2016 i cechował się nieskrępowanym potencjałem inwestycyjnym wynikającym z niewielkich ograniczeń planistycznych i środowiskowych oraz wysokim poziomem wsparcia z zakresu pomocy publicznej udzielanej głównie w systemie aukcyjnym. Okres ten nie został odpowiednio spożytkowany w zakresie dynamiki wzrostu potencjału wytwórczego elektrowni wiatrowych, do czego przyczyniły się niejasne regulacje prawne negatywnie wpływające na możliwość kredytowania inwestycji oraz opór polityczny motywowany m.in. rzekomą radykalną ingerencją wielkoskalowych farm wiatrowych w krajobraz [Smyrgała, 2016, s. 45].

Drugi etap nastąpił wraz z wejściem w życie w 2016 r. *reguły 10H* oznaczającej zakaz lokowania elektrowni wiatrowych w mniejszej odległości niż dziesięciokrotność wysokości turbiny od budynku mieszkalnego. Dodatkowo lokalizacja elektrowni wiatrowych o mocy większej niż moc mikroinstalacji stała się możliwa jedynie w oparciu o miejscowy plan zagospodarowania przestrzennego. Koronnym argumentem za wprowadzeniem tej regulacji były protesty społeczne i twierdzenie, że elektrownie wiatrowe powstawały bez żadnego nadzoru. O ile rzeczywiście przepisy wzorowano na rozwiązaniach niemieckich, to nie wzięto pod uwagę autonomii samorządów lokalnych w zakresie władztwa planistycznego oraz specyfiki polskiej zabudowy mieszkaniowej o charakterze rozproszonym, która sprawiła, że większość terytorium kraju została z dniem 16 lipca 2016 r. całkowicie wykluczona z możliwości prowadzenia inwestycji w elektrownie wiatrowe. Nie bez znaczenia pozostają również kwestie z zakresu ochrony środowiska wynikające z negatywnego wpływu elektrowni wiatrowych na jakość akustyki otoczenia (nadmierna emisja hałasu) oraz destrukcyjny wpływ na siedliska i liczebność ptaków w rejonie eksploatacji farm wiatrowych [Bassi, Bowen, Fankhauser, 2012, s. 22–23]. Obowiązanie *reguły 10H* w praktyce zahamowało rozwój lądowej energetyki wiatrowej w ciągu ostatnich pięciu lat w kraju, niemalże uniemożliwiając osiągnięcie przez Polskę celów w zakresie udziału energii wytwarzanej przez OZE w krajowym miksie energetycznym.

Trzecią fazę rozwoju energetyki wiatrowej, tym razem na lądzie, wyznacza przyjęcie w 2021 r. specustawy wiatrowej w zakresie offshore, której procedowanie zajęło ponad dwa lata, i do której już zgłoszono szereg poprawek wymagających jej nowelizacji przed rozpoczęciem prac budowlanych przez inwestorów. Zmiany legislacyjne idą w parze z deklaracjami strategicznymi poczynionymi na łamach *Polityki energetycznej Polski do 2040 roku* [MKiŚ, 2021], w której wyraźnie wskazano na prymat technologii offshore w odniesieniu do farm wiatrowych w opozycji do lądowych farm wiatrowych, dla których nie przewiduje się już wysokiego potencjału wzrostu krajowych mocy wytwórczych. W dokumencie widoczny jest

wyraźny kontrast w ocenie potencjału obu tych technologii – prognozy dla elektrowni wiatrowych na lądzie przewidują moc zainstalowaną w wysokości ok. 10 GW do 2040 r. (podczas gdy obecnie jest to już ponad 6,5 GW mocy zainstalowanej), natomiast utworzenie nowych 11–12 GW mocy zainstalowanej przypisuje się technologii offshore, która zostanie włączona do krajowej sieci wytwórczej najwcześniej za kilka lat. Co ciekawe, obecne są w debacie publicznej głosy wskazujące na znaczące niedoszacowanie potencjału offshore wynikającego z korzystnych wiatrów pojawiających się stale na Morzu Bałtyckim [Purta, Szarek, Engel, Speelman, van der Pluijm, 2020, s. 53]. Według niektórych z prezentowanych modeli dekarbonizacji polskiej energetyki, morska energia wiatrowa może odpowiadać za moc zainstalowaną na poziomie nawet 30 GW w 2040 r., i zyskać kolejnych 15 GW mocy zainstalowanej do 2050 r. Należy jednak mieć na uwadze, że potencjał ten jest ograniczony rzeczywistą dostępnością terenów morskich – zbudowanie 40–45 GW mocy zainstalowanej w ramach farm wiatrowych offshore na Morzu Bałtyckim oznaczałoby zajęcie ok. połowy całej powierzchni polskiej wyłącznej strefy ekonomicznej, co istotnie zmieniłoby krajobraz i ekosystem regionu północnej części kraju.

Umożliwienie wzrostu udziału mocy wytwórczych pochodzących z odnawialnych źródeł energii wymaga w Polsce podejścia systemowego i kompleksowego w kontekście przewyciężenia rozpoznanych barier wzrostu o charakterze instytucjonalnym, prawnym, ekonomicznym, psychologicznym i społecznym [Wasiuta, 2015, s. 156], a także technologicznym i przestrzennym. Bariery te można podzielić na dwie zasadnicze grupy: instytucjonalno-ekonomiczne oraz psychologiczno-społeczne, spośród których większe znaczenie należałoby przypisać pierwszej z nich. Skupia ona w sobie problemy natury politycznej, prawnej, instytucjonalnej oraz ekonomicznej, związane z zapóźnieniem technologicznym (związanym z pozyskaniem kapitału o niskich kosztach), dostosowaniem regulacji prawnych do oczekiwań inwestorów i wymagań unijnych, oraz negatywnym nastawieniem instytucji państwowych do inwestycji w niekonwencjonalne źródła energii.

Należy zauważyć, że – w przeciwieństwie do sytuacji, która miała miejsce jeszcze kilka lat temu [Wasiuta, 2015, s. 154–155] – obecnie nie istnieją poważniejsze bariery psychologiczno-społeczne w stosunku do rozwoju technologii pozyskiwania energii ze źródeł odnawialnych. Intensywne akcje promocyjne, edukacyjne i o charakterze lobbingsowym przyniosły istotne zmiany w percepcji możliwości, efektywności i wielkoskalowości inwestycji w OZE. Szczególnym dowodem na wzrost zaufania mieszkańców Polski w stosunku do technologii OZE jest stale rosnąca dynamika popularności energetyki prosumenckiej, której rozmiary przekroczyły najbardziej nawet optymistyczne prognozy branży energetycznej [Dyląg, 2021]. Z uwagi na

niebezpieczeństwo dalszego rozwoju energetyki prosumenckiej w sytuacji niewydajności przestarzałej sieci przesyłowej, niedostosowanej do przyjęcia pożądanej ilości mocy przyłączeniowej z mikroinstalacji, należy spodziewać się podjęcia działań legislacyjnych mających na celu ograniczenie preferencyjności przyłączenia do sieci dla prosumentów i zmniejszenie dynamiki inwestycji w energetykę prosumencką w kolejnych latach [Sobczyk-Grygiel, 2021], zarówno w przypadku fotowoltaiki, jak i energetyki wiatrowej. Próby zahamowania rozwoju energetyki prosumenckiej potwierdzają tezę o scentralizowanym i inercyjnym charakterze polskiego systemu energetycznego, który funkcjonuje w sposób przywilejujący koncerny energetyczne kontrolowane przez Skarb Państwa [GLOBEnergia, 2021].

Ciągłe dostosowywanie kierunków polskiej polityki energetycznej do oczekiwań wyrażanych przez koncerny będące dotychczas monopolistami na krajowym rynku energii doprowadziło do powstania zjawiska inercji rynkowej, utrudniającego podejmowanie strategicznych decyzji o charakterze innowacyjnym, które mogłyby działać przeciwko reprezentantom konwencjonalnych grup interesu w sektorze energetycznym [Jędrasik, 2018, s. 163–180]. Umiejętność skutecznego kontestowania silnej pozycji zajmowanej przez scentralizowane, wielkoskalowe przedsiębiorstwa wytwórcze w energetyce konwencjonalnej będzie stanowić jedno z najważniejszych zadań w przeciągu najbliższych lat, gdyż niechęć do zwiększania dynamiki innowacyjności w energetyce stanowi jedną z największych barier wzrostu efektywności polskiej energetyki i redukcji jej emisyjności.

3. Psychologiczno-społeczne bariery ograniczające rozwój morskiej energetyki wiatrowej w Polsce

Rozwój energetyki odnawialnej w Polsce, w tym morskich farm wiatrowych, może w przyszłości doprowadzić do eskalacji pewnych problemów psychologiczno-społecznych, które będą wymagały interwencji ze strony organów władzy państwowej. Realizację inwestycji powinny uwzględniać czynniki również pozaekonomiczne, do których zaliczać się będą takie elementy, jak przywiązanie lokalnej społeczności do naturalnego krajobrazu morskiego, emocje związane z degradacją jakości panoramy Morza Bałtyckiego oraz wrażliwość społeczna na zmiany krajobrazu i zagrożenie dla fauny i flory. Wskazuje się, że sukces energetyki wiatrowej zależeć będzie w dużej mierze od tego, jak skutecznie lokalne społeczności zostaną włączone w procesy decyzyjne, w jaki sposób i jak szeroko będą rozpowszechniane informacje dotyczące inwestycji, a także jaki będzie społeczny punkt widzenia – jest to próba wyznaczenia prognozy konfliktogenności danej inwestycji [Mrozowska, 2020]. Negatywna per-

cepcja krajobrazu nasyconego farmami wiatrowymi jest bowiem elementem, który obecnie ogranicza potencjał rozwoju i znaczenia tej technologii w Polsce. Należy oczekiwać, że doświadczenie zebrane przez inwestorów w toku prac nad rozwojem energetyki wiatrowej na lądzie zaprocentuje również w kontaktach z lokalną społecznością obecną przy procesie realizacji morskich farm wiatrowych, co umożliwi sprawne i kompromisowe rozwiązania inwestycyjne.

W niejako paradoksalnym powiązaniu z barierami psychologiczno-społecznymi pozostają kwestie środowiskowe i przestrzenne. W literaturze coraz częściej dostrzega się negatywny wpływ rozwoju energetyki odnawialnej na środowisko naturalne – w szczególności dotyczy to elektrowni wiatrowych mających negatywny wpływ na populację ptactwa. Choć bowiem wykorzystanie prawie wszystkich źródeł energii odnawialnej jest związane ze znikomym szkodliwym wpływem na środowisko, przynajmniej w miejscu ich instalacji, to należy dostrzec rozproszony charakter generacji energii ze źródeł odnawialnych [Wasiuta, 2015, s. 160]. Wśród społeczności lokalnych, w konfrontacji minimalnego oddziaływania odnawialnych technologii energetycznych na najbliższe otoczenie tracą na znaczeniu, gdyż duża ilość zajętej powierzchni przez instalacje OZE, zarówno wiatrowe, jak i słoneczne, sprawia wrażenie *przytłoczenia* krajobrazu i tworzy postindustrialną monokulturę produkcji energii elektrycznej bez udziału człowieka, co może negatywnie wpływać na poziom aprobaty dalszego rozwoju energetyki odnawialnej na lądzie, w bliskim otoczeniu gospodarstw domowych. Problem barier środowiskowych jest tym większy, że polski ustawodawca do dziś nie rozwiązał wielu kwestii związanych z interpretacją przepisów regulujących kwestie lokalizacji inwestycji na terenach wiejskich i nisko uprzemysłowionych, co przyczynia się do eskalacji konfliktów z mieszkańcami na tle nadmiernej ingerencji instalacji OZE w krajobraz lokalny.

Zwiększenie świadomości społeczeństwa i przełamanie naturalnych oporów przed wprowadzaniem OZE to wciąż jedno z ważniejszych zadań i pożądanych kierunków działań instytucji uczestniczących w procesie transformacji energetycznej. Uwagę należy zwrócić głównie na skuteczną adaptację proponowanych rozwiązań technologicznych i systemowych do warunków polskich [Wasiuta, 2015, s. 163]. Należałoby w tym miejscu w szczególności podkreślić niedocenione przez ustawodawcę znaczenie energetyki prosumenckiej, która jednakże – z uwagi na wskazywany już problem niewystarczającej wydajności krajowych sieci przesyłowych – powinna zostać rozwijana w oparciu o ideę lokalnych wytwórców energii i spółdzielni gminnych, które byłyby zdolne do autonomicznego wytwarzania i przesyłu energii elektrycznej na ograniczonym, niewielkim terytorium, działając kompleksowo m.in. w formule klastrów energii. Rozwiązanie to będzie szczególnie atrakcyjne wraz z upowszechnieniem się technologii magazynowania energii, jednakże jego

implementacja na skalę szerszą niż pilotażowa będzie wymagała pełnego przeformułowania zasad rozliczania się z kosztów energii elektrycznej, na co koncerny energetyczne oraz organy administracji rządowej, w tym Urząd Regulacji Energetyki, nie są w żaden sposób przygotowane [Gryszczuk, 2021].

Podsumowując oczekiwania i obawy ze strony lokalnych społeczności, należy podkreślić, że jedną z barier rozwoju energetyki odnawialnej w skali mikro jest wprowadzony przez ustawodawcę domyślny zakaz prowadzenia przez konsumentów, prosumentów i lokalnych wytwórców energii elektrycznej działalności wytwórczej o charakterze typowo zarobkowym [Ropuszyńska-Surma, Węglarz, 2017, s. 91–94], co skutecznie ogranicza dynamikę rozwoju energetyki odnawialnej skierowanej na zaspokajanie nie osobistego, lecz lokalnego zapotrzebowania. Obwarowanie zarobkowego charakteru dodatkowymi wymogami koncesyjnymi, zezwoleniami i wysoce sformalizowaną wymianą dokumentacji sprawia, że wielu inwestorów nie decyduje się na inwestycję kapitału w sektor małych instalacji OZE na lądzie. Dalsze zaniedbania na tej płaszczyźnie mogą doprowadzić do zmniejszenia poparcia dla rozwoju OZE w Polsce poprzez utrzymywanie przez ustawodawcę wrażenia, że zyski z inwestycji w sektor odnawialnych źródeł energii są prywatyzowane i kumulowane w coraz węższych rękach korporacji międzynarodowych i koncernów energetycznych kontrolowanych przez Skarb Państwa, podczas gdy konsumentom, prosumentom oraz sektorowi małych i średnich przedsiębiorstw pozostawia się jedynie możliwość nieco bardziej preferencyjnego optymalizowania kosztów zużycia energii elektrycznej.

4. Instytucjonalno-ekonomiczne bariery wzrostu i zagrożenia dla rozwoju energetyki wiatrowej w ramach polskiego miksu energetycznego

Choć obecnie nie istnieje alternatywa dla dynamicznego rozwoju energetyki wiatrowej w kontekście celu, jakim jest dążenie do zmniejszenia emisyjności polskiego systemu energetycznego, to rodzi się pytanie o wielkoskalowość prowadzonych inwestycji, ich koszty oraz wpływ na pozostałe procesy technologiczne w zakresie zarządzania efektywnością energetyczną lub elastycznością krajowych i transnarodowych sieci przesyłowych. W przypadku Polski, z dynamicznym rozwojem energetyki wiatrowej na morzu wiązać się będą problemy wynikające z istnienia dużej ilości niestabilnych i nieelastycznych mocy wytwórczych w jednej, niewielkiej części kraju, nieodznaczającej się w dodatku szczególnie wysokim zużyciem energii [NIK, 2018, s. 40].

Dla energetyki wiatrowej nie istnieją w zasadzie zagrożenia o charakterze politycznym na płaszczyźnie legislacji unijnej. Jak każda technologia z zakresu odnawialnych źródeł energii, tak i energetyka wiatrowa cieszy się uznaniem w zakresie dopuszczalności udzielania inwestorom pomocy publicznej w dowolnej skali [Pinkas, 2021, s. 47–60]. Sytuacja inaczej wygląda w Polsce, gdyż na przestrzeni ostatnich lat można obserwować zmiany legislacyjne realizujące interesy polityczne. Wyraźnym przykładem takiego działania w Polsce było przyjęcie regulacji wprowadzających w 2016 r. wspomnianą już *regulę 10H*. Już natomiast po kilku latach, w ślad za zmianą sytuacji politycznej i potrzebą bardziej dynamicznej dywersyfikacji źródeł energii, nastąpiła rezygnacja z *reguły 10H*, której usunięcie z krajowego porządku prawnego powinno pozwolić na wznowienie inwestycji w zakresie elektrowni wiatrowych na lądzie [Teraz Środowisko, 2021b, s. 18]. Wspomniana zmiana nie oznacza pełnej liberalizacji przepisów z punktu widzenia inwestorów, gdyż odstępstwo od wyznaczonej odległości będzie wymagać przeprowadzenia uprzedniej prognozy oddziaływania na środowisko wykonywanej w ramach miejscowego planu zagospodarowania przestrzennego – niezależnie od dokonanej oceny, nie będzie w Polsce możliwości lokalizowania elektrowni wiatrowych na lądzie w odległości mniejszej niż 500 metrów od jakichkolwiek zabudowań mieszkalnych. Obecne zagęszczenie lokalizacji elektrowni wiatrowych na lądzie w powiązaniu z uwarunkowaniami w zakresie rozproszonej zabudowy mieszkaniowej i przeciętnymi warunkami wietrznymi pozwala przyjąć, że potencjał rozwojowy tej technologii w Polsce będzie mógł się realizować przede wszystkim na obszarze Morza Bałtyckiego. Już tylko powyższe przykłady pozwalają dostrzec, że potrzeby polityczne potrafią w Polsce istotnie determinować możliwości rozwoju energetyki odnawialnej, stanowiąc źródło poważnych barier o charakterze regulacyjnym.

Sytuacja kształtuje się zatem odmiennie w przypadku energetyki wiatrowej offshore. Jej preferencyjne traktowanie w Polsce, choć stanowi decyzję o charakterze politycznym, to poparte jest analizą opłacalności ekonomicznej inwestycji z uwagi na warunki wietrzne występujące w północnej części kraju. Należy jednak zauważyć, że początkowe koszty budowy, modernizacji sieci przesyłowej i organizacja systemu zachęt finansowych i logistycznych dla inwestorów stanowią dodatkowy koszt rozwoju energetyki offshore, który nie jest brany pod uwagę w ramach analizy faktycznej opłacalności inwestycji w perspektywie długoterminowej. Jak wskazuje się w literaturze, polityczny charakter inwestycji wynika z wysokiego poziomu zainteresowania rozwojem energetyki offshore przez koncerny energetyczne stanowiące spółki kontrolowane przez Skarb Państwa, które oceniają projekt rozwoju morskich farm wiatrowych jako znacznie korzystniejszy finansowo od rozwoju

jednostek energetyki konwencjonalnej, które cechuje niższy próg wejścia na rynek i tym samym większa konkurencyjność rynku [Żylińska, 2020].

Decyzja o zwiększeniu inwestycji w odnawialne źródła energii wynikała zresztą z prowadzenia przez Unię Europejską konsekwentnej, długofalowej polityki energetyczno-klimatycznej, która poprzez nierynkowe mechanizmy skutecznie doprowadziła do wzrostu kosztów i cen energii elektrycznej produkowanej poprzez spalanie węgla kamiennego oraz brunatnego. Tym samym dominująca rola węgla w Polsce wpłynęła na mechanizmy cenotwórcze w gospodarce w sposób wykraczający daleko poza sektor energetyczny i energochłonny, stając się m.in. przyczyną wzrostu: kosztów energii dla gospodarstw domowych i sektora MŚP, cen towarów przemysłowych i konsumpcyjnych oraz wskaźnika inflacji [Rubene, 2018, s. 45]. Poszerzenie portfela inwestycyjnego spółek kontrolowanych przez Skarb Państwa pozwoli na zwiększenie wartości akcji tychże spółek na notowaniach giełdowych, co obniży koszt pozyskiwania kapitału w przyszłości i umożliwi wykorzystanie efektu dźwigni finansowej przy kolejnych kapitałochłonnych projektach wymagających zaangażowania aktywów państwowych.

Istotnym czynnikiem, który został wzięty pod uwagę w toku podejmowania decyzji odnośnie do roli energetyki typu offshore w Polsce jest wyraźnie scentralizowany charakter polskiego systemu energetycznego – zarówno na płaszczyźnie wytwórczej, jak i przesyłowej. Przeprowadzenie wielkoskalowych inwestycji przy wsparciu lub czynnym udziale kilku największych polskich spółek energetycznych [PKN Orlen, 2021] zapewni im odpowiednią decyzyjność w ramach kształtowania przyszłej polityki energetycznej oraz zwiększy potencjał lobbingsowy rozumiany jako możliwość wywierania skutecznego wpływu na negocjacje dotyczące m.in. poziomu mechanizmów wsparcia z zakresu pomocy publicznej dla odnawialnych źródeł energii.

Kolejnym przejawem scentralizowanego charakteru polskiego systemu energetycznego jest wysoki poziom koncentracji jednostek wytwórczych w portfelu własnościowym spółek kontrolowanych przez Skarb Państwa, ich lokalizacja oraz specyfika połączeń jednostek wytwórczych z sieciami przesyłowymi, które zorganizowane zostały po II wojnie światowej w ten sposób, że ogniskują się głównie wokół historycznych okręgów przemysłowych w trójkącie Katowice – Warszawa – Kraków. Umieszczenie mocy wytwórczych o łącznym potencjale przekraczającym 10 GW mocy zainstalowanej, cechującej się dodatkowo niską elastycznością, niesterownością i nieprzewidywalnością w zakresie wielkości produkcji powoduje konieczność przeformułowania sposobu funkcjonowania całego systemu energetycznego w Polsce i jego modernizacji. Sieć przesyłowa w Polsce nie jest przygotowana do przekierowania działania na oś północ – południe, w szczególności, jeżeli wziąć pod

uwagę zdecydowanie wyższe zużycie energii elektrycznej w południowej części Polski w stosunku do rejonów nadmorskich lub województw północno-wschodnich. Słaba gęstość linii przesyłowych na północy Polski powoduje, że przy nieodpowiedniej generacji z turbin wiatrowych lub awaryjnych wyłączeniach konwencjonalnych jednostek wytwórczych moce przesyłowe są ograniczane poniżej wartości nominalnych [Terlikowski, Paska, 2018, s. 177–178]. Rozbudowa sieci przesyłowej w północnej części kraju pomogłaby zlikwidować ograniczenia sieciowe i pozwoliłaby na pełne wykorzystanie możliwości połączeń asynchronicznych niezależnie od czynników zewnętrznych. Podobne problemy zostały spowodowane przez nadmierną moc przyłączeniową jednostek odnawialnych źródeł energii w innych państwach członkowskich Unii Europejskiej. Z uwagi na czasochłonność i wysoki koszt inwestycji w sieci przesyłowe, problemy te nie zostały jak dotąd skutecznie rozwiązane, co przyczynia się do generowania znacznych strat energii, zwiększenia częstotliwości awarii i wyłączeń sieci oraz wzrostu kosztów zarządzania siecią przesyłową, wpływając w tych aspektach negatywnie również na systemy energetyczne państw ościennych [Deign, 2020].

5. Problem zapewnienia bezpieczeństwa energetycznego Polski w oparciu o model uwzględniający wiodącą rolę energetyki offshore

Z punktu widzenia polskiego systemu energetycznego, dynamiczny wzrost udziału energetyki wiatrowej offshore w krajowym miksie energetycznym będzie natomiast wiązał się z problemami o charakterze regulacyjnym oraz ekonomicznym. Przede wszystkim należy spodziewać się utrzymania tendencji do preferencyjnego traktowania energii pochodzącej z OZE w ramach sieci przesyłowej, która będzie zobowiązana do przyznawania dostępu do sieci mocom wytwórczym wykorzystującym energię wiatrową z uwagi na jej niesterowalność, niedyspozycyjność i nieprzewidywalność. Powyższe będzie wymagało wysokiej sterowalności, dyspozycyjności i jakościowej adaptacji po stronie zarządców sieci przesyłowej, którzy będą zobowiązani do przyjęcia całkowitej ilości energii wytworzonej przez OZE i – równolegle – do obniżenia mocy produkcyjnych wszystkich pozostałych źródeł energii [NIK, 2018, s. 47–48].

Biorąc pod uwagę bezpieczeństwo pracy systemu elektroenergetycznego, istotna jest ilość i moc przyłączonych i pracujących w nim OZE. Także priorytetowy dostęp do sieci daje podłączonym producentom energii z OZE pewność, że w każdej chwili będą mogli sprzedawać i przysyłać wytworzoną energię elektryczną.

Oznacza to obowiązek udzielenia gwarancji, że każda jednostka typu offshore uzyska priorytetowy dostęp do sieci przesyłowej działającej w ramach państwowego Krajowego Systemu Elektroenergetycznego, ale również zapewnienia, że każdej jednostce typu offshore odpowiadać będą odpowiednie moce rezerwowe w zakresie wytwórstwa energii elektrycznej, które mogą możliwie jak najszybciej włączyć się do sieci przesyłowej, jeżeli z jakichkolwiek powodów – najczęściej wynikających z bezwietrznej pogody – niemożliwa będzie generacja energii elektrycznej z farm wiatrowych zlokalizowanych na Morzu Bałtyckim [NIK, 2018, s. 44–46]. W modelowym scenariuszu na każde 10 GW mocy zainstalowanej w ramach OZE przypadać powinno 10 GW mocy zainstalowanej w wysoce dyspozycyjnej jednostce spalającej gaz ziemny lub – w przyszłości – wodór, a także wielkoobjęściowe magazyny energii.

Aktualny stan rozwoju sieci przesyłowej wskazuje, że powyższe warunki nie są obecnie możliwe do spełnienia [NIK, 2014, s. 28]. Już teraz, z powodu negatywnego wpływu technologii OZE na elastyczność i stabilność całego systemu wytwórczego i przesyłowego energii elektrycznej, w Polsce obowiązują zalecane limity w zakresie łącznej mocy zainstalowanej farm wiatrowych – problemem jest brak odpowiedniej struktury sieciowej do przyłączenia dużych mocy OZE (głównie farm wiatrowych) i wynikająca z tego faktu niemożność odbioru tak produkowanej energii. Problemem modelowego optimum związanym z czasem jest niemożliwe do uniknięcia zjawisko wahań w zakresie zarówno produkcji, jak i zapotrzebowania na energię. Moc systemu energetycznego musi bowiem brać pod uwagę szczytowe zapotrzebowanie na energię, natomiast w okresie mniejszego zapotrzebowania moc pozostaje niewykorzystana [Juszczuk, Mazur, 2014, s. 79–81], gdyż elektrowni wiatrowych lub słonecznych nie można wyłączyć bez marnotrawienia produkowanej energii. Obserwacja rozwoju energetyki odnawialnej w innych państwach członkowskich Unii Europejskiej i jej rozwój w kierunku zwiększania potencjału *zastępowalności* OZE sugerują potrzebę baczniejszego przyjrzenia się wymiarom zakłóceń zjawisk rynkowych zachodzących na europejskich rynkach energii. Mimo iż trudno obecnie wyrokować na temat przyszłego progu opłacalności OZE, to nie sposób bagatelizować znaczenia nierynkowych determinant w stosunku do rynkowo uzasadnionych ścieżek rozwoju europejskich sieci energetycznych [Juszczuk, Mazur, 2014, s. 94].

6. Komplementarność energetyki wiatrowej offshore oraz polityki regionalizacji rynków energii w Unii Europejskiej

Odpowiedzią na wyzwania technologiczne i finansowe związane z rosnącym udziałem niestabilnych źródeł OZE w krajowych miksach energetycznych może być dążenie do pogłębienia procesów integracyjnych w zakresie wspólnego zarządzania rynkiem energii elektrycznej w Unii Europejskiej na poziomie regulacji przesyłu energii. Rolę w tym zakresie odgrywają regionalne połączenia transgraniczne pomiędzy krajowymi systemami przesyłowymi, co wpisuje się w unijną strategię tworzenia unii energetycznej [Juszczuk, 2016, s. 40]. Obecnie prowadzenie wymiany handlowej sieciami przesyłowymi pomiędzy poszczególnymi państwami nie jest jeszcze kwestią o krytycznym znaczeniu z uwagi na niedostateczną obecność mocy wytwórczych pochodzących z OZE w poszczególnych państwach członkowskich, jednakże problem ten z każdym rokiem przybierać będzie na wadze. Należy ocenić, że kwestia integracji i koordynacji reguł funkcjonowania infrastruktury przesyłowej pomiędzy poszczególnymi państwami członkowskimi Unii Europejskiej będzie kluczowym zagadnieniem [Thema, 2013, s. 36–39] i najbardziej istotną barierą wzrostu mogącą ograniczać dynamikę rozwoju energetyki wiatrowej, w tym projektów offshore na Morzu Bałtyckim.

Regionalizacja rynków energii jest procesem obserwowanym od kilkunastu lat. Choć początkowo wydawało się, i takie spojrzenie dominowało w nauce, że regionalizacja rynków energii stanowi poboczny element regionalizacji zachodzącej w przestrzeni międzynarodowej, to doświadczenia ostatnich lat skłaniają do uznania, iż proces tworzenia ponadnarodowych, coraz bardziej zintegrowanych rynków produkcji, magazynowania oraz przesyłu energii elektrycznej i gazu ziemnego stanowi coraz ważniejsze zjawisko we współczesnej polityce międzynarodowej [Niedziółka, 2018, s. 111–112]. Jeżeli chodzi o państwa europejskie, to należy zauważyć, że przeorientowanie krajowych polityk przybrało w ostatnich latach na dynamice właśnie dzięki uczestnictwu w poszczególnych procesach globalizacji korporacyjnej oraz regionalizacji inwestycyjnej już w samej Unii Europejskiej [Jasiak, 2017, s. 42–43], wiążąc się z inwestycjami w odnawialne źródła energii, poszukiwaniami nowych złóż paliw kopalnych oraz usprawnieniem łańcucha dostaw pomiędzy rynkami integrowanymi w ramach projektu unii energetycznej.

W świetle nowych ambicji Unii Europejskiej w aspekcie tworzenia jednolitego rynku energii coraz wyraźniej zarysowuje się korzystna dla Polski pewna przewaga konkurencyjna po stronie państw regionu Morza Bałtyckiego, które – posiadając doświadczenie wieloletniej współpracy – potrafią dać świeży impuls zarówno unijnemu projektowi unii energetycznej, jak i wykorzystać unijne cele klimatyczne

dla własnych korzyści. W tej perspektywie sytuacja wokół rynku energii państw Morza Bałtyckiego jawi się jako subregionalizacja w ramach samej Unii Europejskiej. Specyfika rynku energii Morza Bałtyckiego oraz układ krajowych mikśów energetycznych sprawia, że analizowany region jako jeden z pierwszych odznaczył się pewną dozą niezależności i autoregulacji [Niedziółka, 2018, s. 29–32]. Wynika to przede wszystkim z uwarunkowań pogodowych części regionu (silne, stabilne wiatry), korzystnego wpływu samego Morza Bałtyckiego na możliwość wykorzystania potencjału odnawialnych źródeł energii oraz ambicji państw – szczególnie skandynawskich – w zakresie realizacji celów unijnej polityki klimatycznej.

Możliwości, jakie państwom regionu Morza Bałtyckiego oferują okoliczności geopolityczne oraz warunki pogodowe, zdają się zachęcać do podążania w kierunku pogłębiania regionalizacji bałtyckiego rynku energii głównie w ramach instytucji unijnych. Różnorodność źródeł pozyskiwania energii może być wartościowa właśnie z uwagi na nieelastyczność odnawialnych źródeł energii w mikroskali, co pozwoli na efektywniejsze zarządzanie mechanizmem podaży energii dla całego regionu poprzez uruchamianie konwencjonalnych mocy wytwórczych tam, gdzie są one najbardziej potrzebne. Ponadto wielkość regionu sprawia, że sama nieelastyczność źródeł odnawialnych może być mniej uciążliwa z uwagi na wystarczającą różnorodność zjawisk pogodowych sprawiających, że brak wiatru w jednym fragmencie regionu wcale nie oznacza, że w pozostałych miejscach warunki okażą się równie niesprzyjające. Uzupełnieniem tego modelu pozostają hydroelektrownie, które w państwach skandynawskich stanowią istotne źródło energii elektrycznej o przewidywalnym poziomie wahań sezonowych w zakresie wielkości produkcji. W świetle tak rysujących się okoliczności łatwo dostrzec, że największymi wyzwaniem w praktyce są kwestie związane z techniczną przepustowością sieci elektroenergetycznej, kierunkiem i wielkością przesyłu nadpodaży energii oraz logistycznym zarządzaniem regionalnym rynkiem energii w ramach lokalnych grup operacyjnych. Obecnie celem państw nie jest już tylko zwiększenie powiązań międzyrynkowych, ale też dążenie do utworzenia pewnego rodzaju mechanizmu międzynarodowego zarządzania siecią energetyczną, być może na wzór podobny do stosowanego w Stanach Zjednoczonych Ameryki [Baker, 2017].

Oparcie międzynarodowego systemu przesyłowego o zarządzanie logistyczne w ramach regionalnych centrów operacyjnych byłoby prawdopodobnie uznane za przedwczesne, jednakże nie sposób nie dostrzec faktu, że rosnący udział energii ze źródeł odnawialnych oraz szybkie zmiany w zapotrzebowaniu odbiorców na moc wpływają na pracę jednostek wytwórczych centralnie dysponowanych. Obniżają się ich wskaźniki wykorzystania mocy zainstalowanej; jednostki wytwórcze, które do tej pory pracowały w podstawie krzywej zapotrzebowania na moc, zaczynają

funkcjonować jako elektrownie podszczytowe, odstawiane w momencie niższego zapotrzebowania; wzrasta liczba odstawionych czasowo jednostek wytwórczych centralnie dysponowanych w cyklu dobowym. W konsekwencji wyzwaniem stają się kwestie pokrycia szczytowego zapotrzebowania na energię elektryczną i zapewnienie wymaganego poziomu rezerw mocy [Bronk, Magulski, Maćkowiak-Pandera, 2019, s. 11].

O ile w skali całego regionu Morza Bałtyckiego podaż energii elektrycznej w transakcjach dziennych, krótko- lub długoterminowych co do zasady nie jest problemem, gdyż większość państw doświadcza nadpodaży przy korzystnych warunkach pogodowych, to sprawia trudności sama przepustowość sieci transnarodowych, które nie będą w stanie obsłużyć wystarczającej podaży wytworzonej mocy. Warto dostrzec, że w rejonie Morza Bałtyckiego Polska posiada połączenia transgraniczne jedynie ze Szwecją, Niemcami i Litwą. Ograniczona przepustowość połączeń transgranicznych uniemożliwia odpowiednie reagowanie w sytuacji wystąpienia nieplanowanych przepływów energii z rynków zasobnych w chwilową nadwyżkę energii ze źródeł odnawialnych w kierunku rynków opartych o źródła konwencjonalne. Co znamienne, problem ten dotyka rynek polski w znaczącym stopniu – w 2020 r. energia importowana odpowiadała za 10% rocznego krajowego zużycia [Polskie Sieci Elektroenergetyczne (PSE), 2020].

Należy spodziewać się, że rozwój morskich farm wiatrowych w polskiej części Morza Bałtyckiego tylko pogłębi już istniejący problem, a jak dotąd nie wypracowano modelowego rozwiązania w tej kwestii. Wzrost mocy wytwórczych w zakresie OZE wymaga stworzenia kolejnych połączeń transgranicznych w ramach regionalnych rynków energii, które powinny przebiegać w ten sposób, że państwa posiadające wiele jednostek wytwórczych o niskiej emisyjności mają pierwszeństwo w eksporcie energii do państw o wysokiej emisyjności, i są zobowiązane do preferencyjnego zakupu i przyjęcia każdego nadmiarowego wolumenu energii [Eising, Hobbie, Möst, 2020, s. 5]. W przypadku Polski należałoby spodziewać się dalszych inwestycji w sieci przesyłowe pomiędzy Polską a Niemcami, Litwą, Łotwą i Estonią, które również mają problemy z obniżeniem emisyjności swoich systemów energetycznych. Jednocześnie Polska powinna prowadzić działania mające na celu zwiększenie stałego wolumenu importu energii elektrycznej z państw skandynawskich: Norwegii, Szwecji i Finlandii – z uwagi na niskoemisyjność ich jednostek wytwórczych, import energii z tych państw pomoże Polsce w realizacji unijnych celów redukcji emisji dwutlenku węgla oraz zapewni zastępowalność mocy wytwórczych obecnych w przeznaczonych do likwidacji konwencjonalnych elektrowniach węglowych.

Nadchodzące zmiany w ramach postępującej regionalizacji bałtyckiego rynku energii elektrycznej, opierając się o połączenia transnarodowe z przeznaczeniem

handlowym oraz dynamiczny rozwój elektrowni wiatrowych typu offshore, wspierane będą kolejnymi działaniami w kierunku promocji rozwoju mocy wytwórczych opartych o źródła odnawialne także na poziomie rozproszonym i prosumenckim [Cecchinato, 2019, s. 5]. Wobec tego niezbędnym uzupełnieniem kreowanego obecnie regionalnego rynku energii powinno być stworzenie inteligentnie zarządzanego ekosystemu sieci transnarodowych, w ramach którego współpracują krajowi wytwórcy i operatorzy sieci przesyłowych [Niedziółka, 2011, s. 10–12], elastyczni odbiorcy i coraz powszechniejsze systemy magazynowania energii, których przedłużeniem powinna stać się energia skumulowana w magazynach energii oraz w akumulatorach pojazdów elektrycznych (elektromobilność), a także stabilna podstawa mocy wytwórczych w postaci energetyki atomowej.

W tym kontekście niezwykle istotną zmianą będzie przystosowanie rynków energii do reagowania na zmienne warunki nie tylko po stronie podażowej, co przedstawiono powyżej, ale również od strony popytowej. Oparcie regionalnego rynku energetycznego o mechanizmy reakcji strony popytowej (ang. *Demand Side Response*, DSR) polegać będzie na dobrowolnym, czasowym obniżeniu przez odbiorców – przede wszystkim energochłonnych przedsiębiorców przemysłowych lub usługowych – zużycia energii elektrycznej lub przesunięcie w czasie jej szczytowego poboru na polecenie operatora sieci energetycznej w zamian za oczekiwane, ustalone wynagrodzenie. Mechanizmy te pozwolą przyczynić się do zachowania stabilności w ramach regionalnych systemów elektroenergetycznych w przypadku wystąpienia niepożądanego sytuacji bilansowych, w szczególności w godzinach szczytowego zapotrzebowania na energię elektryczną. Jak pokazują analizy, DSR już obecnie wykazuje zalety tego typu usługi, wskazując na ich przewidywalność, elastyczność i niezawodność [Torriti, Hassan, Leach, 2009, s. 4–5]. Zastosowanie mechanizmu DSR na szerszą skalę jako narzędzia stabilizującego system elektroenergetyczny może pozwolić na skuteczniejsze działanie bałtyckiego rynku energii opartego w znaczącej mierze o nieelastyczne moce wytwórcze typu offshore i powinno stanowić, w szczególnym interesie Polski, istotny element długofalowej polityki energetycznej regionu.

7. Potencjał współpracy morskiej energetyki wiatrowej oraz energetyki atomowej w Polsce

Dynamiczny rozwój energetyki wiatrowej oraz fotowoltaicznej w większości państw członkowskich Unii Europejskiej na przestrzeni ostatnich kilkunastu lat sprawił, że dyskusja na temat możliwości włączenia potencjału atomowego do

walki o neutralność klimatyczną regionu została zepchnięta na margines. Dyskusji na temat intensyfikacji rozwoju energetyki atomowej w Europie nie ułatwiły również dwie głośne awarie – w Czarnobyli w 1986 r. oraz w Fukushima w 2011 r., które doprowadziły wręcz do deklaracji o wycofaniu atomowych jednostek wytwórczych z kilku krajowych systemów, przede wszystkim w Niemczech [Forbes, 2011]. Decyzja o rezygnacji z rozwoju energetyki atomowej w Unii Europejskiej ma kilka zasadniczych źródeł, których omówienie przekraczałoby objętość niniejszego rozdziału – należy jednak wskazać na współzależność przyczyn polityczno-społecznych, finansowych i operacyjnych. Z uwagi na tematykę i cel niniejszego tekstu, kluczowa jest analiza przyczyn operacyjnych, gdyż obecna jest w debacie eksperckiej teza o niewystarczającej komplementarności energetyki odnawialnej (tak wiatrowej, jak i fotowoltaicznej) z atomowymi mocami wytwórczymi [Lovering, Yip, Nordhaus, 2016, s. 375–376].

Istniejący w Polsce potencjał wytwórczy energetyki atomowej należy do wyjątkowych w skali europejskiej, w związku z czym powinien rozwijać się z wykorzystaniem równie unikalnej strategii. Z uwagi na niekorzystne warunki pogodowe, hydrologiczne oraz geotermalne, krajowe moce wytwórcze oparte zostały na źródłach konwencjonalnych w postaci elektrowni spalających węgiel kamienny oraz brunatny. Dominująca rola węgla – w szczególności brunatnego – wynika również z wysokiej dostępności złóż tego surowca w kilku regionach Polski. Niewystarczające zasoby innych surowców energetycznych, takich jak ropa naftowa czy gaz ziemny, sprawiły, że niemożliwe było oparcie jednostek wytwórczych o generację energii elektrycznej ze źródeł mniej emisyjnych, tak jak np. miało to miejsce w Wielkiej Brytanii, Norwegii lub Holandii, które opierają swą gospodarkę energetyczną w istotnej mierze na produkcji energii poprzez spalanie gazu ziemnego. Zasadne jest zatem rozważenie podjęcia decyzji o uzupełnieniu krajowego miksu energetycznego o jednostki wytwórcze wykorzystujące energię atomową. Znalazło to zresztą odzwierciedlenie w treści *Polityki energetycznej Polski do 2040 roku*, gdzie wskazano, że atomowe moce wytwórcze powinny powstać w przeciągu najbliższych dwudziestu lat i być odpowiedzialne za stworzenie fundamentu krajowego systemu energetycznego, sukcesywnie przejmując rolę odgrywaną dotychczas przez elektrownie węglowe [MKiŚ, 2021, s. 11].

