

Cena 13,00 zł
(VAT 8%)

Indeks 381306
e-ISSN 2543-8476
PL ISSN 0043-518X

WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

WRZESIEŃ / SEPTEMBER
ROK / VOLUME 67

2022 | 9

GŁÓWNY URZĄD STATYSTYCZNY
STATISTICS POLAND

POLSKIE TOWARZYSTWO STATYSTYCZNE
POLISH STATISTICAL ASSOCIATION



WIADOMOŚCI STATYSTYCZNE

THE POLISH STATISTICIAN

WRZESIEŃ / SEPTEMBER
ROK / VOLUME 67

2022 | 9 (736)

RADA NAUKOWA / SCIENTIFIC COUNCIL

dr Dominik Rozkrut – przewodniczący/chairman (Uniwersytet Szczeciński, Polska), Prof. Anthony Arundel (Maastricht University, Holandia), Eric Bartelsman, PhD, Assoc. Prof. (Vrije Universiteit Amsterdam, Holandia), prof. dr hab. Czesław Domański (Uniwersytet Łódzki, Polska), prof. dr hab. Elżbieta Gołata (Uniwersytet Ekonomiczny w Poznaniu, Polska), Semen Matkovskyy, PhD, Assoc. Prof. (Ivan Franko National University of Lviv, Ukraina), prof. dr hab. Włodzimierz Okrasa (Uniwersytet Kardynała Stefana Wyszyńskiego w Warszawie, Polska), prof. dr hab. Józef Oleński (Polskie Towarzystwo Statystyczne, Polska), prof. dr hab. Tomasz Panek (Szkola Główna Handlowa w Warszawie, Polska), Juan Manuel Rodríguez Poo, PhD, Assoc. Prof. (University of Cantabria, Hiszpania), Iveta Stankovičová, BEng, PhD, Assoc. Prof. (Comenius University in Bratislava, Słowacja), prof. dr hab. Marek Walesiak (Uniwersytet Ekonomiczny we Wrocławiu, Polska), prof. dr hab. Józef Zegar (Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy, Polska)

sekretarz/secretary: Paulina Kucharska-Singh, Główny Urząd Statystyczny, Polska

KOLEGIUM REDAKCYJNE / EDITORIAL BOARD

Tudorel Andrei, PhD, Assoc. Prof. (Bucharest Academy of Economic Studies, Rumunia), mgr Renata Biela (Główny Urząd Statystyczny, Polska), dr Marek Cierpień-Wolan (Uniwersytet Rzeszowski, Polska), dr hab. Grażyna Dehnel, prof. UEP (Uniwersytet Ekonomiczny w Poznaniu, Polska), dr Jacek Kowalewski (Uniwersytet Ekonomiczny w Poznaniu, Polska), dr Jan Kubacki (Polskie Towarzystwo Statystyczne, Polska), dr Grażyna Marciniak (Główny Urząd Statystyczny, Polska), dr hab. Andrzej Młodak, prof. Akademii Kaliskiej (Akademia Kaliska im. Prezydenta Stanisława Wojciechowskiego, Polska), dr hab. Mateusz Pipień, prof. UEK (Uniwersytet Ekonomiczny w Krakowie, Polska), Marek Rojčick, BEng, PhD (University of Economics, Prague, Czechy), Anna Shostya, PhD, Assoc. Prof. (Pace University in New York, Stany Zjednoczone), dr hab. Małgorzata Tarczyńska-Luniewska, prof. US (Uniwersytet Szczeciński, Polska), dr Wioletta Wrzaszcz (Instytut Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy, Polska), dr inż. Agnieszka Zgierska (Główny Urząd Statystyczny, Polska)

ZESPÓŁ REDAKCYJNY / EDITORIAL STAFF

redaktor naczelny / editor-in-chief: Marek Cierpień-Wolan

zastępca redaktora naczelnego / deputy editor-in-chief: Andrzej Młodak

redaktorzy tematyczni / thematic editors: Małgorzata Tarczyńska-Luniewska, Agnieszka Zgierska

redaktor merytoryczny / substantive editor: Wioletta Wrzaszcz

sekretarz/secretary: Małgorzata Zygmunt, Główny Urząd Statystyczny, Polska

ADRES REDAKCJI / EDITORIAL OFFICE ADDRESS

Główny Urząd Statystyczny / Statistics Poland, al. Niepodległości 208, 00-925 Warszawa
tel./phone +48 22 608 32 25, e-mail: redakcja.ws@stat.gov.pl

Redakcja językowa: Wydział Czasopism Naukowych, Główny Urząd Statystyczny

Language editing: Scientific Journals Division, Statistics Poland

Redakcja techniczna, skład i łamanie, opracowanie materiałów graficznych, korekta: Zakład Wydawnictw Statystycznych – zespół pod kierunkiem Wojciecha Szuchta