W związku z wahaniami dobowych poziomów zapotrzebowania na energię, moce wytwórcze podłączone do sieci energetycznej funkcjonują w kolejności według rankingu cenowego – od najtańszych do najdroższych, przy czym na wszystkich europejskich rynkach, również polskim, za cenę uznaje się jedynie cenę energii elektrycznej w momencie produkcji według stanu z ostatniej aukcji. W przypadku Polski rolę mocy podstawowej spełniają wciąż elektrownie węglowe i brunatne.

Zobowiązane są one do elastycznych reakcji i zmniejszania lub nawet wyłączenia swoich mocy, jeżeli warunki atmosferyczne są wyjątkowo korzystne dla energetyki OZE. Gdy podaż energii pozostaje zbyt wielka w stosunku do popytu, a import nie jest możliwy z uwagi na wykorzystane maksimum przepustowości istniejących połączeń, operatorom elektrowni konwencjonalnych pozostają dwie opcje: albo całkowite wygaszenie elektrowni na dany okres, co jest procesem szkodliwym i kosztownym, albo utrzymywanie bloków w gotowości, na minimalnym obciążeniu. Jednakże konwencjonalne, węglowe moce wytwórcze zaprojektowano tak, by pracowały przy praktycznie maksymalnej mocy. Każde z rozwiązań oznacza zatem straty techniczne i finansowe, dodatkowo obniżając atrakcyjność cenową energii produkowanej w tych elektrowniach [Popkiewicz, 2021]. Model działania elektrowni atomowych jest bardzo podobny do obecnych zasad funkcjonowania elektrowni węglowych, w związku z czym zasadne jest pytanie o możliwość współpracy energetyki atomowej i odnawialnej w ramach polskiego miks energetycznego.

Pewnym paradoksem gospodarczym, na który należy zwrócić szczególną uwagę, jest również fakt, że pomimo znaczenia energii elektrycznej trudno jest wskazać obszary działalności inwestycyjnej, w których przesądzony zostałby dodatni wynik finansowy osiągniany przez ten sektor. Niepewność zwrotu poniesionych nakładów dotyczy wszystkich sfer energetyki, zarówno bazujących na źródłach konwencjonalnych, jak i niekonwencjonalnych, jednakże problemy inwestycyjne dotyczą energetyki atomowej w szczególności. Przykłady opóźnień w realizacji elektrowni atomowych, przekraczania zaplanowanych budżetów oraz wcześniejszego wyłączenia mocy atomowych są tego wyraźnym przykładem. Warunki funkcjonowania unijnego systemu energetycznego są nierynkowe i w sposób skrajny preferują wyłącznie rozwój energetyki odnawialnej, co jest czynnikiem utrudniającym stworzenie podstaw do racjonalizacji podejmowanych decyzji inwestycyjnych [Juszczuk, Mazur, 2014, s. 95]. Przykładem narastających konfliktów na linii kreacji osi współpracy OZE – atom jest wieloletnia batalia na szczeblu unijnym o włączenie energetyki atomowej do unijnej taksonomii pozwalającej na inwestowanie w energetykę atomową z wykorzystaniem środków publicznych, bez których stworzenie nowych atomowych mocy wytwórczych w ogóle nie jest możliwe [Dalton, 2021]. Dodatkowym elementem sugerującym zmianę podejścia instytucji unijnych do kwestii komplementarności energetyki atomowej i OZE jest niedawne orzeczenie Trybunału Sprawiedliwości Unii Europejskiej w sprawie z powództwa *Republiki Austrii przeciwko Komisji Europejskiej* z dnia 22 września 2020 r., ECLI:EU:C:2020:742, sygnatura C-594/18 P, dotyczące finansowania brytyjskiej elektrowni atomowej *Hinkley Point C* przy wykorzystaniu mechanizmu pomocy publicznej. Trybunał Sprawiedliwości Unii Europejskiej stwierdził, że projekty inwestycyjne rozwijają-

ce elektrownie atomowe mogą liczyć na objęcie programami pomocy publicznej na zasadach analogicznych, jak energetyka odnawialna w postaci fotowoltaiki i elektrowni wiatrowych. Kolejne przewartościowania na unijnej scenie polityki energetycznej powinny skłonić do przyspieszenia prac nad projektem polskiej elektrowni atomowej jako możliwej do sfinansowania na zasadach przyjętych dla kluczowych projektów energetycznych na poziomie wspólnotowym [Teraz Środowisko, 2021a].

Interesujące z polskiego punktu widzenia byłoby również zaangażowanie się w projekty budowy elektrowni atomowych w państwach sąsiednich, które uczestniczyłyby w handlu energią w ramach regionalnych rynków energii, co w szczególności odnosiłoby się do projektu elektrowni atomowej na Litwie oraz na Ukrainie [Świrski, 2019]. Z uwagi na dostępność istniejących (choć wymagających gruntownej restauracji) mostów energetycznych najwyższego napięcia, perspektywa współpracy polsko-ukraińskiej w zakresie handlu energią elektryczną w celu zapewnienia bezpieczeństwa i stabilności dostaw energii jawi się jako szczególnie interesująca, choć w tym momencie brak jakichkolwiek wiążących ustaleń w kwestii przywrócenia importu ukraińskiej energii elektrycznej pochodzącej z elektrowni atomowej Chmielnicka do Polski [Świrski, 2019]. Dodatkowym impulsem byłyby w tym kontekście wola zapewnienia krajowemu systemowi energetycznemu dostaw energii z zewnętrznych systemów wytwórczych, co ograniczałoby ryzyka związane z awariami krajowych jednostek wytwórczych lub sieci przesyłowych, zwiększając dywersyfikację źródeł energii. Rezerwa operacyjna miałaby z każdym rokiem coraz istotniejsze zadania z uwagi na przestarzałość konwencjonalnych mocy wytwórczych i coraz częstsze awarie. Nie bez znaczenia byłaby również możliwość pozyskiwania energii elektrycznej nieobciążonej kosztami emisji gazów cieplarnianych oraz podwyższenie krajowych wskaźników redukcji emisji CO₂. Problem rosnących cen uprawnień do emisji CO₂, które w latach 2013–2021 wzrosły o ponad 1400%, jest niezwykle istotny z polskiej perspektywy [Krajowy Ośrodek Bilansowania i Zarządzania Emisjami (KOBiZE), 2021, s. 10–11], gdyż obecna, wyraźnie fiskalna formuła systemu handlu emisjami sprawia, że państwo modernizujące miks energetyczny ponosi zarówno całość kosztów inwestycyjnych związanych z utworzeniem nowej elektrowni, jak i całość kosztów związanych z poziomem emisji, za który odpowiedzialna jest jednostka zastępowana. Powyższy model utrudnia Polsce realizację odpowiedniej liczby projektów redukujących emisyjność krajowych mocy wytwórczych z uwagi na brak finansowych mechanizmów stymulujących wyłączenie jednostek konwencjonalnych, opóźniając w ten sposób możliwość osiągnięcia celów redukcji emisji.

8. Podsumowanie

Decyzja o prowadzeniu transformacji energetycznej w sposób zmierzający do osiągnięcia neutralności klimatycznej została już w ramach Unii Europejskiej podjęta i nie sposób oczekiwać, że którekolwiek z państw członkowskich będzie mogło prowadzić niezależną politykę energetyczną w oderwaniu od legislacji unijnej. Odpowiedniemu otoczeniu instytucjonalno-prawnemu towarzyszy zresztą coraz silniejsza presja ekonomiczno-społeczna, która sprawia, że inwestycje w energię odnawialną i niskoemisyjną zyskują zarówno na rentowności, jak i popularności w demokratycznych społeczeństwach. Istotną rolę odgrywa również upowszechnienie technologii związanych z OZE, produkcją wodoru, elektromobilnością oraz magazynami energii, które wspólnie pozwolą dalej obniżyć koszty funkcjonowania systemu energetycznego opartego o odnawialne źródła energii.

Polska polityka energetyczna znajdowała się na swoistym rozdrożu przez ostatnich kilkanaście lat, które można w pewnej mierze opisać jako okres niewykorzystanych szans. Deklarowanemu wsparciu technologii OZE towarzyszyła *reguła 10H* oraz restrykcyjne otoczenie legislacyjne w stosunku do możliwości lokalizacji farm fotowoltaicznych na terenach rolnych, a program polskiej elektrowni jądrowej, choć reaktywowany w 2010 r., przez ponad dziesięć lat nie przyniósł owoców chociażby w postaci wyboru lokalizacji pierwszego reaktora.

Prawdopodobnie realnym przełomem – cezurą w przyszłych opracowaniach z zakresu analizy strategii polskiego bezpieczeństwa energetycznego – będzie kryzys ekonomiczno-społeczny spowodowany pandemią wirusa COVID-19 wraz z jego szerokimi konsekwencjami. Na kanwie zmieniającego się w latach 2019–2021 paradygmatu bezpieczeństwa energetycznego i rosnącej świadomości na temat nieuniknionego charakteru przemian technologicznych, dynamicznie postępującej cyfryzacji oraz decentralizacji miejsc pracy zapadły w Polsce decyzje o charakterze być może fundamentalnym dla oceny kierunku dalszego przebiegu krajowego planu transformacji energetycznej. Przyjęcie *Polityki energetycznej Polski do 2040 roku* [MKiŚ, 2021] oraz programu dotyczącego rozwoju energetyki jądrowej, uchwalenie specustawy w zakresie energetyki offshore, a także liberalizacja przepisów w zakresie inwestycji w lądowe odnawialne źródła energii (m.in. rezygnacja ze sztywnej reguły 10H) sprawiają, że dwa filary bezemisyjnego sektora energetycznego w Polsce zostały, przynajmniej w teorii, ostatecznie ustalone.

Powyższe decyzje programowo-polityczne nie oznaczają, że rozwiązane zostały problemy związane z istniejącymi i pojawiającymi się barierami regulacyjnymi i zagrożeniami dla harmonijnego, jednoczesnego rozwoju energetyki opartej o źródła odnawialne (farmy wiatrowe i fotowoltaiczne) oraz niskoemisyjne (elektrownie atomowe) w Polsce.

Głównym zagrożeniem dla nowocześnie rozumianej koncepcji bezpieczeństwa energetycznego nie będą w przyszłości kwestie surowcowe i przestarzałość krajowych mocy wytwórczych, lecz niestabilność ram prawnych, nieprzewidywalność kierunków inwestycji i podejmowanie decyzji inwestycyjnych w oparciu o koniunkturę polityczną. Wydaje się, że największym wyzwaniem będzie umiejętność wykorzystania rozwoju nowych technologii wodorowych oraz akumulatorowych, jak również dostosowanie sieci przesyłowych do interoperacyjnego zarządzania krajowymi mocami wytwórczymi, które muszą tworzyć jednolity, responsywny system zbudowany na zasadzie *naczyń połączonych*. Niezbędne będzie wzięcie pod uwagę nieliniowego rozwoju energetyki, nieprzewidywalności efektywności technologii będących we wstępnym stadium rozwoju komercyjnego oraz integracja strony popytowej sektora energetycznego, a zatem przemysłu, transportu i konsumpcji [van der Burgt, 2017, s. 10–11] w jeden ponadnarodowy organizm tworzący przewagę konkurencyjną polskiej gospodarki w ramach rynku unijnego.

Jednakże próby odkładania w czasie decyzji o poniesieniu znaczących kosztów inwestycyjnych w rozwiązania wspierające wytwarzanie energii ze źródeł odnawialnych, a – w szerszym ujęciu – całe programy sprzyjające rozwojowi sektora zielonej gospodarki – spowodowały, że gospodarki ponoszą coraz wyższy koszt alternatywny wynikający z zaniechania przeprowadzenia transformacji w szybszym tempie, przez co zmniejszony został potencjał inwestycyjny przedsięwzięć. Inwestycje w ten sektor powinny wynikać obecnie już bezpośrednio z uwarunkowań ekonomicznych, tymczasem okazuje się, że bez mechanizmów wsparcia z zakresu pomocy publicznej niemożliwe jest tworzenie zarówno farm wiatrowych offshore, jak i elektrowni atomowych. Jedynym finansowym mechanizmem quasi-rynkowym sprzyjającym rezygnacji z konwencjonalnych mocy wytwórczych są rosnące koszty uprawnień do emisji dwutlenku węgla, który wzmacnia nieopłacalność polityki energetycznej opartej na źródłach emisyjnych [Komisja Europejska (KE), 2020].

Ponadto stan dyskusji na temat przyszłości polskiej polityki energetycznej wskazuje na brak świadomości, że program implementacji energetyki atomowej do krajowego systemu energetycznego nie koliduje z dynamicznym rozwojem energetyki odnawialnej, lecz stanowi jego istotny fundament poprzez gwarancję stabilnej podstawy mocy wytwórczych. Obecna formuła polskiej polityki energetycznej nie dostrzega nowych wymiarów bezpieczeństwa energetycznego, koncentrując się jedynie na klasycznie rozumianym problemie stabilności dostaw surowcowych i możliwości kontroli mocy krajowych jednostek wytwórczych. Brak również odpowiedniego podejścia do możliwości handlowych i bilansujących, jakie niesie za sobą udział w regionalnych rynkach energii oraz rozwój transgranicznych sieci przesyłowych z państwami Europy Środkowo-Wschodniej (w szczególności Ukrainy).

Podsumowując, największym wyzwaniem stojącym przed polską strategią transformacji energetycznej będzie umiejętność zarządzania nieelastycznością odnawialnych źródeł energii, w szczególności w odniesieniu do scentralizowanych, wielkoskalowych morskich farm wiatrowych (offshore). Problemami mogą być też rosnące koszty utworzenia i utrzymania odnawialnych źródeł energii z uwagi na zmniejszającą się powierzchnię miejsc sprzyjających ich instalacji – dotyczy to zarówno lokalizacji nowych hydroelektrowni, jak i lądowych i morskich elektrowni wiatrowych. Z tą kwestią wiążą się ogólne problemy związane z pozyskaniem dofinansowania ze środków publicznych lub rentownością inwestycji, które nie działają na zasadach rynkowych i nie sposób przewidzieć, jak będzie kształtować się horyzont czasowy zwrotu z inwestycji.

Ostatnim z problemów jest kwestia umiejętnego zarządzania siecią na poziomie transnarodowym, która będzie w stanie reagować na dynamiczne zmiany popytowo-podażowe na rynku energii poprzez integrację regionalnych rynków energii i przygotowanie odbiorców energii elektrycznej do funkcjonowania w oparciu o mechanizmy DSR oraz wykorzystanie magazynów energii. Na drodze do płynnego rozwoju polskiego systemu energetycznego w oparciu o energetykę odnawialną, elektrownie jądrowe i magazyny energii wspierane importem w ramach transgranicznych sieci przesyłowych stoją trudne do przezwyciężenia bariery o charakterze instytucjonalno-ekonomicznym oraz psychologiczno-społecznym, w stosunku do których Polska wciąż nie wypracowała jednolitych rozwiązań. Kompleksowych rozwiązań brakuje również na arenie międzynarodowej, w szczególności na szczeblu unijnym, gdzie partykularne interesy poszczególnych państw lub koncernów z sektora energetyki konwencjonalnej blokują najbardziej efektywne kierunki modernizacji sieci przesyłowych i regionalnych rynków energii, utrudniając realizację celu neutralności klimatycznej i przeciwdziałanie zmianom klimatu.

Trwająca od dnia 24 lutego 2022 r. wojna na terenie Ukrainy drastycznie zmieniła sytuację geopolityczną i ekonomiczną Europy, stając się kolejną cezurą dla dyskusji o przyszłości regionalnej polityki energetycznej i kierunkach jej rozwoju, stawiając na czele – jak można mieć nadzieję – dywersyfikację źródeł wytwórczych.

Z punktu widzenia polskiej polityki energetycznej i stojących przed nią wyzwań, kluczowe stają się natomiast pytania o możliwość jak najszybszej intensyfikacji współpracy regionalnej w ramach rynku energii elektrycznej. Wiele z postulatów podniesionych w niniejszym rozdziale już zaczyna się urzeczywistniać, czego przykładem może być chociażby decyzja o restauracji polsko-ukraińskiego mostu energetycznego najwyższego napięcia i przyjęciu Ukrainy do europejskiej sieci przesyłowej.

Należy spodziewać się, że z uwagi na destabilizację międzynarodowych rynków surowcowych znacząco zwiększy się tempo transformacji energetycznej w Polsce w kierunku energetyki odnawialnej i niskoemisyjnej (atomowej).

Bibliografia

- Baker, P. (2017). *EU must take regionalisation of electricity markets a step further*, Energy Post, <https://energypost.eu/15800-2/> (dostęp: 31.07.2021).
- Bassi, S., Bowen, A., Fankhauser, S. (Eds.). (2012). *The case for and against onshore wind energy in the UK*. London: London School of Economics.
- Bronk, L., Magulski, R., Maćkowiak-Pandera, J. (2019). *Elastyczność krajowego systemu elektroenergetycznego. Diagnoza, potencjał, rozwiązania*. Warszawa: Forum Energii.
- van der Burgt, J. (2017). *Flexibility in the power system. The need, opportunity, and value of flexibility*. Arnhem: DNV GL Group.
- Cecchinato, M. (2019). *Boosting offshore wind energy in the Baltic Sea*. Brussels: Wind Europe.
- Dalton, D. (2021). *Commission announces plans to include nuclear energy*, NUCNET. The Independent Nuclear News Agency, https://www.nucnet.org/news/commission-announces-plans-to-include-nuclear-energy-4-3-2021?fbclid=IwAROKbYnCb5dJFsPfSys-L5fZqth_aWvcd5FS_OIOsEF5C_m0fb3YYIJSpqLY (dostęp: 31.07.2021).
- Deign, J. (2020). *Germany's maxed-out grid is causing trouble across Europe*, Green Tech Media, <https://www.greentechmedia.com/articles/read/germanys-stressed-grid-is-causing-trouble-across-europe> (dostęp: 31.07.2021).
- Dennison, S., Loss, R., Soderstrom, J. (Eds.). (2021). *Europe's Green Moment: How to meet the climate change*, Policy brief no. ECFR/386. London: European Council on Foreign Relations.
- Diesendorf, M. (2016). *Dispelling the nuclear 'baseload' myth: nothing renewables can't do better!*, Renew Economy, <https://reneweconomy.com.au/dispelling-the-nuclear-baseload-myth-nothing-renewables-cant> (dostęp: 31.07.2021).
- Dyląg, B. (2021). *Fotowoltaika rozwija się szybciej, niż zakładał rząd*, Wirtualny Nowy Przemysł – portal gospodarczy, <https://www.wnp.pl/energetyka/fotowoltaika-rozwija-sie-szybciej-niz-zakladal-rzad,466152.html> (dostęp: 31.07.2021).
- Eising, M., Hobbie, H., Möst, D. (2020). *Future wind and solar power market values in Germany – Evidence of spatial and technological dependencies?*, *Energy Economics*, 86.
- Forbes.pl (2011). *Niemcy zatwierdziły odejście od atomu w 2022*, Forbes, <https://www.forbes.pl/wiadomosci/niemcy-zatwierdzily-odejscie-od-atomu-w-2022/zvncnfl> (dostęp: 31.07.2021).
- GLOBEnergia (2021). *System prosumencki negatywnie odbija się na wynikach dużych spółek energetycznych?*, <https://globenergia.pl/system-prosumencki-negatywnie-odbija-sie-na-wynikach-duzych-spolek-energetycznych/> (dostęp: 31.07.2021).
- Gryszczuk, A. (2021). *Pierwsza w Polsce mikrościeć bazująca na OZE i magazynach energii*, Gram w Zielone, <https://www.gramzielone.pl/trendy/104656/pierwsza-w-polsce-mikrosiec-bazujaca-na-oze-i-magazynach-energii> (dostęp: 31.07.2021).

- Jasiak, K. (2017). *Szanse i wyzwania państw Unii Europejskiej w dobie globalizacji korporacyjnej*. Poznań: Uniwersytet Adama Mickiewicza w Poznaniu.
- Jędrasik, I. (2018). Wpływ interesariuszy w polityce energetycznej – przykład kampanii na rzecz energetyki prosumenckiej. W: *Aktorzy i interesy w publikach publicznych w Unii Europejskiej*, U. Kurczewska (red.). Warszawa: Oficyna Wydawnicza SGH.
- Juszczuk, M. (2016). Koncepcja unii energetycznej krajów członkowskich UE jako próba odpowiedzi na potrzeby inwestycyjne w warunkach instytucjonalizacji narzędzi na rzecz rozwoju infrastruktury przesyłowej w sektorze energii elektrycznej. W: *Europejska Unia Energetyczna. Wybrane zagadnienia*, G. Wojtkowska-Łodej (red.). Warszawa: Oficyna Wydawnicza SGH.
- Juszczuk, M., Mazur, B. (2014). Modelowanie kierunków rozwoju systemu elektroenergetycznego wobec uwarunkowań szerokiego otoczenia w Europie, w tym wobec dylematów rozwoju energetyki w Niemczech. W: *Dylematy rozwoju infrastruktury*, J. Gajewski, W. Paprocki (red.). Gdańsk: Europejski Kongres Finansowy.
- Komisja Europejska (2020). *Sprawozdanie Komisji dla Parlamentu Europejskiego, Rady, Europejskiego Komitetu Ekonomiczno-Społecznego i Komitetu Regionów. Ceny i koszty energii w Europie*, sprawozdanie COM (2020) 951 final z dnia 14.10.2020.
- Krajowy Ośrodek Bilansowania i Zarządzania Emisjami (2021). *Raport z rynku CO2 – marzec 2021*, 3.
- Książkowski, K., Kotlewski, D., Maśloch, G. (red.). (2020). Energetyka odnawialna – wyzwanie dla krajów Europy Środkowo-Wschodniej. W: *Raport SGH i Forum Ekonomicznego 2020*. Warszawa.
- Lovering, J., Yip, A., Nordhaus, T. (2016). Historical construction cost of global nuclear power reactors, *Energy Policy*, 91.
- Maćkowiak-Pandera, J., Rubczyński, J. (red.). (2020). *Polska neutralna klimatycznie 2050. Elektryfikacja i integracja sektorów*. Warszawa: Forum Energii.
- Ministerstwo Klimatu i Środowiska (2021). *Polityka energetyczna Polski do 2040 roku*, Załącznik do obwieszczenia Ministra Klimatu i Środowiska z dnia 2 marca 2021 r., Dz. Urz. RP z 2021 r. poz. 264.
- Mirowski, T. (2017). Wybrane problemy związane z wykorzystaniem odnawialnych źródeł energii w Polsce, *Zeszyty Naukowe Instytutu Gospodarki Surowcami Mineralnymi i Energią Polskiej Akademii Nauk*, 98.
- Mrozowska, S. (2020). *Rozwój offshore wind wymaga dobrych procedur komunikacyjnych i partycypacyjnych*, GLOBenergia, <https://globenergia.pl/mrozowska-rozwoj-offshore-wind-wymaga-dobrych-procedur-komunikacyjnych-i-partycypacyjnych/> (dostęp: 31.07.2021).
- Najwyższa Izba Kontroli (2014). *Funkcjonowanie i bezpieczeństwo elektroenergetycznych sieci przesyłowych*, <https://www.nik.gov.pl/plik/id,6632, vp,8426.pdf> (dostęp: 31.07.2021).
- Najwyższa Izba Kontroli (2018). *Rozwój sektora odnawialnych źródeł energii*, KGP.430.022.2017. Warszawa.
- Niedziółka, D. (2011). *Regionalizacja rynków energii*. Warszawa: Oficyna Wydawnicza SGH.
- Niedziółka, D. (2018). *Funkcjonowanie polskiego rynku energii*. Warszawa: Difin.
- Pinkas, A. (2021). Pomoc publiczna skierowana na rozwój morskiej energetyki wiatrowej w Polsce realizacją unijnego celu neutralności klimatycznej do 2050 r., *Samorząd Terytorialny*, 6.

- PKN Orlen (2021). *PKN Orlen selects strategic partner for offshore wind farm project*, <https://www.orklen.pl/EN/PressOffice/pages/pkn-orklen-selects-strategic-partner-for-offshore-wind-farm-project.aspx> (dostęp: 31.07.2021).
- Polskie Sieci Elektroenergetyczne (2020). *Dane systemowe. Raport za 2020 rok*, <https://www.pse.pl/dane-systemowe/funkcjonowanie-rb/raporty-roczne-z-funkcjonowania-rb-za-rok/raporty-za-rok-2020> (dostęp: 31.07.2021).
- Popkiewicz, M. (2021). *OZE czy/i atom? Transformacja dla opornych*, Wysokie Napięcie, źródło: <https://wysokienapiecie.pl/36404-oze-czy-atom-transformacja-dla-opornych/> (dostęp: 31.07.2021).
- Purta, M., Szarek, G., Engel, H., Speelman, E., van der Pluijm, P. (red.). (2020). *Neutralna emisyjnie Polska 2050. Jak wyzwanie zmienić w szansę*. Warszawa: McKinsey & Company.
- Ropuszyńska-Surma, E., Węglarz, M. (2017). Bariery rozwoju rozproszonej energetyki odnawialnej w świetle badań ankietowych, *Przegląd Elektrotechniczny*, 4.
- Rubene, I. (2018). The role of energy prices in recent inflation outcomes: a cross-country perspective, *ECB Economic Bulletin*, 7.
- Smyrgała, D. (red.). (2016). *Między regionalnymi rynkami energii a technologią – przyszłość sektora energetycznego w Polsce*. Warszawa: Collegium Civitas.
- Sobczyk-Grygiel, S. (2021). *Zyska: Zmiany dla prosumentów najwcześniej w połowie 2022 r.* [wywiad], *Dziennik Gazeta Prawna*, https://serwisy.gazetaprawna.pl/energetyka/artykuly/8195234_ireneusz-zyska-sekretarz-stanu-zmiany-dla-prosumentow-najwcześniej-w-polowie-2022-r.html (dostęp: 31.07.2021).
- Sovacool, B. (2020). *Nuclear vs renewables: what's better for climate mitigation?*, Steps-Centre, <https://steps-centre.org/blog/nuclear-vs-renewables-whats-better-for-climate-mitigation/> (dostęp: 31.07.2021).
- Szpak, K. (2020). Polityka klimatyczna Unii Europejskiej w perspektywie 2050 roku. W: *Polityka klimatyczna i jej realizacja w pierwszej połowie XXI wieku*, J. Gajewski, W. Paprocki (red.). Sopot: Centrum Myśli Strategicznych – Europejski Kongres Finansowy.
- Świrski, K. (2019). *Polska i Ukraina są skazane na energetyczną współpracę*, Rzeczpospolita, <https://energia.rp.pl/opinie/17673-polska-i-ukraina-sa-skazane-na-energetyczna-wspolprace> (dostęp: 31.07.2021).
- Teraz Środowisko (2021a). *Energetyka jądrowa w zielonej taksonomii?*, <https://www.teraz-srodowisko.pl/aktualnosci/energetyka-jadrowa-Zielony-Lad-taksonomia-10146.html> (dostęp: 31.07.2021).
- Teraz Środowisko (2021b). *Energetyka wiatrowa w Polsce. Rozwój, wyzwania, perspektywy*. Warszawa.
- Terlikowski, P., Paska, J. (2018). Metodyka wyznaczania transgranicznych zdolności przesyłowego KSE, ze szczególnym uwzględnieniem połączeń transgranicznych, *Przegląd Elektrotechniczny*, 3.
- Thema Consulting Group (2013). *Loop flows – final advice*. Thema Report 2013–36, Thema Consulting Group 2013, https://ec.europa.eu/energy/sites/ener/files/documents/201310_loop-flows_study.pdf (dostęp: 31.07.2021).
- Torriti, J., Hassan, M., Leach, M. (2009). *Demand response experience in Europe: Policies, programmes and implementation*. Guildford: University of Surrey.

- Trybunał Sprawiedliwości Unii Europejskiej (2020). *Republika Austrii przeciwko Komisji Europejskiej*, wyrok z dnia 22.09.2020, ECLI:EU:C:2020:742, C-594/18 P.
- Urząd Regulacji Energetyki (2020). *Raport krajowy Prezesa Urzędu Regulacji Energetyki 2020*. Warszawa.
- US Environmental Protection Agency (2021). *Overview of Greenhouse Gases*, <https://www.epa.gov/ghgemissions/overview-greenhouse-gases> (dostęp: 31.07.2021).
- Wasiuta, A. (2015). Identyfikacja i oszacowanie barier wzrostu energetyki z odnawialnych źródeł energii. W: *Wybrane problemy administracji publicznej. Prawo – Zarządzanie – Polityka*, W. Mikołajczewska, P. Kierończyk (red.). Gdańsk: Wydawnictwo Gdańskiej Szkoły Wyższej.
- Wojtkowska-Łodej, G. (2014). Wyzwania klimatyczne i energetyczne a polityka Unii Europejskiej, *Polityka Energetyczna*, 17(3).
- Żylińska J. (2020). *Offshore celem dla państwowych koncernów*, Dziennik Gazeta Prawna, https://serwisy.gazetaprawna.pl/energetyka/artykuly/1469663_offshoer-energetyka-koronawirus-wegiel-gornictwo-oze.html (dostęp: 31.07.2021).

Development of Offshore Wind Energy in Poland. Key Regulatory Barriers and Threats

Abstract

The aim of the article is to assess the regulatory barriers to the development of offshore wind energy in Poland from the point of view of the energy transformation strategy and challenges in the process of developing low-emission sources. The study analyses the development of offshore technology in Europe and its growing role in the domestic production capacities with the use of comparative studies based on current research results and sectoral reports. The risk of regulatory barriers was assessed as high due to the lack of transmission infrastructure capable of accommodating additional generation capacity, a centralised model of energy infrastructure management, competitiveness of conventional generation capacities and insufficient development of cross-border networks. It was indicated that a long-term strategy for the energy transformation should be pursued, with the awareness of unequal economic potential of the countries participating in the transformation process. To implement the expected model of energy transformation, it will be necessary to connect national energy markets (regionalisation) and remove regulatory barriers inhibiting the development of renewable and low-emission energy.

Keywords: energy security, renewable energy sources (RES), energy policy, regionalisation of energy markets, energy transformation

Weighting Schemes for Momentum Trading Strategies in the CEE Currency Universe

Radosław Jawor

Abstract

In this article, we construct portfolios using momentum strategies and we analyse the impact of different weighting schemes and buy/sell signals on the performance of these strategies. The strategies are based on a set of four Central and Eastern European currencies: the Polish zloty (PLN), the Czech koruna (CZK), the Hungarian forint (HUF), and the Romanian leu (RON). We take the perspective of a Eurozone investor, so the base currency is the euro (EUR) instead of the US dollar. We find out that the momentum strategy based on the lagged total return, with a formation period of 12 months and with the most recent month omitted, performs best. When a constant weighting scheme is applied, such a portfolio reaches a mean annualised return of 1.99% and a Sharpe ratio of 0.58. Our results show that more complex weighting schemes do not cause any essential improvement of any strategy performance. We also find that momentum strategies based on lagged FX returns are not profitable.

Keywords: CEE currencies, Eurozone, foreign exchange, momentum strategy, portfolio weighting schemes

1. Introduction

Momentum strategies are trend-following trading strategies assuming that past high returns will continue to be high and past low returns will continue to be low. As a result, investors' decision whether to buy or sell a particular asset depends on its past performance. In the foreign exchange (FX) market, investors buy (take a long position) on currencies which appreciated or sell (take a short position) on currencies which depreciated in the strategy formation period, also known as the look-back period. The strategy formation period is usually set to values ranging from 1 to 12 months.

The implementation of an investment strategy requires, apart from the algorithm which generates buy/sell signals, the choice of a weighting scheme. The simplest constant scheme sorts currencies according to their buy/sell signal strengths given by past returns. Then weights $1/n$ are assigned to n currencies with the highest returns and weights $-1/n$ are assigned to n currencies with the lowest returns. However, there are many other possible weighting schemes. In this article, we analyse three different weighting schemes and their impact on the performance of momentum strategies.

Compared with the literature, where most liquid and heavily traded G10 currencies¹ are usually analysed from the perspective of a US dollar investor [Brunnermeier et al., 2008; Lustig et al., 2011; Novy-Marx, 2012; Menkhoff et al., 2012; Asness et al., 2013; Jurek, 2014], our analysis is non-standard. We use four currencies of Central and Eastern European (CEE) countries – the Polish zloty (PLN), the Czech koruna (CZK), the Hungarian forint (HUF), and the Romanian leu (RON) – and assume a euro investor perspective. Since all four countries are members of the European Union, their currencies are usually traded against the euro, not the US dollar. Moreover, their economies are mainly linked to West European economies, especially German, which mostly use the euro as their home currencies. Finally, all four countries are closely related, as Poland, the Czech Republic and Hungary, together with Slovakia, constitute the Visegrad Group (V4) that facilitates mutual international cooperation. Romania also maintains a close relationship with other CEE countries, e.g. via the Three Seas Initiative. These four currencies were already analysed by e.g. Jaworski [2021], but from the exchange rates forecasting perspective or also included in broader studies, e.g. by Menkhoff et al. [2012] on a set of total 48 currencies. However, to the best of our knowledge, the euro-based FX trading strategies have never been evaluated on this currency set.

Several similar non-US dollar based studies on currency trading strategies can be found in the literature. For instance, Mikova et al. [2020] implement strategies with the Russian ruble (RUB) as a base currency instead of the US dollar. They conduct an analysis using a currency set composed of G10 currencies and 16 emerging markets currencies (PLN and HUF included). Also, Hsu and Chen [2021] study momentum strategies with the Chinese yuan (RMB) as a base currency and analyse the currencies of countries from China's Belt and Road Initiative – including two out of four our currencies (CZK and RON).

¹ The G10 currencies are: the Australian dollar (AUD), the Canadian dollar (CAD), the euro (EUR), the Japanese yen (JPY), the New Zealand dollar (NZD), the Norwegian krone (NOK), the pound sterling (GBP), the Swedish krona (SEK), the Swiss franc (CHF) and the US dollar (USD).

With regard to the results on momentum strategies, Menkhoff et al. [2012] find that the momentum strategy with a one-month formation and holding period has the best performance in the sample covering years 1976–2010. This strategy achieved an average return of 9.46% and a Sharpe ratio of 0.95. Asness et al. [2013] construct portfolios with constant and ranking weighting schemes setting the holding period to 12 months and argue that omitting the most recent month in signal computation increases the profitability. They also state that momentum portfolios with a constant weighting scheme are slightly better than portfolios with a ranking weighting scheme. The former portfolio reaches an average return of 3.5% and a Sharpe ratio of 0.34, while the latter reaches a return of 3.0% and a Sharpe ratio of 0.32. It is worth noting that the portfolio with a ranking weighting scheme has also a lower standard deviation of returns. Momentum portfolios with the Chinese yuan as a base currency investigated by Hsu and Chen [2021] have the best performance when both formation and holding period are set to one month. This portfolio reaches an average return of 5.16% and a Sharpe ratio of 0.57. Mikova et al. [2020] apply momentum strategies with the Russian ruble as a base currency and with a formation period set to 1, 3, 6, 9, and 12 months. They conclude that a one-month formation period has the best performance. However, momentum portfolios are profitable only during a period of a fixed-rate ruble (2003–2014). When portfolios are evaluated on a floating ruble period (2015–2018), losses are generated.

There is a twofold contribution of our article to the literature. First, we comprehensively investigate how the choice of the weighting schemes and the way of signal computation affect the performance of momentum portfolios. Second, we analyse the trading potential of a regional set of four CEE currencies from the perspective of the euro investors. We find out that the choice of signal computation method does make a significant impact on strategies performance. On the other hand, complex weighting schemes do not improve strategies performance compared to a simple one. Momentum strategies can provide profit opportunities, which, however, depends strongly on the type of signal and the choice of formation period.

The remaining part of the article is structured as follows. Section 2 presents data used in our analysis. Section 3 describes the adopted methodology – calculation of excess returns, buy/sell signals, portfolio construction, weighting schemes, and performance evaluation. Section 4 presents the performance of momentum portfolios and analyses the impact of the use of different weighting schemes. Section 5 concludes the study.

2. Data

Our sample data sourced from Refinitiv Eikon Datastream consist of daily bid and ask spot rates and bid and ask spreads between spot and one-month forward rates of the Polish zloty, the Czech koruna, the Hungarian forint, and the Romanian leu against the euro from 15 June 2012 to 28 February 2021. The rates are expressed as foreign currency units per one euro. We conduct our analysis on monthly data; so, in order to change the data frequency, we take the last observation of every month as e.g. in [Burnside et al., 2011a]. The advantage of this approach is that the data are not averaged over each month but instead represent the rate on the last trading day of every month [Filippou et al., 2018].

The mid rates, i.e. arithmetic means of bid and ask quotes, are calculated as

$$S_t = \frac{S_t^a + S_t^b}{2}, \quad \text{spread}_t = \frac{\text{spread}_t^a + \text{spread}_t^b}{2},$$

where S_t is the mid spot rate in time t expressed as foreign currency units per one euro, and spread_t – the mid spread between spot and one-month forward rate in time t expressed in *pips*. Superscripts a and b indicate ask and bid quotes, respectively. We calculate forward mid rates

$$F_t = S_t + \frac{1}{10000} \text{spread}_t.$$

Thus, F_t is a one month forward mid rate in time t expressed as foreign currency units per one euro. The use of one-month forward rates is a standard in the literature, see e.g. [Burnside et al., 2011b, Menkhoff et al., 2012, Barroso and Santa-Clara, 2015].

Finally, to adopt a Eurozone investor perspective, i.e. the euro is the base currency, we invert spot and forward rates

$$s_t = \frac{1}{S_t}, \quad f_t = \frac{1}{F_t},$$

resulting in s_t and f_t being spot and forward mid rates expressed as the euro units per one unit of a foreign currency. Table 1 reports descriptive statistics of single currency returns calculated using s_t and f_t according to (1). Two currencies – PLN and RON have positive mean returns, while CZK and HUF, when going long, on average, generate losses. RON has the smallest standard deviation of returns and is the only currency, whose returns have positive skewness.

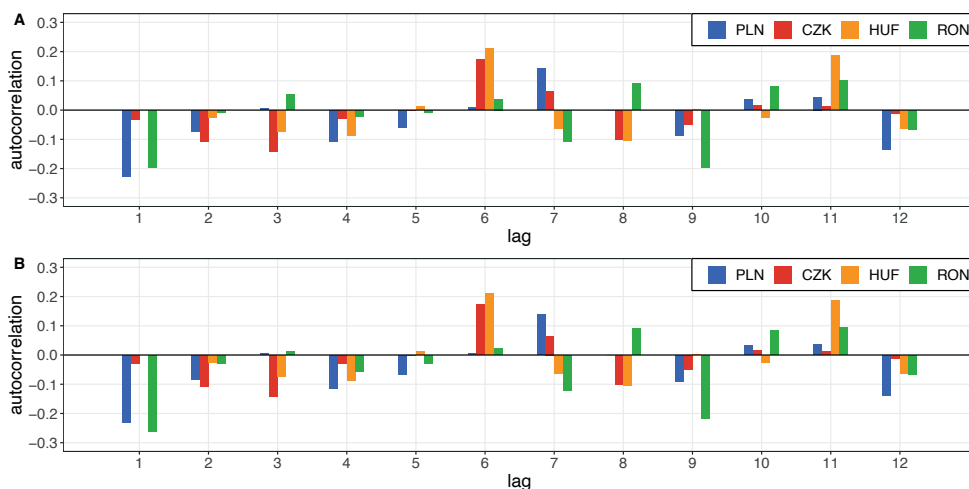
Table 1. Descriptive statistics and normality test of single currency returns

	PLN	CZK	HUF	RON
Minimum	-61.33	-80.54	-78.05	-29.33
Maximum	40.74	38.93	35.88	34.90
Mean	1.57	-0.15	-2.54	1.64
Std dev.	5.30	4.25	5.72	2.84
Skewness	-12.24	-74.67	-20.37	14.08
Kurtosis	4.23	107.40	9.37	13.46
J-B test stat.	4.27	843.73	14.25	15.58
J-B test p-value	0.12	0.00	0.00	0.00

Note: All descriptive statistics are annualised and expressed in percentages. The J-B test stands for the Jarque-Bera test for normality. We report the test statistics and p-value.

Source: Author’s own elaboration.

Figure 1. Autocorrelation of single currency total returns (A) and FX returns (B)



Note: The negative first order autocorrelation of PLN, CZK and HUF makes clear why momentum buy/sell signals are more effective if we skip most recent month in signals computation.

Source: Author’s own elaboration.

3. Methodology

In this part of the article, we present how we calculate excess returns and buy/sell signals. Subsequently, we describe weighting schemes and portfolio construction methods. Finally, we show the performance evaluation measures used in this work.

3.1. Excess returns

We calculate ordinary excess return r_t in time t of a single currency

$$r_t = \frac{s_t - f_{t-1}}{s_{t-1}} \quad (3.1)$$

which can be easily decomposed into two components. The first is the IR return – representing profits or losses from the interest rate differential between Euro-zone and a particular country (assuming the covered interest parity – CIP), and the second – the FX return – stands for profits or losses from changes of spot rate. We denote these two components of return as r^{IR} and r^{FX} , respectively

$$r_t = \frac{s_t - f_{t-1}}{s_{t-1}} = \frac{s_{t-1} - f_{t-1}}{s_{t-1}} + \frac{s_t - s_{t-1}}{s_{t-1}} = r_t^{IR} + r_t^{FX}.$$

3.2. Buy/sell signals

Buy/sell signals in momentum strategies are based on lagged returns. However, shown before, returns can be decomposed into two components. Therefore, we consider two types of buy/sell signals: the first type based on lagged total returns, and second type based on lagged FX returns. We do not compute signals based on IR returns because this would result in creating the carry trade type strategy and not the momentum strategy. Having two types of signals, we can adjust their formation periods. Menkhoff et al. [2012] showed that the momentum strategies with a holding period set to one month achieve the best performance in terms of average returns and the Sharpe ratio. Following their findings, we set the holding period to 1 month. With regard to the formation period, we set it to 1, 3, 6, 9, and 12 months, so the momentum signals are computed as cumulative lagged returns. The signal of currency k in time t for the momentum strategy MOM_T with T -month formation period is thus

$$\text{signal}_{k,t}^{MOM_T} = \prod_{i=1}^T (r_{k,t-i} + 1) - 1,$$

where $T \in \{1, 3, 6, 12\}$ is the formation period length, i.e. the number of monthly returns from the past that are used, and r_k can be a total or FX return. Asness et al. [2013] drop the most recent month returns to avoid a one-month reversal in returns, which also exists in our sample, see Figure 1. We follow their approach and create

other types of signals – based again on a total or FX return but with most recent month skipped – denoted by MOM_T^{-1}

$$\text{signal}_{k,t}^{\text{MOM}_T^{-1}} = \begin{cases} \text{signal}_{k,t-1}^{\text{MOM}_1} = r_{k,t-2}, & \text{if } T = 1, \\ \text{signal}_{k,t-1}^{\text{MOM}_{T-1}} = \prod_{i=2}^T (r_{k,t-i} + 1) - 1, & \text{if } T \neq 1. \end{cases}$$

Let us stress that here T is no longer the number of monthly returns from the past as the most recent return is omitted. These signals use $T - 1$ monthly returns, with the exception of $T = 1$, when the buy/sell signal is just a single monthly return lagged by two periods.

3.3. Portfolio construction and weighting schemes

Once we have computed buy/sell signals for all the currencies, we sort them decreasingly. We construct euro-neutral portfolios, i.e. portfolios which have a zero initial net value in the euro. In the case of constant and ranking weighting schemes, we always take a long position in two currencies with the highest signal values and a short position in two with the lowest signal values. In the case of proportional weighting scheme, there is a possibility that there are three currencies in one portfolio leg and only one currency in the other.

The *constant* weighting scheme assigns weights $1/2$, $1/2$, $-1/2$, and $-1/2$ according to the order of signals. Weights in both portfolio legs sum up to 1 and -1 , respectively, resulting in a portfolio net value of zero.

$$w_{k,t}^{\text{CONST}} = \begin{cases} 1/2 & \text{if } \text{rank}(\text{signal}_{k,t}) \in \{1,2\}, \\ -1/2 & \text{if } \text{rank}(\text{signal}_{k,t}) \in \{3,4\}, \end{cases}$$

where $\text{rank}(\cdot)$ indicates the position of a currency after sorting, so the rank of the currency with the highest signal value is 1, and the rank of the currency with the lowest signal value is 4.

Apart from the simple constant weighting scheme, we consider the *ranking* weighting scheme as described by Asness et al. [2013]. This weighting scheme was applied also by Kojien et al. [2013], Menkhoff et al. [2017], Dahlquist and Hasseltoft [2020]. The weight on currency k at time t is then

$$w_{k,t}^{\text{RANK}} = c_t \left(\text{rank}(\text{signal}_{k,t}) - \frac{1}{N} \sum_k \text{rank}(\text{signal}_{k,t}) \right), \quad (3.2)$$

where N is the number of assets and c_t is a parameter scaling portfolio to one base currency unit long and one base currency unit short. For our four-currency universe, (3.2) simplifies into

$$w_{k,t}^{\text{RANK}} = \frac{1}{2} \left(\text{rank}(\text{signal}_{k,t}) - \frac{1}{4} \sum_k \text{rank}(\text{signal}_{k,t}) \right),$$

so the ranking weights are $3/4, 1/4, -1/4, -3/4$ indicating more exposure to currencies with stronger signals.

The third and the last weighting scheme depends on the deviation from the average value of signal in time t . This method is similar to the one of Daniel et al. [2017] used for carry trade strategies. We call this scheme *proportional*

$$w_{k,t}^{\text{PROP}} = 2 \frac{\text{signal}_{k,t} - \overline{\text{signal}}_t}{\sum_{k=1}^4 |\text{signal}_{k,t}|},$$

where $\overline{\text{signal}}_t$ is the average value of signal in time t

$$\overline{\text{signal}}_t = \frac{1}{4} \sum_{k=1}^4 \text{signal}_{k,t},$$

and a multiplication of weights by 2 is required in order to set the weights sums of both portfolio long and short legs as 1 and -1 – it is a scaling parameter. As mentioned before, when using the proportional weighting scheme, there is a possibility that the number of currencies in long and short legs will not be equal. Moreover, there is a very small, but non-zero probability of an exclusion of a currency from the portfolio structure. This could take place when a signal of a particular currency would be exactly equal to the average value of all four signals.

3.4. Performance evaluation

Apart from the descriptive statistics like mean, standard deviation, skewness, and kurtosis, we also evaluate portfolio performance using three commonly used measures: the Sharpe ratio, the Sortino ratio and the maximum drawdown.

The Sharpe ratio is defined as the excess return of portfolio divided by the standard deviation of returns [Sharpe, 1966]

$$\text{Sharpe} = \frac{\bar{r}}{\sigma},$$

where \bar{r} is the mean return of portfolio and σ is the standard deviation of portfolio returns. In the analysis of stock market portfolios, the excess return in the numerator of the Sharpe ratio is usually expressed as the portfolio return reduced by the risk-free rate. Nevertheless, in the case of foreign exchange and zero-investment portfolios, all the accrued profit is an excess return. As a consequence, there is no need to subtract the risk-free rate from the returns.

The Sortino ratio is a variation of the Sharpe ratio. Instead of the risk-free rate, Sortino and Price [1994] used the target return (or minimum acceptable return, MAR) of a portfolio, and instead of the standard deviation of returns, they use the downside deviation (DD) which rises if and only if the return is below the target return

$$\text{Sortino} = \frac{\bar{r} - \text{MAR}}{\text{DD}}.$$

We set the MAR to 0, as for zero-investment portfolios and excess returns, any positive value of average portfolio returns means we seize a profit opportunity by well-read buy/sell signals.

Finally, the maximum drawdown measures the decline from the most recent peak to the trough of a portfolio before a new peak is achieved. It represents the maximum loss an investor can suffer by buying at the highest point and selling at the lowest [Bacon, 2008].

4. Results

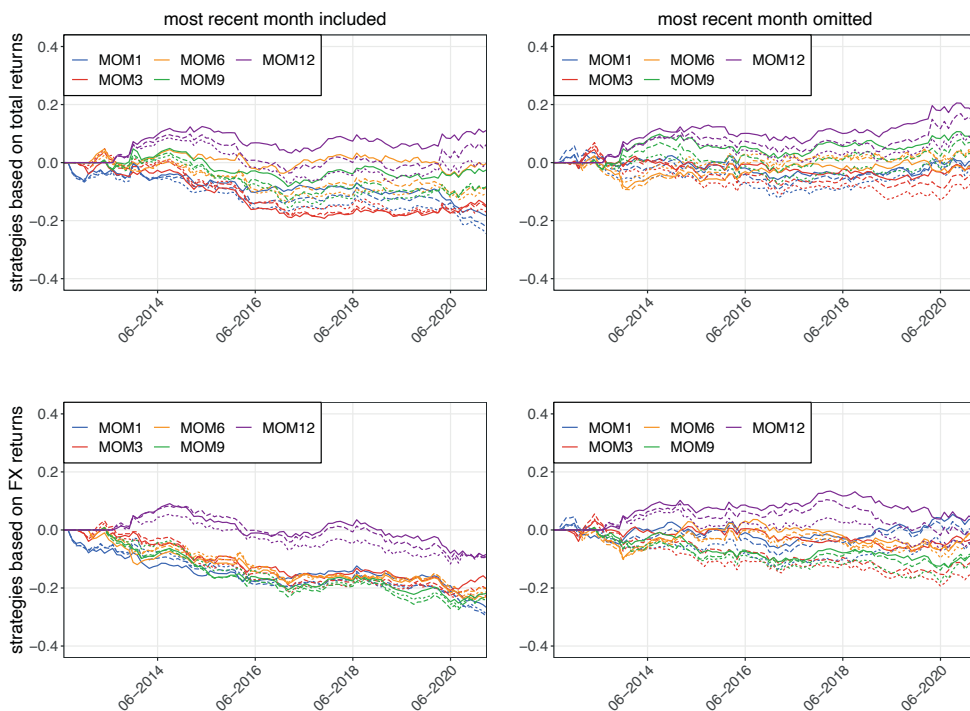
4.1. Performance of the momentum strategies

Descriptive statistics and performance measures of analysed portfolios are reported in Tables 2 and 3. The standard deviation of returns varies between values of 3 and 5 – portfolio returns volatility is relatively stable and does not depend on the choice of strategy formation period. Best performing strategies are positively skewed, while others do not show any particular pattern of skewness values. Returns of almost all the strategies tend to be leptokurtic.

Strategies based on signals from total returns presented in Table 2 tend to achieve average returns which increase with the length of strategy formation period T . Moreover, momentum strategies based on total returns perform better when the most recent month is omitted. These observations are also confirmed by the top panels of Figure 2 comparing strategies MOM_T and MOM_T^{\dagger} with signals derived from total returns.

From strategies MOM_T which use the most recent month, only 4 achieved a positive average return, while for strategies MOM_T^{-1} omitting the most recent month, 10 out of 15 generate profits. The best strategy characterised by the highest mean return (1.99%), the Sharpe ratio (0.58) and the Sortino ratio (1.09) is MOM_{12}^{-1} with the constant weighting scheme. This is not in line with the results of Menkhoff et al. [2012], Hsu and Chen [2021], Mikova et al. [2020] and may be partially caused by the use of a small set of currencies.

Figure 2. Cumulative returns of portfolios



Note: Solid line indicates constant weighting scheme, long-dashed line – ranking weighting scheme, and dashed line – proportional weighting scheme.

Source: Author’s own elaboration.

Table 3 shows that momentum strategies do not work when signals are based on the past changes of spot exchange rates. All portfolios based on signals with the most recent month included generate losses. The worst strategy is MOM_1 with the ranking weighting scheme and losses of almost 4%. Because of strong negative first-order autocorrelation, see Figure 1 Panel B, this strategy reaches a Sharpe ratio of -0.95 . Taking absolute values, its Sharpe ratio is almost twice as big as in the case of the MOM_{12}^{-1} strategy based on total returns. The absolute values of

the Sortino ratios of these strategies are almost equal (-1.07 and 1.09 , respectively). Therefore, in this case, the *anti-momentum* mean reversion strategy seems to be beneficial. Moreover, as shown in the bottom left panel of Figure 2, portfolios with formation periods from 1 up to 9 months perform similarly, generating losses at a steady pace. The only strategy that differs from others is $MOM_{1,2}$, with cumulative returns fluctuating around zero. Also, it is worth pointing out that when the most recent month is omitted, the Sharpe ratios of such strategies are mostly close to 0. This means that signals lack informative content about future returns. It is clearly visible in the bottom right panel of Figure 2, where no upward or downward trends in cumulative portfolio returns can be seen.

4.2. Decomposition of returns

The last two columns in Tables 2 and 3 show the shares of IR and FX returns in total portfolio return. The best strategy $MOM_{1,2}$ based on lagged total returns with constant weighting scheme is also the most balanced strategy between both returns components, as the shares of IR and FX returns are equal to 58% and 42%, respectively. Both IR and FX returns are positive, and they sum up resulting in the highest total return.

The top left panel of Figure 3 shows how cumulative IR and FX returns change over time and how they contribute to the total cumulative returns. Negative returns of strategies based on lagged FX returns are mainly driven by the FX returns themselves. Returns from MOM_1 portfolios, irrespective of the weighting scheme, are on average independent of the IR returns. In the case of the MOM_1 portfolio with ranking weights, the share of IR returns is equal exactly to 0. This is shown in the bottom left panel of Figure 3, where the line representing IR cumulative returns is flat and does not deviate significantly from zero at any period of time, while the line representing FX cumulative returns almost overlaps the total cumulative returns line.

Table 3 also indicates a relationship between the formation period length and returns. The longer is the formation period, the closer to zero are average returns. Also, the longer is the formation period, the more negative is the IR returns share.

Some portfolios have extremely high shares with different signs of IR and FX returns in total returns. This happens when the portfolio average total returns are close to zero, so the IR and FX returns are cancelling each other out, resulting in high values in relation to the average returns. This is shown in the top right panel of Figure 3.

In general, Figure 3 shows differences in characteristics of IR and FX returns. IR returns – interest rate differentials between countries – are essentially known

at the time of the investment decision. Moreover, interest rates change slowly and gradually – this is the reason why lines showing IR cumulative returns are smooth. On the other hand, changes in spot exchange rates are less predictable, so FX cumulative returns lines have many peaks and troughs.

Figure 3. Total cumulative returns, interest rate cumulative returns, and FX cumulative returns of selected portfolios



Source: Author's own elaboration.

Table 2. Descriptive statistics of returns of portfolios based on lagged total returns

Portfolio	Min	Max	Mean	Sd	Skew	Kurt	Sharpe	Sortino	MD	IR	FX
<i>portfolios based on signals with the most recent month included</i>											
MOM1 const	-39.29	53.69	-2.24	3.94	14.22	18.31	-0.57	-0.74	220.55	-0.15	1.15
MOM1 rank	-42.55	63.42	-2.79	4.32	13.58	28.74	-0.65	-0.82	267.89	-0.14	1.14
MOM1 prop	-43.93	68.74	-3.12	4.46	14.98	34.99	-0.70	-0.88	295.21	-0.16	1.16
MOM3 const	-31.40	28.72	-1.78	3.70	-0.23	-3.13	-0.48	-0.61	236.58	-0.36	1.36
MOM3 rank	-34.66	52.68	-1.84	4.02	10.77	10.75	-0.46	-0.60	256.30	-0.47	1.47
MOM3 prop	-39.93	37.39	-2.07	4.19	-0.23	0.60	-0.49	-0.63	261.73	-0.48	1.48
MOM6 const	-31.21	28.96	0.03	3.56	-2.55	-2.40	0.01	0.01	108.43	27.90	-26.90

Portfolio	Min	Max	Mean	Sd	Skew	Kurt	Sharpe	Sortino	MD	IR	FX
MOM6 rank	-34.66	33.04	-0.96	3.74	-1.12	0.57	-0.26	-0.34	179.15	-1.12	2.12
MOM6 prop	-41.72	32.38	-1.18	3.83	-6.77	2.36	-0.31	-0.40	192.14	-1.03	2.03
MOM9 const	-32.72	53.69	-0.22	3.62	20.22	21.74	-0.06	-0.09	150.45	-4.73	5.73
MOM9 rank	-26.51	43.07	-0.91	3.81	10.07	3.15	-0.24	-0.33	190.08	-1.35	2.35
MOM9 prop	-32.42	49.95	-0.88	4.26	9.33	5.46	-0.21	-0.29	211.54	-1.48	2.48
MOM12 const	-26.75	53.69	1.27	3.60	31.32	29.39	0.35	0.63	97.59	0.79	0.21
MOM12 rank	-30.83	62.28	0.79	4.12	27.59	23.76	0.19	0.32	146.96	1.43	-0.43
MOM12 prop	-33.13	50.85	0.09	4.19	15.43	10.50	0.02	0.03	174.26	14.06	-13.06
<i>portfolios based on signals with the most recent month omitted</i>											
MOM1 const	-33.84	28.96	0.20	3.65	-3.64	-1.77	0.06	0.08	101.09	1.70	-0.70
MOM1 rank	-42.89	28.53	0.10	4.28	-15.20	1.67	0.02	0.03	170.84	4.19	-3.19
MOM1 prop	-55.05	34.81	-0.02	4.58	-20.39	8.30	0.00	0.00	171.49	-34.42	35.42
MOM3 const	-31.40	36.74	-0.04	3.61	4.49	2.36	-0.01	-0.01	121.11	-14.57	15.57
MOM3 rank	-34.66	34.48	-0.39	4.11	-0.17	-0.17	-0.10	-0.13	175.79	-1.71	2.71
MOM3 prop	-36.10	36.08	-0.81	4.43	-0.81	-1.38	-0.18	-0.25	215.37	-0.93	1.93
MOM6 const	-53.69	28.96	-0.10	3.79	-19.00	13.92	-0.03	-0.04	131.90	-6.88	7.88
MOM6 rank	-43.96	42.34	0.57	4.28	-4.36	3.51	0.13	0.19	116.09	1.69	-0.69
MOM6 prop	-50.91	47.03	0.63	4.48	-4.74	6.63	0.14	0.20	105.55	1.77	-0.77
MOM9 const	-28.40	53.69	1.14	3.45	21.01	23.57	0.33	0.55	91.11	0.73	0.27
MOM9 rank	-27.63	43.07	0.90	3.99	9.57	2.93	0.22	0.35	117.02	1.20	-0.20
MOM9 prop	-33.89	42.36	0.53	4.31	5.39	1.44	0.12	0.19	123.07	2.31	-1.31
MOM12 const	-26.75	53.69	1.99	3.40	32.69	37.46	0.58	1.09	68.47	0.58	0.42
MOM12 rank	-27.63	62.28	1.73	4.05	24.43	23.71	0.43	0.73	82.53	0.69	0.31
MOM12 prop	-31.73	47.60	1.04	4.18	9.77	6.16	0.25	0.39	105.46	1.19	-0.19

Note: Excluding the last two columns, all the reported values are annualised. Descriptive statistics and maximum draw-down are expressed in percent. Kurtosis is reported as the excess kurtosis. MD stands for maximum drawdown. The last two columns show interest rate return (IR) and FX return (FX) contribution to the mean total return. Values IR and FX were computed as the ratio of the average IR (FX) return to the average total return, so they sum up to 1.

Source: Author's own elaboration.

Table 3. Descriptive statistics of returns of portfolios based on lagged FX returns

Portfolio	Min	Max	Mean	Sd	Skew	Kurt	Sharpe	Sortino	MD	IR	FX
<i>portfolios based on signals with the most recent month included</i>											
MOM1 const	-39.29	24.60	-3.49	3.67	-8.22	-0.35	-0.95	-1.07	321.12	0.03	0.97
MOM1 rank	-42.55	46.31	-3.97	4.17	3.19	12.73	-0.95	-1.11	359.11	0.00	1.00
MOM1 prop	-43.95	63.90	-3.84	4.31	13.68	32.76	-0.89	-1.08	349.48	-0.04	1.04
MOM3 const	-32.72	24.60	-2.04	3.78	-5.73	-3.02	-0.54	-0.66	255.77	-0.14	1.14
MOM3 rank	-34.66	24.01	-2.54	3.77	-3.44	-2.59	-0.67	-0.82	298.62	-0.15	1.15
MOM3 prop	-42.50	24.95	-2.91	3.94	-9.24	2.37	-0.74	-0.87	303.44	-0.10	1.10
MOM6 const	-53.69	28.96	-2.90	3.94	-14.48	8.17	-0.74	-0.86	287.61	-0.11	1.11

Portfolio	Min	Max	Mean	Sd	Skew	Kurt	Sharpe	Sortino	MD	IR	FX
MOM6 rank	-43.96	33.04	-2.43	3.99	-6.78	2.71	-0.61	-0.74	270.23	-0.16	1.16
MOM6 prop	-42.08	30.00	-2.65	4.15	-9.32	0.91	-0.64	-0.76	284.16	-0.16	1.16
MOM9 const	-32.72	25.61	-2.76	3.56	-4.81	-2.59	-0.77	-0.91	310.66	-0.11	1.11
MOM9 rank	-40.69	31.55	-2.90	3.91	0.80	0.71	-0.74	-0.90	330.23	-0.11	1.11
MOM9 prop	-37.83	40.50	-3.06	4.04	1.08	4.94	-0.76	-0.91	320.84	-0.13	1.13
MOM12 const	-31.21	53.69	-0.94	3.50	16.27	24.98	-0.27	-0.37	204.99	-0.29	1.29
MOM12 rank	-29.97	43.96	-0.94	3.74	11.78	5.59	-0.25	-0.35	204.29	-0.33	1.33
MOM12 prop	-29.91	47.51	-1.01	3.88	13.63	12.12	-0.26	-0.36	195.25	-0.32	1.32
<i>portfolios based on signals with the most recent month omitted</i>											
MOM1 const	-33.84	28.96	0.56	3.68	-4.28	-1.59	0.15	0.22	104.71	-0.18	1.18
MOM1 rank	-42.89	28.53	0.25	4.22	-15.53	2.43	0.06	0.08	154.36	0.11	0.89
MOM1 prop	-53.24	34.76	-0.31	4.53	-19.38	6.94	-0.07	-0.09	182.85	-0.54	1.54
MOM3 const	-31.40	36.74	-0.39	3.56	9.49	2.57	-0.11	-0.16	122.99	-0.35	1.35
MOM3 rank	-48.09	34.48	-1.34	4.23	-3.04	4.30	-0.32	-0.42	237.81	-0.18	1.18
MOM3 prop	-46.96	37.98	-1.82	4.44	-4.84	2.26	-0.41	-0.53	266.88	-0.13	1.13
MOM6 const	-53.69	28.96	-0.10	4.04	-12.56	8.76	-0.02	-0.03	127.92	-2.37	3.37
MOM6 rank	-43.96	42.34	-0.08	4.32	0.22	3.93	-0.02	-0.03	122.25	-3.23	4.23
MOM6 prop	-52.58	47.76	-0.31	4.71	-6.83	8.12	-0.07	-0.09	121.29	-1.08	2.08
MOM9 const	-31.21	25.61	-1.00	3.25	-7.77	-1.03	-0.31	-0.39	165.40	-0.32	1.32
MOM9 rank	-32.31	36.45	-0.80	3.97	3.56	-1.34	-0.20	-0.28	176.47	-0.45	1.45
MOM9 prop	-36.44	42.81	-1.26	4.32	6.29	4.13	-0.29	-0.40	220.60	-0.32	1.32
MOM12 const	-31.21	53.69	0.56	3.42	22.29	26.77	0.16	0.26	108.80	0.16	0.84
MOM12 rank	-29.97	43.96	0.02	3.91	13.98	4.90	0.01	0.01	134.75	6.30	-5.30
MOM12 prop	-31.66	40.34	-0.44	4.22	7.06	3.97	-0.11	-0.15	158.80	-0.52	1.52

Note: reporting convention as in Table 2.

Source: Author's own elaboration.

4.3. Impact of weighting schemes on portfolio performance

Tables 4 and 5 present the number of long and short positions of currencies and their average weights in portfolio. The compositions of MOM₁ portfolios are very volatile – currency weights change substantially month by month. A number of long and short positions of all the currencies is similar. Also, the average weights of the currencies are small and close to zero.

As the strategy formation period gets longer, long-term trends are becoming visible. In the case of the portfolios based on lagged total return, PLN is 55 times on a long position in MOM₁, but when the formation period is set to 12 months this number rises to 77. On the contrary, HUF in MOM₁ is 42 times on a long position,

and in MOM₁₂ HUF is observed on a long position only 9 times. Similar observations can be made for CZK and RON.

Surprisingly, for portfolios based on lagged FX returns these long-term trends are visible only for CZK and HUF. PLN and RON, irrespective of the formation period, are approximately half of the time a funding currency (in short positions) and half of the time an asset currency (in long positions). This is also reflected in their average weights, which neither exceed 0.1 nor fall below -0.1.

Table 4. Number of long and short positions of currencies and average currency weights in portfolios based on lagged total returns

Portfolio	# long				# short				average weight			
	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON
<i>portfolios based on signals with the most recent month included</i>												
MOM1 const	55	52	42	59	49	52	62	45	0.03	0.00	-0.10	0.07
MOM1 rank	55	52	42	59	49	52	62	45	0.09	-0.01	-0.10	0.02
MOM1 prop	60	55	47	60	44	49	57	44	0.09	-0.02	-0.11	0.03
MOM3 const	60	53	32	59	42	49	70	43	0.09	0.02	-0.18	0.08
MOM3 rank	60	53	32	59	42	49	70	43	0.10	0.03	-0.26	0.12
MOM3 prop	61	55	32	62	41	47	70	40	0.12	0.01	-0.28	0.15
MOM6 const	66	49	23	60	33	50	76	39	0.16	0.00	-0.25	0.10
MOM6 rank	66	49	23	60	33	50	76	39	0.18	0.02	-0.33	0.13
MOM6 prop	62	41	20	60	37	58	79	39	0.18	-0.01	-0.34	0.17
MOM9 const	70	40	15	67	26	56	81	29	0.21	-0.08	-0.31	0.18
MOM9 rank	70	40	15	67	26	56	81	29	0.22	-0.03	-0.40	0.20
MOM9 prop	67	42	17	65	29	54	79	31	0.21	-0.01	-0.42	0.22
MOM12 const	70	42	10	64	23	51	83	29	0.22	-0.04	-0.35	0.17
MOM12 rank	70	42	10	64	23	51	83	29	0.22	0.00	-0.39	0.17
MOM12 prop	72	42	9	66	21	51	84	27	0.22	-0.01	-0.43	0.22
<i>portfolios based on signals with the most recent month omitted</i>												
MOM1 const	54	52	43	59	50	52	61	45	0.02	0.00	-0.09	0.07
MOM1 rank	54	52	43	59	50	52	61	45	0.08	-0.01	-0.10	0.02
MOM1 prop	59	55	47	59	44	48	56	44	0.09	-0.02	-0.10	0.03
MOM3 const	59	49	40	56	43	53	62	46	0.08	-0.02	-0.10	0.05
MOM3 rank	59	49	40	56	43	53	62	46	0.09	-0.02	-0.15	0.08
MOM3 prop	61	52	37	61	41	50	65	41	0.09	0.02	-0.20	0.09
MOM6 const	66	46	28	58	33	53	71	41	0.16	-0.03	-0.20	0.08
MOM6 rank	66	46	28	58	33	53	71	41	0.15	-0.02	-0.26	0.13
MOM6 prop	64	44	23	57	35	55	76	42	0.16	-0.02	-0.30	0.16
MOM9 const	73	41	22	56	23	55	74	40	0.24	-0.07	-0.25	0.08
MOM9 rank	73	41	22	56	23	55	74	40	0.24	-0.03	-0.34	0.13

Portfolio	# long				# short				average weight			
	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON
MOM9 prop	69	43	17	62	27	53	79	34	0.22	-0.02	-0.40	0.19
MOM12 const	77	34	10	65	16	59	83	28	0.29	-0.12	-0.35	0.18
MOM12 rank	77	34	10	65	16	59	83	28	0.25	-0.05	-0.38	0.18
MOM12 prop	72	39	9	65	21	54	84	28	0.23	-0.02	-0.42	0.21

Note: Four average weights sum up to zero, as these are zero-investment portfolios. The average weight in the case of constant and ranking weights is constrained to a range $[-0.5; 0.5]$ and $[-0.75; 0.75]$, respectively, while the average weight of currency in portfolio with proportional weights can take any value between -1 and 1 .

Source: Author's own elaboration.

Table 5. Number of long and short positions of currencies and average currency weights in portfolios based on lagged FX returns

Portfolio	# long				# short				average weight			
	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON	PLN	CZK	HUF	RON
<i>portfolios based on signals with the most recent month included</i>												
MOM1 const	50	60	49	49	54	44	55	55	-0.02	0.08	-0.03	-0.03
MOM1 rank	50	60	49	49	54	44	55	55	0.02	0.06	-0.04	-0.04
MOM1 prop	57	60	49	53	47	44	55	51	0.04	0.05	-0.06	-0.04
MOM3 const	55	59	40	50	47	43	62	52	0.04	0.08	-0.10	-0.01
MOM3 rank	55	59	40	50	47	43	62	52	0.02	0.12	-0.16	0.02
MOM3 prop	50	64	42	57	52	38	60	45	-0.01	0.15	-0.16	0.03
MOM6 const	51	65	30	52	48	34	69	47	0.01	0.15	-0.19	0.02
MOM6 rank	51	65	30	52	48	34	69	47	0.02	0.17	-0.20	0.01
MOM6 prop	54	64	31	52	45	35	68	47	0.02	0.18	-0.22	0.01
MOM9 const	56	61	29	46	40	35	67	50	0.08	0.12	-0.18	-0.02
MOM9 rank	56	61	29	46	40	35	67	50	0.07	0.19	-0.22	-0.03
MOM9 prop	54	64	29	53	42	32	67	43	0.05	0.21	-0.27	0.01
MOM12 const	54	68	16	48	39	25	77	45	0.07	0.20	-0.29	0.01
MOM12 rank	54	68	16	48	39	25	77	45	0.07	0.21	-0.28	0.00
MOM12 prop	52	67	24	53	41	26	69	40	0.04	0.23	-0.28	0.01
<i>portfolios based on signals with the most recent month omitted</i>												
MOM1 const	49	60	50	49	55	44	54	55	-0.03	0.08	-0.02	-0.03
MOM1 rank	49	60	50	49	55	44	54	55	0.01	0.06	-0.03	-0.04
MOM1 prop	56	60	49	52	47	43	54	51	0.04	0.05	-0.05	-0.04
MOM3 const	52	60	44	48	50	42	58	54	0.01	0.09	-0.07	-0.03
MOM3 rank	52	60	44	48	50	42	58	54	0.02	0.10	-0.11	-0.01
MOM3 prop	52	57	41	51	50	45	61	51	0.02	0.12	-0.13	-0.01
MOM6 const	55	60	37	46	44	39	62	53	0.05	0.10	-0.12	-0.03
MOM6 rank	55	60	37	46	44	39	62	53	0.04	0.12	-0.13	-0.04
MOM6 prop	54	65	33	51	45	34	66	48	0.01	0.15	-0.17	0.01

	# long				# short				average weight			
MOM9 const	58	61	25	48	38	35	71	48	0.10	0.12	-0.22	0.00
MOM9 rank	58	61	25	48	38	35	71	48	0.07	0.16	-0.23	0.00
MOM9 prop	58	60	28	52	38	36	68	44	0.05	0.19	-0.26	0.02
MOM12 const	54	67	22	43	39	26	71	50	0.07	0.20	-0.23	-0.03
MOM12 rank	54	67	22	43	39	26	71	50	0.05	0.22	-0.25	-0.02
MOM12 prop	48	66	25	49	45	27	68	44	0.03	0.21	-0.25	0.01

Note: notation convention as in Table 4.

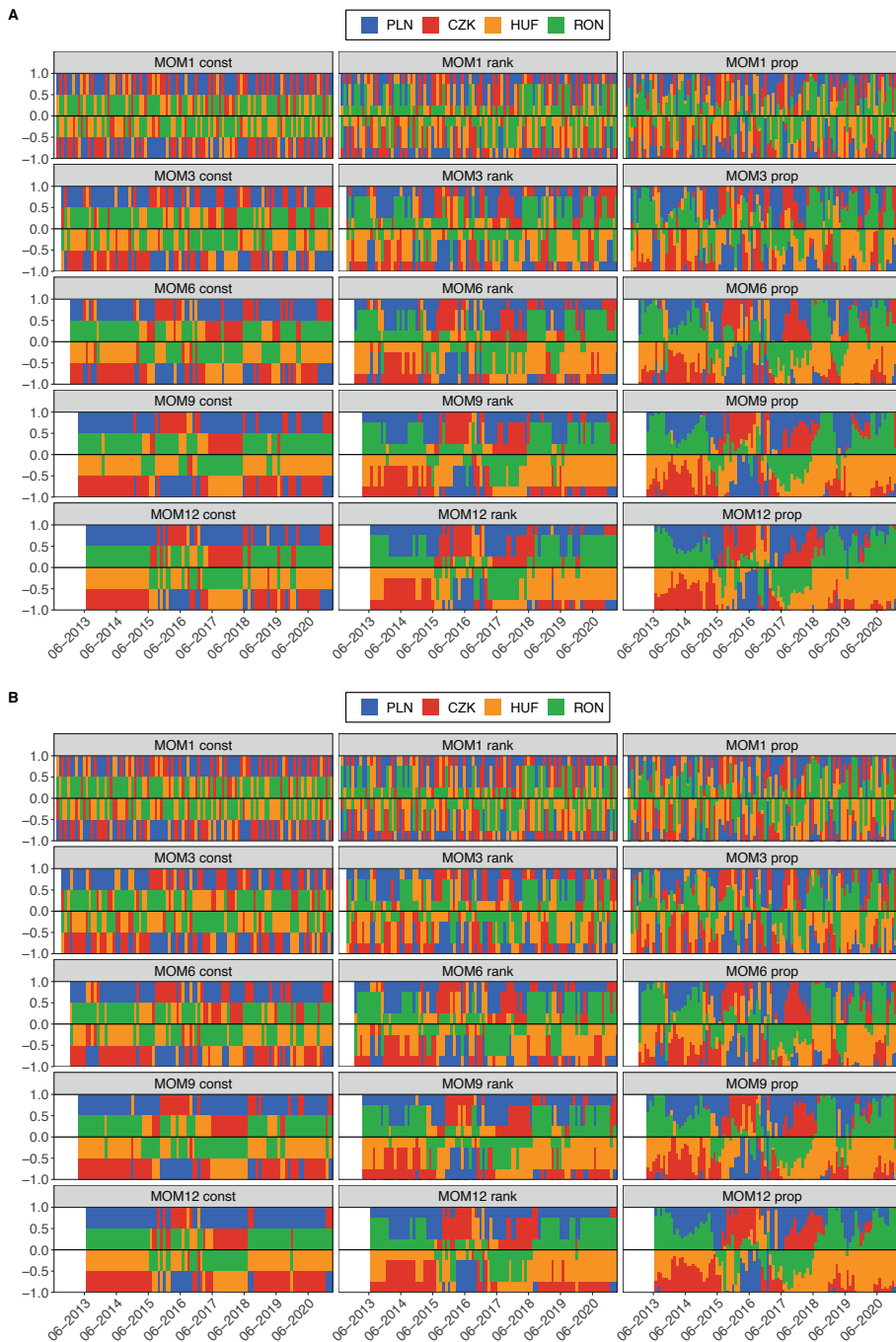
Source: Author's own elaboration.

In Figures 4 and 5, the emergence of long-term trends is clearly visible. While weights of currencies in portfolios with short formation periods seem to change chaotically, weights in portfolios with long formation periods remain stable over time. Moreover, subfigures presenting portfolios with the proportional weighting scheme provide information on slow, gradual changes in buy/sell signal values. Another advantage of proportional weights is that they can give informative content about extraordinary deviations from the average buy/sell signal in a particular period.

Average weights presented in Tables 4 and 5 are constrained in the case of constant and ranking weights to a range $[-0.5, 0.5]$ and $[-0.75, 0.75]$, respectively, while the average weight of currency in portfolio with proportional weights can take any value between -1 and 1 . This should lead to higher average weights in absolute terms when a portfolio is constructed using the proportional weighting scheme. Interestingly, this does not take place. There are only small changes of average weights values, in both directions, between portfolios constructed using proportional and other two weighting schemes.

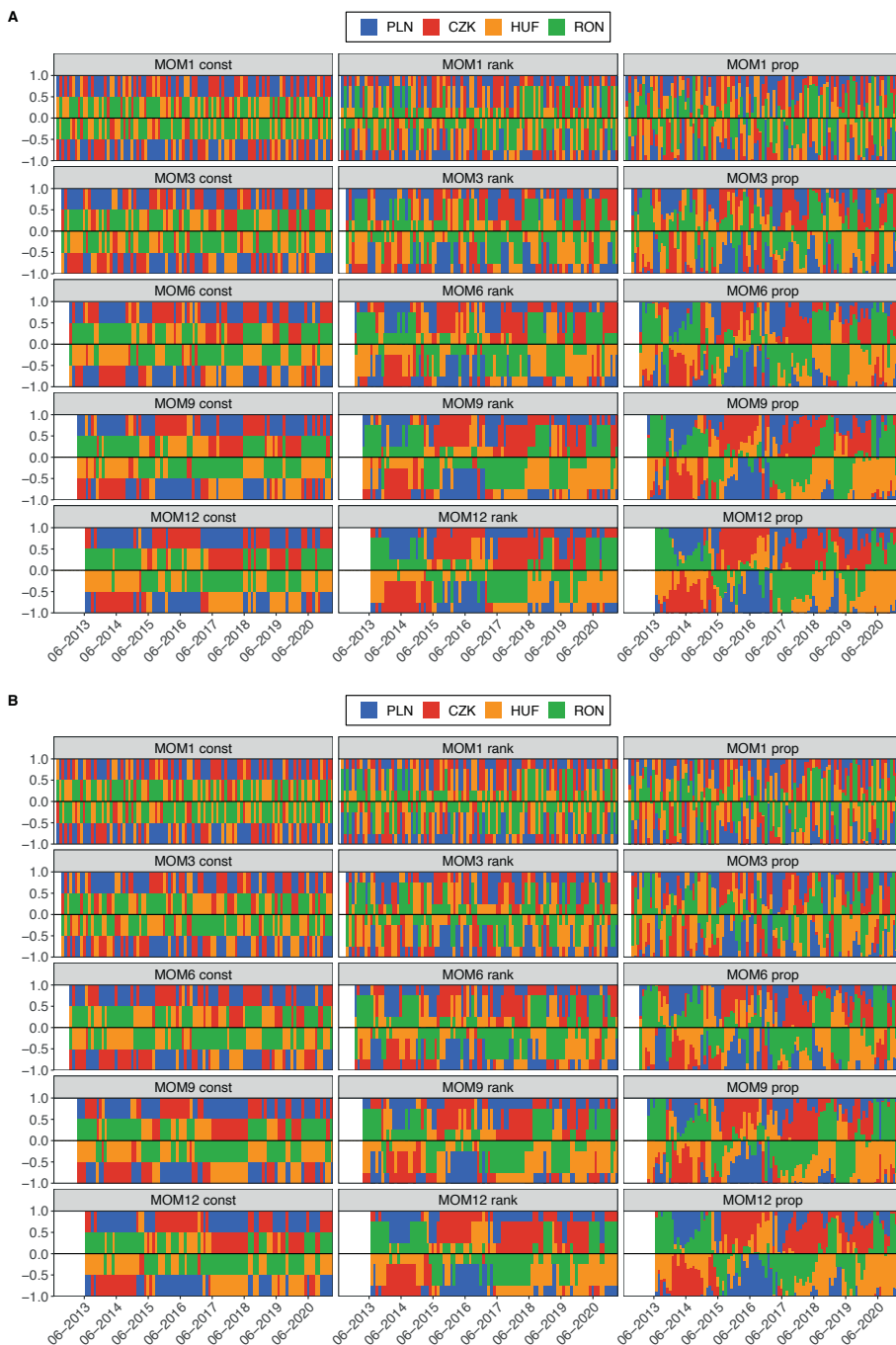
The impact of applying different weighting schemes is shown in Figure 6. The scatterplot presenting portfolio average returns and standard deviations of portfolio returns illustrates the relationship between the weighting scheme and portfolio performance. By analysing 60 portfolios (20 per weighting scheme), it is easy to observe that the choice of weighting scheme does not have any impact on the portfolio average return. However, the use of ranking and proportional weighting schemes increases the standard deviation of portfolio returns. As a result, ranking and proportional weighting schemes do not improve the momentum strategies performance, providing similar returns, albeit at a higher risk.

Figure 4. Weights of portfolios based on lagged total return with the most recent month included (A) and omitted (B)



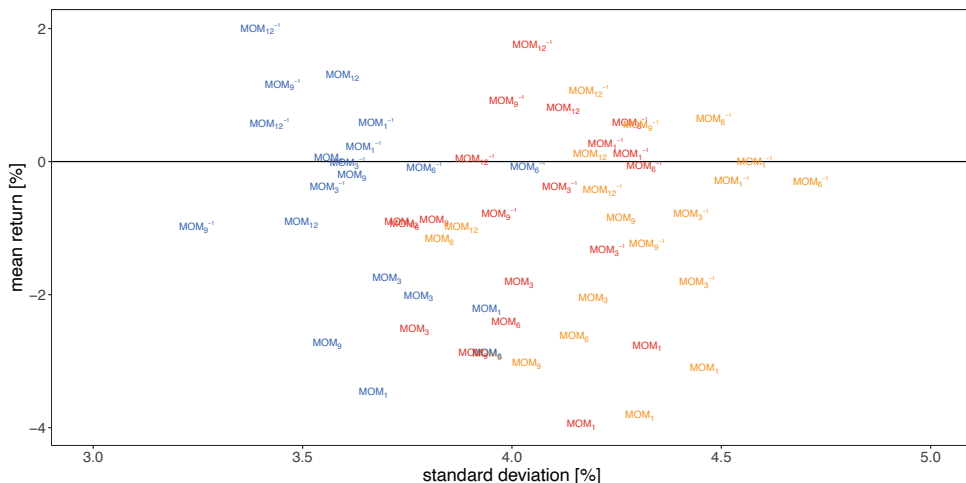
Source: Author's own elaboration.

Figure 5. Weights of portfolios based on lagged FX return with the most recent month included (A) and omitted (B)



Source: Author's own elaboration.

Figure 6. Mean returns and standard deviations of constructed portfolios



Note: blue, red, and orange indicate constant, ranking, and proportional weighting schemes, respectively.
 Source: Author's own elaboration.

5. Conclusions

Our study shows that from the perspective of a Eurozone investor, the momentum strategy based on lagged cumulative long term total return, with the most recent return dropped (MOM_{12}^{-1}) performs best in the CEE currency universe. A portfolio using this strategy with a constant weighting scheme reaches an average return of 1.99% and a Sharpe ratio of 0.58. Compared to the performance of the same strategy in Asness et al. [2013], our portfolio has an average return lower by 1.5 percentage points and the Sharpe ratio higher by 0.24. MOM_{12}^{-1} portfolios with ranking and proportional weighting schemes present inferior performance compared to MOM_{12}^{-1} with constant weighting scheme. The empirical results show that more complex weighting schemes like ranking and proportional weights do not cause any essential performance improvement.

Ranking and proportional weights affect the standard deviation of returns, increasing their volatility, but do not improve average returns, compared to constant weights. This finding is in line with Asness et al. [2013] stating that portfolios with constant weights have a slightly better performance than portfolios with ranking weights.

The analysis of momentum strategies performance shows that past high returns are not always followed by returns of the same magnitude. We find in our data a one-month reversal of returns. Thus, the signals which do not take into account the most

recent monthly return lead to improved performance. Many other momentum portfolios generate negative returns, especially those based on lagged FX returns with the most recent month included.

Among portfolios with positive average returns, the IR returns are the main driver of profits. Returns decomposition shows that in the best performing strategies, IR returns account for more than 50% of total returns. Losses generated by other portfolios are mainly due to the FX returns. Average IR returns, regardless of the strategy, are almost always positive. Since interest rates in the Eurozone are lower than in the CEE countries, this seems to be an effect of the choice of currency set.

The findings presented above indicate that momentum strategies based on lagged total returns and with long formation periods should be preferred. Moreover, negative first-order autocorrelation of returns justifies skipping the most recent return signals computation, while the comparison of different weighting schemes shows that there is no improvement from using more complex weights over the constant ones.