Technical editing, typesetting, preparation of graphic materials, proofreading: Statistical Publishing Establishment – team supervised by Wojciech Szuchta



Zakład Wydawnictw
Statystycznych

Druk i oprawa / Printed and bound by:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment
al. Niepodległości 208, 00-925 Warszawa, zws.stat.gov.pl

Wersja elektroniczna, stanowiąca wersję pierwotną czasopisma, jest dostępna na ws.stat.gov.pl
The primary version of the journal, issued in electronic form, is available at ws.stat.gov.pl

© Copyright by Główny Urząd Statystyczny and the authors, some rights reserved. CC BY-SA 4.0 licence



Indeks 381306

Informacje w sprawie sprzedaży czasopisma / Sales of the journal:

Zakład Wydawnictw Statystycznych / Statistical Publishing Establishment

tel./phone +48 22 608 32 10, +48 22 608 38 10

Prenumerata jest prowadzona przez / Subscription is available at RUCH S.A.

Zamówienia na prenumeratę można składać na stronie / Subscriptions can be ordered at www.prenumerata.ruch.com.pl

SPIS TREŚCI
CONTENTS

Od redakcji	IV
From the editorial team	
Zaproszenie do nadsyłania artykułów	VI
Call for papers	
Statystyka w praktyce	
Statistics in practice	
Mariusz Hamulczuk, Marcin Idzik	
Ocena aktywności ekonomicznej w polskim sektorze bankowym w okresie pandemii COVID-19 na podstawie testów koniunktury	1
Assessment of the economic activity in the banking sector during the COVID-19 pandemic on the basis of business tendency survey	
Aneta Ptak-Chmielewska, Karolina Baszniak, Jarosław Kurpanik	
Wpływ pandemii COVID-19 na stan zdrowia psychicznego społeczeństwa	24
Influence of the COVID-19 pandemic on the mental health of society	
Studia interdyscyplinarne. Wyzwania badawcze	
Interdisciplinary studies. Research challenges	
Magdalena Mleczek, Przemysław Slesiński, Tomasz Milewski, Artur Łączyński, Dariusz Miziołek, Edyta Woźniak, Jędrzej Bojanowski	
Satelitarne rozpoznawanie upraw i szacowanie ich powierzchni w ramach systemu SATMIROL	53
Satellite-based crop identification and area estimation under the SATMIROL system	
In memoriam	
Anatol Kula (1938–2022)	75
Dyskusje. Recenzje. Informacje	
Discussions. Reviews. Information	
Joanna Sadowy	
Wydawnictwa GUS. Sierpień 2022	77
Publications of Statistics Poland. August 2022	
Dla autorów	79
For the authors	
Działy „WS” – tematyka artykułów	90
WS sections – topics of the articles	

OD REDAKCJI

We wrześniowym numerze „Wiadomości Statystycznych. The Polish Statistician” proponujemy Państwu lekturę trzech artykułów naukowych. W dwóch z nich przedstawiono wyniki badań, które łączy zagadnienie pandemii COVID-19. Tematyka trzeciego wpisuje się w obszar wyzwań badawczych, z którymi mierzy się statystyka publiczna.

W otwierającej wydanie pracy *Ocena aktywności ekonomicznej w polskim sektorze bankowym w okresie pandemii COVID-19 na podstawie testów koniunktury* dr hab. Mariusz Hamulczuk i dr Marcin Idzik badają percepcję koniunktury gospodarczej w bankowości w czasie kryzysu wywołanego przez pandemię oraz uwarunkowania obserwowanych zmian. W swoim badaniu wykorzystują testy koniunktury prowadzone przez GUS oraz Kantar Polska w latach 2004–2021. Na podstawie składowej trendo-cyklicznej otrzymanej metodą X-13-ARIMA ustalają, że w czasie kryzysu wywołanego przez pandemię koniunktura w sektorze bankowym pogorszyła się bardziej niż podczas kryzysu finansowego z lat 2008–2009. Bankowość nie należy też do sektorów, które najszybciej zareagowały osłabieniem koniunktury w czasie pandemii. Wcześniej jej pogorszenie uwidoczniło się w budownictwie, przemyśle, transporcie oraz na rynku nieruchomości.

Artykuł *Wpływ pandemii COVID-19 na stan zdrowia psychicznego społeczeństwa* autorstwa dr hab. Anety Ptak-Chmielewskiej, prof. SGH, mgr Karoliny Baszniak i mgr. inż. Jarosława Kurpanika dotyczy z kolei oceny wpływu pandemii COVID-19 na kondycję psychiczną członków gospodarstw domowych. W badaniu posłużono się danymi pochodzącymi z ankiety *COVID Impact Survey*, którą organizacja Data Foundation przeprowadziła w 2020 r. w Stanach Zjednoczonych wśród osób dorosłych. Analizie poddano 6768 obserwacji. Wykorzystano metody i modele statystyczne oraz data mining, takie jak: regresja logistyczna, drzewa decyzyjne, wzmacnianie gradientowe, metoda k -najbliższych sąsiadów, sztuczne sieci neuronowe i metoda wektorów wspierających. Z badania wynika, że izolacja, zdalny tryb nauczania i pracy oraz mniejsza aktywność fizyczna przyczyniają się do pogarszania się stanu zdrowia psychicznego.

Dr inż. Magdalena Mleczeko, mgr Przemysław Slesiński, mgr Tomasz Milewski, mgr Artur Łączyński, mgr Dariusz Miziołek, dr Edyta Woźniak i dr Jędrzej Bojanowski w pracy *Satelitarne rozpoznawanie upraw i szacowanie ich powierzchni w ramach systemu SATMIROL* prezentują innowacyjny system satelitarnej identyfikacji i monitorowania upraw rolnych, który został opracowany przez Główny Urząd Statystyczny, Instytut Geodezji i Kartografii oraz Centrum Badań Kosmicznych Polskiej Akademii Nauk na potrzeby statystyki rolnictwa. System SATMIROL jest jednym z mechanizmów uzyskiwania strategicznych danych z zakresu rolnictwa, pozwalającym odbiorcom na aktualną ocenę sytuacji i bieżące wspieranie działań w ramach polityk rolnych, w tym rolnośrodowiskowych. Poza omówieniem wybranych modułów systemu autorzy przedstawiają także wyniki szacunku powierzchni upraw w skali kraju w trzech okresach pomiarowych w latach: 2019, 2020 i 2021.

W dziale In memoriam żegnamy Anatola Kulę, wieloletniego sekretarza „Wiadomości Statystycznych”, który zmarł w sierpniu tego roku.

Wydanie kończy się omówieniem nowości wydawniczych GUS przygotowanym przez Joannę Sadowy.

Zapraszamy do lektury.

FROM THE EDITORIAL TEAM

The September issue of *Wiadomości Statystyczne. The Polish Statistician* contains three scientific articles. Two of them present the results of research connected with the COVID-19 pandemic, while the third paper focuses on the challenges faced by official statistics.

The opening article *Assessment of the economic activity in the banking sector during the COVID-19 pandemic on the basis of business tendency surveys* by Mariusz Hamulczuk, PhD, DSc, and Marcin Idzik, PhD, examines the perception of the economic situation in the banking sector during the crisis caused by the pandemic, and the determinants of the observed changes. The authors use business tendency surveys conducted by Statistics Poland and Kantar Polska for the years 2004–2021. On the basis of the trend-cycle component obtained by means of the X-13-ARIMA method, they conclude that the economic situation in the banking sector deteriorated more during the crisis caused by the COVID-19 pandemic than during the 2008–2009 financial crisis. Another conclusion is that the banking sector was not the first to experience an economic downturn due to the pandemic, as it was preceded by a slowdown observed in the construction, industry, transport and the real estate sectors.

In the paper *Influence of the COVID-19 pandemic on the mental health of society* by Aneta Ptak-Chmielewska, PhD, DSc, professor at the SGH Warsaw School of Economics, Karolina Baszniak, MSc, and Jarosław Kurpanik, BEng, MSc, the authors assess the influence of the COVID-19 pandemic on the mental health of members of households. The study is based on data from the *COVID Impact Survey*, carried out among adult respondents by the Data Foundation think tank in 2020 in the US. The survey used 6,768 observations. The authors apply statistical and data mining methods and models, such as logistic regression, decision trees, Extreme Gradient Boosting, the *k*-nearest neighbours method, artificial neural networks and the support vector machine. The results indicate that isolation, remote education and work, and limited physical activity contribute to the worsening of the mental health of the population.

Magdalena Mleczko, BEng, PhD, Przemysław Slesiński, MSc, Tomasz Milewski, MSc, Artur Łączyński, MSc, Dariusz Miziołek, MSc, Edyta Woźniak, PhD, and Jędrzej Bojanowski, PhD, in their paper entitled *Satellite-based crop identification and area estimation under the SATMIROL system* present an innovative system of satellite-based crop identification and crop growth monitoring, created by Statistics Poland, the Institute of Geodesy and Cartography and the Space Research Centre of the Polish Academy of Sciences to meet the needs of agricultural statistics. The SATMIROL system is one of the mechanisms of obtaining strategic agricultural data which allows an updated assessment of the current situation and a continuous support for activities undertaken in the framework of agricultural policies, including agri-environmental policies. In addition to discussing selected modules of the system, the authors present a nationwide crop area estimation performed in three periods: 2019, 2020 and 2021.