References

- Asness, C.S., Moskowitz, T.J., Pedersen, L.H. (2013). Value and momentum everywhere. *The Journal of Finance*, 68(3), pp. 929–985.
- Bacon, C.R. (2008). *Practical portfolio performance measurement and attribution*, vol. 546. John Wiley & Sons.
- Barroso, P. and Santa-Clara, P. (2015). Beyond the carry trade: Optimal currency portfolios. *Journal of Financial and Quantitative Analysis*, 50(5), pp. 1037–1056.
- Brunnermeier, M.K., Nagel, S., and Pedersen, L.H. (2008). Carry trades and currency crashes. *NBER macroeconomics annual*, 23(1), pp. 313–348.
- Burnside, C., Eichenbaum, M., Kleshchelski, I., Rebelo, S. (2011a). Do peso problems explain the returns to the carry trade? *The Review of Financial Studies*, 24(3), pp. 853–891.
- Burnside, C., Eichenbaum, M., and Rebelo, S. (2011b). Carry trade and momentum in currency markets. *Annu. Rev. Financ. Econ.*, 3(1), pp. 511–535.
- Dahlquist, M. and Hasseltoft, H. (2020). Economic momentum and currency returns. *Journal of Financial Economics*, 136(1), pp. 152–167.
- Daniel, K., Hodrick, R.J., Lu, Z., et al. (2017). The carry trade: Risks and drawdowns. *Critical Finance Review*, 6(2), pp. 211–262.
- Filippou, I., Gozluklu, A.E., and Taylor, M.P. (2018). Global political risk and currency momentum. *Journal of Financial and Quantitative Analysis*, 53(5), pp. 2227–2259.
- Hsu, C-C. and Chen, M.-L. (2021). Currency momentum strategies based on the Chinese yuan: Timing of foreign exchange volatility. *Journal of International Financial Markets, Institutions and Money*, 72, p. 101315.

- Jaworski, K. (2021). Forecasting exchange rates for central and eastern European currencies using country-specific factors. *Journal of Forecasting*, 40(6), pp. 977–999.
- Jurek, J.W. (2014). Crash-neutral currency carry trades. *Journal of Financial Economics*, 113(3), pp. 325–347.
- Koijen, R.S., Moskowitz, T.J., Pedersen, L.H., and Vrugt, E.B. (2013). Carry. Working Paper 19325, National Bureau of Economic Research.
- Lustig, H., Roussanov, N., and Verdelhan, A. (2011). Common risk factors in currency markets. *The Review of Financial Studies*, 24(11), pp. 3731–3777.
- Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2012). Currency momentum strategies. *Journal of Financial Economics*, 106(3), pp. 660–684.
- Menkhoff, L., Sarno, L., Schmeling, M., and Schrimpf, A. (2017). Currency value. *The Review of Financial Studies*, 30(2), pp. 416–441.
- Mikova, E., Teplova, T., and Munir, Q. (2020). Puzzling premiums on FX markets: Carry trade, momentum, and value alone and strategy diversification. *Emerging Markets Finance and Trade*, 56(1), pp. 126–148.
- Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics*, 103(3), pp. 429–453.
- Sharpe, W.F. (1966). Mutual fund performance. *The Journal of Business*, 39(1), pp. 119–138.
- Sortino, F.A. and Price, L.N. (1994). Performance measurement in a downside risk framework. *The Journal of Investing*, 3(3), pp. 59–64.

Metody ważenia aktywów w strategiach momentum dla walut Europy Środkowej i Wschodniej

Streszczenie

W niniejszym rozdziale konstruowano portfele przy użyciu strategii momentum oraz analizowano wpływ stosowania różnych metod ważenia aktywów i sygnałów kupna/sprzedaży na wyniki tych strategii. Strategie są budowane na czterech walutach krajów Europy Środkowej i Wschodniej: złotego polskiego (PLN), korony czeskiej (CZK), forinta węgierskiego (HUF) i leja rumuńskiego (RON). Przyjęto perspektywę inwestorów ze strefy euro, więc walutą bazową jest euro (EUR), a nie dolar amerykański. Z przeprowadzonych badań wynika, że najbardziej efektywna jest strategia momentum oparta na opóźnionym całkowitym zwrocie, z okresem formowania wynoszącym 12 miesięcy i pomijającym ostatni miesiąc. Przy zastosowaniu stałego schematu ważenia taki portfel osiąga średni roczny zwrot 1,99% i wskaźnik Sharpe'a 0,58. Uzyskane wyniki pokazują, że bardziej złożone metody ważenia nie powodują istotnej poprawy wyników jakiegokolwiek z analizowanych strategii. Stwierdzono także, że strategie momentum oparte na opóźnionej zmianie kursu walutowego nie są opłacalne.

Słowa kluczowe: waluty CEE, strefa euro, waluty, strategia momentum, metody ważenia aktywów

Jaka powinna być nowa polska partia? Maksymalizacja wyniku wyborczego partii politycznej za pomocą hybrydowego algorytmu Simulated Annealing – Multi-Start Tabu Search

Radosław Jawor, Szymon Dziubak

Streszczenie

Wybór preferowanego ugrupowania podczas wyborów parlamentarnych jest skomplikowanym zagadnieniem z dziedziny nauk społecznych. Uzupełnieniem empirycznych i teoretycznych badań nad tym zjawiskiem może być zastosowanie metod matematycznych. Bazując na ekstrapolacji, wyznaczono rozkład poglądów politycznych Polaków. Następnie z użyciem probabilistycznego algorytmu wyżarzania symulowanego (ang. *Simulated Annealing*) wyznaczono macierze kowariancji poglądów partii politycznych, minimalizując różnicę pomiędzy rzeczywistym poparciem politycznym a wyznaczonym za pomocą modelu. Następnie zaproponowano hybrydowy algorytm metaheurystyczny Simulated Annealing – Multi-start Tabu Search do znalezienia parametrów poglądów nowej partii politycznej maksymalizujących jej wynik wyborczy. Uzyskane wyniki wskazują, że nowa partia polityczna ma szansę zdobyć 19,4% głosów w wyborach i może zostać trzecią największą siłą w Sejmie. W rozdziale zaprezentowano również kierunki dalszych badań i rozwoju modelu.

Słowa kluczowe: metaheurystyki, Tabu Search, Simulated Annealing, partie polityczne, poparcie polityczne

1. Wstęp

Za początek zyskiwania popularności przez metaheurystyki uznaje się wcześnie lata 80. XX w. [Osman, Laporte, 1996]. Heurystyki definiuje się jako metody służące do przyspieszania i ułatwiania procesu podejmowania decyzji poprzez

redukcję liczby informacji niezbędnej do przetworzenia [Dale, 2015]. Heurystyki nie są jednak algorytmami dokładnymi, to znaczy rozwiązanie, które znajdują, jest przybliżone, ale rzadko dokładne. Ich stosowanie ma sens tam, gdzie stosowanie algorytmów dokładnych jest kosztowne obliczeniowo [Sörensen, Glover, 2013, s. 1]. Problem maksymalizacji wyniku wyborczego partii politycznej jest problemem ciągłym w przestrzeni pięciowymiarowej. Zastosowanie metod optymalizacji globalnej, metaheurystyk, ma na celu zmniejszenie kosztu obliczeniowego związanego z rozwiązaniem problemu. Jak udowodniono później, hybrydowy algorytm Simulated Annealing – Multi-start Tabu Search przeszukuje dużą przestrzeń rozwiązań, przez co jest mniej wydajny czasowo niż standardowe metaheurystyki, jednak umożliwia znalezienie lepszego rozwiązania. Stanowi zadowalający kompromis pomiędzy kosztem obliczeniowym a dokładnością. Do rozwiązania problemu wykorzystano wyniki wyborcze koalicji oraz partii politycznych, których przedstawiciele zostali zaprzysiężeni na posłów na Sejm RP IX kadencji.

Zainteresowanie modelami matematycznymi w analizie decyzji politycznych i wyborczych jest obecne w literaturze od dawna. Decyzje wyborców położonych w przestrzeni ideologicznej wraz z partiami politycznymi rozważane są już przez Downsa [1957], który opisuje jednowymiarową przestrzeń ideologiczną. Dostrzega on wpływ rozkładu poglądów społeczeństwa na optymalne – maksymalizujące wynik wyborczy – położenie partii politycznych w przestrzeni. Downs prezentuje jedynie model teoretyczny, który jednak stanowi podstawę do wielu kolejnych, pogłębionych oraz empirycznych badań.

Misra [2012] proponuje matematyczny model do tworzenia strategii partii politycznej w zależności od jej członków, wyborców i poglądów politycznych w systemie dwupartyjnym. Używa przy tym modelu stosowanego w epidemiologii, adaptując go do realiów politycznych. Misra postuluje istnienie trzech klas w populacji – klasy wyborców oraz dwóch klas członków istniejących partii. Kontakty między poszczególnymi klasami mogą skutkować wstąpieniem do partii, wystąpieniem z partii lub zamianą jednej partii na drugą, a liczba członków danej partii wpływa na jej potencjał wyborczy.

V. Ostapenko, O. Ostapenko, Belyaeva i Stupnitskaya [2012] przedstawiają model alokacji funduszy partii politycznych w trakcie kampanii wyborczej pomiędzy regiony, w których partie mają różne poparcie. Optymalna alokacja funduszy ma za zadanie maksymalnie powiększyć elektorat danej partii. Model rozbudowany jest do nieskończonej ilości partii politycznych. Autorzy zauważają ponadto, że ten sam model może zostać wykorzystany do, przykładowo, optymalizacji alokacji kapitału przedsiębiorstwa w warunkach konkurencji na wielu rynkach.

Yang, Abrams, Kernell i Motter [2020] proponują wyjaśnienie mechanizmu powodującego postępującą polaryzację partii politycznych w Stanach Zjednoczonych. Model opisujący ten mechanizm stał się bazą do stworzenia modelu polskiego systemu partyjnego opisanego w niniejszym rozdziale.

Askari, Younas i Saeed [2020] przedstawiają model oparty na metaheurystykach, którego celem jest optymalizacja decyzji dotyczących głównych kwestii, z jakimi mierzy się partia polityczna.

Algoritmy optymalizacyjne są również obecne w literaturze. Gendreau [2003] definiuje algorytm przeszukiwania z listą tabu (ang. *Tabu Search*) jako algorytm pozwalający metodom optymalizacji lokalnej na wyjście poza optimum lokalne. Jest to możliwe dzięki wprowadzeniu przejścia do rozwiązania, które nie polepsza wartości funkcji celu. Wykorzystywana jest przy tym pamięć rozwiązań już przeszukanych, nazywana listą tabu (ang. *Tabu List*).

Poza wykorzystaniem pamięci, Glover i Laguna [1998] uważają, że algorytm Tabu Search różni się od algorytmu genetycznego i algorytmu wyżarzania symulowanego (ang. *Simulated Annealing*) tym, że nie polega on w dużym stopniu na częściowej losowości, która jest zaimplementowana za pomocą losowania. Przedstawiają także szeroki zakres teoretycznych i praktycznych zastosowań algorytmu Tabu Search.

Idea fazy intensyfikacji w algorytmie Tabu Search oparta jest na podejściu człowieka do przeszukiwania obszaru rozwiązań. Polega na przerwaniu, w pewnych odstępach czasu, normalnego przeszukiwania i zmianie obszaru przeszukiwań, a następnie powrocie do standardowego procesu [Aarts, Korst, Michiels, 2005].

Multi-start Tabu Search ma na celu przerwanie dotychczasowego obszaru przeszukiwań i jego zmianę, gdy spełnione zostaną założone z góry warunki [James, 2009]. Autorzy definiują kilka podejść do perturbacji algorytmu Tabu Search – jednym z nich jest Random Restart Approach. Random Restart Tabu Search (RRTS) restartuje przeszukiwanie obszaru rozwiązań, rozpoczynając nowe przeszukiwanie od losowego punktu. Restart dokonywany jest dopiero, gdy spełnione są założone z góry warunki. Martí [2003] podkreśla, że własność restartu pomaga zdywersyfikować przeszukiwanie obszaru rozwiązań i zwiększa prawdopodobieństwo znalezienia rozwiązania optymalnego.

Pojęcie wyżarzania zostało wprowadzone pierwotnie do problemu optymalizacji kombinatorycznej [Aarts i in., 2005]. Użycie temperatury w algorytmach zaproponował Metropolis [Metropolis, M.N. Rosenbluth, A.W. Rosenbluth, A.H. Teller, E. Teller, 1953]. Jego algorytm może zostać użyty do wygenerowania sekwencji rozwiązań kombinatorycznego problemu optymalizacji. Następnie dodany został parametr kontrolny, który nadzoruje temperaturę. Te algorytmy i pomysły umożliwiły stworzenie algorytmu Simulated Annealing.

Kombinacja algorytmów hybrydowych jest znana w literaturze. Lenin [Lenin, Reddy, Suryakalavathi, 2016] zastosował hybrydowy algorytm Tabu Search – Simulated Annealing do optymalizacji problemu mocy biernej. Udowadnia, że algorytm hybrydowy może szybciej zbiegać do rozwiązania optymalnego oraz wyznaczać lepszą funkcję celu niż algorytmy Tabu Search i Simulated Annealing osobno.

Mishra [Mishra, Prakash, Tiwari, Shankar, Chan, 2005] stosuje natomiast zagadnienie hybrydowego algorytmu Tabu-Simulated Annealing do rozwiązania problemu dotyczącego asortymentu produktów z wieloma ograniczeniami. Stwierdza, że algorytm przeszukuje dużą przestrzeń rozwiązań i ma małą wydajność, dlatego jego rozwiązania produkcyjne są ograniczone.

Swarnkar [Swarnkar, Tiwarim, 2004] proponuje rozwiązanie problemu nakładu pracy maszyn w elastycznym systemie produkcyjnym. Dochodzi do podobnego wniosku, że algorytm hybrydowy przeszukuje dużą przestrzeń rozwiązań i dlatego jest mało wydajny. Algorytm znajduje jednak rozwiązanie optymalne lub bliskie optymalnemu.

Celem badania jest sprawdzenie, jakie parametry poglądów nowej partii politycznej mogą zapewnić jej maksymalny wynik wyborczy w wyborach do Sejmu w Polsce.

2. Model systemu partyjnego

Przedstawiony poniżej model został stworzony na podstawie podejścia zaprezentowanego przez Yang i in. [2020]. Konieczne było jednak wprowadzenie pewnych zmian, ze względu na różne charakterystyki amerykańskiego i polskiego systemu partyjnego. Yang i in. [2020] operują na modelu jednowymiarowym, co jest rozsądnym podejściem w przypadku amerykańskiego systemu dwupartyjnego. W celu zastosowania modelu na polskiej scenie politycznej konieczne było jego rozszerzenie do dwóch wymiarów. Z drugiej strony używany model jest uproszczeniem modelu Yang i in. [2020], gdyż nie są w nim rozważani wyborcy niedecydujący się na udział w wyborach i oddanie głosu. Ponadto w przeciwieństwie do podejścia Yang i in. [2020] w niniejszej pracy rozważany jest model statyczny, dający wyniki jedynie w dwóch punktach czasu – przed i po wprowadzeniu nowej partii do systemu.

W modelu każda partia zajmuje konkretny punkt (x,y) w dwuwymiarowej przestrzeni ideologicznej. Pierwszy wymiar określa poglądy ekonomiczne na osi lewica – prawica. Drugi wymiar ogólnie oznacza poglądy niezwiązane z ekonomią *sensu stricto*, głównie społeczne. Za Cześnikiem, Kotnarowskim i Żerkowską-Balas [2015] można wyróżnić podejścia: Ingleharta [1977, materializm – postmaterializm],

Kriesiego i in. [2006, integracja – demarkacja], Hooghe, Marksa i Wilson [2002, GAL-TAN, ang. *Green/Alternative/Libertarian – Traditional/Authoritarian/Nationalist*] oraz Cześnika i in. [2015, solidaryzm – liberalizm]. W tej pracy, ze względu na dostępność zbiorów danych, zostały użyte ostatnie dwa wymienione podejścia.

W tej samej przestrzeni ideologicznej swoje poglądy określają wyborcy. Mogą oni też określić podobieństwo swoich poglądów do pozycji ideologicznych partii. W tym celu zostały zdefiniowane funkcje podobieństwa poglądów między wyborcą a partią:

$$f_{(\mu, \Sigma)}(x, y), \quad (1)$$

gdzie: $\mu = [x_i y_i]$, $\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{bmatrix}$, a (x, y) to poglądy wyborcy.

Funkcje f są funkcjami gęstości dwuwymiarowego rozkładu normalnego o wartości oczekiwanej μ równej pozycji danej partii w przestrzeni ideologicznej oraz odpowiednio skalibrowanej macierzy kowariancji Σ , której elementy mogą być interpretowane jako poziom rozmycia poglądów partii. Im większe wartości wariancji w danym wymiarze, tym trudniejsze określenie jest konkretnej pozycji partii w wymiarze ekonomicznym, czy też społecznym. W związku z tym, *ceteris paribus*, wyborcy z większego obszaru przestrzeni ideologicznej są skłonni oddać na daną partię głos. Na podstawie tak zdefiniowanych funkcji podobieństwa poglądów można określić prawdopodobieństwo oddania głosu na partię przez danego wyborcę. Osoba o poglądach (a, b) zagłosuje na k -tą spośród n partii z prawdopodobieństwem $P_k(a, b)$:

$$P_k(a, b) = \frac{f_k(a, b)}{\sum_{i=1}^n f_i(a, b)}, \quad (2)$$

gdzie f_i to funkcja podobieństwa poglądów względem partii i .

Rozkład wyborców na przestrzeni ideologicznej nie jest jednostajny – poglądy skrajne są znacznie rzadsze od poglądów umiarkowanych, w pobliżu centrum. W związku z tym została wprowadzona funkcja rozkładu wyborców $d(x, y)$, która określa jak wielu wyborców znajduje się w danym miejscu w przestrzeni. Zatem wynik wyborczy k -tej partii V_k może zostać określony za pomocą całki iloczynu funkcji rozkładu wyborców i prawdopodobieństwa oddania głosu na k -tą partię po przestrzeni ideologicznej S :

$$V_k = \iint_S d(x, y) P_k(x, y) dx dy. \quad (3)$$

Znając wszystkie wartości V_i , można w prosty sposób wyznaczyć procentowe poparcie k -tej partii wśród wyborców p_k :

$$p_k = \frac{V_k}{\sum_{i=1}^n V_i} * 100\%. \quad (4)$$

Przedstawiony powyżej model uzależnia decyzję wyborcy o oddaniu głosu na daną partię tylko i wyłącznie od stopnia podobieństwa poglądów wyborcy z programem partii, który jest reprezentowany poprzez położenie partii w przestrzeni ideologicznej. W rzeczywistym świecie tak jednak nie jest, dlatego konieczne jest przyjęcie założeń co do zachowania wyborców. Zostały one przedstawione poniżej.

Wyborcy posiadają jednoznacznie określone poglądy. Są oni w stanie określić swoje położenie w przestrzeni ideologicznej.

Wyborcy są doskonale zorientowani i poinformowani na temat aktualnej sytuacji politycznej. W szczególności znają oni programy wszystkich partii w systemie politycznym i ich położenie w przestrzeni ideologicznej.

Wszyscy wyborcy określają prawdopodobieństwo oddania swojego głosu na daną partię za pomocą tych samych funkcji podobieństwa poglądów.

Wyborcy niezależnie od swoich poglądów i niezależnie od partii występujących w systemie politycznym mają taką samą skłonność do wzięcia udziału w wyborach. Wobec tego zbędne staje się modelowanie, jaka część populacji nie odda głosu, ponieważ nie wpływa to na udziały oddanych głosów na poszczególne partie.

Wyborcy głosują z większym prawdopodobieństwem na partię o bliższych im poglądach niż na partię o dalszych, nie stosują głosowania taktycznego, nie są *single-issue voters*.

3. Źródła danych

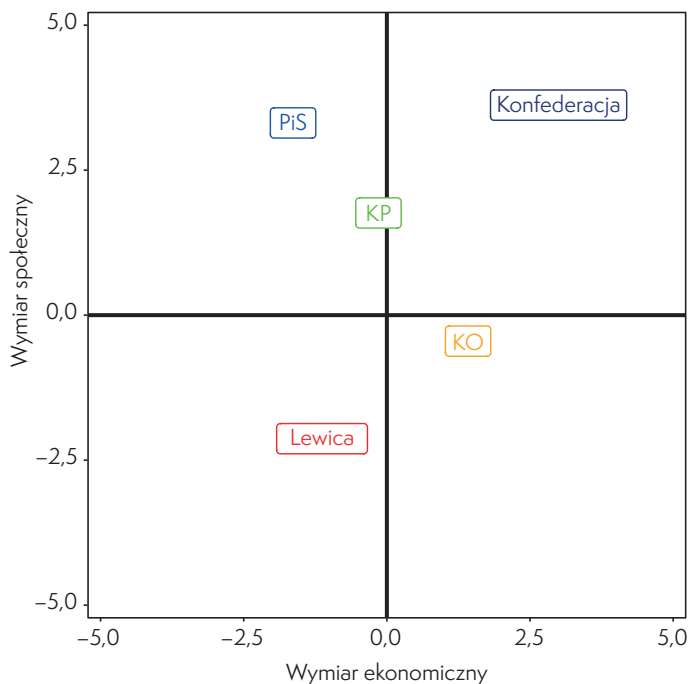
Z powyższej specyfikacji modelu wynika, że wykorzystuje on trzy typy parametrów:

- położenie partii w przestrzeni,
- rozkład wyborców w przestrzeni,
- macierze kowariancji partii.

Położenie pięciu polskich partii w przestrzeni ideologicznej zostało określone na podstawie danych *Chapel Hill Expert Survey*, CHES 2019 [Bakker i in. 2020], które przedstawia rysunek 1. Oś pozioma określa poglądy ekonomiczne między dwoma biegunami: ekonomiczna lewica (lewa połowa) – ekonomiczna prawica (prawa połowa). Oś pionowa określa poglądy społeczne szczegółowo opisywane

jako zieloni, alternatywni, libertarianie (GAL, dolna połowa) oraz tradycjoniści, autorytarni, nacjoniści (TAN, górna połowa). Zarówno poglądy ekonomiczne, jak i społeczne partii zostały ocenione przez ankietowanych ekspertów w 11-stopniowej skali, od -5 do 5 . W związku z tym rozważana dwuwymiarowa przestrzeń ideologiczna jest kwadratem o bokach w punktach $(-5,5)$, $(5,5)$, $(5,-5)$ oraz $(-5,-5)$.

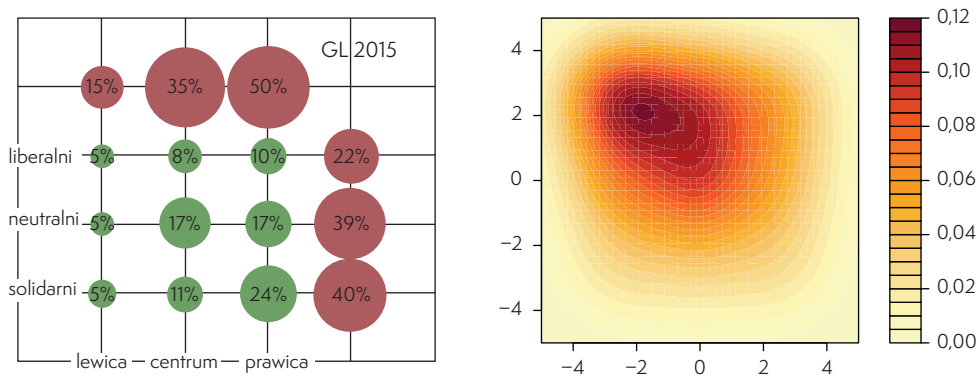
Rysunek 1. Położenie polskich partii politycznych w przestrzeni ideologicznej



Źródło: opracowanie własne na podstawie Bakker i in. [2020].

Do określenia rozkładu wyborców w przestrzeni zostały wykorzystane dane Polskiego Generalnego Studium Wyborczego 2015. W źródłowym zbiorze danych nazwy wymiarów przestrzeni ideologicznej zostały określone w inny sposób niż w przypadku CHES 2019. Flis i Kwiatkowska [2018] określają te dwa wymiary jako lewica – prawica oraz Polska solidarna – Polska liberalna. Flis i Kwiatkowska [2018] twierdzą, że podstawowym wymiarem jest oś lewica – prawica rozumiana jako oś poglądów społecznych, natomiast poglądy ekonomiczne są zoperacjonalizowane jako pozycja na osi Polska solidarna – Polska liberalna. Po przekształceniu tego systemu współrzędnych na współrzędne wykorzystane w CHES 2019 i ekstrapolacji danych prezentowanych na rysunku 2. A otrzymano funkcję rozkładu wyborców przedstawioną na rysunku 2. B.

Rysunek 2. A. Samoidentyfikacje deklarujących głosowanie w wyborach
B. Funkcja rozkładu wyborców $d(x, y)$



Źródło: A. Flis i Kwiatkowska [2018] na podstawie PGSW 2015; B. opracowanie własne na podstawie Flis i Kwiatkowska [2018].

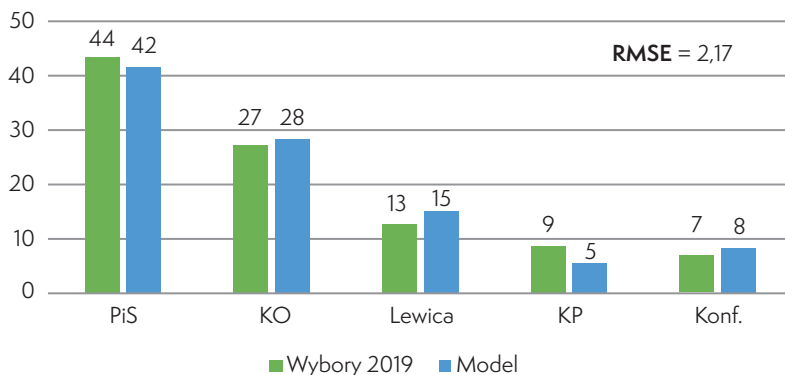
Jako że do tej pory model tego typu nie był używany do opracowania polskiego systemu politycznego, konieczne było dokonanie samodzielnej kalibracji macierzy kowariancji partii. W celu ustalenia wartości tych parametrów, mając za ustalone położenie partii i rozkład wyborców, poszukiwano takich wartości macierzy kowariancji, które minimalizowały pierwiastek z błędu średniokwadratowego procentowych wyników partii otrzymywanych za pomocą modelu względem rzeczywistych wyników wyborów parlamentarnych w 2019 r.:

$$\min RMSE = \min \sqrt{\frac{1}{5} \sum_{i=1}^5 (p_i - p_i^*)^2}, \quad (5)$$

gdzie p_i^* to procentowy wynik i -tej partii w wyborach parlamentarnych w 2019 r.

Do optymalizacji użyto algorytmu symulowanego wyżarzania, wykorzystując przekształcenia [Atwood, Foster, 1973] optymalizowanych zmiennych (10 wariancji i 5 korelacji), tak aby mieć zapewnione dodatnie określoności macierzy kowariancji tworzonych przez te zmienne. Otrzymane wyniki zostały przedstawione na rysunku 3. Różnice między rzeczywistymi wynikami wyborczymi a wynikami uzyskanymi za pomocą modelu po kalibracji macierzy kowariancji nie przekraczają 3 pp., a pierwiastek błędu średniokwadratowego (RMSE) wynosi jedynie 2,17. Model zatem dobrze przybliży rzeczywiste poparcie partii politycznych.

Rysunek 3. Porównanie wyników wyborów parlamentarnych w 2019 r. z wynikami wskazanymi przez model po kalibracji macierzy kowariancji (w %)



Źródło: opracowanie własne na podstawie obliczeń.

4. Problem optymalizacyjny

Problem optymalizacyjny rozważany w tym rozdziale to maksymalizacja wyniku wyborczego nowej partii politycznej. Mając dany model systemu partyjnego i istniejące w nim rzeczywiste partie polityczne, poszukiwane są takie parametry nowej partii, które umożliwią jej uzyskanie maksymalnie dużego poparcia wyborców. Optymalizowanych jest pięć parametrów, które jednoznacznie charakteryzują daną partię. Są to:

- położenie w przestrzeni ideologicznej $\mu = [x, y]$,
- macierz kowariancji Σ , która obliczana jest na podstawie trzech zmiennych:
 - wariancji partii w wymiarze ekonomicznym σ_x^2 ,
 - wariancji partii w wymiarze społecznym σ_y^2 ,
 - korelacji pomiędzy wymiarami ρ_{xy} .

Sformalizowana postać problemu optymalizacyjnego zaprezentowana jest wzorem 6.

$$\max_{\mu, \Sigma} \int_{-5}^5 \int_{-5}^5 d(x, y) P_6(x, y) dx dy \tag{6}$$

5. Opis algorytmu

Do rozwiązania problemu zaproponowano hybrydowy algorytm Tabu Search – Simulated Annealing z fazą intensyfikacji i restartu algorytmu za pomocą podejścia RRTS. Ustalona ilość iteracji algorytmu to 1000. Za punkty początkowe przyjęto wartości ze wzoru 7.

$$x_0 = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad M_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad (7)$$

gdzie x_0 to punkt początkowy położenia, a M_0 to początkowa macierz kowariancji. Punkt położenia kandydujący i parametry macierzy kowariancji kandydujące losowane są z kuli o promieniu $r = 0,15$ i środka w aktualnym punkcie, tj. punkcie składającym się z aktualnego punktu położenia i parametrów macierzy kowariancji.

Warunkiem przejścia do fazy intensyfikacji jest numer iteracji podzielny przez 10. Następuje wtedy przejście do nowego punktu takiego, że jego współrzędne losowane są z rozkładu jednostajnego, gdzie $x, y \in [-3, 3]$. Dodatkowo zwiększany jest promień sąsiedztwa do 0,1, wobec standardowego wynoszącego 0,05.

Restart dokonywany jest z prawdopodobieństwem $p = 0,8$, gdy spełniony jest następujący warunek zapisany wzorem 8.

$$|f(x_k) - f(x_{k-2})| < 10^{-3} \wedge |f(x_{k-1}) - f(x_{k-3})| < 10^{-3} \wedge |f(x_{k-2}) - f(x_{k-4})| < 10^{-3} \quad (8)$$

Restart oznacza losowe wybranie punktu położenia o współrzędnych $x, y \in [-3, 3]$ i zmianie promienia sąsiedztwa do 0,08. Faza restartu polega na kontynuacji przeszukiwania przestrzeni rozwiązań z wylosowanego punktu, w przeciwieństwie do fazy intensyfikacji, która jest jednorazową zmianą obszaru przeszukiwań powracającą do poprzedniej trajektorii. Funkcja aktywacji Metropolisia wyrażona jest we wzorze 9.

$$A(x_{k+1}^c, x_k, t_k) = \min(1, \exp((f(x_{k+1}^c) - f(x_k)) / t_k)), \quad (9)$$

gdzie x_{k+1}^c to punkt kandydujący.

Maksymalna długość listy tabu została ustalona na 150 rekordów. Przekroczenie długości listy oznacza dodanie nowego rekordu oraz usunięcie najstarszego za pomocą metody FIFO (ang. *First In, First Out*). Lista tabu i jej aktualizacja zapisana jest we wzorze 10.

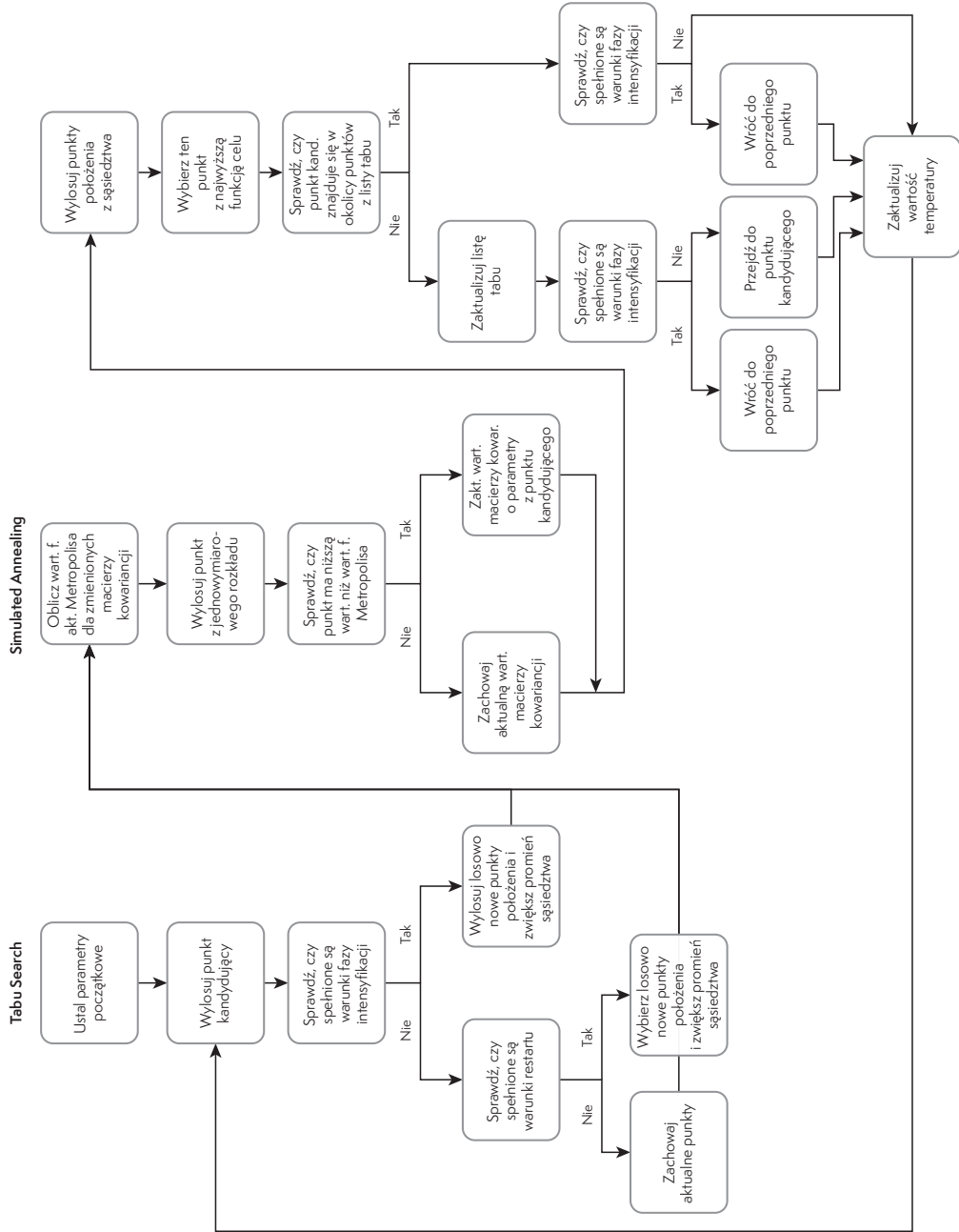
$$TL_{k+1} = FIFO(TL_k, x_{k+1}) \quad (10)$$

TL_k – lista tabu

$|TL_k| = n$

$TL_k = x_k, x_{k-1}, x_{k-2}, \dots, x_{k-n+1}$, gdzie k to numer iteracji.

Rysunek 4. Struktura proponowanego podejścia hybrydowego



Źródło: opracowanie własne na podstawie: Gendreau [2003, s. 165–186]; Bertsimas i Tsitsiklis [1993, s. 10–15] oraz Lenin [2016].

Punkt położenia z okolicy jednego z rekordów z listy tabu musi spełniać warunki dotyczące współrzędnych takie, że odległość euklidesowa współrzędnej x punktu położenia i odległość współrzędnej x punktu z listy tabu jest mniejsza lub równa 0,005 oraz odległość euklidesowa współrzędnej y punktu położenia i odległość współrzędnej y punktu z listy tabu jest mniejsza lub równa 0,005.

Początkowy parametr temperatury, t , ustalony został jako 100, a parametr kontrolny, α , ustalony został jako 0,99. Aktualizacja temperatury określona jest wzorem $t_{k+1} = \alpha \cdot t_k$. Po n -iteracjach wartość temperatury będzie miała wartość $t_n = \alpha^n \cdot t_0$.

Przebieg proponowanego hybrydowego algorytmu Simulated Annealing – Multi-start Tabu Search zaprezentowany został na rysunku 4.

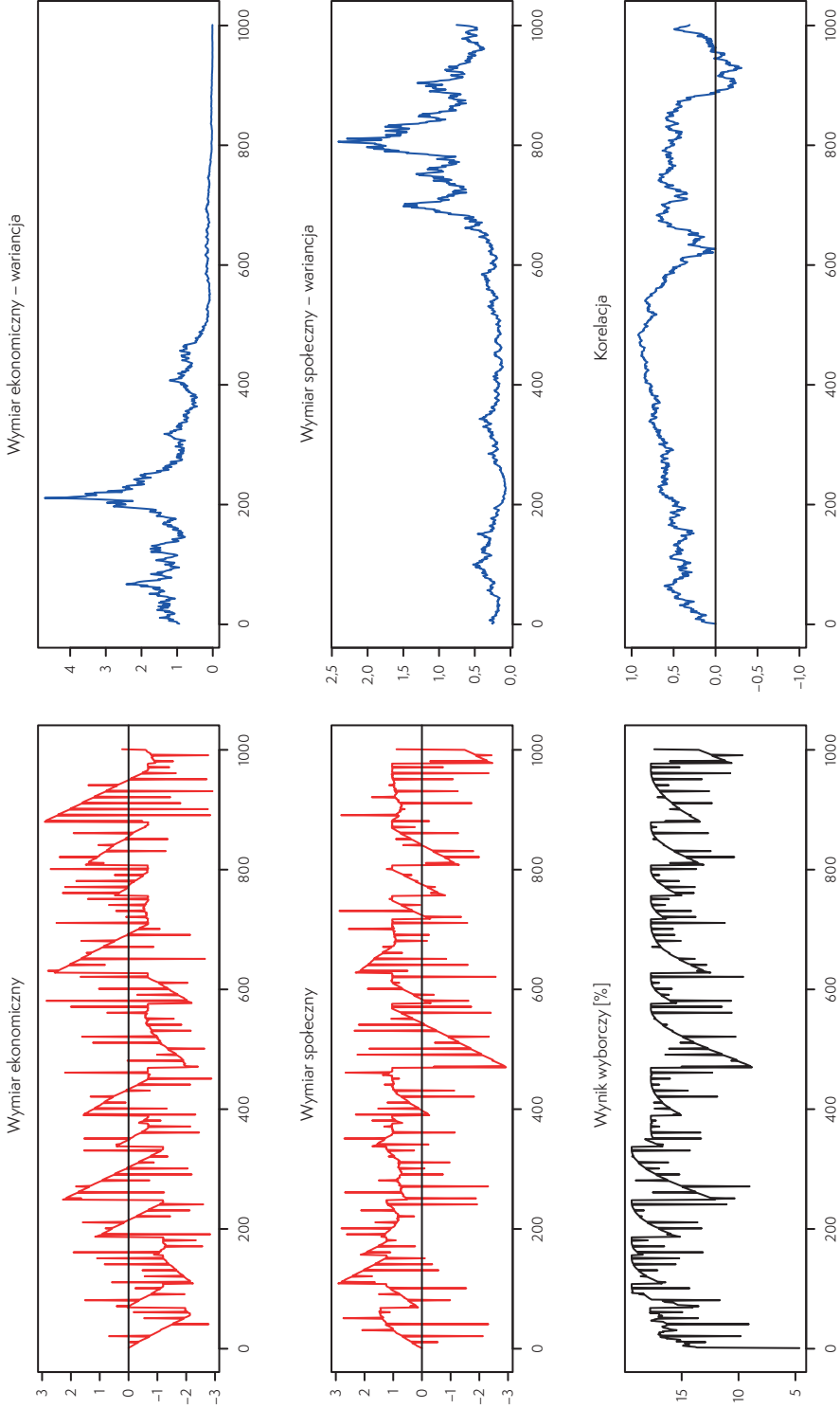
Schemat pokazuje opisane wyżej podejście w podziale na pojedyncze procesy. Zaprezentowaną pętlę wykonuje się określoną liczbę razy – w przypadku opisywanego problemu jest to 1000 iteracji.

6. Wyniki

Rysunek 5 przedstawia parametry nowej partii politycznej ze względu na numer iteracji. Położenie ekonomiczne (A) i społeczne (B) partii pokazuje, że w każdej fazie restartu wartość zbiega do punktu bliskiego optymalnemu punktowi położenia. Wykres funkcji celu (C) po każdym restarcie zbiega do okolic maksymalnej wartości funkcji celu. Widoczna jest również faza intensyfikacji charakteryzująca się jednostkowym dużym odchyleniem od trajektorii przeszukiwań w danym czasie oraz faza restartu cechująca się skokową zmianą wartości funkcji poprzez zmianę punktów położenia, a następnie gładkim przejściem w kierunku rozwiązania optymalnego. Druga kolumna przedstawia wykresy odpowiednio dla wariacji położenia ekonomicznego (D) i społecznego (E) partii, a następnie korelację pomiędzy poglądami ekonomicznymi a społecznymi partii (F).