In the *In Memoriam* section, we remember Anatol Kula, a long-term secretary of *Wiadomości Statystyczne. The Polish Statistician*, who passed away in August of this year.

A discussion of Statistics Poland's most recent publications, prepared by Joanna Sadowy, closes the issue.

We wish you pleasant reading.

Zaproszenie do nadsyłania artykułów do działu Spisy powszechnie – problemy i wyzwania

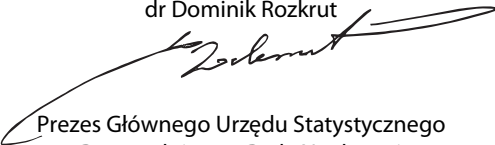
Szanowni Państwo,

z okazji przypadających na początek dekady 2020 spisów powszechnych w „Wiadomościach Statystycznych. The Polish Statistician” został utworzony dział Spisy powszechnie – problemy i wyzwania. Powstał on jako miejsce twórczej wymiany doświadczeń statystyków i badaczy podejmujących tę tematykę, jedną z najważniejszych w statystyce publicznej. W związku z tym mamy przyjemność zaprosić Państwa do nadsyłania artykułów naukowych, które będą przedstawiać propozycje efektywnych rozwiązań dotyczących spisów – zarówno organizacyjnych, jak i metodologicznych – a także zawierać analizę danych statystycznych uzyskanych w wyniku spisów, prowadzoną z zastosowaniem nowoczesnych narzędzi teoretycznych i informatycznych. Mile widziane są opracowania skupiające się na praktycznych aspektach związanych z przeprowadzaniem spisów, w tym dotyczących obciążenia odpowiedzi i ochrony tajemnicy statystycznej.

Ostatnie spisy powszechnie stały się szczególnym wyzwaniem ze względu na niespodziewane problemy organizacyjno-metodologiczne, w ogromnej mierze wynikające z wybuchu pandemii COVID-19 oraz rosnących niepokojów społecznych i gospodarczych obserwowanych na całym świecie. W rezultacie spisy nie tylko dostarczyły konkretnych danych statystycznych, lecz także ujawniły potrzebę zastosowania nadzwyczajnych środków organizacyjnych i technicznych podczas zbierania i przetwarzania danych, co powinno być uwzględnione w ocenie jakości wyników. Paradoksalnie zdarzenia losowe, które wymagają opracowania alternatywnych scenariuszy działań i przygotowania na ryzyko w dalszych badaniach statystycznych, mogą się przyczynić do postępu w ich realizacji. Wymienione wyżej zagadnienia wymagają kompleksowych analiz, które za pośrednictwem „Wiadomości Statystycznych. The Polish Statistician” mają szansę trafić do szerokiego grona odbiorców.


Zachęcamy do zapoznania się z informacjami na temat czasopisma i wskazówkami dla autorów, dostępnymi na stronie ws.stat.gov.pl. Artykuły w języku polskim lub angielskim należy przesyłać na adres: redakcja.ws@stat.gov.pl.

dr Dominik Rozkrut



Prezes Głównego Urzędu Statystycznego
Przewodniczący Rady Naukowej
„Wiadomości Statystycznych.
The Polish Statistician”

dr Marek Cierpiął-Wolan



Redaktor Naczelny
„Wiadomości Statystycznych.
The Polish Statistician”

Issues and challenges in census taking **– call for papers**

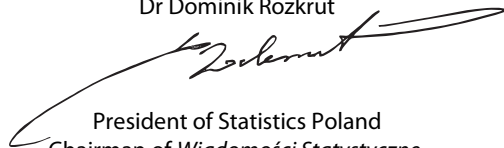
Dear Readers,

In response to the censuses carried out at the beginning of the current decade, a new section entitled *Issues and challenges in census taking* has been introduced to the *Wiadomości Statystyczne. The Polish Statistician* journal. The section is meant as a platform for the creative exchange of experiences among statisticians and researchers dealing with national censuses, which are one of the most important undertakings of official statistics. In this connection, we have the pleasure of inviting you to send in scientific papers proposing effective census-related solutions, both organisational and methodological, as well as works featuring analyses of data obtained through censuses, performed with the use of innovative theoretical and IT tools. We also welcome papers focusing on the practical aspects related to carrying out censuses, including response burden and the protection of statistical confidentiality.

Censuses have recently become especially challenging – this mainly being the effect of unforeseen organisational and methodological problems related to the outbreak of the COVID-19 pandemic and intensifying social and economic instabilities observed throughout the globe. The most recent censuses have therefore not only provided concrete statistical data, but also necessitated the application of extraordinary organisational and technical measures in collecting and processing data, which should be taken into account while assessing the quality of the results. Paradoxically, random occurrences which require devising alternative procedures and preparation for risk in the further research activity are likely to contribute to the progress of research. All the above-mentioned issues, as we can see, require comprehensive analyses, which now, with the help of *Wiadomości Statystyczne. The Polish Statistician*, stand a chance of reaching large audiences.

We would like to invite you to learn more about our journal and the guidelines for authors, which are available on our website: ws.stat.gov.pl. Articles written in Polish or English should be sent to: redakcja.ws@stat.gov.pl.

Dr Dominik Rozkrut



President of Statistics Poland
Chairman of *Wiadomości Statystyczne.*
The Polish Statistician Scientific Council

Dr Marek Cierpiat-Wolan



Editor-in-Chief
of *Wiadomości Statystyczne.*
The Polish Statistician

Ocena aktywnoŝci ekonomicznej w polskim sektorze bankowym w okresie pandemii COVID-19 na podstawie testów koniunktury

Mariusz Hamulczuk^a, Marcin Idzik^b

Streszczenie. Jednym z obszarów gospodarki, na kt3re pandemia COVID-19 wywarła wplyw, jest sektor bankowy. Celem badania omawianego w artykule jest uzyskanie odpowiedzi na pytania, jaka była percepcja koniunktury w sektorze bankowym w Polsce w czasie kryzysu spowodowanego przez pandemię COVID-19 w ŝwietle wyników testów koniunktury oraz jakie były uwarunkowania zmian koniunktury w tym okresie. Badanie przeprowadzono na podstawie testów koniunktury prowadzonych przez GUS oraz Kantar Polska w latach 2004–2021. Wnioski sformułowano na podstawie składowej trendo-cyklicznej otrzymanej metodą X-13-ARIMA. Analizy wykazały negatywny związek między siłą obostrzeń pandemicznych a koniunkturą w bankowoŝci, szczególnie w pierwszej fazie kryzysu. Testy koniunktury pokazują, że w czasie kryzysu wywołanego przez pandemię koniunktura w sektorze bankowym pogorszyła się bardziej niź podczas kryzysu finansowego z lat 2008–2009. Na rynku kredytowym można było zaobserwować zarówno doŝyć gwałtowny spadek, jak i bardzo szybkie wzrosy, co spowodowało, że w 2021 r. wskaźniki koniunktury rynku kredytowego powróciły do poziomu sprzed wybuchu pandemii. Pogorszenie koniunktury na rynku depozytowym w początkowej fazie kryzysu nie było tak silne jak na rynku kredytowym, niemniej jednak nie zaobserwowano jej znaczącej poprawy w 2021 r.

Słowa kluczowe: bankowoŝć, koniunktura, testy koniunktury, pandemia COVID-19

JEL: G21, E32

Assessment of the economic activity in the Polish banking sector during the COVID-19 pandemic on the basis of business tendency surveys

Abstract. One of the areas of the economy affected by the COVID-19 pandemic is the banking sector. The aim of the study presented in the article is to answer the questions: what was the perception of the economic situation in the banking sector in Poland during the crisis caused by the COVID-19 pandemic in the light of the results of business tendency surveys, and what determined its changes over time. The research was carried out on the basis of business tendency surveys conducted by Statistics Poland and Kantar Polska in the years 2004–2021.

^a Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, Instytut Ekonomii i Finansów / Warsaw University of Life Sciences, Institute of Economics and Finance. ORCID: <https://orcid.org/0000-0002-4956-8516>.
E-mail: mariusz_hamulczuk@sggw.edu.pl.

^b Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, Instytut Ekonomii i Finansów / Warsaw University of Life Sciences, Institute of Economics and Finance. ORCID: <https://orcid.org/0000-0002-7632-2941>. Autor korespondencyjny / Corresponding author, e-mail: marcin_idzik@sggw.edu.pl.

The conclusions were based on the trend-cycle component obtained by means of the X-13-ARIMA method. The analyses showed a negative relationship between the extent of the pandemic restrictions and the economic activity in banking, especially in the first phase of the crisis. Business tendency surveys show that during this crisis, the economic situation in the banking sector deteriorated more than during the financial crisis of 2008–2009. A relatively sharp decline as well as rapid rebounds were observed in the credit market, which caused its indicators to return in 2021 to the levels recorded before the outbreak of the pandemic. In the initial phase of the crisis, the economic downturn observed in the deposit market was not as sharp as in the credit market. On the other hand, however, the economic situation in the deposit market failed to improve significantly in 2021.