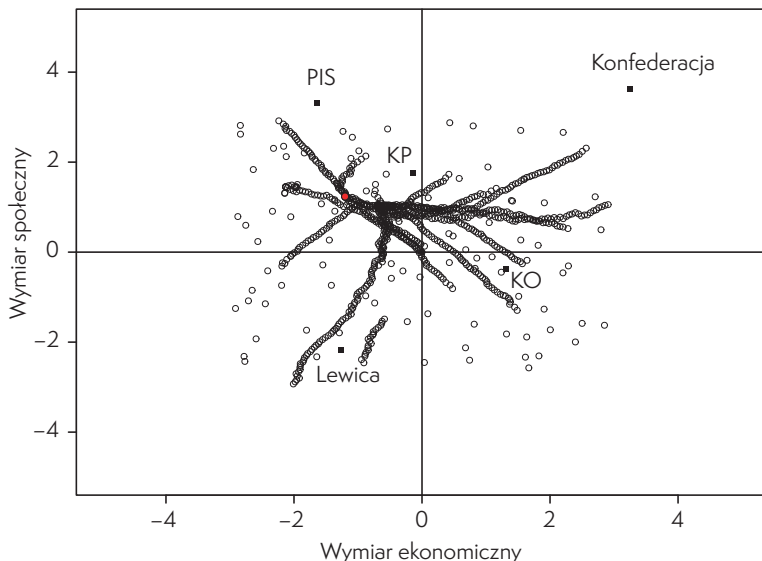
Rysunek 6 przedstawia przeszukiwane punkty położenia partii. Czerwonym kolorem zaznaczony został punkt położenia optymalny. Widoczna jest tutaj faza intensyfikacji (pojedyncze punkty) oraz faza restartu (rozpoczęcie przeszukiwań od nowych, wylosowanych punktów). Trajektorja przeszukiwań każdej fazy restartu zbiega do punktu położenia optymalnego. Duży obszar przeszukiwań i faza intensyfikacji dodatkowo potwierdzają tezę o optymalności znalezionej punktu położenia. Oznacza to, że znaleziony przez algorytm optymalny punkt położenia jest punktem położenia bliskim optymalnemu w rzeczywistości.

Rysunek 5. Parametry poglądów nowej partii politycznej w zależności od iteracji, (A) poglądy ekonomiczne, (B) poglądy społeczne, (C) wartości funkcji celu, (D) wariacja poglądów ekonomicznych, (E) wariacja poglądów społecznych, (F) wartość korelacji między poglądami



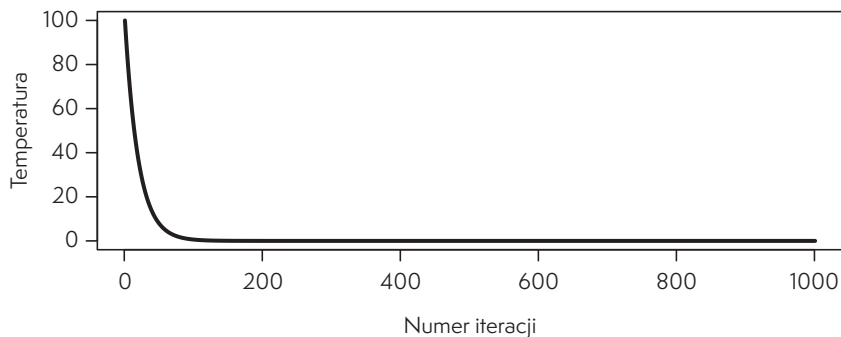
Źródło: opracowanie własne na podstawie obliczeń.

Rysunek 6. Przeszukiwane punkty położenia partii



Źródło: opracowanie własne na podstawie obliczeń.

Rysunek 7. Temperatura dla algorytmu Simulated Annealing



Źródło: opracowanie własne na podstawie obliczeń.

Rysunek 7 pokazuje spadek temperatury dla algorytmu Simulated Annealing. Widoczne zbieganie temperatury do niskich wartości bliskich 0 potwierdza, że algorytm przeszedł z fazy przeszukiwania losowego (ang. *random search*) do fazy zachłannej (ang. *greedy search*).

Zgodnie z uzyskanymi wynikami, nowa partia polityczna ma szansę na uzyskanie 19,4% poparcia. Przepływ elektoratu wygląda następująco:

- od partii Lewica: 2 pp.,
- od partii Koalicja Obywatelska: 4 pp.,

- od partii Koalicja Polska: 1 pp.,
- od partii PiS: 12 pp.,
- od partii Konfederacja: 0 pp.

Tabela 1 pokazuje wyniki i położenie partii politycznych. Nowa partia zajmuje trzecie miejsce pod względem poparcia, będąc zaklasyfikowaną za ugrupowania-
mi PiS i Koalicja Obywatelska.

Tabela 1. Wyniki i położenie partii politycznych

Partia	Ekonomiczne	Społeczne	Poparcie (%)
Lewica	-1,2599	-2,1725	13,2
KO	1,3095	-0,3648	24,2
KP	-0,1467	1,7618	3,9
PiS	-1,6381	3,3143	30,0
Konfederacja	3,2471	3,6211	7,7
Nowa partia	-0,7376	1,0199	19,4

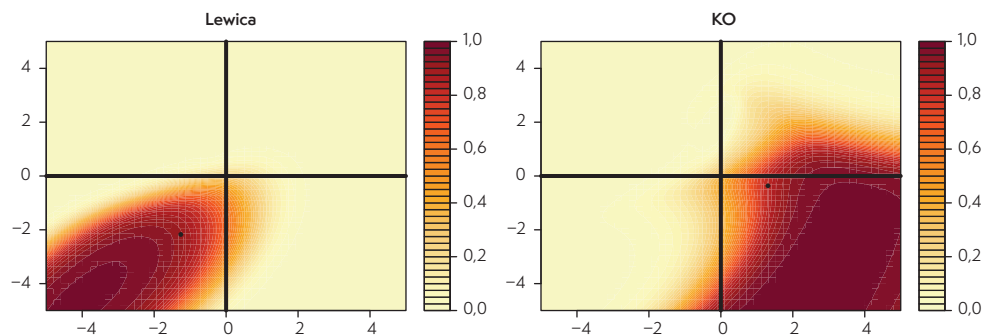
Źródło: opracowanie własne na podstawie obliczeń.

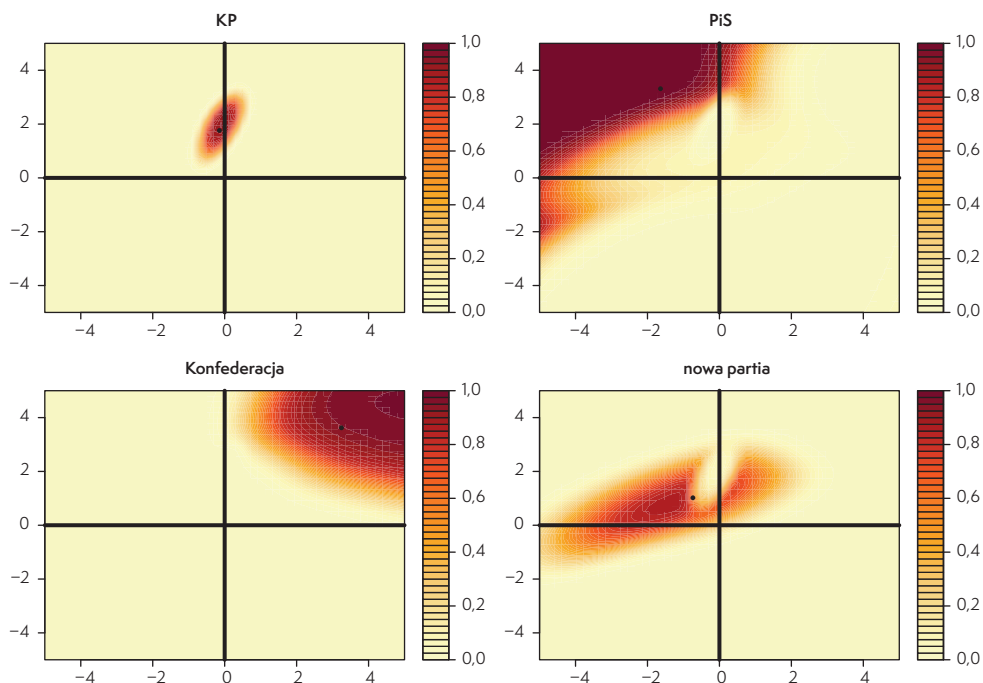
Punkt bliski maksimum globalnego poparcia osiągnięty został w trakcie 109 iteracji. Optymalny punkt położenia $x_o = \begin{bmatrix} -0,7376 \\ 1,0199 \end{bmatrix}$, a optymalna macierz

kowariancji:

$$M_o = \begin{bmatrix} 1,6 & 0,5 \\ 0,5 & 0,62 \end{bmatrix}.$$

Rysunek 8. Poparcie poszczególnych partii w zależności od położenia wyborców dla ugrupowania (A) Lewica, (B) KP, (C) Konfederacja, (D) KO, (E) PiS, (F) nowa partia





Źródło: opracowanie własne na podstawie obliczeń.

Rysunek 8 przedstawia rozkład poparcia poszczególnych partii politycznych w zależności od położenia grup wyborców. Wyższe poparcie w danej grupie zaznaczone jest kolorem brązowym. Atrakcyjność poszczególnych obszarów zależy od funkcji gęstości wyborców. Czarnym punktem zaznaczone jest położenie danej partii.

7. Podsumowanie

Problem badawczy w tym rozdziale dotyczył maksymalizacji wyniku wyborczego nowej partii politycznej w Polsce za pomocą hybrydowego algorytmu Simulated Annealing – Multi-start Tabu Search. W tym celu stworzono model odzwierciedlający zachowanie poparcia polskich partii politycznych oraz hybrydowy algorytm optymalizacyjny. Analiza dostarcza następujących wniosków:

- maksymalne poparcie, które może zdobyć nowa partia polityczna, to 19,4%;
- największą ilość wyborców utraci ugrupowanie PiS;
- nowa partia polityczna powinna mieć poglądy prawicowe w kwestiach światopoglądowych i lewicowe w kwestiach gospodarczych.

Zaproponowane rozwiązanie może zostać zastosowane w innych warunkach politycznych dzięki modyfikacji parametrów. Kwestiami do przyszłych badań jest wykorzystanie problemu do podobnej analizy w innych krajach. Istotne jest również wykorzystanie innego algorytmu optymalizacyjnego do rozwiązania problemu. Warta dalszych badań będzie również próba uchylecia części lub wszystkich założeń. W celu odniesienia się do zaproponowanych, a także innych, podobnych rozważań, warto wyjść od analizy przeprowadzonej w niniejszym rozdziale.

Bibliografia

- Aarts, E., Korst, J., Michiels, W. (2005). *Simulated Annealing*. Springer.
- Askari, Q., Younas, I., Saeed, M. (2020). Political Optimizer: A novel socio-inspired meta-heuristic for global optimization, *Knowledge-Based Systems*, 195.
- Atwood, G.R., Foster, W.W. (1973). Transformation of bounded variables in simplex optimization techniques, *Industrial and Engineering Chemistry Process Design and Development*, 12(4), s. 485–486.
- Bakker, R., Hooghe, L., Jolly, S., Marks, G., Polk, J., Rovny, J., Steenbergen, M., Vachudova, M.A. (2020). 2019 Chapel Hill Expert Survey. Chapel Hill, NC: University of North Carolina, Chapel Hill.
- Bertsimas D., Tsitsiklis, J. (1993). Simulated Annealing, *Statistical Science* 8(1), s. 10–15.
- Cześniak, M., Kotnarowski, M., Żerkowska-Balas, M. (2015). New dimensions of political competition: Liberal versus Solidary Poland, Paper at ECPR Joint Sessions.
- Dale, S. (2015). Heuristics and biases: The science of decision-making, *Business Information Review*.
- Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper.
- Flis, J., Kwiatkowska, A. (red.). (2018). *Polski system partyjny: idee i identyfikacje*. Warszawa: Oficyna Wydawnicza ASPRA-JR.
- Gendreau, M. (2003). *An Introduction to Tabu Search*. Springer.
- Glover, F., Laguna, M. (1998). *Tabu Search*. Kluwer Academic Publishers.
- Hooghe, L., Marks, G., Wilson, C.J. (2002). Does left/right structure party positions on European integration?, *Comparative Political Studies*, 35(8), s. 965–989.
- Inglehart, R. (1977). Long term trends in mass support for European unification, *Government and Opposition*, 12(2), s. 150–177.
- James, T. (2009). Multistart tabu search and diversification strategies for the quadratic assignment problem, *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 39(3), s. 579–596.
- Kriesi, H., Grande, E., Lachat, R., Dolezal, M., Bornschier, S., Frey, T. (2006). Globalization and the transformation of the national political space: Six European countries compared, *European Journal of Political Research*, 45(6), s. 921–956.

- Lenin, K., Reddy, B.R., Suryakalavathi, M. (2016). Hybrid Tabu search-simulated annealing method to solve optimal reactive power problem, *International Journal of Electrical Power & Energy Systems*, 82(1), s. 87–91.
- Martí, R. (2003). Multi-Start Methods. W: *Handbook of Metaheuristics* (s. 355–368), F. Glover, G.A. Kochenberger (Eds.). Boston, MA: Springer US.
- Metropolis, N., Rosenbluth, M.N., Rosenbluth, A.W., Teller, A.H., Teller, E. (1953). Equation of state calculations by fast computing machines, *The Journal of Chemical Physics*, 21(6), s. 1087–1092.
- Mishra, N., Prakash, Tiwari, M.K., Shankar, R., Chan, F. (2005). Hybrid tabu-simulated annealing based approach to solve multi-constraint product mix decision problem, *Expert Systems with Applications*, 29(2), s. 446–454.
- Misra, A.K. (2012). A simple mathematical model for the spread of two political parties, *Non-linear Analysis: Modelling and Control*, 17(3), s. 343–354.
- Osman, I.H., Laporte, G. (1996). Metaheuristics: A bibliography, *Annals of Operations Research*, 63(5), s. 511–623.
- Ostapenko, V., Ostapenko, O., Belyaeva, E., Stupnitskaya, Y. (2012). Mathematical models of the battle between parties for electorate or between companies for markets, *Cybernetics and Systems Analysis*, 48(6), s. 814–822.
- Sörensen, K., Glover, F. (2013). Metaheuristics. W: *Encyclopedia of Operations Research and Management Science*, s. 960–970.
- Swarnkar, R., Tiwarim, M.K. (2004). Modeling machine loading problem of FMSs and its solution methodology using a hybrid tabu search and simulated annealing-based heuristic approach, *Robotics and Computer-Integrated Manufacturing*, 20(3), s. 199–209.
- Yang, V.C., Abrams, D.M., Kernell, G., Motter, A.E. (2020). Why are U.S. parties so polarized? A “satisficing” dynamical model, *SIAM Review*, 62(3), s. 646–657.

What Should a New Polish Party Be like? Maximising the Election Result of a Political Party using a Hybrid Algorithm Simulated Annealing – Multi-Start Tabu Search

Abstract

The choice of a preferred party during parliamentary elections is a complex social science issue. The use of mathematical methods may complement empirical and theoretical research in this field. The distribution of Polish people’s political views was computed based on extrapolation. Then, using Simulated Annealing, which is a probabilistic algorithm, the covariance matrices of political parties’ views were determined, minimising the difference between the actual and the model-base support. Hybrid metaheuristic algorithm Simulated Annealing – Multi-start Tabu Search was proposed

to find the parameters of a new political party views which might maximise its election result. The results indicate that a new political party has a chance to get 19.4% of votes in the elections, which is the third highest result among political parties in the Polish Parliament (Sejm). The article also presents the directions of further research and development of the proposed model.

Keywords: metaheuristics, Tabu Search, Simulated Annealing, political parties, political support

Modelowanie procesów w kontekście projektów robotyzacji procesów biznesowych

Jakub Stelmaszczyk

Streszczenie

Przedmiotem pracy jest eksploracja aktualnych tendencji w kontekście modelowania procesów w projektach wykorzystujących technologię robotyzacji procesów biznesowych (ang. *Robotic Process Automation* – RPA), poprzez analizę publikacji naukowych dotyczących tematyki RPA. Dodatkowo wykonano badanie ankietowe wśród polskich specjalistów branży RPA w celu poznania ich opinii oraz preferencji na temat modelowania procesów. Tematyka modelowania procesów w projektach RPA jest istotna, ponieważ diagramy procesów są nieodzownym elementem analizy wykonywanej na potrzeby projektów robotyzacji procesów. Ze względu na powszechne stosowanie diagramów mogą mieć one istotny wpływ na wyniki projektów RPA, przez co, przy braku standaryzacji lub niskiej jakości diagramów, na etapie analizy mogą powstawać luki analityczne powodujące opóźnienia w projektach robotycznych oraz dostarczanie niepełnych funkcjonalności.

Słowa kluczowe: robotyzacja procesów biznesowych, RPA (*Robotic Process Automation*, RPA), modelowanie procesów, diagramy procesów, BPMN, UML, schemat blokowy, analiza biznesowa, automatyzacja

1. Wstęp

W ostatnich latach technologia robotyzacji procesów biznesowych (ang. *Robotic Process Automation* – RPA) jest na etapie intensywnego wzrostu, jeśli chodzi o przychody dostawców tej klasy narzędzi. Coraz więcej firm dostrzega możliwości, jakie daje technologia RPA, umożliwiającą otrzymanie znaczących korzyści (np. ograniczenie kosztów, przyspieszenie procesów) w krótkim czasie. Technologia ta może służyć także jako rozwiązanie taktyczne niektórych problemów firm (jak np. braki kadrowe, rosnące wynagrodzenia, jakość danych) w długofalowej strategii IT przed-

siębiorstwa. Zwiększone inwestycje w tę technologię przekładają się na wiele projektów wdrożeniowych, a także na tworzenie tzw. centrów doskonałości (ang. *Centers of Excellence* – CoE), co z kolei przekłada się na zwiększone zapotrzebowanie na specjalistów znających rozwiązania RPA. W ramach tych wszystkich przedsięwzięć, ci specjaliści będą spotykać się z diagramami procesów, które są nieodzownymi elementami analizy procesów przeznaczonych do robotyzacji, tworzenia koncepcji rozwiązania przy użyciu technologii RPA oraz dokumentacji efektów tych projektów. Aktualnie, ze względu na młody wiek branży RPA, nie został wykształcony żaden ogólnie przyjęty standard modelowania procesów w projektach, zarówno na etapie analizy i rozpoznawania procesu biznesowego wykonywanego przez pracowników, jaki i dokumentacji tego, jak docelowo będzie wyglądał proces wykonywany przez robota.

Celem tej publikacji jest eksploracja aktualnych tendencji przy wyborze notacji do diagramów wykorzystywanych w trakcie modelowania procesów w projektach RPA, a także zbadanie preferencji specjalistów w kwestii tworzenia diagramów. W tym celu dokonano przeglądu literatury poświęconej tematyce RPA pod kątem wykorzystanych w niej diagramów procesów, a także przeprowadzono badanie ankietowe wśród polskich specjalistów branży RPA, dotyczące ich aktualnych doświadczeń oraz preferencji w obszarze modelowania procesów. Na podstawie przeglądu i badania ankietowego autor artykułu przedstawił aktualny obraz modelowania procesów w projektach RPA, a także wskazał propozycję obszarów, które należałoby poprawić w celu usprawnienia przebiegów projektów wykorzystujących tę technologię.

2. Podstawy teoretyczne modelowania procesów w projektach RPA

Modelowanie procesów jest bardzo szerokim tematem. Diagramy procesów są m.in. tworzone w ramach:

- dokumentacji tego, jak wyglądają dane procesy biznesowe w przedsiębiorstwie;
- spełnienia wymagań w celu pozyskania certyfikatu ISO;
- dokumentacji projektów IT w celu zrozumienia otoczenia biznesowego, w jakim rozwiązanie informatyczne będzie implementowane;
- dokumentacji projektów IT w celu określenia wymagań biznesowych odnośnie do oprogramowania.

W tej publikacji modelowanie procesów będzie rozpatrywane w ramach bardzo wąskiego wycinka tej dziedziny, czyli pod kątem wykorzystania modelowania procesów w projektach RPA. W tego typu projektach produktem pracy anality-

ków biznesowych są często diagramy prezentujące procesy w formie zastanej (ang. *as-is*, pol. jak jest) oraz w formie docelowej (ang. *to-be*, pol. jak będzie). W tej części pracy podjęta zostanie próba zdefiniowania zarówno technologii RPA, jak i modelowania procesów w jej kontekście, tak aby w jasny sposób określić tematykę niniejszej pracy.

2.1. Czym jest RPA

W celu zdefiniowania technologii robotyzacji procesów biznesowych posłużyć się publikacją organizacji IEEE z 2017 r., przytoczoną przez uznane autorytety w branży RPA, Mary C. Lacity i Leslie P. Willcocks w książce *Robotic Process and Cognitive Automation The Next Phase* [2018]. W publikacji IEEE *Przewodnik po terminach i pojęciach w inteligentnej automatyzacji* (ang. *Guide on Terms and Concepts in Intelligent Automation*) [2017] robotyzacja procesów biznesowych została zdefiniowana następująco: „Robotyzacja procesów biznesowych: wstępnie skonfigurowana instancja oprogramowania, która wykorzystuje reguły biznesowe i wstępnie zdefiniowaną choreografię działań do ukończenia kombinacji procesów, działań, transakcji i zadań w jednym lub większej liczbie niepowiązanych systemów oprogramowania w celu dostarczenia wyniku lub usługi z zaangażowaniem ludzi jedynie w sytuacjach wyjątkowych”.

Z powyższej definicji wynika, że RPA jest technologią wchodzącą w szerszy obszar technologii automatyzacji. Dodatkowo w celu sprecyzowania czym jest technologia RPA, można przytoczyć wnioski prof. Andrzeja Sobczaka, który w publikacji *Budowa modelu zarządzania robotyzacją procesów biznesowych* [Sobczak, 2019] wyłonił dla tej technologii cztery następujące wyróżniki:

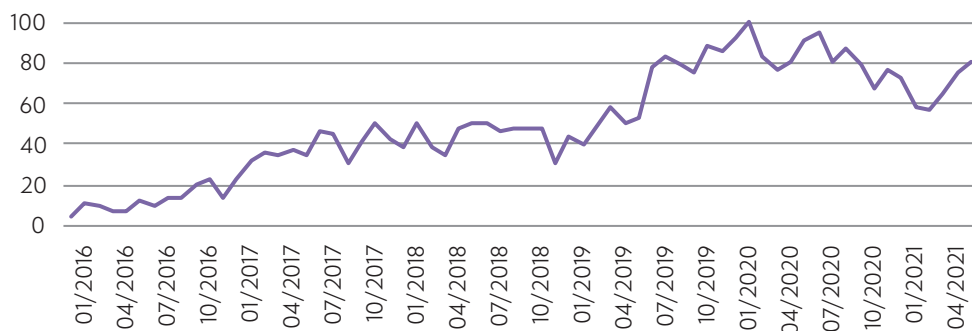
- jest to oprogramowanie działające w oparciu o zadany algorytm, ale coraz częściej wzbogacane o pewne elementy sztucznej inteligencji, dzięki czemu jest w stanie podejmować bardziej złożone decyzje (w szczególności uczyć się na podstawie dostarczonych danych – zarówno ustrukturyzowanych, jak i nieustrukturyzowanych);
- służy do zautomatyzowania operacji wielkoskalowych (masowych) (tj. wykonywanych wielokrotnie w założonej jednostce czasu – np. w ciągu miesiąca lub roku), realizowanych dotychczas przez operatora, najczęściej poprzez ich wierne odtworzenie;
- oprogramowanie wykonuje operacje bezpośrednio na graficznym interfejsie użytkownika (ang. *Graphical User Interface* – GUI) systemów informatycznych – tak samo, jak robi to człowiek-operator (choć niektóre narzędzia RPA umożliwiają wykonywanie operacji bezpośrednio na warstwie danych);

- oprogramowanie wykorzystuje logikę biznesową stanowiącą integralną część aplikacji, z którymi będzie pracował robot programowy, co eliminuje problem odtwarzania takiej logiki, który występuje w tradycyjnych modelach integracji systemów lub rozwoju.

2.2. Aktualna pozycja technologii RPA na rynku

Technologia RPA jest relatywnie nowym zjawiskiem w branży IT. Szczyt popularności (ang. *hype*) tej technologii przypadł na początek 2020 r. W ostatnim czasie widać, że RPA przechodzi z fazy wczesnej adopcji (ang. *early adopters*) w fazę bardziej dojrzałą, przez co zainteresowanie nieznacznie wyhamowało. Wyhamowanie może być także związane z obniżeniem nakładów inwestycyjnych w projekty IT w wyniku pandemii COVID-19. Na rysunku 1 zaprezentowana została popularność hasła RPA w wyszukiwarce Google od 2016 r. Jak widać wzrost w drugim kwartale 2021 jest bardzo dynamiczny – może to wynikać z faktu zwiększonego popytu na robotyzację procesów, wynikającego z nowej potrzeby uodpornienia kluczowych procesów w przedsiębiorstwach na niedobór pracowników, o którym kadry zarządzające mogły się przekonać w czasie zamknięcia gospodarki związanego z reżimem sanitarnym i przeniesienia pracowników na pracę zdalną.

Rysunek 1. Popularność frazy „Robotic Process Automation” w wyszukiwarce Google



Źródło: Google Trends, zainteresowanie hasłem „Robotic Process Automation” w wyszukiwarce Google od 2016 r.¹

Według raportu Gartnera z września 2020 r. rynek dostawców RPA w 2019 odnotował przychody rzędu 1,4 mld USD, natomiast przewidywania na temat dalszych

¹ Informacja nt. wartości na osi Y dostępna na stronie Google Trends: „Liczby reprezentują poszczególne zainteresowania w wyszukiwaniu względem najwyższego punktu na wykresie. Wartość 100 oznacza najwyższą popularność hasła. Wartość 50 oznacza, że popularność hasła była dwukrotnie mniejsza. Wartość 0 wskazuje, że dla danego hasła nie ma wystarczających danych”, <https://trends.google.pl/trends/?geo=PL> (dostęp 29.04.2021).

wzrostów wyestymowano na 11,9% do poziomu blisko 1,6 mld USD w 2020 r. oraz kolejnych 19,5% wzrostu w 2021 r. do poziomu blisko 1,9 mld USD [Costello, Rimol, 2020]. Powyższe estymacje zostały zaprezentowane w poniższej tabeli.

Tabela 1. Aktualne i przewidywane przychody dostawców narzędzi RPA

	2019	2020	2021
Przychody (\$M)	1411,1	1579,5	1888,1
Wzrost (%)	62,93	11,94	19,53

Źródło: Gartner (wrzesień 2020).

2.3. Modelowanie procesów w kontekście projektów RPA

Cykl życia projektu RPA można sprowadzić do sześciu kluczowych etapów [Enriquez, Jimenez-Ramirez, Dominguez-Mayo, Garcia-Garci, 2020; Jimenez-Ramirez, Reijers, Barba, Del Valle, 2019].

- 1) **Etap analizy** – na tym etapie rozpoznane zostają przebiegi potencjalnych procesów możliwych do zrobotyzowania. Tworzony jest diagram procesu w formie zastanej (ang. *as-is*), często jako element szerszego dokumentu nazywanego dokumentem definicji procesu (ang. *Process Definition Document* – PDD).
- 2) **Etap projektowania** – po zakwalifikowaniu danego procesu do projektu robotyzacyjnego, następuje etap projektowania, którego wynikiem ma być koncepcja rozwiązania wykorzystującego technologię RPA w danym procesie. W ramach tego etapu tworzone są diagramy procesu w jego docelowej formie (ang. *to-be*), czyli tego, jak ma wyglądać dany proces po wdrożeniu rozwiązania RPA. Diagramy procesu w formie docelowej często są elementami szerszego dokumentu nazywanego dokumentem projektu rozwiązania (ang. *Solution Design Document* – SDD). Diagramy i dokumenty stworzone na tym etapie często stanowią dokumentację rozwiązania RPA.
- 3) **Etap wytwarzania** – na tym etapie RPA deweloperzy tworzą instrukcję dla robota przy wykorzystaniu jednego lub wielu narzędzi od dostawców RPA. RPA deweloperzy wytwarzają roboty głównie w oparciu o dokument SDD i diagramy procesów tam zawartych. Na tym etapie zarówno dokument SDD, jak i diagramy mogą ewoluować, ponieważ podczas wytwarzania robota często nie wszystkie początkowe założenia co do konstrukcji rozwiązania są możliwe do spełnienia. Możliwe jest również, że RPA deweloper będzie w stanie zaproponować bardziej efektywne rozwiązanie, niż to określone w projekcie. W takiej sytuacji deweloperzy oraz analitycy modyfikują dokument SDD, tak aby wiernie

odzwierciedlał on konstrukcję finalnego rozwiązania, dzięki czemu dokument będzie mógł pełnić funkcję dokumentacji po zakończeniu projektu.

- 4) **Etap wdrożenia** – po etapie wytworzenia, działające roboty wdrażane są w indywidualnych środowiskach (najczęściej są to maszyny wirtualne (ang. *Virtual Machines* – VM)), tak aby przetestować, jak radzą sobie ze wszystkimi programami i aplikacjami obecnymi w procesie. Jest to etap, który ograniczy negatywne skutki ewentualnych błędów, które mogłyby zakłócić przebieg procesu w środowisku produkcyjnym.
- 5) **Etap testowania i kontrolowania** – zwany także etapem szczególnej opieki lub hiperopieki (ang. *hypercare*). Na tym etapie robot zostaje wdrożony w środowisko produkcyjne i zaczyna procesować przebiegi realnego procesu. Wykrywane są ścieżki przebiegu, nieuwzględnione w poprzednich etapach, i wytwarzany jest sposób ich obsługi. Prace na tym etapie prowadzone są do momentu osiągnięcia „stabilności” pracy robota, czyli akceptowalnego poziomu ilości zatrzymań robota oraz odsetka obsłużonych przebiegów procesu.
- 6) **Etap operacyjny i utrzymaniowy** – na tym etapie robot staje się pełnoprawnym pracownikiem i wykonuje zadane przebiegi procesów według określonego harmonogramu. Monitorowana jest wydajność robota oraz ewentualne przestoje lub błędy. Możliwe staje się również rozpoczęcie nowego cyklu analizy i projektowania w celu ulepszenia pracy robota.

Zagadnienie modelowania procesów jest obecne w szczególności w pierwszym i drugim etapie cyklu projektu RPA. To właśnie na tych etapach powstają diagramy procesów w formie zastanej i docelowej. Warto jednak zauważyć, że na etapach od trzeciego do piątego diagramy w formie docelowej mogą jeszcze ewoluować i być zmieniane w zależności od potrzeb implementacyjnych i nowych nieprzewidzianych przebiegów, niewykrytych na etapie analizy. Natomiast diagramy w formie zastanej po pierwszym etapie projektu co do zasady pozostają niezmienione. W niniejszej pracy, w ramach rozważań nad podejściem do modelowania procesów w projektach RPA, brane będą pod uwagę właśnie diagramy, które powstają na pierwszym i drugim etapie projektu robotycznego.

3. Materiały i metody ich pozyskania

Mając podstawy teoretyczne zarówno z fundamentalnych informacji i pojęć branży RPA oraz modelowania procesów w kontekście projektów robotycznych, w niniejszej części publikacji analizie zostaną poddane prace naukowe poświęcone

tematyce RPA oraz badanie ankietowe przeprowadzone wśród polskich ekspertów z branży RPA. Poniżej opisane zostały zebrane materiały oraz metody ich pozyskania.

3.1. Przegląd literatury naukowej z dziedziny RPA

Podczas przeglądu literatury fachowej przeszukane zostały trzy następujące bazy: Springer Link², ResearchGate³ i Google Scholar⁴.

Ze względu na fakt, że technologia RPA jest relatywnie nowym zjawiskiem w branży IT i liczba publikacji w tym obszarze nie jest jeszcze duża, autor zdecydował się na wyszukiwanie jedynie anglojęzycznych pozycji. We wszystkich trzech bazach wyszukiwane były następujące frazy: „robotic process automation”, „robotic process automation case study”, „RPA case study” oraz „process modelling”.

W przypadku bazy Springer Link oraz ResearchGate do następnego etapu uwzględniono po ok. 50 pierwszych wyników, natomiast w bazie Google Scholar ok. 100 pierwszych wyników, co dało początkową pulę ok. 200 publikacji.

Kolejnym krokiem było wyodrębnienie tylko publikacji zawierających jakiegokolwiek graficzne reprezentacje procesów (niezależnie czy w formie zastanej, docelowej, poglądowej itp.), które można interpretować jako diagramy procesów w dowolnej notacji (wliczając w to także notacje własne). Po tym etapie wyodrębniono 28 publikacji.

Następnie odrzucono publikacje, które zawierały diagramy procesów w języku innym niż angielski, co uniemożliwiało jednoznaczne zrozumienie prezentowanych przez diagramy procesów. Na tym etapie odrzucono trzy publikacje.

Ostatnim kryterium było, aby diagramy w publikacji odzwierciedlały rzeczywiste środowisko biznesowe projektów RPA, tzn. mogłyby być efektem prac prowadzonych na pierwszym lub drugim etapie cyklu życia projektu RPA. Wszelkie diagramy o charakterze czysto poglądowym zostały odrzucone. Na tym etapie odrzucono trzy publikacje, co oznaczało, że w finalnej analizie przeglądu literaturowego uwzględnione zostały 22 publikacje.

3.2. Badanie ankietowe wśród specjalistów branży RPA

W celu zbadania aktualnych tendencji oraz preferencji dotyczących modelowania procesów wśród specjalistów RPA, przeprowadzono badanie ankietowe w formie online. Forma ta w dużej mierze została podyktowana obostrzeniami sanitarnymi

² <https://link.springer.com/> (dostęp: 1.03.2021–31.03.2021).

³ <https://www.researchgate.net/> (dostęp: 1.03.2021–31.03.2021).

⁴ <https://scholar.google.com/> (dostęp: 1.03.2021–31.03.2021).

wprowadzonymi w związku z pandemią COVID-19. Ankieta została przeprowadzona w języku polskim i składa się z czterech części, mających następujące cele:

- 1) określenie doświadczenia zawodowego i roli respondenta w projektach RPA,
- 2) zbadanie dotychczasowych doświadczeń respondentów związanych z modelowaniem procesów w projektach RPA,
- 3) zbadanie poziomu wiedzy respondentów oraz jej źródeł odnośnie do modelowania procesów,
- 4) zbadanie preferencji respondentów dotyczących modelowania procesów w projektach RPA.

W badaniu ankietowym wzięło udział 33 respondentów, spośród których jeden oznaczył, że nie należy do grupy specjalistów branży RPA, więc ostatecznie analiza została przeprowadzona na 32 wypełnionych ankietach. Liczba odpowiedzi nie jest wystarczająca, aby stanowić wiarygodne źródło analizy statystycznej, dlatego analiza wyników badania ankietowego w niniejszej pracy ma jedynie charakter poglądowy oraz wspomagający, jako dopełnienie analizy literatury fachowej.

4. Rezultaty przeglądu literatury naukowej

W ramach przeglądu literatury została wykonana tabela 2, w której przedstawiono publikacje uwzględnione do analizy (selekcja publikacji została opisana w poprzedniej części artykułu). W tabeli wyróżniono rok publikacji, użyte notacje w ramach diagramów znajdujących się w publikacji oraz jej typ. Klasyfikacja według powyższych parametrów prezentuje się następująco:

Tabela 2. Klasyfikacja publikacji z podziałem na użytą notację oraz rodzaj publikacji

Autorzy publikacji	Rok	Użyta notacja				Rodzaj publikacji		
		BPMN	UML	Schemat blokowy	Notacja własna	Artykuł naukowy	Praca magisterska	Praca licencjacka
Y. Rizk i in.	2020	X				X		
M. König i in.	2020	X	X			X		
T. Chakraborti i in.	2020	X				X		
S. Aguirre i in.	2017	X				X		
C. Flechsig i in.	2019	X				X		
D. Tran, T. Ho	2018			X	X			X
B. Amini	2019	X			X			X
J. Holma	2020	X						X
T. Qvick	2019	X		X			X	

Autorzy publikacji	Rok	Użyta notacja				Rodzaj publikacji		
		BPMN	UML	Schemat blokowy	Notacja własna	Artykuł naukowy	Praca magisterska	Praca licencjacka
J. Hindel i in.	2020	X	X			X		
M. Romão i in.	2019	X				X		
A. Wróblewska i in.	2018				X	X		
M. Cernat i in.	2020				X	X		
D. Šimek i in.	2019				X	X		
J. Myllymäki	2019				X		X	
V. Gružasuskas i in.	2020	X				X		
F. Johnson i in.	2018				X		X	
A. Kasic	2020	X					X	
S. Anagnoste	2018				X	X		
J. Härd i in.	2020			X	X		X	
F. da Silva Santos	2019	X					X	
T. Malathi i in.	2020			X		X		

Źródło: opracowanie własne.

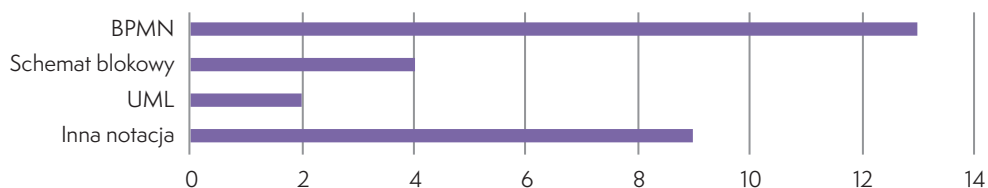
Warto zwrócić uwagę na lata publikacji. Najstarsza zakwalifikowana publikacja pochodzi z 2017 r., natomiast blisko połowa wszystkich publikacji została wydana w roku poprzedzającym niniejszy artykuł. Wskazuje to na fakt, że RPA jest bardzo młodą technologią i dorobek naukowy w tej dziedzinie nie jest jeszcze duży. Lata publikacji potwierdzają także trend popularności wyszukiwań frazy „robotic process automation” zaprezentowany na rysunku 1.

Kolejną konsekwencją faktu, iż RPA jest nową technologią, jest to, że aż 9 z 22 publikacji zakwalifikowanych do analizy to prace dyplomowe (6 – magisterskie, 3 – licencjackie). Istotnym czynnikiem, który miał wpływ na liczbę prac dyplomowych zakwalifikowanych do analizy, może być też bardziej praktyczna forma prac dyplomowych, która często prezentuje studium przypadku użytej technologii (ang. *case study*), co przekłada się na potrzebę zaprezentowania diagramów obrazujących robotyzowany proces. Prace dyplomowe pochodzą z baz Google Scholar oraz ResearchGate, które uwzględniają prace publikowane przez uczelnie w krajach, w których taka praktyka jest powszechna.

Najważniejszą statystyką w tej części artykułu jest liczba wystąpień wybranych notacji w analizowanych publikacjach. Wyróżniono trzy najczęściej występujące notacje: BPMN, schemat blokowy oraz UML. W przypadku zastosowania innej notacji niż te wymienione powyżej lub użycia przez autorów notacji własnej, publikacja

została zakwalifikowana jako „Inna notacja”. Wyniki opisanej klasyfikacji zostały zaprezentowane na rysunku 2.

Rysunek 2. Liczba wystąpień wybranych notacji



Źródło: opracowanie własne.

Pod względem oficjalnych notacji zdecydowaną przewagą wyróżnia się notacja BPMN (ang. *Business Process Model and Notation*), która została użyta w ponad połowie analizowanych publikacji. Prawdopodobnie decydujący wpływ na wybór tej notacji ma jej procesowy charakter, który umożliwia przedstawienie najistotniejszych aspektów robotyzowanych procesów. Sama nazwa notacji, zawierająca słowo „proces” (ang. *process*), może sprawiać, że wydaje się ona bardziej adekwatna do projektów wykorzystujących technologię o nazwie RPA, która także zawiera słowo „proces” w swoim akronimie. Autorzy publikacji używali notacji BPMN zarówno do prezentowania wysokopoziomowych diagramów procesów, których to część będzie poddawana robotyzacji, jak i modelowaniu dokładnej instrukcji (krok po kroku) dla robota. Jako że notacja BPMN okazała się najpopularniejszą notacją pod względem wystąpień diagramów stworzonych przy użyciu tej notacji w analizowanych publikacjach, wykonany został przegląd pod kątem poprawności użycia notacji BPMN. Okazało się, że we wszystkich 13 publikacjach, w których użyto notacji BPMN można było wyróżnić mniejsze lub większe błędy, które finalnie zostały skategoryzowane do pięciu głównych grup błędów zaprezentowanych w poniższej tabeli:

Tabela 3. Kategorie błędów w przypadku użycia notacji BPMN

Autorzy publikacji	Rok	Kategorie błędów				
		Baseny i tory	Zdarzenia	Bramki	Przeływ	Poziom algorytmu
Y. Rizk i in.	2020		X	X	X	
M. König, L. Bein	2020		X	X		
T. Chakraborti i in.	2020		X			
S. Aguirre, A. Rodriguez	2017	X	X		X	
C. Flechsig i in.	2019		X			X

Autorzy publikacji	Rok	Kategorie błędów				
		Baseny i tory	Zdarzenia	Bramki	Przepływ	Poziom algorytmu
B. Amini	2019	X		X		X
J. Holma	2020	X	X		X	
T. Qvick	2019	X	X	X		
J. Hindel i in.	2020	X	X			
M. Romao i in.	2019	X	X			
V. Gružasuskas, D. Ragavan	2020		X			
A. Kasic	2020		X	X		
F. da Silva Santos	2019	X				X

Źródło: opracowanie własne.

Na drugim miejscu znalazło się wystąpienie schematu blokowego (ang. *flow-chart*). Jest to bardzo prosta notacja, co może być uznane za duży atut w kontekście projektów RPA. Wynika to z faktu, że roboty programowe tworzone przy pomocy prostych narzędzi niskokodowych [Pasałowski, 2017] (ang. *low code*), przechodzą zazwyczaj dość prostą ścieżkę procesu (ang. *happy path*) z niewielką ilością odstępstw, a jeśli takowe wystąpią, to są one kierowane do ludzkiego operatora. W przypadku tych prostych algorytmów będących instrukcjami dla robota, zamodelowanie przebiegu takiego zadania przy pomocy schematu blokowego może okazać się bardzo skuteczne. Jednak ze względu na swoje ograniczenia, schemat blokowy może być mało przydatny przy modelowaniu poglądowego diagramu procesu biznesowego, tworzonego w celu ogólnego zrozumienia procesu, którego część zadań będzie poddana robotyzacji.

Trzecią pod względem popularności oficjalną notacją jest UML (ang. *Unified Modeling Language*), który wystąpił tylko w dwóch publikacjach. Warto jednak zauważyć, że UML nie został wykorzystany w celu zamodelowania procesu biznesowego ani pojedynczego algorytmu pracy robota. W obu przypadkach został użyty tylko jeden rodzaj diagramu UML, czyli diagram sekwencji i dwukrotnie został on użyty jako uzupełnienie do diagramu BPMN w celu podkreślenia kolejności przesyłanych komunikatów [Wrycza, Marcinkowski, Wyrzykowski, 2006] w ramach robotyzowanego procesu.