Keywords: banking, business cycle, business tendency surveys, COVID-19 pandemic

1. Wprowadzenie

Pandemia spowodowana rozprzestrzenianiem się wirusa SARS-CoV-2 wpłynęła w zasadzie na wszystkie sfery aktywności społeczno-gospodarczej. Obostrzenia gospodarcze, restrykcje, ograniczenia mobilności ludzi i towarów wprowadzane w odpowiedzi na wybuch pandemii doprowadziły m.in. do spadku produkcji, obniżenia wymiany handlowej i zaburzeń w konsumpcji towarów i usług. Dotyczyło to wielu sektorów i miało charakter globalny (Hossain, 2021; Hurley i Walker, 2021; Weersink i in., 2021). Okres ten określa się mianem kryzysu COVID-19. Wirus SARS-CoV-2 okazał się zagrożeniem dla gospodarki realnej, a także dla rynków finansowych. Pandemia ujawniła liczne powiązania – zarówno przyczynowo-skutkowe, jak i symptomatyczne – w różnych obszarach życia gospodarczego i społecznego w wymiarze krajowym oraz międzynarodowym.

Efektom pierwszej fazy pandemii był bezprecedensowy spadek aktywności gospodarczej. Według szacunków International Monetary Fund (2021) w 2020 r. w relacji rocznej globalny PKB obniżył się o 3,3%, a obroty handlowe – o 8,5%. Według Organisation for Economic Co-Operation and Development (OECD, 2020) globalna recesja była najgłębsza od czasu wielkiego kryzysu dwudziestolecia międzywojennego, z pewnymi różnicami między krajami w zależności od działań podjętych w celu wspierania narodowych gospodarek. Można zatem uznać, że globalny kryzys COVID-19, który wybuchł na początku 2020 r., spowodował gwałtowne osłabienie koniunktury gospodarczej na świecie (Dayong i in., 2020; Związek Przedsiębiorców i Pracodawców, 2020). Według dziennika „Rzeczpospolita” lockdown ogłoszony w Polsce w marcu 2020 r. kosztował ponad 25 mld zł (*Koronawirus. Kolejny lockdown...*, 2021). Z kolei zgodnie z danymi Federacji Przedsiębiorców Polskich w pierwszym roku pandemii, w efekcie związanych z nią ograniczeń, polska gospodarka straciła 185 mld zł (Związek Przedsiębiorstw Finansowych w Polsce, 2021).

Wprowadzenie obostrzeń pandemicznych w Polsce przyniosło gwałtowne załamanie aktywności gospodarczej (Rokicki, 2020). Negatywne skutki pandemii ujawni-

niły się zarówno w ilościowych indeksach charakteryzujących aktywność gospodarczą, jak i w statystykach jakościowych opisujących zmiany aktywności przedsiębiorstw czy też zmiany sytuacji ekonomicznej gospodarstw domowych (Główny Urząd Statystyczny [GUS], 2021b; Męcina i Potocki, 2020; Związek Przedsiębiorstw Finansowych w Polsce, 2021). Według barometru koniunktury GUS (2021a) bieżący wskaźnik ufności konsumenckiej w pierwszej fazie pandemii spadł do poziomu dotąd nienotowanego – do $-28,7$ w II kwartale 2020 r., a światowy indeks niepewności pandemicznej kształtował się w tym czasie na najwyższym poziomie w historii.

Obostrzenia pandemiczne wywołały szok popytowy w obszarze inwestycji, jak również w konsumpcji prywatnej. Z badań przeprowadzonych w kwietniu 2020 r. przez Instytut Rozwoju Gospodarczego Szkoły Głównej Handlowej w Warszawie (IRG SGH) wynika, że zaledwie 9% polskich przedsiębiorstw budowlanych, handlowych i przemysłu przetwórczego nie odczuło negatywnych skutków restrykcji, natomiast co czwarte uznało je za dotkliwe (Adamowicz i in., 2020). Z tych samych badań wynika, że w reakcji na obostrzenia w prowadzeniu działalności gospodarczej i wynikający z nich spadek przychodów 52% przedsiębiorstw dokonało redukcji kosztów pracowniczych, 50% skróciło czas pracy, w 27% firm obniżono wynagrodzenie pracowników, a co piąte przedsiębiorstwo ograniczyło zatrudnienie.