Ostatnią wyróżnioną metodą w analizie publikacji jest kategoria „Inna notacja”. Zawarto w niej wszystkie notacje, które nie zostały rozpoznane przez autorów jako konkretna notacja lub były one wymyślone przez autorów publikacji. Istotną część tych diagramów posiadała elementy zapożyczone z innych, oficjalnych notacji, głównie z EPC (ang. *Event-Driven Process Chain*), BPMN oraz UML. Diagramy

w tej kategorii często miały charakter poglądowy lub wręcz marketingowy i z powodzeniem mogłyby być prezentowane np. podczas procesu sprzedaży projektu RPA przez kadrę menedżerską wyższego szczebla.

5. Rezultaty badania ankietowego

W tej części artykułu przedstawione zostaną wyniki badania ankietowego przeprowadzonego wśród polskich specjalistów branży RPA. Łączna liczba wypełnionych ankiet uwzględnionych do badania wyniosła 32.

5.1. Doświadczenie zawodowe respondentów

W pierwszej części ankiety respondenci odpowiedzieli jaką rolę pełnią w projektach RPA oraz jak długo działają w branży. Zdecydowana większość (21 osób) respondentów to osoby, których główną rolą jest wytwarzanie rozwiązań informatycznych (RPA Developer – 15 oraz RPA Tech Lead – 6). Pozostali respondenci pełnili funkcje wspierające wytwarzanie robotów programowych (RPA Manager – 3, Konsultant RPA – 3, Inna – 3, Analityk RPA – 2).

Ponad połowa respondentów (18 osób) posiada nie więcej niż trzy lata doświadczenia, natomiast zaledwie dwie osoby mają ponad 5 lat doświadczenia w branży RPA. Tak jak zostało to opisane we wstępie, technologia RPA jest stosunkowo młoda i duży wzrost popularności osiągnęła dopiero w latach 2018–2019, co doskonale pokrywa się z doświadczeniem respondentów (ankieta została przeprowadzana w drugim kwartale 2021). Krótki staż zawodowy respondentów jest spodziewany także ze względu na wnioski płynące z przeglądu literatury, gdzie zdecydowana większość prac pochodziła z 2020 r.

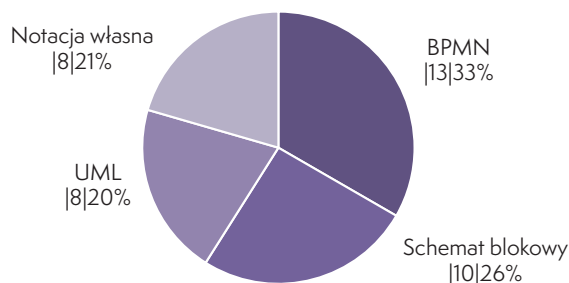
5.2. Dotychczasowe doświadczenia respondentów związane z modelowaniem procesów w projektach RPA

W drugiej części ankiety respondenci odpowiedzieli na serię pytań dotyczących ich dotychczasowych doświadczeń związanych z modelowaniem procesów w projektach RPA. Tylko jedna odpowiedź wskazała na kompletny brak styczności z diagramami procesów, natomiast pozostałych 31 respondentów spotyka się z diagramami w pracy z technologią RPA, z czego 23 respondentów dodatkowo uczestniczy w ich tworzeniu lub współtworzeniu. Powyższe wyniki potwierdzają, jak powszechnym zagadnieniem jest modelowanie procesów w projektach RPA.

Wśród 31 respondentów, którzy mają styczność z diagramami procesów, ponad połowa (19) określiła, że diagramy w projektach RPA, w których brali oni udział, zostały stworzone w konkretnej notacji lub notacjach.

Następnie respondenci, którzy na poprzednie pytanie odpowiedzieli twierdząco, otrzymali pytanie, gdzie mogli wybrać jedną lub wiele notacji, które były wykorzystywane do tworzenia diagramów na projektach RPA. Odpowiedzi na to pytanie zostały zaprezentowane na poniższym wykresie kołowym. Warto zwrócić uwagę na fakt, że w przypadku trzech oficjalnych notacji wybranych przez respondentów (BPMN, Schemat blokowy i UML), tendencja odnośnie do wykorzystywania tychże notacji pokrywa się z wynikami pochodzącymi z przeglądu literatury i zaprezentowanymi na rysunku 2. W obu przypadkach najbardziej popularna okazała się notacja BPMN, następnie schemat blokowy, a na trzecim miejscu notacja UML.

Rysunek 3. W jakich notacjach tworzone są diagramy w projektach RPA?



Źródło: opracowanie własne.

Na rysunkach 4 i 5 zaprezentowane zostały odpowiedzi respondentów na pytanie w jakim stopniu zgadzają się oni z poszczególnymi stwierdzeniami. Z wyników zaprezentowanych na rysunku 10 można wyciągnąć wniosek, że pomimo ogólnej powszechności diagramów, ich tworzeniu nie jest poświęcane wiele uwagi. Mimo przykładania niewielkiej uwagi do tworzenia diagramów, większość respondentów określiła, że jakość diagramów z którymi mieli do czynienia w ramach projektów RPA była raczej dobra. Na trzecie pytanie dotyczące wystąpienia negatywnego wpływu diagramów niskiej jakości na rezultaty projektów RPA, aż 24 respondentów odpowiedziało twierdząco. W kontekście powyższych odpowiedzi może to oznaczać, że stosunkowo mała liczba błędów (większość respondentów uważa, że diagramy w ich projektach były raczej dobrej jakości) przekłada się na negatywne rezultaty projektów. Do występowania tych błędów może przyczyniać się poświęcanie niewielkiej uwagi tworzeniu diagramów.

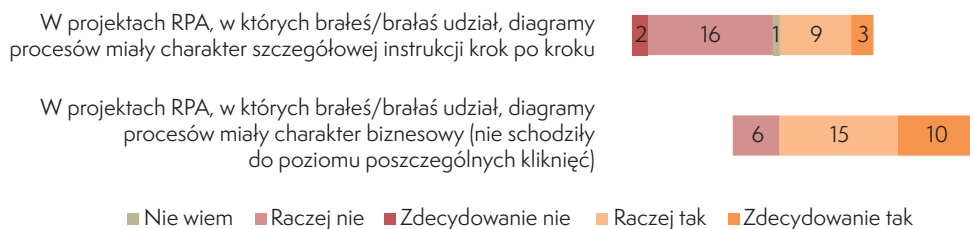
Rysunek 4. Odpowiedzi respondentów na pytanie czy zgadzają się z poniższymi stwierdzeniami



Źródło: opracowanie własne.

Pierwsze miejsce notacji BPMN pod względem popularności w projektach RPA zarówno na podstawie przeglądu literatury, jak i badania ankietowego, może być wyjaśnione odpowiedziami zaprezentowanymi na rysunku 5. Widać na nim, że diagramy procesów tworzone w ramach projektów RPA mają raczej charakter biznesowy (nie są one instrukcją krok po kroku dla robota), co uzasadnia wybór notacji BPMN do ich tworzenia. Instrukcja krok po kroku, na której podstawie tworzony jest algorytm robota programowego, w większości przypadków prawdopodobnie jest spisywana w innej formie (np. listy kroków w dokumencie tekstowym).

Rysunek 5. Odpowiedzi respondentów na pytanie czy zgadzają się z poniższymi stwierdzeniami



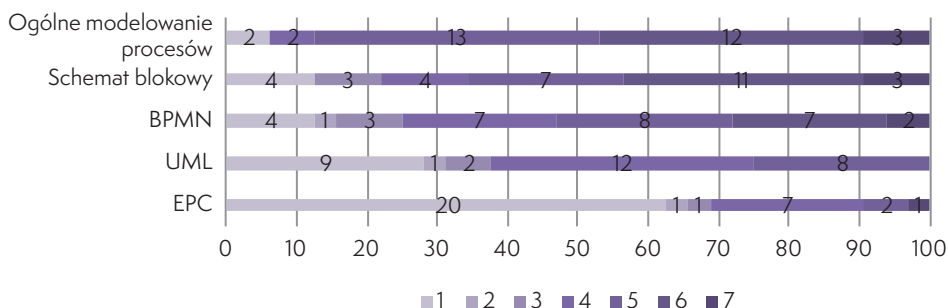
Źródło: opracowanie własne.

5.3. Poziom wiedzy respondentów odnośnie do modelowania procesów

W trzeciej części ankiety respondenci otrzymali serię pytań w celu określenia ich poziomu oraz źródła wiedzy w zakresie modelowania procesów. W pierwszej kolejności respondenci określili swój poziom umiejętności w skali od 1 do 7, gdzie 1 oznacza kompletny brak znajomości danej notacji, natomiast 7 oznacza bycie eks-

pertem w ramach danej notacji. Na rysunku 6 zaprezentowano odpowiedzi respondentów. Osoby ankietowane wysoko oceniły swoje umiejętności w ramach ogólnej dziedziny modelowania procesów (88% oceniło swoje umiejętności na 5 lub więcej), a także w ramach notacji schematu blokowego (67% na 5 lub więcej) i notacji BPMN (63% na 5 lub więcej). Lepsze opanowanie schematu blokowego od notacji BPMN i UML jest wynikiem spodziewanym, głównie ze względu na stopień skomplikowania wszystkich trzech notacji, gdzie schemat blokowy jest zdecydowanie notacją najprostszą i najbardziej intuicyjną. Notacja EPC została oceniona bardzo nisko, ponieważ aż 63% respondentów odpowiedziało, że zupełnie nie zna tej notacji. Jest to prawdopodobnie związane z faktem, że szczyt jej popularności przypadł dawno, przed wzrostem popularności technologii RPA, w wyniku czego ta notacja i technologia minęły się w czasie.

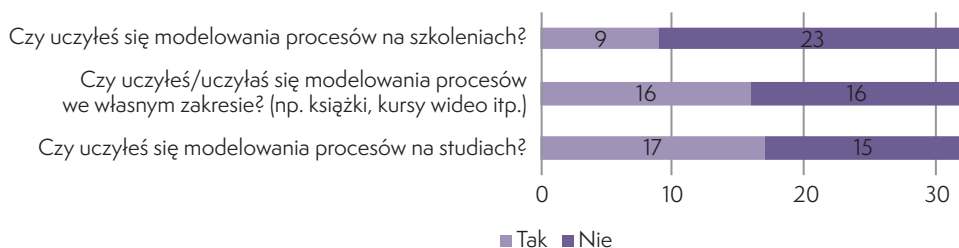
Rysunek 6. Poziom umiejętności w ramach poszczególnych notacji według respondentów, gdzie 1 oznacza „nie znam wcale”, a 7 oznacza „jestem ekspertem” (w %)



Źródło: opracowanie własne.

Następnie respondenci musieli odpowiedzieć, z jakich źródeł czerpali wiedzę na temat modelowania procesów. W ankiecie zadano pytanie o naukę w ramach szkoleń, studiów oraz we własnym zakresie. Z odpowiedzi zaprezentowanych na rysunku 7 wynika, że przeważającymi źródłami wiedzy odnośnie do modelowania procesów są studia i nauka we własnym zakresie (poprzez np. książki, kursy wideo itp.). Oznacza to, że pracodawcy mają duże pole do poprawy pod względem szkoleń firmowych z tematyki modelowania procesów, która mogłaby przyczynić się do ograniczenia negatywnych skutków w projektach RPA, wynikających z tworzenia diagramów niskiej jakości.

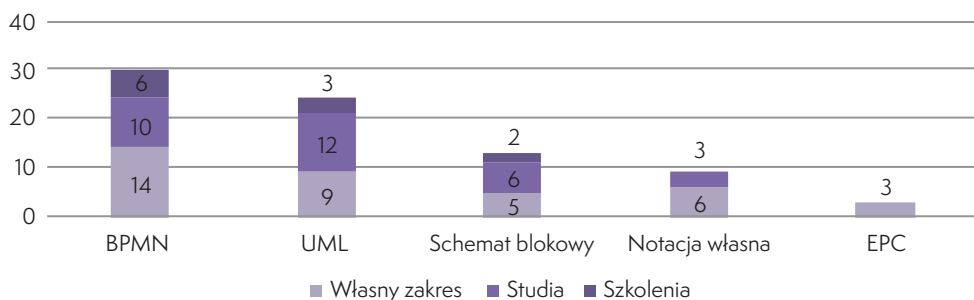
Rysunek 7. Źródła wiedzy respondentów w ramach modelowania procesów



Źródło: opracowanie własne.

Na rysunku 8 widać jakich notacji uczyli się respondenci z wyżej wymienionych źródeł. Na pierwszym miejscu ponownie uplasowała się notacja BPMN (94% ankietowanych uczyło się tej notacji), a zaraz za nią notacja UML (75% ankietowanych uczyło się tej notacji).

Rysunek 8. Źródła wiedzy respondentów w ramach poszczególnych notacji



Źródło: opracowanie własne.

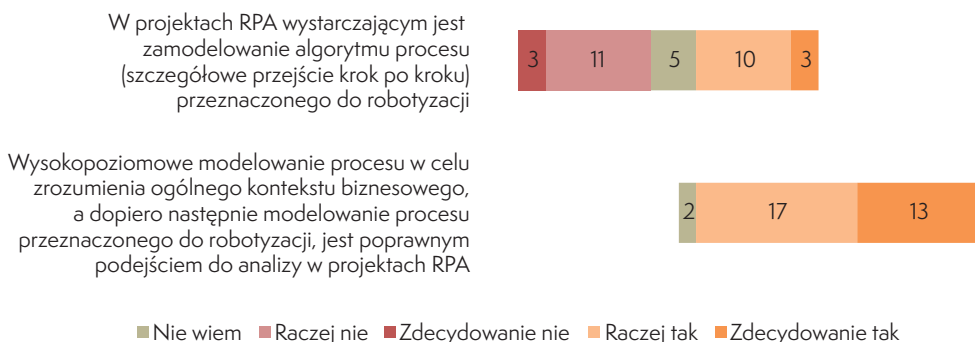
5.4. Preferencje specjalistów RPA odnośnie do modelowania procesów

Ostania część badania ankietowego została poświęcona zbadaniu preferencji respondentów odnośnie do modelowania procesów w projektach RPA.

Na rysunku 9 zaprezentowano odpowiedzi na pytania dotyczące preferencji w kwestii charakteru diagramów tworzonych na potrzeby projektów RPA. Respondenci w niewielkiej przewadze uznali, że zamodelowanie samego algorytmu robotyzowanego procesu nie jest wystarczającym efektem prac analitycznych prowadzonych w ramach projektu RPA. Dodatkowo wszyscy respondenci, którzy wyrazili swoją opinię, poparli słuszność praktyki polegającej na początkowym zamodelowaniu procesu biznesowego w formie wysokopoziomowej, a dopiero następnie szczegółowego zamodelowania fragmentu procesu przeznaczonego do robotyzacji. Odpowiedzi

na temat preferencji odnośnie do charakteru diagramów tworzonych na potrzeby projektów RPA pokrywają się z odpowiedziami na temat ich aktualnego charakteru zaprezentowanymi na rysunku 5.

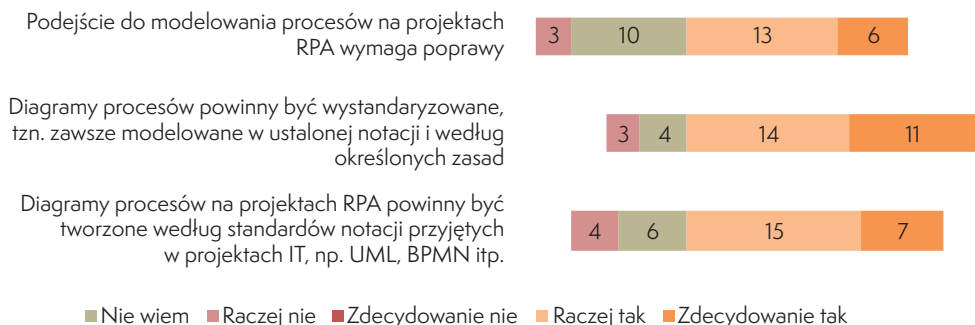
Rysunek 9. Odpowiedzi respondentów na pytanie czy zgadzają się z poniższymi stwierdzeniami



Źródło: opracowanie własne.

Na rysunku 10 widać, że 19 z 32 (59%) respondentów widzi potrzebę poprawy podejścia do modelowania procesów w ramach projektów RPA, a jedynie 3 (9%) respondentów ma przeciwne zdanie. Jedynym z możliwych rozwiązań na poprawę podejścia do modelowania procesów w projektach RPA jest wystandaryzowanie diagramów do ustalonej notacji (jednej lub wielu) i określenie jasnych zasad odnośnie do tworzenia diagramów. Za potrzebą standaryzacji opowiedziało się 25 (78%) respondentów, natomiast 22 respondentów poparło propozycję, aby w ramach standaryzacji diagramy procesów na projektach RPA były tworzone przy użyciu notacji ogólnie przyjętych w projektach branży IT jak np. BPMN czy UML.

Rysunek 10. Odpowiedzi respondentów na pytanie czy zgadzają się z poniższymi stwierdzeniami



Źródło: opracowanie własne.

Niewiele ponad połowa respondentów (53%) określiła, że ich zdaniem diagram procesów nie jest najważniejszym elementem analizy procesu w projektach RPA. Wynik ten pokrywa się niejako z wynikiem zaprezentowanym na rysunku 5, gdzie respondenci określili, że na projektach, w których brali udział, nie poświęcano zbyt wiele czasu na tworzenie diagramów procesu.

Rysunek 11. Istotność diagramów procesów według respondentów

Diagram procesu jest najważniejszym elementem analizy procesu w projektach RPA

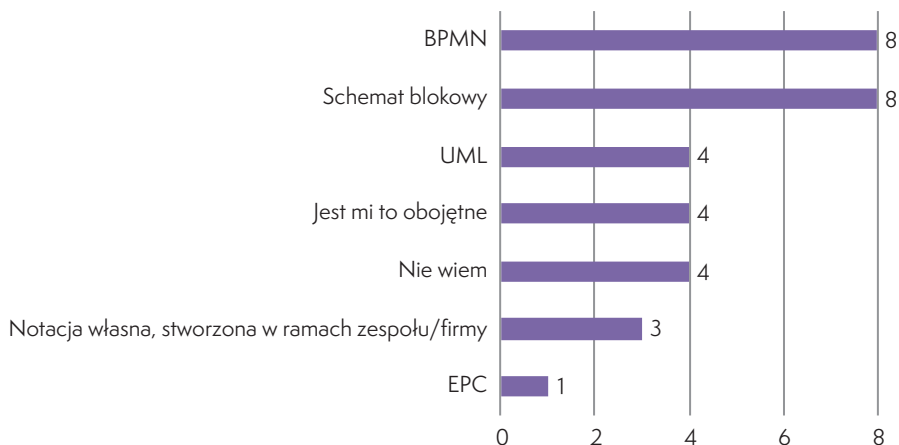


■ Nie wiem ■ Raczej nie ■ Zdecydowanie nie ■ Raczej tak ■ Zdecydowanie tak

Źródło: opracowanie własne.

Ostatnie pytanie dotyczyło najchętniej wybieranej notacji w ramach projektów RPA. Ponownie w czołówce uplasowały się notacje BPMN oraz schemat blokowy, które otrzymały po 8 odpowiedzi, natomiast na trzecim miejscu znalazła się notacja UML. Łącznie 8 respondentów stwierdziło, że notacja wykorzystywana na projektach RPA, w których uczestniczą, jest im obojętna lub nie wiedzą, którą notację preferują.

Rysunek 12. Notacje, z jakich respondenci najchętniej korzystają w projektach RPA



Źródło: opracowanie własne.

6. Wnioski i rekomendacje

Podsumowując rozważania na temat modelowania procesów w ramach projektów RPA, zdaniem autora, można dojść do następujących wniosków, które zostały krótko omówione w poniższych akapitach.

Diagramy procesów są nieodzownym elementem zdecydowanej większości projektów RPA. Wynika to ze specyfiki technologii RPA, która służy do zastępowania w wybranych fragmentach procesu biznesowego pracy ludzkiej pracą wykonywaną przez robota programowego. Fakt powszechności diagramów procesów w ramach projektów RPA potwierdziło badanie ankietowe.

Rozwiązania RPA są na tyle nowe, że jeszcze nie zdążyły się wykształcić powszechne praktyki modelowania procesów w projektach wykorzystujących tę technologię. Pomimo tego, na podstawie zarówno przeglądu literatury fachowej, jak i badania ankietowego, można wyróżnić pewne tendencje, jak np.:

- przewagę notacji BPMN i schematu blokowego (zarówno w badaniu literaturowym, jak i ankietowym);
- przewagę diagramów o charakterze biznesowym nad diagramami o charakterze algorytmicznym;
- uzupełnianie diagramów stworzonymi w jednej notacji innymi, np. diagramu wysokopoziomowego procesu biznesowego w BPMN diagramem na poziomie wykonywalnym w UML lub schemacie blokowym.

Niska jakość diagramów tworzonych na potrzeby projektów RPA oraz całościowe podejście do modelowania procesów w projektach RPA może mieć negatywny wpływ na rezultaty tychże projektów. Do typowych konsekwencji można zaliczyć np.: luki analityczne, opóźnienia w realizacji projektu, niespełnienie wymagań biznesowych.

Diagramy BPMN stworzone w ramach literatury fachowej poświęconej tematyce RPA posiadają liczne błędy. Możliwe przyczyny błędnego wykorzystywania notacji BPMN w ramach projektów RPA mogą być następujące:

- specjaliści RPA mylą proces biznesowy z procedurą (algorytmem), w efekcie czego próbują oni za pomocą notacji BPMN, która została stworzona na potrzeby modelowania procesów biznesowych, modelować wykonywalne instrukcje dla robota programowego;
- autorzy diagramów traktują roboty programowe jak oddzielnych uczestników procesów biznesowych;
- osoby odpowiedzialne za tworzenia diagramów w projektach RPA nie posiadają odpowiedniej wiedzy i przeszkolenia z notacji BPMN;
- ze względu na młody wiek technologii RPA, przyciąga ona specjalistów niedoświadczonych w modelowaniu procesów biznesowych w ramach projektów IT.

W celu poprawy jakości diagramów procesów tworzonych na potrzeby projektów RPA, autor rekomenduje podjęcie następujących kroków:

Jasne określenie standardów i zasad modelowania procesów na projektach RPA; na potrzeby analizy należy rozróżnić proces biznesowy, w ramach którego część zadań będzie poddana robotyzacji, od procedury (algorytmu krok po kroku), którą robot programowy będzie wykonywać w celu realizacji poszczególnych zadań.

Przy modelowaniu procesów biznesowych najlepszym wyborem będzie użycie notacji BPMN, która do tego celu została stworzona i jest najbardziej powszechną notacją aktualnie stosowaną w ramach projektów RPA.

Przy modelowaniu procedury dla robota programowego najlepszym wyborem będzie notacja schematu blokowego, która bardzo dobrze nadaje się do tego celu i jest drugą najbardziej powszechną notacją aktualnie stosowaną w ramach projektów RPA.

Szkolenia wewnątrzfirmowe z tematyki modelowania procesów biznesowych są obszarem, który wymaga większej uwagi firm prowadzących projekty RPA. Przeszkolenie pracowników z umiejętności modelowania procesów, ograniczonej do zakresu zagadnień związanych z RPA, nie powinno być procesem ani długim, ani kosztownym, a prawdopodobne jest, że przyczyniłoby się do znacznej poprawy jakości diagramów tworzonych na potrzeby projektów RPA, a co za tym idzie, ograniczenia powstawania negatywnych skutków diagramów niskiej jakości.

7. Podsumowanie

Modelowanie procesów jest bardzo istotnym zagadnieniem w kontekście projektów RPA ze względu na powszechność diagramów w ramach analizy na tychże projektach, jak i na możliwe konsekwencje biznesowe wykorzystania diagramów niskiej jakości. W ramach pracy ustalono aktualne tendencje dotyczące modelowania procesów w branży RPA. Na podstawie przeglądu literatury fachowej oraz badania ankietowego można wyłonić dwie najczęściej stosowane notacje w ramach projektów RPA, mianowicie BPMN oraz schemat blokowy. Obie te notacje odpowiadają na dwie zupełnie różne potrzeby odnośnie do modelowania procesów w projektach RPA. Notacja BPMN, zdaniem autora, powinna być stosowana do całościowego i wysokopoziomowego zamodelowania procesu biznesowego, którego to konkretne zadania będą dopiero robotyzowane. Natomiast procedury opisujące jak dane zadanie wykonać, powinny zostać zamodelowane za pomocą schematu blokowego.

W ramach przeglądu literatury autor zidentyfikował liczne błędy przy próbie wykorzystania notacji BPMN. We wnioskach wyszczególnione zostały te błędy,

których przyczyną może być charakterystyka projektów RPA oraz nawyki specjalistów w tej branży.

Na podstawie przeglądu literatury i badania ankietowego wyłoniona została także lista rekomendacji, które mogą przyczynić się do poprawy jakości diagramów tworzonych w ramach projektów RPA, a co za tym idzie, ograniczyć negatywne skutki błędnych diagramów.

Ze względu na nieliczną literaturę fachową dotyczącą poruszonego tematu badawczego oraz liczebność próby w ankiecie, wnioski i rekomendacje zawarte w niniejszej publikacji mają charakter pogładowy. W celu dokładniejszego zbadania praktyk modelowania procesów w ramach projektów RPA, należałoby rozważyć analizę diagramów powstałych na potrzeby realnych projektów komercyjnych oraz ocenę wpływu ewentualnych błędów na rezultaty tychże projektów. Dodatkowo można kontynuować pracę w temacie badań preferencji specjalistów RPA odnośnie do modelowania procesów, poprzez przeprowadzenie bardziej szczegółowej ankiety w języku angielskim, co z pewnością zwiększyłoby liczbę respondentów i dałoby możliwość wyciągnięcia statystycznie istotnych wniosków.

Bibliografia

- Aguirre, S., Rodriguez, A. (2017). Automation of a Business Process Using Robotic Process Automation (RPA): A Case Study. W: *Applied Computer Sciences in Engineering. 4th Workshop on Engineering Applications, WEA 2017, Cartagena, Colombia* (s. 65–71), J. Figueroa-García, E. López-Santana, J. Villa-Ramírez, R. Ferro-Escobar (Eds.). Springer.
- Amini, B. (2019). *Robotic Process Automation Implementation within an organization* (bachelor's thesis).
- Chakraborti, T., Isahagian, V., Khalaf, R.Y., Khazaeni, Y., Muthusamy, V., Rizk, Y., Unuvar, M. (2020). From Robotic Process Automation to Intelligent Process Automation – Emerging Trends. W: *Business Process Management: Blockchain and Robotic Process Automation Forum*, M. Müller, N. Ostern, M. Rosemann (Eds.). Springer.
- Costello, K., Rimol, R. (2020). *Gartner Says Worldwide Robotic Process Automation Software Revenue to Reach Nearly \$2 Billion in 2021*, <https://www.gartner.com/en/newsroom/press-releases/2020-09-21-gartner-says-worldwide-robotic-process-automation-software-revenue-to-reach-nearly-2-billion-in-2021> (dostęp: 12.05.2021).
- Enriquez, J.G., Jimenez-Ramirez, A., Dominguez-Mayo, F.J., Garcia-Garci, J.A. (2020). Robotic Process Automation: A Scientific and Industrial Systematic Mapping Study, *IEEE Access*, 8.
- Flechsigt, C., Lohmer, J., Lasch, R. (2019). Realizing the Full Potential of Robotic Process Automation Through a Combination with BPM. W: *Logistics Management* (s. 104–119), C. Bierwirth, T. Kirschstein, D. Sackmann (Eds.). Springer.

- Gružasuskas, V., Ragavan, D. (2020). Robotic Process Automation for document processing: A case study of a logistics service provider, *Journal of Management*, 36(2), s. 119–126.
- Guide on Terms and Concepts in Intelligent Automation (2017). *IEEE*.
- Hindel, J., Cabrera, L.M., Stierle, M. (2020). *Robotic Process Automation: Hype or Hope?*, https://library.gito.de/wp-content/uploads/2021/08/R6_Hindel-Robotic_Process_Automation-512_c.pdf (dostęp: 31.03.2021).
- Holma J. (2020). *SAP Automation Using Robot Process Automation* (bachelor's thesis).
- Jimenez-Ramirez, A., Reijers, H.A., Barba, I., Del Valle, C. (2019). A Method to Improve the Early Stages of the Robotic Process Automation Lifecycle. W: *Advanced Information Systems Engineering* (s. 446–461), P. Giorgini, B. Weber (Eds.). Springer Link.
- Kasic, A. (2020). *Robotic Process Automation of Tasks on the Example of Back Office Processes* (master's thesis).
- König, M., Bein, L. (2020). Integrating Robotic Process Automation into Business Process Management. W: *Business Process Management: Blockchain and Robotic Process Automation Forum* (s. 132–146), A. Asatiani i in. (Eds.). Springer.
- Lacity, M.C., Willcocks, L.P. (2018). *Robotic Process and Cognitive Automation. The Next Phase*. Warwickshire: SB Publishing.
- Malathi, T., Selvamuthukumar, D., Diwaan Chandar, C.S., Niranjana, V., Swashtika, A.K. (2020). *An Experimental Performance Analysis on Robotics Process Automation (RPA) With Open Source OCR Engines: Microsoft Ocr And Google Tesseract OCR*, OP Conference Series: Materials Science and Engineering.
- Myllymäki, J. (2019). *Developing Organization's Processes With Robotic Process Automation – A Case Study* (master's thesis).
- Pasławski, K. (2017). *Nadchodzi czas IT bez fachowców?*, <https://www.crn.pl/aktualnosci/nadchodzi-czas-it-bez-fachowcow> (dostęp: 2.06.2021).
- Qvick, T. (2019). *Understanding Business Process Complexity for Robotic Process Automation* (master's thesis).
- Rizk, Y., Isahagian, V., Boag, S., Khazaeni, Y., Unuvar, M., Muthusamy, V., Khalaf, R. (2020). *A Conversational Digital Assistant for Intelligent Process Automation* International Conference on Business Process Management 2020 RPA Forum.
- Romao, M., Costa, J., Costa, C.J. (2019). *Robotic Process Automation: A Case Study in the Banking Industry*, 14th Iberian Conference on Information Systems and Technologies (CISTI).
- Santos, F. da Silva (2019). *Using a Test Automation Tool for Robotic Process Automation: an Empirical Study* (master's thesis).
- Šimek, D., Šperka, R. (2019). How robot/human orchestration can help in an HR department: A case study from a pilot implementation, *Organizacija*, 52(3).
- Sobczak, A. (2019). Budowa modelu zarządzania robotyzacją procesów biznesowych, *Informatyka Ekonomiczna*, 2(52), s. 87–88. Wrocław: Wydawnictwo Uniwersytetu Ekonomicznego.
- Svensson J. (2020). *How Organizational Structures Affect the Implementation of Robotic Process* (master's thesis).
- Tran, D., Ho, T. (2018). *Workflow Methodology Development of RPA Solution for A Vietnamese Bank: A Case Study of Korkia Oy* (bachelor's thesis).

Wróblewska, A., Stanisławek, T., Prus-Zajączkowski, B., Garncarek, Ł. (2018). Robotic Process Automation of Unstructured Data with Machine Learning, *Annals of Computer Science and Information Systems*, 16.

Wrycza, S., Marcinkowski, B., Wyrzykowski, K. (2006). *Język UML 2.0 w modelowaniu systemów informatycznych*. Gliwice: Helion.

Process Modeling in the Context of RPA Projects

Abstract

The subject of the work is the exploration of current trends in the context of modeling processes in projects using the Robotic Process Automation (RPA) technology through the analysis of scientific publications on RPA. Additionally, a survey was conducted among Polish RPA specialists to learn their opinions and preferences with regard to process modeling. The subject of process modeling in RPA projects is important because process diagrams are an indispensable element of the analysis performed for the needs of process robotisation projects. Due to the widespread use of diagrams, they can have a significant impact on the results of RPA projects, which, in the absence of standardisation or low-quality of diagrams, may result in analytical gaps at the stage of analysis causing delays in robotic projects and providing incomplete functionality.

Keywords: Robotic Process Automation (RPA), process modeling, process diagrams, BPMN, UML, flowchart, business analysis, automation

Behawioralne przyczyny zgonów związanych z pandemią COVID-19. Przestrzenna analiza ekonometryczna śmiertelności w Unii Europejskiej w 2020 roku

Mateusz Szysz

Streszczenie

Dlaczego w niektórych regionach pandemia COVID-19 doprowadziła do wzrostu śmiertelności o kilkadziesiąt procent, podczas gdy w innych miejscach liczba zgonów pozostawała niemalże bez zmian? Celem pracy było zbadanie determinantów przyrostu liczby zgonów w 2020 r. w 198 regionach Unii Europejskiej za pomocą narzędzi przestrzennego modelowania ekonometrycznego. W szczególności sprawdzono znaczenie charakterystyk behawioralnych mieszkańców poszczególnych regionów, takich jak zaufanie do instytucji publicznych i autorytetów medycznych oraz zgodność z wartościami, jakimi kierują się decydenci w sprawie wprowadzania obostrzeń. Z przeprowadzonej analizy wynika, że zarówno większe zaufanie do działań rządu i zaleceń ekspertów z zakresu medycy, jak i przekonanie, że restrykcje w większym stopniu przynoszą pozytywne skutki zdrowotne niż straty dla gospodarki i ograniczenia wolności osobistej, prowadzi do mniejszego przyrostu zgonów związanych z chorobą COVID-19. Ponadto wykazano, że mniejsza wydolność służby zdrowia wyrażona wielkością wydatków na zdrowie i liczbą lekarzy, większa gęstość zaludnienia, wyższy poziom produktu krajowego w przeliczeniu na mieszkańca oraz większa liczba zgonów w blisko położonych regionach przyczyniały się do wzrostu śmiertelności w danym regionie.

Słowa kluczowe: pandemia COVID-19, śmiertelność, behawioralne przyczyny zgonów, przestrzenna analiza ekonometryczna

1. Wstęp

Choroba COVID-19 została po raz pierwszy rozpoznana i opisana w listopadzie 2019 r. w środkowych Chinach [Ma, 2020]. Następnie rozprzestrzeniła się po całym kraju, a w styczniu 2020 zanotowano pierwsze przypadki zachorowań w innych państwach [WHO, Novel Coronavirus (2019-nCoV), Situation Report, 2020].

11 marca 2020 r. przypadki zakażenia wirusem SARS-CoV-2 były już odnotowane w 114 krajach. Ponadto do tego dnia stwierdzono ponad 4000 zgonów spowodowanych infekcją wirusa, a przewidywania na następne tygodnie wskazywały na dalszy szybki wzrost liczby nowych zakażeń i zgonów. W związku z tym tego dnia Światowa Organizacja Zdrowia uznała COVID-19 za pandemię [WHO Director-General's opening remarks at the media briefing on COVID-19, 2020].

W następnych dniach liczba nowych przypadków i zgonów na COVID-19 bardzo szybko rosła, szczególnie w Europie. W związku z tym już 13 marca Dyrektor Generalny Światowej Organizacji Zdrowia Tedros Adhanom Ghebreyesus ogłosił, że Europa stała się epicentrum pandemii [Nebehay, 2020].

W sumie od marca do grudnia 2020 r. liczba dodatkowych zgonów w Unii Europejskiej w stosunku do średniej dla analogicznego okresu z lat 2016–2019 wyniosła ok. 580 tys. [Eurostat, 2021]. Jednak przyrost śmiertelności nie był równomierny, ponieważ w niektórych regionach liczba zgonów praktycznie nie zmieniła się, a w innych wzrosła o kilkadziesiąt procent.

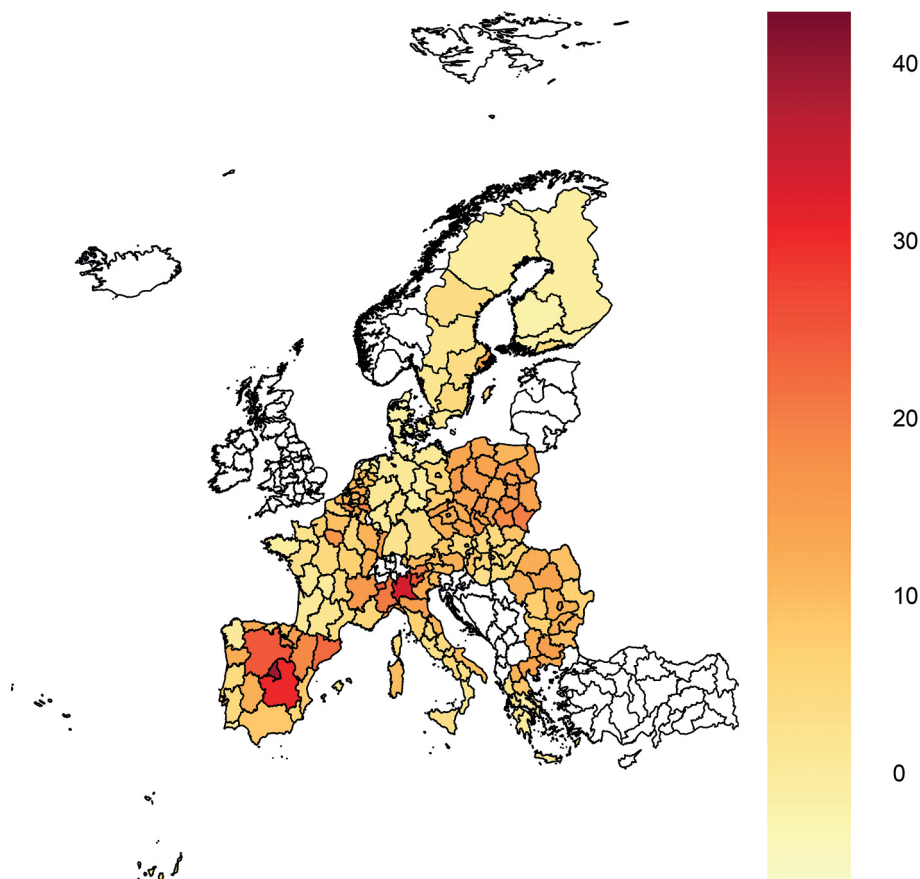
Zróznicowanie regionalne w przyroście liczby zgonów zostało przedstawione na rysunku 1. W miejscach takich jak Madryt czy Lombardia przyrost zgonów przekroczył 30%. Z kolei w południowo-zachodniej Francji czy w Finlandii nie przekroczył kilku procent. Co więcej, można zauważyć, że w regionach sąsiadujących z regionami o wyższym poziomie śmiertelności przyrost zgonów był większy, co sugeruje przestrzenny charakter zależności pomiędzy przyrostami zgonów w poszczególnych miejscach.

Ponadto obszar Unii Europejskiej jest bardzo zróżnicowany pod względem wielu charakterystyk, takich jak poziom rozwoju gospodarczego, struktura demograficzna czy też gęstość zaludnienia. Jednocześnie różne państwa stosowały odmienne środki, których intencją było ograniczenie rozwoju pandemii COVID-19.

Celem pracy było zbadanie za pomocą narzędzi ekonometrycznego modelowania przestrzennego zależności pomiędzy wspomnianymi charakterystykami różnych regionów i przyjętą polityką państwa a przyrostem liczby zgonów w czasie pandemii COVID-19. Dodatkowo wzięte zostały pod uwagę aspekty behawioralne, takie jak zaufanie do instytucji publicznych i autorytetów medycznych w poszczególnych państwach oraz zgodność z wartościami, jakimi kierują się decydenci przy wybie-

raniu obowiązujących restrykcji. Zgodnie z badaną hipotezą, większe zaufanie do instytucji publicznych i autorytetów medycznych prowadzi do tego, że mieszkańcy regionu objętego restrykcjami są w większym stopniu przekonani o słuszności ich wprowadzania, a przez to chętniej ich przestrzegają. Podobnie przekonanie większej części społeczeństwa o tym, że ograniczenia przynoszą więcej pozytywnych skutków dla zdrowia niż strat dla gospodarki i wolności osobistych, może przyczynić się do bardziej skrupulatnego przestrzegania obostrzeń, a przez to do wolniejszego rozprzestrzeniania się wirusa i mniejszego przyrostu liczby zgonów.

Rysunek 1. Procentowy przyrost zgonów w regionach NUTS 21 w pierwszych 52 tygodniach 2020 r. w stosunku do średniej z lat 2016–2019



Źródło: opracowanie własne na podstawie danych z Eurostatu [2021].

¹ Na mapie przedstawione są jedynie dane dla regionów, które zostały wzięte pod uwagę w dalszej analizie. Ponadto w przypadku Niemiec wzięto pod uwagę informacje dotyczące regionów NUTS 1 z uwagi na brak bardziej szczegółowych danych.

2. Determinanty śmiertelności związanej z pandemią COVID-19

Od wybuchu pandemii powstało wiele prac analizujących różne aspekty jej rozwoju. W szczególności przeprowadzonych zostało wiele badań dotyczących między państwowego i międzyregionalnego zróżnicowania liczby zakażeń i zgonów. Coccia [2021] porównał 161 krajów pod kątem śmiertelności związanej z COVID-19, wydatków na zdrowie, stopnia zanieczyszczenia środowiska oraz PKB *per capita*. Z przeprowadzonej analizy wynikało, że zarówno poziom zamożności, jak i wielkość wydatków na zdrowie były ujemnie skorelowane ze współczynnikiem umieralności. Z kolei w przypadku zanieczyszczenia powietrza relacja była odwrotna. W szczególności państwa, w których wielkość wydatków na zdrowie przekraczała 7,5% PKB oraz wynosiła ponad 2300 dolarów, charakteryzowały się niższym współczynnikiem umieralności od pozostałych. Podobnie w krajach, w których przeciętne narażenie na działanie pyłu zawieszonego PM_{2,5} nie przekraczało dopuszczalnych norm, śmiertelność była mniejsza.