W 2021 r. w gospodarce światowej nastąpiło wyraźne ożywienie wynikające z realizacji odłożonego popytu. Niemniej jednak narastała presja inflacyjna, związana m.in. z wysokimi cenami energii i surowców, obniżeniem stóp procentowych przez banki centralne oraz uruchomieniem różnych form wsparcia dla przedsiębiorstw (Miniszewski, 2021). W efekcie mieliśmy do czynienia z dosyć znacznymi wahaniem nastrojów gospodarczych, a co za tym idzie – wahaniem koniunktury gospodarczej.

Sektor finansowy również doświadczył skutków utrzymującej się niepewności związanej z kryzysem (European Banking Authority, 2020; S&P Global Ratings, 2020), czego konsekwencją był spadek aktywności klientów indywidualnych oraz przedsiębiorców na rynku bankowym, szczególnie w pierwszej fazie pandemii. Potwierdza to łączny zysk sektora bankowego w Polsce na koniec września 2020 r., który wyniósł 5,9 mld zł i był niższy o 49,2% od zysku w analogicznym okresie poprzedniego roku, kiedy wyniósł 11,6 mld zł (Rudke, 2020). W rezultacie nastąpiła zmiana struktury źródeł zysków banków. Związek Banków Polskich (2021) poinformował, że w 2021 r. jako główne źródło zysku 40% banków w Polsce wskazało redukcję kosztów (spadek o 32 p.proc. rok do roku), 13% – wzrost przychodów odsetkowych (wzrost o 9 p.proc. rok do roku), a 47% – wzrost przychodów z opłat i prowizji (wzrost o 26 p.proc. rok do roku). W przypadku 53% banków za zmianę poziomu przychodów odpowiadało zwiększenie sprzedaży produktów dotychczasowym klientom (wzrost o 19 p.proc. rok do roku), 9% – zwiększenie liczby nowo

pozyskanych klientów (wzrost o 6 p.proc. rok do roku), a 38% – wzrost cen usług bankowych (spadek o 12 p.proc. rok do roku).

Pandemia miała też duży wpływ na aktywność klientów na rynku kredytowym. W sierpniu 2020 r. banki w Polsce udzieliły o 30% mniej kredytów gotówkowych i o 22% mniej kredytów mieszkaniowych niż rok wcześniej. Wyjątek stanowiły kredyty ratalne, których udzielono o 10% więcej w relacji rok do roku. W przypadku kredytów hipotecznych nastąpił wzrost o 5,3% rok do roku (Narodowy Bank Polski, 2021). Kryzys gospodarczy spowodował zwiększenie udziału kredytów zagrożonych na świecie. Polityka przyjęta przez rządy i banki centralne niewątpliwie pomogła złagodzić wyzwania stojące przed bankami w wyniku kryzysu COVID-19 (OECD, 2021). Jak wskazują Korzeb i Niedziółka (2020), duże banki były odporniejsze niż małe na zagrożenia pojawiające się podczas pandemii.

Pandemia nie tylko doprowadziła do pogorszenia sytuacji w zakresie zjawisk mikro- i makroekonomicznych, lecz także wpłynęła na nastroje konsumentów i przedsiębiorców oraz na aktywność klientów na rynku bankowym (Cieślik i in., 2021; Kluza i Krzesicki, 2021). Pomiar wpływu pandemii na aktywność gospodarczą w tym kontekście umożliwia analiza zmian wartości wskaźników koniunktury uzyskiwanych metodą testu koniunktury. W Polsce badanie koniunktury gospodarczej, z uwzględnieniem zakresu podmiotowego i przedmiotowego obowiązującego we Wspólnym Zharmonizowanym Programie Unii Europejskiej Badań Koniunktury Gospodarczej i Konsumentów, realizuje GUS (2011). Od 2007 r. badanie to obejmuje comiesięczny monitoring w sekcji działalność finansowa i ubezpieczeniowa. Badania koniunktury w bankowości są prowadzone również przez IRG SGH, Kantar Polska¹ – na zlecenie Związku Banków Polskich – oraz zespół badawczy Uniwersyteku Ekonomicznego w Poznaniu. Badania te nie mają jednak zunifikowanej formuły.

Test koniunktury jest specyficznym badaniem opinii. Pozwala na zbieranie informacji od podmiotów gospodarczych na temat sytuacji w ich własnych organizacjach i ich otoczeniu (Bieć, 1996). W odróżnieniu od statystyki ilościowej badania koniunktury gospodarczej metodą testu koniunktury są oparte na subiektywnych opiniach ankietowanych dyrektorów i kierowników przedsiębiorstw, którzy dokonują oceny bieżącej sytuacji gospodarczej i jej prawdopodobnych zmian. Tym samym metoda testu koniunktury jest sposobem pomiaru koniunktury obejmującej subiektywną ocenę zarówno bieżących zmian kategorii ekonomicznych, jak i ogólnego klimatu koniunktury, rozumianego jako bieżące warunki prowadzenia działalności