Z drugiej jednak strony, z badań przeprowadzonych przez Chodicka i Weil [2020], opartych o dane z Izraela i 39 państw europejskich wynikało, że istnieje dodatnia korelacja pomiędzy śmiertelnością na COVID-19 a poziomem PKB *per capita*. Praca została wykonana jeszcze w 2020 r., czyli przed pojawieniem się efektywnych środków zapobiegających zakażeniom, takich jak szczepienia. Autorzy dodatnią zależność pomiędzy poziomem rozwoju gospodarczego kraju a śmiertelnością tłumaczą faktem, że bogatsze kraje charakteryzują się większą liczbą turystów i podróży biznesowych m.in. z Chin, które były epicentrum pandemii. Dlatego większa liczba kontaktów z zagranicą w takich krajach jak Wielka Brytania, Francja, Holandia czy Hiszpania prowadziła do większej liczby zakażeń. Z kolei to prowadziło do większej liczby zgonów zarówno bezpośrednio, czyli przez to, że większy odsetek osób narażonych na śmierć w związku z chorobą COVID-19 był zakażony, jak i pośrednio, czyli przez to, że wydolność służby zdrowia przekroczyła maksymalny limit, w związku z czym wiele osób nie otrzymało potrzebnej pomocy.

Podobne wnioski dotyczące zależności pomiędzy dostępnością do opieki zdrowotnej a śmiertelnością na COVID płyną z pracy Ji, Ma, Peppelenboscha i Pan [2020], którzy porównali zróżnicowanie śmiertelności w różnych regionach Chin. Zgodnie z przewidywaniami, w miejscach z większą dostępnością do opieki zdrowotnej liczba zgonów była mniejsza niż w pozostałych regionach o podobnej liczbie stwierdzonych przypadków zakażeń.

Innym czynnikiem, który może mieć wpływ na liczbę nowych przypadków COVID-19 i liczbę zgonów, jest gęstość zaludnienia. Przykładowo Bhadra, Mukher-

jee i Sarkar [2021] sprawdzili wpływ gęstości zaludnienia w danym regionie Indii na śmiertelność. Zgodnie z przedstawionymi wynikami badana zależność wystąpiła, ale była umiarkowana, co stoi w pewnej sprzeczności z badaniami przeprowadzonymi przez badaczy z Uniwersytetu Johns Hopkinsa [2020], którzy stwierdzili, że znaczenie ma nie sama gęstość zaludnienia, a wielkość obszaru metropolitalnego dotkniętego pandemią.

Nie mniej istotnym czynnikiem, który determinuje liczbę zakażeń i śmiertelność w danym regionie, jest liczba nowych zachorowań i zgonów w sąsiadujących regionach, co wynika z faktu, że ludzie mieszkający na stosunkowo blisko położonych względem siebie obszarach mają częściej styczność, przez co dochodzi do szybszej transmisji wirusa. Taki efekt został potwierdzony w badaniach Toktaşa [2020], który sprawdzał zależności przestrzenne pomiędzy poszczególnymi krajami Europy. Z tych badań wynika, że istnieje dodatnia autokorelacja przestrzenna na poziomie państw. W związku z tym można przypuszczać, że podobna analogiczna zależność istnieje pomiędzy mniejszymi regionami, takimi jak województwa w Polsce czy też landy w Niemczech.

Liczne badania wykazały, że ryzyko śmierci na COVID-19 rośnie wraz z wiekiem osoby zakażonej [Albitar, Ballouze, Ping Ooi, Maisharah Sheikh Ghadzi, 2020]. W związku z tym można się spodziewać, że im większy odsetek osób starszych w danej populacji, tym większa jest śmiertelność w związku z zakażeniami. Przykładowo Rinaldi i Paradisi w swoim badaniu [2020] przeprowadzonym na bazie danych z pierwszej fali zakażeń we Włoszech wykazali, że w przeciwieństwie do innych miejsc w społecznościach, w których odsetek osób powyżej 60. r.ż. przekroczył 30%, całkowity współczynnik śmiertelności nie mógł być mniejszy niż 0,5%.

Istotnym elementem polityki poszczególnych państw w 2020 r. było wprowadzanie restrykcji w celu ograniczania liczby zakażeń i śmiertelności, w szczególności *lockdownów*. Jednak wpływ ograniczeń na liczbę zgonów jest niejasny. Na przykład ogłoszona we Włoszech praca Borriego, Draga, Santantonio, Sobbria [2020] wskazała, że *lockdown* pozwolił na zmniejszenie liczby zgonów w kwietniu 2020 o ok. 5000, głównie wśród osób starszych. Z kolei Balmford, Annan i Hargreaves [2020] na podstawie porównania różnych krajów stwierdzili, że opóźnienia we wprowadzaniu restrykcji mogły przyczynić się do zwiększenia liczby zgonów. Inne wnioski płyną z pracy Bjørnskova [2021], który na podstawie porównania 24 państw Europy stwierdził, że nie było bezpośredniego związku pomiędzy polityką państwa w kwestii restrykcji a śmiertelnością. Podobnie Meo i in. [2020] na podstawie porównania losowo wybranych państw z całego świata stwierdzili, że *lockdown* miał wpływ na spadek liczby nowych przypadków zakażeń, ale spadek liczby zgonów nie był istotny statystycznie.

Znaczący wpływ na skuteczność wprowadzanych restrykcji ma zaufanie do instytucji, które je ustalają. Oksanen i in. w swojej pracy [2020] porównali 25 państw europejskich pod kątem zaufania do instytucji publicznych i nadmiernej śmiertelności na COVID-19. Dane, na podstawie których przeprowadzono analizę, pochodzą z Europejskiego Badania Społecznego [European Social Survey] sprzed pandemii. Pytano w nim mieszkańców różnych państw o zaufanie do parlamentu, polityków, partii politycznych itd. Ostatecznie zauważono ujemną zależność pomiędzy zagregowanym wskaźnikiem zaufania do instytucji a śmiertelnością na COVID-19 w czasie pierwszych tygodni pandemii. Autorzy argumentują, że ustawodawcy w społeczeństwach o wyższym stopniu zaufania instytucjonalnego mogli sobie pozwolić na wprowadzanie łagodniejszych obostrzeń niż w krajach takich jak Włochy czy Hiszpania, przy jednoczesnym uzyskaniu podobnych efektów. Niemal identyczne wnioski płyną z artykułu Charrona, Lapuente i Rodrigueza-Pose [2020], w którym przeanalizowano 153 regiony Europy pod kątem zaufania społecznego i polaryzacji pomiędzy zwolennikami i przeciwnikami rządu. Z badania wynika, że zarówno niskie zaufanie społeczne, jak i silna polaryzacja wpływają na zwiększoną liczbę zgonów związanych z COVID-19.

3. Metodyka badania międzyregionalnego zróżnicowania śmiertelności

W analizie wzięto pod uwagę 198 regionów NUTS 2², dla których dane o procentowym przyroście zgonów w pierwszych 52 tygodniach 2020 r. zostały zaprezentowane na rysunku 1. W każdym modelu zmienną objaśnianą był procentowy przyrost zgonów w danym regionie w stosunku do średniej liczby zgonów z lat 2016–2019. Warto zaznaczyć, że przyczyną nadmiarowej śmiertelności nie musiało być w każdym przypadku bezpośrednie zakażenie wirusem SARS-CoV-2, ponieważ mogła ona wynikać z ograniczenia dostępu do świadczeń medycznych dla pacjentów z innymi chorobami. Ponadto pomiędzy krajami występowały różnice w sposobie raportowania i strategii testowania pacjentów. Przykładowo w niektórych krajach testowano osoby zmarłe, a w innych nie, co mogło wpływać na różnice w statystykach [Rzeczpospolita, 2020]. Dlatego dla większej porównywalności pomiędzy regionami położonymi w różnych krajach korzystano z danych o przyroście zgonów ogółem. Zmienne objaśniające zostały zaprezentowane w tabeli 1.

² W przypadku Niemiec wzięte zostały pod uwagę dane dotyczące NUTS 1 z uwagi na brak bardziej szczegółowych informacji.

Dane dotyczące ich wszystkich poza zmiennymi o surowości restrykcji pochodzą z bazy danych Eurostatu.

Tabela 1. Opis zmiennych objaśniających

Zmienna objaśniająca	Obszar, którego dotyczy
Liczba lekarzy w przeliczeniu na 100 tys. mieszkańców w danym regionie	Jakość i wydolność służby zdrowia
Wydatki na służbę zdrowia w przeliczeniu na mieszkańca w parytecie siły nabywczej w danym kraju	
Liczba łóżek szpitalnych w danym regionie w przeliczeniu na 100 tys. osób	
Gęstość zaludnienia w danym regionie	Prawdopodobieństwo częstszych kontaktów, szybszej transmisji wirusa i większej liczby zakażeń
PKB <i>per capita</i> w parytecie siły nabywczej w danym regionie	Zamożność i częstość interakcji z innymi regionami
Wskaźnik zagrożenia ubóstwem	Poziom biedy
Mediana wieku	Struktura demograficzna
Odsetek ludności w wieku 60/65/70/75/80/85+	
Średni poziom surowości restrykcji	Polityka państwa w obszarze walki z pandemią
Zmienność poziomu surowości restrykcji	
Zaufanie do autorytetów i instytucji publicznych	Efekty behawioralne
Zgodność z wartościami, jakimi kierują się decydenci w sprawach restrykcji	

Źródło: opracowanie własne.

Wszystkie zmienne objaśniające zostały pogrupowane zgodnie z obszarami, których dotyczą. Zarówno liczba lekarzy w danym regionie, liczba łóżek szpitalnych w odniesieniu do liczby mieszkańców, jak i wydatki na zdrowie obrazują możliwości służby zdrowia do skutecznego leczenia pacjentów dotkniętych wirusem oraz pozostałych. Ostatnia z tych zmiennych, czyli liczba łóżek, nie była istotna w żadnym z modeli. W związku z tym nie jest ona analizowana w dalszej części pracy.

Kolejne zmienne to gęstość zaludnienia w danym regionie, która obrazuje prawdopodobieństwo częstszych interakcji pomiędzy mieszkańcami danego obszaru i PKB *per capita* w parytecie siły nabywczej, agregujący podatność mieszkańców danego regionu na częstszą interakcję z osobami z innych miejsc.

Następne trzy grupy zmiennych, dotyczące poziomu ubóstwa w danym regionie, odsetka osób starszych, które były bardziej narażone na zgon w związku z zakażeniem i przeciętnego stopnia surowości restrykcji jak i ich zmienności nie okazały się istotne w żadnym z modeli. Jako przeciętny stopień surowości restrykcji wzięta

została pod uwagę średnia wartość wskaźnika „COVID-19 Government Response Stringency Index”, opublikowanego przez Our World In Data [Our World in Data, Covid Stringency Index]. Obejmował on odpowiedź państwa na rozwój pandemii poprzez zamykanie szkół i biur, odwoływanie wydarzeń masowych, ograniczenia w podróżowaniu, prowadzenie kampanii informacyjnych, politykę dotyczącą testowania itd. Z kolei jako zmienność restrykcji wzięta została wariancja opisanego wskaźnika od początku pandemii do końca badanego okresu.

Zmienne z ostatniej grupy zostały utworzone na podstawie badania przeprowadzonego przez Kantar pod koniec kwietnia 2020 r. dla Parlamentu Europejskiego [Parlament Europejski, Eurobarometr, Public opinion in the EU in time of coronavirus crisis, 2020]. Polegało ono na tym, że mieszkańcy różnych krajów Unii Europejskiej odpowiadali na pytania dotyczące ich opinii na temat działań podejmowanych przez Unię Europejską oraz państwa członkowskie w kontekście walki z pandemią oraz zaufania do instytucji publicznych. Część z tych pytań, która opisana jest w tabeli 2, została wykorzystana do przeprowadzenia analizy czynnikowej.

Tabela 2. Opis zmiennych wykorzystanych do konstrukcji zmiennych behawioralnych

Zmienna	Pytanie	Sposób agregacji	Obszar, którego dotyczy
Przekonanie o przewyższeniu strat gospodarczych korzyściami zdrowotnymi	Do którego stwierdzenia dotyczącego konsekwencji ograniczeń podjętych przez władzę w celu ograniczenia rozwoju pandemii jest Panu/Pani bliżej? 1) Korzyści zdrowotne są większe niż szkody ekonomiczne. 6) Szkody ekonomiczne są większe niż korzyści zdrowotne.	odsetek osób, które na 6-stopniowej skali Likerta wybrały opcję 1 lub 2	przekonanie społeczeństwa o tym, że restrykcje przyniosą więcej korzyści niż strat gospodarczych
Przyzwolenie na ograniczenia własnej wolności	Do którego z poniższych stwierdzeń jest Panu/Pani bliżej? 1) Walka z koronawirusem w pełni uzasadnia ostatnie ograniczenia moich wolności osobistych. 6) Jestem przeciwny/a ograniczaniu mojej wolności niezależnie od występowania koronawirusa.	odsetek osób, które na 6-stopniowej skali Likerta wybrały opcję 1 lub 2	przekonanie ludności o tym, że skuteczna walka z pandemią jest ważniejsza niż problem ograniczenia własnej wolności
Zaufanie do autorytetów z zakresu zdrowia	Z listy wybierz osoby i instytucje, którym ufasz najbardziej w kwestii pozyskiwania informacji o pandemii (maksymalnie 3).	odsetek osób, które wskazały krajowe autorytety medyczne	zaufanie do zaleceń lekarzy i specjalistów z zakresu medycyny
Zaufanie do rządu		odsetek osób, które wskazały rząd	zaufanie do informacji przekazywanych przez organy państwowe i przedstawicieli władzy

Zmienna	Pytanie	Sposób agregacji	Obszar, którego dotyczy
Poparcie dla rządu	Ogólnie rzecz biorąc, w jakim stopniu jest Pan/Pani zwolennikiem/zwolenniczką bądź przeciwnikiem/przeciwniczką obecnych władz państwowych?	odsetek osób, które wybrały „w pełni popieram” albo „raczej popieram”	ogólne poparcie dla rządzących
Satysfakcja z działań podjętych przez rząd	Jak bardzo jest Pan/Pani zadowolony/a z działań podjętych dotychczas przez rząd przeciwko rozprzestrzenianiu się pandemii koronawirusa?	odsetek osób, które wybrały „bardzo zadowolony/a” albo „raczej zadowolony/a”	poparcie dla działań rządzących podejmowanych w celu walki z pandemią

Źródło: opracowanie własne na podstawie badania przeprowadzonego przez Kantar.

Tabela 3. Macierz rotowanych składowych

Zmienna	Składowa 1	Składowa 2
Przekonanie o przewyższaniu strat gospodarczych korzyściami zdrowotnymi		0,95
Przyzwolenie na ograniczenia własnej wolności		0,807
Zaufanie do autorytetów z zakresu zdrowia	0,82	
Zaufanie do rządu	0,911	
Poparcie dla rządu	0,91	
Satysfakcja z działań podjętych przez rząd	0,91	

Źródło: opracowanie własne.

Analiza czynnikowa doprowadziła do redukcji liczby wymiarów do dwóch. Skupiły one aż ok. 84% całkowitej zmienności w próbie. Podsumowanie uzyskanych składowych przedstawione jest w tabeli 3, gdzie zawarte zostały wartości współczynnika korelacji pomiędzy odpowiedziami na poszczególne pytania a składowymi. W tabeli pozostawione zostały jedynie wartości większe niż 0,5.

W ten sposób stworzone zostały dwie ostatnie zmienne przedstawione w tabeli 1, czyli zaufanie do autorytetów i instytucji publicznych (składowa 1) oraz zgodność z wartościami, jakimi kierują się decydenci (składowa 2).

Opisane zmienne zostały użyte w analizie za pomocą narzędzi przestrzennego modelowania ekonometrycznego. Ogólna postać rozważanych modeli przedstawiona jest za pomocą poniższego równania:

$$y = \rho W y + X \beta + \lambda W \epsilon + u,$$

gdzie y to wektor zmiennej objaśnianej, W to macierz wag przestrzennych, X to macierz zmiennych objaśniających, ρ i λ to estymowane współczynniki przestrzenne, a β to wektor estymowanych współczynników regresji.

Po nałożeniu restrykcji $\lambda = 0$ otrzymywany jest model SAR, a w przypadku $\rho = 0$ model SEM. Z kolei dla powyższego modelu bez żadnych restrykcji otrzymywany jest model SARAR, a przy nałożeniu obydwu restrykcji jednocześnie model liniowy bez zależności przestrzennych.

Macierz wag przestrzennych została stworzona w oparciu o odwrotności kwadratów odległości pomiędzy środkami poszczególnych regionów. W ten sposób regiony położone bliżej siebie oddziałują wzajemnie silniej niż regiony, które są od siebie oddalone.

4. Wyniki

Analiza reszt z modelu liniowego za pomocą testu Morana I wyraźnie wskazała na występowanie autokorelacji przestrzennej (por. z tabelą 6). W związku z tym wskazane jest włączenie do modelu komponentu przestrzennego. Finalnym modelem jest model SARAR, którego podsumowanie znajduje się w tabeli 4.

Tabela 4. Oszacowanie parametrów modelu SARAR

Zmienna	Oszacowanie współczynnika
Wyraz wolny	7,0828**
Liczba lekarzy w przeliczeniu na 100 tys. mieszkańców w danym regionie	-0,00796**
Wydatki na służbę zdrowia w przeliczeniu na mieszkańca w parytecie siły nabywczej w danym kraju	-0,00183***
Gęstość zaludnienia w danym regionie	0,000964**
PKB <i>per capita</i> w parytecie siły nabywczej w danym regionie	0,0001218**
Zaufanie do autorytetów i instytucji publicznych	-1,203176**
Zgodność z wartościami, jakimi kierują się decydenci w sprawach restrykcji	-1,005545*
ρ	0,59641**
λ	0,60298**

Uwaga: *, **, *** wskazują istotność na poziomie 10%, 5% i 1%.

Źródło: opracowanie własne.

Oszacowania współczynników przy zmiennych dotyczących stanu służby zdrowia, takie jak liczba lekarzy w przeliczeniu na liczbę osób i wydatki na zdrowie, okazały się istotnie mniejsze od zera, co jest zgodne z przewidywaniami, że lepiej rozwinięta służba zdrowia pomaga ograniczyć śmiertelność.

Kolejne zmienne, czyli gęstość zaludnienia i PKB liczone parytetem siły nabywczej *per capita*, również są istotne, ale z dodatnimi znakami. Oznacza to, że przy innych czynnikach niezmiennych większej gęstości zaludnienia towarzyszyła większa śmiertelność w czasie pandemii. Z kolei dodatnia zależność między poziomem zamożności a wielkością przyrostu liczby zgonów związanych z COVID-19 może wynikać z przeciętnie większej liczby kontaktów mieszkańców bogatszych regionów z osobami z zewnątrz.

Kluczowe z punktu widzenia analizy zmienne behawioralne, takie jak zaufanie do autorytetów i instytucji publicznych oraz zgodność z wartościami, jakimi kierują się decydenci w sprawach restrykcji również są istotne. W pierwszym przypadku jest to istotność na poziomie 5% (p-value równe 0,03), a w drugim przypadku na poziomie 10% (p-value równe 0,07). Oznacza to, że ważnym czynnikiem ograniczania śmiertelności w czasie poza samym nakładaniem restrykcji jest to, żeby objęte nimi społeczeństwo miało przekonanie co do ich pozytywnych skutków. Po pierwsze, mieszkańcy regionu objętego obostrzeniami muszą mieć przekonanie, że rzeczywiste cele ograniczeń proponowanych przez instytucje publiczne i autorytety medyczne są zgodne z deklarowanymi celami, którymi są ograniczanie transmisji wirusa i zmniejszanie śmiertelności. Po drugie, samo zaufanie do autorytetów może być niewystarczające w ograniczaniu liczby zgonów, gdy wartości i cele wyznawane przez ludzi są inne niż cele, jakimi kierują się rządzący. Taka sytuacja może mieć miejsce, gdy znaczna część społeczeństwa uważa, że przestrzeganie obostrzeń może doprowadzić do większych strat gospodarczych niż ewentualnych korzyści związanych ze zmniejszeniem liczby kontaktów społecznych. Podobnie część ludności jest skłonna wierzyć ekspertom, że restrykcje mogą zmniejszyć liczbę zachorowań, ale mimo wszystko może nie być to dla nich wystarczający powód, aby pozwolić na ograniczenie swojej wolności.

Ostatnie dwa współczynniki, czyli ρ i λ dotyczą komponentów przestrzennych modelu. Każdy z nich jest istotnie większy od zera. Wynika z tego, że wybrana postać funkcyjna pozwala wyjaśnić charakter zależności przestrzennych dla analizowanego zjawiska. Dodatnie znaki przy oszacowaniach parametrów można interpretować w taki sposób, że większemu przyrostowi zgonów w danym regionie powinien towarzyszyć wzrost zgonów w sąsiednich regionach.

Analizowane były również alternatywne postacie modelu, których podsumowanie jest pokazane w tabeli 5.

Część zmiennych w przypadku niektórych modeli nie jest istotna statystycznie. Warto jednak zauważyć, że wszystkie współczynniki odpowiadające komponentom przestrzennym, o ile występowały w danym modelu, były istotne na poziomie 1%,

co sugeruje, że zależność przestrzenna dla analizy zgonów związanych z chorobą COVID-19 jest bardzo ważnym elementem analizy zjawiska.

Tabela 5. Oszacowania współczynników dla alternatywnych modeli

	Model liniowy	SAR	SEM
Wyraz wolny	9,572***	3,1186	12,684***
Liczba lekarzy w przeliczeniu na 100 tys. mieszkańców w danym regionie	-0,007907*	-0,0065597*	-0,0072771*
Wydatki na służbę zdrowia w przeliczeniu na mieszkańca w parytecie siły nabywczej w danym kraju	-0,001926***	-0,0015413***	0,0023499***
Gęstość zaludnienia w danym regionie	0,0001353	0,00062372	0,00083142*
PKB <i>per capita</i> w parytecie siły nabywczej w danym regionie	0,0002503***	0,00016516***	0,00010899**
Zaufanie do autorytetów i instytucji publicznych	-2,485***	-1,2406***	-1,4427**
Zgodność z wartościami, jakimi kierują się decydenci w sprawach restrykcji	-0,965*	-0,67662	-1,3026**
ρ	-	0,79949***	-
λ	-	-	0,83053***

Uwaga: *, **, *** wskazują istotność na poziomie 10%, 5% i 1%.

Źródło: opracowanie własne.

Ponadto należy zwrócić uwagę na fakt, że pierwsza zmienna behawioralna, czyli zaufanie do autorytetów i instytucji publicznych, była istotna na poziomie 1% w modelu liniowym i modelu SAR. W przypadku modelu SEM była ona istotna na poziomie 5%. Z kolei druga zmienna behawioralna, czyli zgodność z wartościami, jakimi kierują się decydenci w sprawach restrykcji, była istotna na poziomie co najmniej 10% w przypadku dwóch z trzech modeli. Jednak warto odnotować, że w modelu SAR wartość empirycznego poziomu istotności wynosiła 0,106, co oznacza, że również była blisko poziomu, dzięki któremu można ją uznać za istotną. Podsumowując, można stwierdzić, że obydwie komponenty behawioralne są istotnymi zmiennymi objaśniającymi niezależnie od przyjętej postaci modelu.

Tabela 6. Empiryczny poziom istotności w teście Morana I dla reszt z różnych modeli

Model	liniowy	SAR	SEM	SARAR
p-value	$1,06 \cdot 10^{-12}$	0,001125	0,01248	0,2129

Źródło: opracowanie własne.

Każdy z omawianych modeli przestrzennych znacznie ogranicza występowanie autokorelacji przestrzennej reszt, co zostało pokazane w tabeli 6. Bliska zeru wartość empirycznego poziomu istotności dla modelu liniowego wskazuje na bardzo silną autokorelację przestrzenną, która znacząco jest ograniczana przez modele SAR i SEM. Jednak żaden z tych dwóch modeli nie niweluje jej w całości. Dopiero w przypadku modelu SARAR wartość empirycznego poziomu istotności jest wyraźnie większa niż np. 0,1.

Tabela 7. Wartość kryterium Akaike dla różnych modeli

Model	liniowy	SAR	SEM	SARAR
AIC	1275,5	1230,1	1230,5	1220,4

Źródło: opracowanie własne.

Tabela 8. Porównanie modeli za pomocą testu mnożników Lagrange'a

Test	Wartość p-value
LM: err	$3,512 \cdot 10^{-11}$
LM: lag	$6,271 \cdot 10^{-11}$
RLM: err	0,06487
RLM: lag	0,1316
SARMA	$9,556 \cdot 10^{-11}$

Źródło: opracowanie własne.

W tabelach 7 i 8 przedstawione są dalsze porównania wszystkich opisanych modeli. Wartość kryterium Akaike jest zdecydowanie największa dla modelu liniowego. W przypadku pozostałych modeli wartości są bardziej porównywalne, ale mimo wszystko zgodnie z tą miarą model SARAR powinien zostać uznany za najbardziej preferowany. Z kolei porównania różnych postaci zależności przestrzennej za pomocą mnożników Lagrange'a już nie dają bezpośrednich podstaw do stwierdzenia, że któryś z modeli jest wyraźnie preferowany, co wynika z faktu, że każdy z nich wyraźnie lepiej opisuje analizowane zjawisko niż model liniowy.

5. Podsumowanie

Przeprowadzona analiza pokazała, że efekty behawioralne były istotnym czynnikiem objaśniającym zmienność w przyrostach zgonów pomiędzy regionami NUTS 2 w 2020 r. w czasie pandemii koronawirusa. Zarówno większe zaufanie społeczeństwa

do rządzących i autorytetów medycznych, jak i wyższy stopień zgodności z wartościami, jakimi kierują się decydenci, prowadziły do mniejszej liczby zgonów związanych z chorobą COVID-19. Przedstawione wnioski mają znaczenie zarówno z punktu widzenia krótko- jak i długookresowej polityki państwa.

Rządzący, decydując się na wprowadzenie restrykcji, muszą brać pod uwagę fakt, że część społeczeństwa może im nie ufać i być mniej skłonna do przestrzegania obostrzeń. Jednak nawet jeżeli znaczna część społeczeństwa nie ma zastrzeżeń co do wiarygodności autorytetów medycznych i decydentów, to i tak może inaczej widzieć bilans rachunku korzyści i strat związanych z restrykcjami i celowo je łamać. W takich przypadkach państwo, chcąc osiągnąć zmniejszenie liczby interakcji między ludźmi, musi bardziej zdecydowanie penalizować brak przestrzegania obostrzeń albo nakładać jeszcze większe restrykcje i zmniejszać liczbę możliwych kontaktów społecznych w innych aspektach życia.

Warto zauważyć, że jednym z elementów, na podstawie których stworzono zmienną dotyczącą zaufania do instytucji publicznych i autorytetów, było ogólne zaufanie do rządu. Patrząc z perspektywy długookresowej, władze państwa, chcąc poprawić efektywność swojej polityki i ograniczyć negatywne skutki ewentualnych przyszłych pandemii, powinny dbać o stopniowe budowanie zaufania do instytucji publicznych.

Istotność oszacowań komponentów przestrzennych wszystkich modeli, wyniki testu Morana I, wartości kryteriów informacyjnych czy też wyniki testów mnożnika Lagrange'a wyraźnie wskazują, że model przestrzenny z macierzą wag opartą o kryterium odległości geograficznej dużo lepiej wyjaśnia analizowane zjawisko niż model liniowy. W związku z tym można stwierdzić, że śmiertelność w regionach znajdujących się w pobliżu danego miejsca ma kluczowe znaczenie w kontekście analizowania w nim tego samego zjawiska. Z punktu widzenia polityki państwa jest to o tyle istotne, że różne działania w kierunku ograniczania rozwoju pandemii powinny być podejmowane w oparciu o dane z sąsiednich regionów.

Przedstawione modele pokazują, że istotnym elementem ograniczania śmiertelności związanej z pandemią jest także jakość służby zdrowia. Chcąc zwiększyć jej możliwości, państwo musi podnosić wydatki na zdrowie i starać się o to, aby stopniowo podnosić liczbę medyków.

Z kolei dodatni wpływ gęstości zaludnienia na przyrost liczby zgonów może sugerować, że w regionach o większym zagęszczeniu ludności potrzebne są surowsze ograniczenia prowadzące do większego ograniczenia bezpośrednich interakcji między ludźmi. Podobnie jest z bogatszymi regionami, których mieszkańcy są bardziej narażeni na kontakty z osobami zakażonymi z innych państw. W tym przypadku

również mogą być potrzebne surowsze restrykcje, wymierzone przede wszystkim w możliwość przemieszczania się pomiędzy regionami i państwami.

Składam podziękowania Panu dr. hab. Andrzejowi Torójowi, prof. SGH za cenne uwagi wniesione do niniejszej pracy.

Bibliografia

- Albitar, O., Ballouze, R., Ping Ooi, J., Maisharah Sheikh Ghadzi, S. (2020). Risk factors for mortality among COVID-19 patients, *Diabetes Research and Clinical Practice*, 166.
- Balmford, B., Annan, J.D., Hargreaves, J.C. i in. (2020). Cross-country comparisons of Covid-19: Policy, politics and the price of life, *Environmental and Resource Economics*, 76, s. 525–551.
- Bhadra, A., Mukherjee, A., Sarkar, K. (2021). Impact of population density on Covid-19 infected and mortality rate in India, *Modeling Earth Systems and Environment*, 7, s. 623–629.
- Bjørnskov, C. (2021). Did lockdown work? An economist's cross-country comparison, *CESifo Economic Studies*, 67(3), s. 318–331.
- Borri, N., Drago, F., Santantonio, C., Sobbrino, F. (2020). *The 'Great Lockdown': Inactive workers and mortality by COVID-19*, CESifo Working Papers, 8584, www.cesifo.org/en/publikationen/2020/working-paper/great-lockdown-inactive-workers-and-mortality-covid-19 (dostęp: 1.07.2021).
- Charron, N., Lapuente, V., Rodriguez-Pose, A. (2020). *Uncooperative Society, Uncooperative Politics or Both? How Trust, Polarization and Populism Explain Excess Mortality for COVID-19 across European regions*. University of Gothenburg.
- Chodick, G., Weil, C. (2020). *COVID-19 Death toll: the role of the nation's economic development*, "medRxiv.org", www.medrxiv.org/content/10.1101/2020.07.18.20156778v1 (dostęp: 1.07.2021).
- Coccia, M. (2021). High health expenditures and low exposure of population to air pollution as critical factors that can reduce fatality rate in COVID-19 pandemic crisis: a global analysis, *Environmental Research*, 199.
- European Social Survey, www.europeansocialsurvey.org/
- Eurostat (2021). *580 000 excess deaths between March and December 2020*, www.ec.europa.eu/eurostat/web/products-eurostat-news/-/ddn-20210312-2 (dostęp: 1.07.2021).
- Eurostat, Baza danych, <https://ec.europa.eu/eurostat/data/database>
- Eurostat, *Deaths by week, sex, 5-year age group and NUTS 3 region*, [//appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_mweek3](http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_mweek3) (dostęp: 1.07.2021).
- Ji, Y., Ma, Z., Peppelenbosch, M.P., Pan, Q. (2020). Potential association between COVID-19 mortality and health-care resource availability, *The Lancet*, 8(4).
- Johns Hopkins Bloomberg School of Public Health (2020). *Urban density not linked to higher coronavirus infection rates – and is linked to lower COVID-19 death rates*, www.jhsph.edu/news/newsreleases/2020/urban-density-not-linked-to-higher-coronavirus-infection-rates-and-is-linked-to-lower-covid-19-death-rates.html (dostęp: 1.07.2021).

- Ma, J. (2020). *Coronavirus: China's first confirmed Covid-19 case traced back to November 17*, South China Morning Post, www.scmp.com/news/china/society/article/3074991/coronavirus-chinas-first-confirmed-covid-19-case-traced-back (dostęp: 1.07.2021).
- Meo, S.A., Abukhalaf, A.A., Alomar, A.A. i in. (2020). Impact of lockdown on COVID-19 prevalence and mortality during 2020 pandemic: observational analysis of 27 countries, *European Journal of Medical Research*, 25(56).
- Nebehay, S. (2020). *Europe is epicenter of coronavirus pandemic*: WHO, Reuters, www.reuters.com/article/us-health-coronavirus-who-idUSKBN2102Q0 (dostęp: 1.07.2021).
- Oksanen, A., Kaakinen, M., Latikka, R., Savolainen, I., Savela, N., Koivula, A. (2020). Regulation and Trust: 3-Month Follow-up, Study on COVID-19 Mortality in 25 European Countries, *JMIR Public Health Surveill* 6(2).
- Our World In Data, Covid Stringency Index, ourworldindata.org/grapher/covid-stringency-index.
- Parlament Europejski, Eurobarometr (2020). *Public opinion in the EU in time of coronavirus crisis*, www.europarl.europa.eu/at-your-service/en/be-heard/eurobarometer/public-opinion-in-the-eu-in-time-of-coronavirus-crisis (dostęp: 1.07.2021).
- Rinaldi, G., Paradisi, M. (2020). *An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak*, "medRxiv.org", www.medrxiv.org/content/10.1101/2020.04.18.20070912v2 (dostęp: 1.07.2021).
- Rzeczpospolita, *Skąd różnice w umieralności na COVID-19 pomiędzy różnymi państwami?*, www.rp.pl/Koronawirus-SARS-CoV-2/303299950-Skad-roznice-w-umieralnosci-na-COVID-19-pomiedzy-roznymi-panstwami.html (dostęp: 1.07.2021).
- Toktaş, Y. (2020). Spatial analysis of relationship between COVID-19 cases and globalization in Europe, *Zeszyty Naukowe SGGW, Polityki Europejskie, Finanse i Marketing*, 24(73), s. 255–265.
- WHO Director-General's opening remarks at the media briefing on COVID-19–11 March 2020, www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19-11-march-2020 (dostęp: 1.07.2021).
- WHO, Novel Coronavirus (2019-nCoV), Situation Report – 1, 21 January 2020, www.who.int/docs/default-source/coronaviruse/situation-reports/20200121-sitrep-1-2019-ncov.pdf (dostęp: 1.07.2021).

Behavioural Causes of Deaths Related to the COVID-19 Pandemic. Spatial Econometric Analysis of Mortality in the European Union in 2020

Abstract

Why did the COVID-19 pandemic lead to increased mortality of several dozen percent in some regions, while the death-toll in other places remained almost unchanged? The aim of the study was to investigate the determinants of the increased number of deaths in 198 regions of the European Union in 2020 using spatial econometric modeling tools. In particular, the study analyses the significance of behavioural characteristics of the inhabitants of every region, such as trust in public institutions and medical authorities as well as the compliance with the values followed by decision-makers with regard to the introduction of restrictions. The study shows that both greater trust in the government actions and recommendations of medical experts, as well as the belief that the positive health effects of restrictions outweigh the losses for the economy and limitations of personal freedom, lead to a lower increase in the deaths related to COVID-19. Moreover, it has been shown that a lower capacity of the health service in terms of health expenditure and the number of doctors, higher population density, higher domestic product per capita and higher number of deaths in nearby regions contributed to the increase in mortality in a given region.

Keywords: COVID-19 pandemic, mortality, behavioral causes of deaths, spatial econometric analysis

Dokąd zmierzamy? Wartości w reklamach w XXI wieku

Anna Kozłowska, Agnieszka Wojtukiewicz

Streszczenie

Pandemia COVID-19 w znaczącym stopniu wpłynęła na życie ludzi na całym świecie. Ta bezprecedensowa sytuacja znalazła swoje odzwierciedlenie w mediach, w tym w przekazach reklamowych. Celem niniejszej pracy jest identyfikacja wpływu pandemii COVID-19 na treść reklamową w obszarze prezentowanych wartości. Pierwsza część pracy poświęcona jest badaniom literaturowym na temat pojęcia wartości w różnych naukach. W drugiej części zostały przeanalizowane raporty CBOS dotyczące wartości cenionych przez Polaków. Trzecia część pracy to jakościowa analiza zawartości przekazów reklamowych. Przebadano 112 spotów reklamowych dla czterech rodzajów reklam – informacyjnych, emocjonalnych, tworzących nawyk oraz dających satysfakcję.

Słowa kluczowe: wartości, reklama, analiza zawartości treści, pandemia COVID-19

1. Wstęp

Wartość to pojęcie bardzo niejednoznaczne. Może się odnosić do aspektów ekonomicznych, psychologicznych, jak i duchowych. Rozważania na temat wartości oraz ich hierarchii obecne są w różnych naukach społecznych. Do wartości odwołują się przekazy medialne, w tym reklamowe, dzięki czemu mogą wywrzeć silny wpływ na odbiorcę i działać na jego proces decyzyjny.

Celem niniejszej pracy jest identyfikacja wpływu pandemii COVID-19 na treść reklamową w obszarze prezentowanych wartości. W związku z tym zostało postawione pytanie ogólne o to, jakie wartości były obecne w przekazach reklamowych przed oraz w trakcie pandemii COVID-19. Poszczególne wartości, których obecność była wskazywana w materiałach, zostały wyróżnione na podstawie raportów CBOS na temat systemu wartości Polaków oraz w oparciu o najpowszechniejsze wątki w danej kategorii produktowej. Zakłada się, że w czasie pandemii COVID-19 w wartościach pokazywanych w reklamie zaszły zmiany, które miały na celu dosto-

sowanie przekazów reklamowych do warunków pandemicznych. Z racji tego, że pandemia COVID-19 znacząco wpłynęła na codzienność ludzi na całym świecie i wiązała się ze znacznymi ograniczeniami, uznano za zasadne stwierdzenie, iż w reklamach zmiany te znalazły swoje odzwierciedlenie. Cel opracowania zostanie osiągnięty poprzez jakościową analizę zawartości treści reklamowej.

Część pierwsza to operacjonalizacja pojęcia wartości i przedstawienie go z perspektywy psychologicznej, socjologicznej, nauk o języku, a także nauk o kulturze. Część druga poświęcona jest analizie zmian, jakie zaszły w systemie wartości Polaków w ostatnich 15 latach, ze szczególnym uwzględnieniem czasu pandemii COVID-19. W tym celu wykorzystane zostały raporty CBOS „Aktualne problemy i wydarzenia”, zrealizowane w latach 2005–2019, oraz raport CBOS „Wartości w czasach zarazy” z 2020 r. Część trzecia to jakościowa analiza zawartości treści reklamowej. W zakres badanego materiału weszło 112 spotów reklamowych dla czterech kategorii produktowych: meble, piwo, środki czystości i perfumy. W celu analizy poszczególnych komunikatów reklamowych został stworzony specjalny klucz kategoryzacyjny.

2. Pojęcie wartości w różnych naukach

Temat wartości obecny jest w różnych naukach od wieków. Wartości badane są na wielu polach: przez psychologów, socjologów, antropologów, ekonomistów, pedagogów, etyków, teologów czy filozofów. Nie zawsze to pojęcie jest interpretowane w ten sam sposób [Feather, 2002, s. 52]. Wręcz uderzająca jest niejednoznaczność tego terminu w różnych naukach. Terminu „wartości” używa się, w zależności od kontekstu, do opisu zainteresowań, przyjemności, sympatii, preferencji, obowiązków, zobowiązań moralnych, żądz, pragnień, celów czy potrzeb [Williams, 1979, s. 16]. Wartości łączy się, lub nawet myli, z takimi pojęciami, jak: postawa, cecha, norma czy potrzeba. Kluczowym problemem dla nauki stało się ujednoczenie definicji, sposobów podjęcia tematu oraz środków do pomiaru wartości [Hitlin, Piliavin, 2004]. Co więcej, już po wyróżnieniu poszczególnych wartości, nie zawsze są one rozumiane w ten sam sposób – i tak sprawiedliwość, tolerancja, równość czy dobro to wartości, co do których nie ma jasnej zgody, co właściwie oznaczają.

W psychologii pojęcie wartości jest oddzielone od jego społecznego wymiaru – badacze koncentrują się na jednostce. Wartość traktuje się jako zjawisko o charakterze poznawczym, doświadczane indywidualnie, i definiuje w kontekście przekonań, motywów, potrzeb i zachowań człowieka [Dubis, 2014, s. 36–37]. Jest to integralna część myśli każdej osoby. Koncentrowanie się na badaniu umysłu człowieka przy jed-

noczesnym zaniedbywaniu świata wewnętrznego byłoby zaprzeczeniem podejścia holistycznego w psychologii i traktowania człowieka jako całości [Obacz, 2015, s. 63].

Pierwsze próby wprowadzenia pojęcia wartości w psychologii można odnaleźć w pracach Lewina z lat 50. XX w. Choć Lewin [Lewin, 1951, za: Feather, 2022, s. 63–64] koncentrował się głównie na pojęciu motywacji, a pojęcie wartości uważał za raczej niejasne, to jednak brał je pod uwagę w swojej działalności. Jego zdaniem wartości to właściwości jednostki, które mają wpływ na zachowanie, ale nie mają charakteru celu, np. osoba nie stara się „osiągnąć” uczciwości jako wartości, ale raczej uczciwość to coś, co kieruje zachowaniem. French i Kahn [French, Kahn, 1962, za: Feather, 2022, 66–67], podążając za myślą Lewina, twierdzą, że wartości mają zdolność do prowokowania działań zorientowanych na jakiś cel poprzez wywoływanie wartościowań wobec obiektów, zachowań lub stanów [Feather, 2002, s. 66–67]. Zmianę w podejściu do człowieka w psychologii przyniosły lata 60. i 70. XX w. Bühler, w przeciwieństwie do Lewina, postrzega wartości właśnie jako cele, których osiągnięcie ma przybliżyć ludzi do urzeczywistnienia ich tendencji życiowych, aspiracji, potrzeb, pragnień [Bühler, 1979, za: Obacz, 2015, s. 66–67].

Pojęcie wartości w socjologii podejmowane jest w szczególności w wymiarze społecznym – wiąże się z takimi pojęciami, jak: normy, wzorce czy standardy zachowań wspólne dla grup społecznych, zaś jako ich źródło identyfikuje się podobne doświadczenia [Dubis, 2014, s. 37]. Przy definiowaniu wartości kilka utrudnień wyszczególnił Hechter [Hechter, 1993, za: Hitlin, Piliavin 2004] w swojej pracy z 1993 r.:

- są one bardzo trudne do identyfikacji;
- współczesne badania nie dają wskazówek ku zrozumieniu, jak wartości kształtują zachowania;
- wyjaśnienia teorii behawioralnej są nieprzekonujące, kiedy nie znamy procesu, jaki stoi za generowaniem wartości;
- wartości są trudne do zmierzenia.

Hitlin i Piliavin kilkanaście lat później zauważają ponadto, że wartości:

- często są połączone z innymi zjawiskami psychologicznymi;
- różnią się historycznie i kulturowo, jeśli chodzi o ich znaczenie [Hitlin, Piliavin, 2004].

Jedną z ważniejszych definicji wartości zaproponował Kluckhohn w latach 50. XX w., stwierdzając, że wartość to koncepcja tego, co przez człowieka pożądane i chciane, charakterystyczna dla jednostki lub grupy, mająca wpływ na wybory spośród możliwych sposobów, środków i rezultatów działań. Może być wyraźna lub domniemana [Kluckhohn, 1951, s. 395]. Z kolei Rokeach w swojej klasycznej pracy z 1973 r. twierdził, że wartości to trwałe przekonanie, że pewien sposób postępowania jest indywidualnie lub społecznie bardziej preferowany niż przeciwny czy

odmienny sposób postępowania [Rokeach, 1973, s. 5]. Kolejną definicję wartości, podobną do sformułowanej przez Kluckhohna, zaproponowali Schwartz i Bilsky [1987, s. 551] – ich zdaniem wartości to koncepcje lub przekonania o pożądanym rezultatach czy zachowaniach, które wykraczają poza specyficzne sytuacje, kierują wyborami lub oceną działań i wydarzeń, a także są uporządkowane według istotności.

We współczesnych badaniach nad wartościami często analizowane są relacje wartości i języka. Z punktu widzenia przeciętnego człowieka język pomaga ludziom lepiej zrozumieć, czym jest w ogóle pojęcie wartości, a także jak rozumieć poszczególne wartości [Wartecka, 2015, s. 186], jaka jest hierarchia wartości i relacje między nimi oraz jaki jest kanon wartości w danej kulturze [Bartmiński, Grzeszczak, 2014]. Język jest narzędziem naszego myślenia, a co za tym idzie – już samo nazywanie może prowadzić do określonych reakcji emocjonalnych [Puzynina, 2014, s. 9]. Bartmiński i Grzeszczak [2014] wyróżniają trzy najważniejsze funkcje języka dla wartości:

- pełni wobec wartości funkcję informacyjną – dzięki niemu można w ogóle mówić o wartościach jako takich;
- utrwała wartości i stanowi narzędzie do ich przekazywania;
- umożliwia wartościowanie.

Problemy z definicją wartości zauważa również Puzynina [2014, s. 4], np. słowo „wartość” łączy się z innym pojęciem – „wartościowaniem” – i wtedy oznacza to, co uważamy za dobre (wartość pozytywna) lub to, co uważamy za złe (wartość negatywna). W języku pojawiło się również słowo „antywartość”, a wtedy „wartość” pozostaje już tylko określeniem pozytywnym. Trudno jednak zidentyfikować spektrum wartości – czasem za wartości uznaje się wszystko, co jest pozytywne, w tym wartości witalne czy hedonistyczne, a czasem wartościami są tylko wartości wyższe – duchowe.

3. Zmiany w systemie wartości Polaków w obliczu pandemii COVID-19

Centrum Badania Opinii Społecznej (CBOS) od 2005 r. co kilka lat publikuje raporty dotyczące wartości i norm w życiu Polaków. Analizowane w pracy badania zostały przeprowadzone w latach 2005, 2010, 2013, 2019¹ oraz 2020 [CBOS, 2005; CBOS, 2010; CBOS, 2013; CBOS, 2019; CBOS, 2020]. Pierwszym wnioskiem, który nasuwa się przy analizie raportów sprzed pandemii COVID-19 w zestawieniu z raportem z listopada 2020 r., jest zmiana na pierwszym miejscu listy. We wcześniejszych

¹ W przypadku raportu z 2019 r. nie jest dostępna pełna wersja, a jedynie skrócona publikacja wyników badania.

raportach nigdy żadna z wartości nie była powyżej rodziny i szczęścia rodzinnego. Tymczasem pandemia przyniosła znaczącą zmianę na tej pozycji. Polacy w 2020 r. jako najważniejszą wartość wskazali zdrowie. Tę opcję wybrało aż 47% ankietowanych. Można przypuszczać, że jest to bezpośredni efekt pandemii COVID-19, kiedy to zdrowie stało się kluczową kwestią i troską wielu Polaków. Drugie miejsce i 39% głosów w 2020 r. uzyskała rodzina. Dodatkowo 2% respondentów wskazało jako najważniejszą wartość dobro rodziny lub pomyślność rodziny. Nawet po zsumowaniu tych odpowiedzi nadal pozostaje to wynik odległy od pierwszego miejsca o 6 punktów procentowych [CBOS, 2020].

Analizie warto poddać również odpowiedzi, które znalazły się na kolejnych miejscach w rankingu. W latach 2005–2013 trzecią lokatę zajmowało uczciwe życie, a odsetek respondentów wskazujących tę odpowiedź rósł. Na kolejnych miejscach znajdowały się naprzemiennie w różnych latach takie wartości, jak praca zawodowa i spokój. Już w 2019 r. widać znaczące różnice w tej kategorii: na trzecim miejscu uplasował się spokój, dalej uczciwe życie i wiara religijna. Praca zawodowa zajmowała dopiero ósme miejsce [CBOS, 2020]. Tymczasem zaledwie nieco ponad rok później to właśnie pracę, pewność zatrudnienia i stabilne zatrudnienie wskazało 8% ankietowanych, co było trzecim najwyższym wynikiem. Kolejna pozycja również ma znaczenie ekonomiczne: 6% respondentów powiedziało, że najważniejsze w ich życiu są: dobrobyt, dostatek, bezpieczeństwo finansowe. Na kolejnym miejscu znalazł się spokój. Zajmujące w poprzednich latach wysokie lokaty uczciwość i wiara religijna zdobyły kolejno zaledwie 1% i 3% [CBOS, 2020].

Na podstawie badań opublikowanych przez CBOS można stwierdzić, że pandemia COVID-19 zmieniła hierarchię wartości Polaków (ciekawym jest, jak długo utrzyma się taka tendencja). Polacy bardziej niż w poprzednich latach cenią sobie zdrowie oraz pracę i wartości materialne, co może być efektem niestabilnej sytuacji na świecie. Praca i regularny wpływ środków finansowych mogą dawać poczucie bezpieczeństwa ekonomicznego, a przez to pewnej stabilności i przygotowania na różne koleje losu. Istotnymi wartościami w życiu Polaków są też rodzina, spokój i ojczyzna, które mogą stanowić wyraz dążeń do utrzymywania tradycji i bezpieczeństwa psychicznego.

4. Metodologia badań

Jako metodę badawczą przyjęto jakościową analizę zawartości treści, która daje możliwość nie tylko identyfikacji stosowanych sposobów oddziaływania na odbiorcę, ale również określenia, jakie były intencje nadawcy przekazów reklamowych.

Materiał badawczy został dobrany na podstawie modelu FCB, w którym wyróżnia się cztery strategie komunikacji reklamowej: informacyjną, emocjonalną, tworzącą nawyk oraz dającą satysfakcję. W przypadku każdej strategii wybrano jeden rodzaj produktu, kolejno: meble, perfumy, środki czystości i piwo. W badaniu brano pod uwagę spoty trwające między 16 a 60 sekund. Materiały do badania zostały wybrane spośród filmów proponowanych przez algorytm Google po zawężeniu wyników wyszukiwania do określonych kryteriów. Przeanalizowano łącznie 112 przekazów reklamowych z sześciu przedziałów czasowych:

- przed pandemią – styczeń 2018–grudzień 2019 (ze względu na niską świadomość obecności pandemii w badaniu nie brano pod uwagę stycznia i lutego 2020);
- początek pandemii – marzec–maj 2020 – powszechna izolacja, czas najbardziej restrykcyjnego zamknięcia w domach;
- letnie luzowanie obostrzeń – czerwiec–sierpień 2020 – czas wakacji i niższej liczby zakażeń koronawirusem, poluzowania obostrzeń;
- druga fala pandemii – wrzesień–listopad 2020 – zaostrzenie restrykcji spowodowane gwałtownym wzrostem zachorowań w sezonie jesienno-zimowym;
- czas świąteczny – grudzień 2020 – szczególny czas dla reklamodawców;
- trzecia fala pandemii – styczeń–marzec 2021 – ponowne wprowadzenie restrykcji, moment, gdy już przez ok. rok żyliśmy w rzeczywistości pandemicznej i kolejna fala nie była już zaskoczeniem.

Celem analizy badawczej został opracowany klucz kategoryzacyjny. Klucz zawiera 12 pytań ogólnych, a do każdego z nich przyjęto zestaw odpowiedzi, dzięki czemu kodowanie wyników badania w dużej mierze ograniczyło się do zaznaczeń tak/nie.

Tabela 1. Klucz kategoryzacyjny

Pytania ogólne	Pytania szczegółowe
Tytuł reklamy	
Link	
1.1. A. Do jakich wartości odwołuje się reklama?	rodzina
	▪ przyjaźń*
	▪ honor*
	▪ dzieci
	▪ zdrowie
	▪ praca
	▪ dobrobyt, bezpieczeństwo finansowe
	▪ spokój
▪ ojczyzna	

Pytania ogólne	Pytania szczegółowe
	<ul style="list-style-type: none"> ▪ uczciwe życie
	<ul style="list-style-type: none"> ▪ natura, ekologia*
	<ul style="list-style-type: none"> ▪ przyjemność*
	<ul style="list-style-type: none"> ▪ czystość, ład, porządek*
2.1. A. Czy reklama pokazuje, że ludzie mogą przezwyciężyć wszelkie trudności dzięki pozytywemu nastawieniu?	Czy reklama pokazuje, że ludzie mogą przezwyciężyć wszelkie trudności dzięki danej marce?
	Jak marka wspomaga ludzi w przezwyciężaniu trudności?
3.1. A. Czy w reklamie łączy się markę z czasem?	<ul style="list-style-type: none"> ▪ przeszłość (tradycja, tradycyjne przepisy, tradycyjna kuchnia, dawna technologia, stara receptura, oryginalny od... roku) ▪ teraźniejszość (jesteśmy tu i teraz, to jest Twój dzień, dziś, obecnie, ten moment) ▪ przyszłość (nowe technologie, idziemy do przodu, rozwijamy się, zmieniamy się)
4.1. A. Jak przedstawia się relacje międzyludzkie?	Czy reklama pokazuje skupiska ludzi?
	Czy ludzie zachowują dystans?
	Czy ludzie się przytulają?
	Czy ludzie są pokazywani w intymnych sytuacjach, blisko siebie?
	W jaki sposób się witają?
	W jaki sposób spędzają czas wolny?
	W jaki sposób ludzie się ze sobą komunikują, np. przez Internet czy raczej osobiście?
	Czy noszą maseczki?
5.1. A. Jaki sposób radzenia sobie z problemem pojawia się w reklamie?	Usunięcie problemu: problem – irytacja – produkt – usunięcie problemu – ulga
	Uniknięcie problemu: problem – strach – produkt – uniknięcie problemu – uspokojenie
	Niepełna satysfakcja: rozczarowanie – produkt – poczucie satysfakcji
	Ambivalentna postawa: niepokój – produkt – postawa – spokój
	Brak produktu: niski poziom irytacji – produkt – wygoda
6.1. A. Jaki sposób radzenia sobie z negatywnymi emocjami pojawia się w reklamie?	Gratyfikacja zmysłowa: bylejakość (lub neutralność) – produkt – radość
	Stymulacja intelektualna: obawa (lub neutralność) – produkt – szczęście
	Dążenie do mistrzostwa: naiwność (lub neutralność) – produkt – poczucie kompetencji
	Uznanie społeczne: lęk społeczny (lub neutralność) – produkt – pochlebstwo
	Konformizm: wstyd (lub neutralność) – produkt – duma

cd. tabeli 1.1

Pytania ogólne	Pytania szczegółowe
7.1. A. Czy reklama podkreśla fakt wspierania bezpieczeństwa transakcji (np. zakupy bezgotówkowe), do jakiego typu bezpieczeństwa się odwoływano, np. bezpieczeństwo finansowe, epidemiczne?	
8.1. A. Czy w reklamie występuje nawiązanie do domu jako miejsca bezpiecznego?	<ul style="list-style-type: none"> ▪ dom jako miejsce bezpieczne ▪ dom to rodzina ▪ zostać w domu ▪ praca w domu ▪ rozrywka w domu (np. oglądanie filmów, gry itp.) ▪ dom to szczęście, wspólne spędzanie czasu
8.2. A. Czy w reklamie podkreśla się potrzebę spędzania czasu poza domem?	<ul style="list-style-type: none"> ▪ imprezy plenerowe ▪ spędzanie wolnego czasu na powietrzu, np. jazda na rowerze ▪ wspólne wyjścia do kina ▪ wydarzenia sportowe ▪ praca poza domem ▪ spotkania u przyjaciół ▪ spędzanie czasu na łonie natury (park, las, łąka) ▪ wyjścia do baru, restauracji, klubu ▪ wyjścia do sklepu, na zakupy
9.1. A. Czy w reklamie pojawiają się informacje o zaangażowaniu przedsiębiorstwa w działalność społeczną? Jaką?	<ul style="list-style-type: none"> ▪ wolontariat pracowniczy ▪ zysk ze sprzedaży produktu zasila organizację charytatywną ▪ wsparcie wolontariatu i fundacji ▪ wspieranie społeczności lokalnej
9.2. A. Czy w reklamie pojawiają się jasne i klarowne informacje o wysiłkach, jakie marka podejmuje w walce z jakimś problemem społecznym? Dlaczego ten problem społeczny? Czy jest powiązany z marką? Czy jest powiązany z grupą docelową?	<ul style="list-style-type: none"> ▪ bieda ▪ bezdomność ▪ choroby cywilizacyjne ▪ przemoc ▪ alkoholizm ▪ samotność osób starszych ▪ problemy ze środowiskiem
9.3. A. Czy w reklamie pojawiają się informacje o wspieraniu konkretnych grup społecznych, np. lekarzy, swoich pracowników, osób starszych?	

* Wartości widoczne w reklamach, ale spoza raportów CBOS.

Źródło: opracowanie Anna Kozłowska.

5. Analiza zawartości przekazów reklamowych

Pierwszą grupę spotów reklamowych stanowią reklamy informacyjne poświęcone kategorii mebli. Przeanalizowano 10 spotów sprzed pandemii (styczeń 2018–grudzień 2019) i 18 reklam z okresu marzec 2020–marzec 2021. Wartością, do której odwoływano się w największej liczbie przekazów reklamowych, był spokój. Wpisuje się on w narrację, gdzie meble gwarantują użytkownikom komfort i wygodę. Przed pandemią do wygody odwoływano się rzadziej niż po jej wybuchu. W czasie pandemii częściej jako wartość wskazywano pracę i zdrowie, co pokrywa się z wynikami zaprezentowanymi w raportach CBOS. Rzadziej pojawiały się nawiązania do rodziny czy dzieci, czego możliwą przyczyną jest fakt, że w trakcie pandemii wielu ludzi straciło swoich bliskich i takie odwołanie dla niektórych odbiorców mogłoby być bolesne.

W reklamie mebli stosunkowo rzadko mówiono o tym, że wszelkie trudności można przezwyciężyć dzięki pozytywnemu nastawieniu lub dzięki danej marce, a w czasie pandemii odsetek reklam o takim przekazie jeszcze spadł. Przed pandemią w spotach mówiono o tym, że marka oferuje dopasowane do klienta produkty, odpowiada na jego potrzeby i zapewnia wygodę. W reklamie w trakcie pandemii odwoływano się do zdrowia – promowano zachowania prozdrowotne i ochronę przed koronawirusem, np. zachowywanie dystansu czy pozostawanie w domu. Ponadto pokazywano zmiany w wystroju wnętrz jako odskocznię od codzienności, a dom jako miejsce, gdzie można czuć się spełnionym i wolnym. Marki mebli zazwyczaj łączono z przeszłością i przyszłością, rzadko pojawiały się odniesienia do teraźniejszości. W materiale badawczym odniesień do czasu jest niewiele, a tylko dwa takie odniesienia pojawiły się w czasie pandemii. Może to wskazywać na wyjątkowość okresu pandemicznego i chęć oderwania marki od trudnych warunków zewnętrznych.

W materiale badawczym dominowały dwa sposoby radzenia sobie z problemem – „brak produktu: niski poziom irytacji – produkt – wygoda” oraz „usunięcie problemu: problem – irytacja – produkt – usunięcie problemu – ulga”. Taki model może wynikać ze specyfiki produktu – meble gwarantują użytkownikom komfort oraz podwyższają ogólny standard życia. Pozostałe trzy sposoby radzenia sobie z problemem pojawiały się raczej sporadycznie, z podobną częstotliwością.

Kolejny etap analizy obejmował wskazanie sposobu radzenia sobie z negatywnymi emocjami. W tym przypadku dominował model „gratyfikacji zmysłowej: bylejakość (lub neutralność) – produkt – radość”. Meble wyraźnie mają nam poprawić humor – wprowadzić w dobry nastrój. Tego rodzaju produkt pozwala też konsumentom dążyć do mistrzostwa, do poczucia kompetencji, a także szczęścia.

Tylko jedna reklama spełniała założenia modelu „uznanie społeczne: lęk społeczny (lub neutralność) – produkt – pochlebstwo”. Nigdzie nie wystąpił model „konformizm: wstyd (lub neutralność) – produkt – duma”. Wyraźnie meble mają zaspokajać nasze potrzeby indywidualne – nie ma dla nas znaczenia docenienie przez innych czy akceptacja społeczna. Meble kupujemy dla siebie, dla własnej wygody i poczucia satysfakcji.

W reklamie mebli w czasie pandemii zaszły widoczne zmiany. O ile zarówno w trakcie pandemii, jak i przed nią w reklamach pokazywano skupiska ludzi, np. w sklepach meblowych, to już od marca 2020 r. w żadnej reklamie ludzie nie witają się poprzez uścisk ręki czy objęcie (podczas gdy we wcześniejszym okresie taka forma powitania się zdarzała). Co więcej, w niektórych przekazach reklamowych wyraźnie pokazane jest zachowywanie dystansu społecznego. W okresie pandemii pojawiły się też wątki komunikacji online, kiedy to członkowie rodziny i znajomi rozmawiają ze sobą przez komunikatory internetowe. Aż połowa reklam zawierała informacje o możliwości zakupu online z dostawą do domu lub z wygodnym odbiorem w sklepie. W reklamie nikt jednak nie nosił maseczki, nawet przebywając w sklepie meblowym lub markecie budowlanym, choć to właśnie zakupy były najczęściej promowaną aktywnością podejmowaną poza domem. Przed pandemią oprócz wyjść na zakupy najczęstszą czynnością pokazywaną w reklamach były spotkania w restauracjach, barach i klubach, z kolei po marcu 2020 r. taki wzorzec zachowania nie został ukazany w ani jednej reklamie. W ponad połowie reklam to dom stał się miejscem pracy, odpoczynku, rozrywki, spędzania czasu z bliskimi, co było naturalnym odzwierciedleniem sytuacji życiowej wielu Polaków.

Właściciele marek meblowych rzadko decydują się na poruszanie w komunikacji reklamowej problemów społecznych. Przed pandemią w grudniu 2019 r. jedna z marek zachęcała do spotkań z seniorami w trakcie świąt Bożego Narodzenia. Podczas pandemii pojawiła się reklama promująca kampanię społeczną „Cała Polska walczy z koronawirusem”. W przekazie zawarto informacje o wspieraniu konkretnej fundacji. W innej reklamie, grudniowej, poruszono problem marnowania żywności w okresie świątecznym. To znowu pokazuje, że meble łączone są wyraźnie z wartościami indywidualnymi, które z rzadka przekładają się na potrzeby społeczne.

Podsumowując, w reklamie mebli wątki związane z pandemią COVID-19 były wyraźnie obecne. Zmiany w wartościach promowanych w reklamach były zgodne z trendami, które pojawiały się w raportach CBOS. W treści przekazów reklamowych szczególnie częstym było promowanie zakupów online. Odpowiedzią na sytuację pandemii było także pokazywanie w reklamach rozmów przez komunikatory internetowe oraz ograniczonego kontaktu fizycznego przy powitaniach. Niestety, żaden z twórców reklam nie odważył się pokazać rzeczywistości sklepów meblo-

wych i marketów budowlanych, gdzie ludzie mogą przebywać jedynie w maseczkach ochronnych. Taki zabieg mógłby pozytywnie wpłynąć na zachowania ludzi i zachęcić ich do zakrywania nosa i ust w zamkniętej przestrzeni. W dostępnym materiale badawczym pojawiło się również bardzo mało kampanii społecznych, szczególnie tych związanych z pandemią COVID-19. Uważamy, że to również niewykorzystana szansa na promocję pozytywnych postaw społecznych.

Kolejną grupą komunikatów reklamowych są reklamy emocjonalne dotyczące perfum. Dla tej kategorii produktowej dostępny był najobszerniejszy materiał badawczy. Dokonano analizy 10 spotów reklamowych sprzed pandemii (styczeń 2018–grudzień 2019) i 19 reklam z okresu marzec 2020–marzec 2021. Co ciekawe, w tym wypadku najczęściej pojawiającymi się wartościami były te nieujęte w raportach CBOS – przyjemność i niezależność. W reklamie perfum rzadko pojawiały się inne wartości, zarówno przed pandemią, jak i w trakcie. Tylko w kilku przekazach odwoływano się do przyjaźni czy dobrobytu finansowego. Prawdopodobnie jest to efekt narracji prowadzonej w reklamie produktów nastawionych na budowanie osobowości konsumenta, gdzie konkretny zapach i marka dają użytkownikom siłę do przezwyciężania trudności oraz przyjemność płynącą z doznań zmysłowych i estetycznych. Co istotne, nie zmienia się to w trakcie pandemii. Często w spotach występują znane aktorki i aktorzy – są przepelnieni energią, nie ma dla nich rzeczy niemożliwych, a perfumy dodatkowo wzmacniają ich siłę i pomagają w obliczu przeciwności losu. Marki motywują bohaterów do podjęcia działań, przyciągają ludzi, dają poczucie wolności i niezależności, a każdy może być tym, kim chce. Fabuła spotów zazwyczaj rozgrywa się poza czasem, rzadko padają określenia nawiązujące do przeszłości i przyszłości – bohater reklamowy żyje tu i teraz. W żadnym spocie reklamowym nie pojawiły się odniesienia do teraźniejszości, również do czasu pandemii.

Najczęstszym sposobem radzenia sobie z problemem jest model niepełnej satysfakcji, gdzie na początku występuje rozczarowanie, a następnie, po skorzystaniu z produktu, poczucie satysfakcji. W tym wypadku nie występuje irytacja, a skutki skorzystania z produktu nie kończą się na wygodzie. Perfumy stanowią dodatkowy element codzienności, który stanowi pewne jej uzupełnienie, a efekty ich stosowania mają dawać przyjemność – rzadziej usuwać konkretne niedogodności. Nie dziwi zatem fakt, że najczęstszym sposobem radzenia sobie z negatywnymi emocjami jest w tym wypadku gratyfikacja zmysłowa, gdzie produkt przynosi użytkownikowi radość. Co ciekawe, na drugim miejscu znalazł się model konformistyczny, gdzie na początku występuje wstyd lub neutralność, a następnie dzięki produktowi konsument odczuwa dumę. Trzecim najczęstszym sposobem radzenia sobie z negatywnymi emocjami było uznanie społeczne zwieńczone pochlebstwem.

Wydaje się jednak, że chodzi tutaj raczej o uznanie społeczne osoby wyróżniającej się na tle innych.

Treść reklam perfum jest skoncentrowana wokół wolności oraz troski o samego siebie. Większość spotów ukazuje skupiska ludzi, ale w żadnym spocie z czasów pandemii ludzie nie zachowują dystansu społecznego ani nie noszą maseczek. Komunikują się ze sobą tylko osobiście, nikt nie używa komunikatorów internetowych. Nigdzie również nie podejmowano problemów społecznych. Tylko w jednym spocie zachęcano do robienia zakupów bez wychodzenia z domu. Bohaterowie spotów reklamowych spędzają czas na łonie natury, na przyjęciach, w restauracjach czy w gronie przyjaciół. Tylko w trzech reklamach pokazano rozrywkę w domu. Bohaterowie często nawiązują też kontakt fizyczny, a ich relacje z płcią przeciwną mają wyraźnie romantyczny charakter.

Podsumowując, pandemia COVID-19 w reklamie perfum nie istnieje. Akcja spotów rozgrywa się zupełnie poza światem rzeczywistym, a bohaterowie reklam są piękni, wyzwoleni i adorowani przez innych. Jednym z powodów takiej sytuacji może być fakt, iż reklamy perfum są często elementami globalnych kampanii zaprojektowanych z dużym wyprzedzeniem czasowym i w momencie ich nagrywania pandemia jeszcze nie wpłynęła na świat. Ponadto reklamy luksusowych produktów przyzwyczyły konsumentów do przenoszenia ich w inny, lepszy świat. Piękne kadry, eleganckie stroje bohaterów i nastrojowa muzyka pozwalają oderwać się od rzeczywistości pandemii, i to również mogło być jednym z celów twórców kampanii reklamowych.

Kolejną grupą spotów reklamowych są reklamy tworzące nawyk, promujące środki czystości, takie jak proszki do prania, ściereczki czy płyny do naczyń. Ze względu na podobieństwo między kolejnymi reklamami oraz dostępność materiału badawczego kilka rodzajów produktów zostało uwzględnionych w jednej kategorii. W zakres materiału badawczego weszło osiem spotów reklamowych sprzed pandemii (styczeń 2018–grudzień 2019) i 17 reklam z okresu marzec 2020–marzec 2021. Nie dziwi zapewne fakt, że najczęściej promowaną wartością w tym wypadku była rodzina i dzieci, kolejną – przyjemność (w tym również komfort, wygoda), a wszystko to powiązane zostało z potrzebą homeostazy – czystością, ładem i porządkiem. Taki rozkład akcentów wynika ze specyfiki kategorii produktowej – często o czystość dba się z przyczyn estetycznych i higienicznych, a także w trosce o bezpieczeństwo rodziny, w szczególności dzieci, które nie mogą jeszcze zatroszczyć się o siebie.

W reklamie środków czystości stosunkowo często mówi się o tym, że pozytywne nastawienie może pomóc w przezwyciężeniu wszelkich trudności, chociaż w trakcie pandemii taka narracja pojawiała się rzadziej. Produkty tego rodzaju ułatwiają codzienne życie, zapewniają komfort, wygodę i czystość. W tym wypadku markę

łączy się z przyszłością. Najczęściej pojawiają się frazy takie jak: nowa technologia, innowacyjne/a..., nowość. Jednak tylko jedna reklama nawiązywała do sytuacji pandemii i potrzeby skutecznej dezynfekcji.

W reklamie środków czystości najczęstszym schematem radzenia sobie z problemem jest „usunięcie problemu: problem – irytacja – produkt – usunięcie problemu – ulga”. Można przypuszczać, że popularność takiego schematu dla kategorii środków czystości wynika z tego, iż scenariusze takich reklam często przedstawiają prostą fabułę, gdzie dzięki konkretnemu produktowi znika problem, jakim jest plama lub inne zanieczyszczenia. Środki czystości pozwalają też uniknąć problemu, a w przypadku wcześniejszego braku produktu – zapewniają wygodę po użyciu konkretnego specyfiku. Najrzadziej pojawiał się model niepełnej satysfakcji, kiedy to na początku występuje rozczarowanie, a dzięki produktowi konsument osiąga poczucie satysfakcji. W żadnej z reklam natomiast nie pojawił się schemat ambiwalentnej postawy, kiedy to produkt zapewnia spokój. Pokrywa się to z wynikami analizy wartości ukazywanych w reklamach, gdyż spokój pojawiał się tam zdecydowanie rzadko. Wśród prezentowanych sposobów radzenia sobie z negatywnymi emocjami dominował model gratyfikacji zmysłowej. Drugim najczęściej pojawiającym się modelem było dążenie do mistrzostwa poprzez użycie odpowiednich specyfików. Prawdopodobną przyczyną popularności tego schematu może być fabuła spotów środków czystości – osoby odpowiedzialne za prowadzenie domu początkowo nie znają jakiegoś rozwiązania i gdy już wprowadzą nowy produkt do swojego życia, czują się kompetentne, a ich życie staje się łatwiejsze. Mniej popularnymi sposobami radzenia sobie z negatywnymi emocjami dla tej kategorii produktowej były modele stymulacji intelektualnej, konformizmu i uznania społecznego.

Charakterystycznym dla reklam środków czystości jest fakt, że rzadko występują w nich grupy ludzi. Środki czystości to produkty, z których korzysta się zazwyczaj samemu lub w niewielkim gronie osób, które dzielą to samo gospodarstwo domowe. Ten aspekt znajduje swoje odzwierciedlenie w reklamie. Takie treści, jak spotkania u przyjaciół, wyjścia na zakupy czy spędzanie czasu na świeżym powietrzu pojawiają się w tym wypadku stosunkowo rzadko, przy czym w żadnym spocie bohaterowie nie nosili maseczek ochronnych ani nie komunikowali się z bliskimi za pomocą Internetu. Informacja o konieczności dezynfekcji pojawiła się tylko w jednej reklamie. W analizowanych przekazach to dom jest bezpiecznym miejscem, gdzie można spędzać czas z rodziną, dobrze się bawić i być szczęśliwym. Tylko jedna reklama miała charakter kampanii społecznej – dotyczyła ochrony środowiska naturalnego i problemu braku wody.

Podsumowując, pandemia w reklamie środków czystości jest niemalże nieobecna. Tylko jeden przekaz odnosił się wprost do rzeczywistości pandemii i zwracał

uwagę na potrzebę dezynfekcji. Możliwymi przyczynami braku tematu pandemii jest specyfika środków czystości – w reklamie prezentowane jest tylko działanie chemiczne produktu, bez społecznego wymiaru sprzątania. Nie pojawia się również zapewnienie o działaniu specyfiku na wirusy, zapewne w obawie przed oskarżeniem o manipulację strachem przed pandemią.

Na koniec przeanalizowane zostały spoty reklamowe piwa, zaliczane do kategorii reklam dających satysfakcję. Przeanalizowano 10 reklam sprzed pandemii (styczeń 2018–grudzień 2019) i 16 spotów z okresu marzec 2020–marzec 2021. W reklamie piwa najczęściej odwoływano się do pragnienia przyjemności oraz do wartości przyjaźni. Spożywanie napojów alkoholowych ma sprzyjać miłej, radosnej atmosferze i przebywaniu w gronie znajomych. Co ciekawe, w trakcie pandemii zwiększyła się częstość zarówno odwołań do przyjemności, jak i do przyjaźni – najwyraźniej kontakty międzyludzkie w czasie izolacji stały się szczególnie cenne. Najczęściej promowaną wartością spośród tych uwzględnionych w raportach CBOS jest spokój, co łączy się z narracją, gdzie spożywanie piwa sprzyja odpoczynkowi i relaksowi. W czasie pandemii pojawiły się także nawiązania do pracy i zdrowia jako wartości. Treści te nie były obecne w badanym materiale sprzed pandemii. W reklamie piwa nie przedstawiono jako wartości uczciwego życia, dzieci i rodziny – są one mało istotne z punktu widzenia tej kategorii produktowej.

Reklama pokazuje, że dzięki danej marce i pozytywnemu nastawieniu można przewyżczać trudności. Jest to widoczne zarówno w spotach przed pandemią, jak i w tych z czasu pandemii. Spożywanie piwa ułatwia nawiązywanie nowych znajomości i podejmowanie odważnych, czasem nawet szalonych, decyzji. W wielu przekazach reklamowych pojawiały się nawiązania do czasu – najczęściej mówiono o tradycyjnych technologiach wytwarzania piwa i tradycjach browarniczych, co daje poczucie kontroli rzeczywistości.

W reklamie piwa był głównie przedstawiony sposób radzenia sobie z problemem poprzez „brak produktu: niski poziom irytacji – produkt – wygoda”. Może to wynikać z faktu, iż piwo nie jest raczej traktowane jako produkt niezbędny w codziennym życiu – ma uprzyjemniać życie, dawać satysfakcję. Często spotykane były również modele „usunięcie problemu: problem – irytacja – produkt – usunięcie problemu – ulga”, zazwyczaj w scenariuszu reklam, gdzie grupa znajomych nie miała pomysłu na spędzenie wolnego czasu i dzięki piwu zdecydowała się podjąć jakąś akcję. Typowe jest też „uniknięcie problemu: problem – strach – produkt – uniknięcie problemu – uspokojenie”, kiedy fabuła pokazuje rywalizację pomiędzy ludźmi oraz „niepełna satysfakcja: rozczarowanie – produkt – poczucie satysfakcji”, kiedy bohaterowie są znudzeni i poszukują wrażeń. W reklamie piwo nie daje uspokojenia, co pokrywa się ze stosunkowo niskim przedstawieniem spokoju jako wartości.

W reklamie piwa jako sposób radzenia sobie z negatywnymi emocjami zdecydowanie dominowała „gratyfikacja zmysłowa: bylejakość (lub neutralność) – produkt – radość”, gdyż piwo ma dawać po prostu satysfakcję – natychmiastową przyjemność. Drugim najczęstszym modelem było dążenie do mistrzostwa i w efekcie stosowania produktu – poczucie kompetencji. Każdy z kolejnych trzech modeli był już rzadziej obecny. Piwo to wyraźnie produkt realizujący indywidualne potrzeby konsumenta – chociaż może być konsumowany w otoczeniu innych i przynosić poczucie spełnienia w kontaktach międzyludzkich, to przede wszystkim ma przynosić satysfakcję pijącemu.

Treść spotów piwa nie zmieniła się znacząco w czasie pandemii – nadal pojawiają się skupiska ludzkie. W żadnej z reklam ludzie nie noszą maseczek, a tylko w pojedynczych zachowywany jest dystans między uczestnikami spotkania. Bohaterowie reklam często się przytulają lub nawiązują ze sobą bliski kontakt fizyczny. Tylko w dwóch reklamach pokazuje się komunikację przez Internet. Co ważne, reklamy te pochodziły z pierwszego okresu pandemii, czyli marca–maja 2020, kiedy to izolacja społeczna była najbardziej restrykcyjnie przestrzegana. Wśród spotów, które powstały w czasie pandemii, znalazło się wiele takich, gdzie bohaterowie wychodzą do baru, klubu lub restauracji, spędzają czas na świeżym powietrzu, uczestniczą w imprezach plenerowych, spotykają się u przyjaciół, a nawet biorą udział w wydarzeniach sportowych. Dom zazwyczaj nie jest ukazywany jako miejsce rozrywki i zabawy – taki wątek był widoczny zaledwie w czterech komunikatach reklamowych. W reklamach piwa nie ma również informacji o zaangażowaniu w wolontariat. Jeden przekaz reklamowy promował region, z którego pochodzi dana marka piwa, zaś dwie inne reklamy zwracały uwagę na problemy związane ze środowiskiem naturalnym – nie wydaje się to jednak w żaden sposób powiązane z pandemią.

Podsumowując, z wyjątkiem dwóch przypadków koronawirus w reklamie piwa nie istnieje. Promowane wartości przed pandemią i w jej trakcie niewiele się różnią. Również treść przekazów reklamowych jest oparta na podobnych scenariuszach: pokazywana jest dobra zabawa w gronie przyjaciół, zazwyczaj w miejscach publicznych, bez zachowania środków bezpieczeństwa. Nigdzie nie zachęcano do pozostania w domu czy izolacji społecznej, nie promowano także zakupów online z dostawą do domu. Ze względu na specyfikę kategorii produktowej trudno jednak oczekiwać takich działań od producentów piwa. W kulturze źle widzianym jest spożywanie alkoholu w samotności. Pandemia koronawirusa nie wpisuje się w tę narrację i prawdopodobnie z tego powodu praktycznie nie występuje w przekazach.

6. Wnioski

Pandemia COVID-19 znacząco wpłynęła na nasze poczucie pewności i bezpieczeństwa. Jej oddziaływanie widoczne jest w wielu dziedzinach życia – również w wartościach wyznawanych przez ludzi. Zmiany te znajdują swoje odzwierciedlenie w mediach oraz w kulturze. Celem niniejszej pracy było zbadanie poprzez jakościową analizę treści, czy zmieniał się przekaz komunikatów reklamowych w obszarze wartości. Zgodnie z założeniem wartości przedstawiane w reklamach faktycznie zmieniły się podczas pandemii, aczkolwiek występowały różnice pomiędzy przekazami dla różnych kategorii produktowych. Największe zmiany zaszły w reklamie mebli, należącej do kategorii reklam informacyjnych. W trakcie pandemii cenioną wartością stał się spokój; nie zabrakło też odwołań do pracy czy zdrowia, które były rzadziej obecne przed pandemią. Najbardziej odporne na zmiany okazały się za to reklamy perfum, w których wątki związane z koronawirusem, dystansem społecznym czy izolacją praktycznie zupełnie nie występują. Ze względu na mniejszą próbę badawczą trudniej porównać stan przed pandemią i w trakcie pandemii w przypadku reklam środków czystości, chociaż i w tym wypadku widać większe nastawienie na rodzinę i zdrowie. W reklamie piwa z kolei przed pandemią nie ukazywano pracy i zdrowia jako wartości, zaś w okresie od marca 2020 r. do marca 2021 r. wątki te zaczęły się pojawiać. Tego rodzaju różnice w narracji reklamowej wydają się być zasadne – zależne od kategorii produktowej.

W analizowanym materiale badawczym tylko raz poinformowano o wsparciu dla fundacji zajmującej się pomocą w walce z koronawirusem. Wydaje się być to niewykorzystana szansa dla marek na zrobienie czegoś, co mogło mieć realny wpływ na konsumentów, a przez to na tworzenie wizerunku firmy społecznie odpowiedzialnej.

Co ciekawe, wiele spotów podczas pandemii COVID-19 wydaje się być odezwanych od codzienności. Podczas gdy Polaków dotykały najbardziej restrykcyjne obostrzenia, nie publikowano spotów, w których bohaterowie stosują się do panujących warunków. O ile w przypadku reklam perfum, gdzie klient dąży do ideału, jest to uzasadnione, o tyle w wypadku spotów środków czystości wydaje się istotne pokazywanie, jak pandemia odbija się na naszym codziennym życiu.

W przyszłości warto byłoby poszerzyć badanie o większą próbę badawczą i przeprowadzić badania o charakterze ilościowym. Około 30 spotów dla każdej kategorii jest tylko minimalną liczbą, gdzie pewne kierunki zmian są zasygnalizowane, ale trudno zweryfikować, czy mają one zastosowanie dla całych analizowanych kategorii produktowych. Ponadto istotnym aspektem, choć trudnym do weryfikacji, jest data planowania i realizacji kampanii reklamowej – być może dostrzegane nie-

dostosowanie do panującej sytuacji pandemicznej wzięło się z braku możliwości natychmiastowej reakcji nadawcy.

Bibliografia

- Bartmiński, J., Grzeszczak, M. (2014). Jak rekonstruować kanon wartości narodowych i europejskich?, *Etnolingwistyka*, 26, s. 21–44.
- Bühler Ch. (1979). Humanistic psychology as a personal experience, *Journal of Humanistic Psychology*, 19(1), s. 5–22.
- Dubis, M. (2014). Wartości i style życia młodzieży, *Jagiellońskie studia socjologiczne*, 1, s. 35–45.
- Feather, N. (2002). *Values, Achievement, and Justice. Studies in the Psychology of Deservingness*. Kluwer Academic Publishers.
- French, J.R.P., Kahn, R. L. (1962). A programmatic approach to studying the industrial environment and mental Health, *Journal of Social Issues*, 18(3), s. 1–47.
- Hechter M. (1993). Values Research in the Social and Behavioral Sciences. W: *The Origin of Values*, M. Hechter, L. Nadel, R.E. Michod (Eds.). Aldine Transaction.
- Hitlin, S., Piliavin, J. (2004). Values: Reviving a dormant concept, *Annual Review of Sociology*, 30, s. 359–393.
- Kluckhohn, C. (1951). Values and Value-Orientations in the Theory of Action: An Exploration in Definition and Classification. W: *Toward a General Theory of Action* (s. 388–433), T. Parsons, E. Shils (Eds.). Cambridge: Harvard University Press.
- Lewin, K. (1951). *Field Theory in Social Science: Selected Theoretical Papers*. New York: Harper & Row.
- Obacz, P. (2015). Psychologia humanistyczna: humanistyczne ideały i wartości w psychologii. W: *Idee i wartości w języku i kulturze* (s. 62–71), I. Matusiak-Kempa, A. Naruszewicz-Duchlińska (red.). Olsztyn: Instytut Filologii Polskiej Uniwersytetu Warmińsko-Mazurskiego.
- Puzynina, J. (2014). Kłopoty z nazwami wartości (i wartościami), *Ruch literacki. Dwumiesięcznik*, 1(322), s. 1–12.
- Raport CBOS: „Wartości w czasach zarazy”, 2020.
- Raporty CBOS: „Aktualne problemy i wydarzenia”, zrealizowane w latach 2005–2019.
- Rokeach, M. (1973). *The Nature of Human Values*. Free Press.
- Schwartz, S., Bilsky, M. (1987). Toward a universal psychological structure of human values, *Journal of Personality and Social Psychology*, 53, s. 550–562.
- Wartecka, A. (2015). Analiza aksjologiczna wybranych wypowiedzi chwalebnych w programie telewizyjnym typu talent show. W: *Idee i wartości w języku i kulturze* (s. 186–198), I. Matusiak-Kempa, A. Naruszewicz-Duchlińska (red.). Olsztyn: Instytut Filologii Polskiej Uniwersytetu Warmińsko-Mazurskiego.
- Williams, Jr., R.M. (1979). Change and Stability in Values and Value Systems: A Sociological Perspective. W: *Understanding Human Values* (s. 15–46), M. Rokeach (Ed.). The Free Press.

Where Are We Going to? Values in Advertising in the 21st century

Abstract

The COVID-19 pandemic has a significant impact on people's lives around the world. This unparalleled situation was reflected in the media as well as advertising. This study aims to analyse the impact of the COVID-19 pandemic on the content of advertisements in the area of presented values. The first part of the study reviews the literature focusing on the definition of value in different fields of research. The second part analyses reports made by CBOS (the Centre for Public Opinion Research) on the values most appreciated by Poles. The last part provides a qualitative content analysis of 112 commercials divided into 4 groups based on the FCB model (informative, affective, habitual, satisfaction).

Keywords: values, advertising, content analysis, COVID-19 pandemic