¹ Agencja badawcza powstała z wyniku połączenia agencji TNS Polska i Millward Brown. Od 1993 r. prowadzi badania koniunktury skoncentrowane tylko na działalności bankowej. Wyniki badań Kantar Polska są publikowane w comiesięcznych raportach Monitor Bankowy na stronie Związku Banków Polskich: <https://www.zbp.pl/raporty-i-publikacje/raporty-cykliczne/PENGAB>.

gospodarczej (Skikiewicz, 2020). Wyniki uzyskane w ten sposób niekoniecznie znajdują pełne odzwierciedlenie w danych ze sprawozdawczości gospodarczej.

Opracowania naukowe na temat koniunktury w bankowości w Polsce czy na świecie w czasie kryzysu COVID-19 są nieliczne. Prezentacja wyników badań koniunktury w bankowości ogranicza się do cyklicznych raportów zawierających zestawienia tworzone z wyników pomiarów miesięcznych lub kwartalnych (np. Kluza i Krzesicki, 2021; Związek Banków Polskich, 2021). Brakuje pogłębionych analiz, które uwzględniałyby porównania w czasie czy między sektorami. W tym kontekście badanie omawiane w niniejszym artykule zapełnia istniejącą lukę badawczą. Jego celem jest uzyskanie odpowiedzi na pytania, jaka była percepcja koniunktury w sektorze bankowym w Polsce w czasie kryzysu spowodowanego pandemią COVID-19 w świetle wyników testów koniunktury oraz jakie były uwarunkowania zmian koniunktury w tym okresie.

2. Metoda badania

W omawianym badaniu przeanalizowano zmiany koniunktury w bankowości na tle innych sektorów, porównano reakcję przedsiębiorców podczas kryzysu COVID-19 i kryzysu finansowego z lat 2008–2009 oraz oceniono wpływ kryzysu COVID-19 na poszczególne segmenty rynku bankowego w Polsce. Badanie oparto na danych na temat koniunktury gospodarczej pochodzących z badania koniunktury gospodarczej prowadzonego przez GUS oraz z badania koniunktury w bankowości prowadzonego przez Kantar Polska. Były to szeregi czasowe o częstotliwości miesięcznej z okresu od stycznia 2004 r. do października 2021 r. Poddano je dekompozycji metodą X-13-ARIMA.

Aby zrealizować cel badawczy, wykonano kolejno następujące czynności:

- scharakteryzowano podłoże i przebieg kryzysu COVID-19;
- dokonano ogólnej oceny koniunktury gospodarczej w poszczególnych sektorach polskiej gospodarki (według PKD 2007) w latach 2004–2021 z wykorzystaniem testów koniunktury gospodarczej GUS;
- porównano wyniki testów koniunktury Kantar Polska za lata 2004–2021, szacowanej na podstawie syntetycznego miernika Pengab, z wynikami testów koniunktury gospodarczej GUS dotyczącymi działalności finansowej i ubezpieczeniowej w Polsce;
- dokonano szczegółowej analizy zmian koniunktury w bankowości (na podstawie wartości wskaźnika Pengab) na tle koniunktury w poszczególnych sektorach gospodarki polskiej (na podstawie danych GUS) w latach 2019–2021;

- na podstawie wyników testów koniunktury Kantar Polska oceniono wpływ kryzysu COVID-19 na poszczególne segmenty sektora bankowego w Polsce.

Badania koniunktury metodą testu koniunktury są prowadzone przy użyciu ankiet zawierających pytania o podstawowe wielkości ekonomiczne, istotne dla oceny sytuacji przedsiębiorstw. Są to pytania diagnostyczne, dotyczące bieżącego miesiąca lub ostatnich trzech miesięcy, oraz prognostyczne, odnoszące się do najbliższych trzech miesięcy. Zarówno w przypadku badania GUS, jak i badania Kantar Polska zakres zadawanych pytań jest bardzo szeroki i dotyczy różnych aspektów działalności przedsiębiorstw. Pytania te są podstawą do stworzenia cząstkowych wskaźników diagnostycznych i prognostycznych. Wskaźniki cząstkowe są później agregowane (uśredniane) do wskaźników obrazujących ogólny klimat koniunktury gospodarczej w poszczególnych sektorach oraz w całej gospodarce (w przypadku badania Kantar Polska dotyczy to sektora bankowego i jego segmentów). Metody badania przyjęte przez GUS i Kantar Polska są tożsame pod względem sposobu formułowania pytań i obliczania indeksów.

Oceny i prognozy formułowano poprzez stwierdzenie, że wystąpił (wystąpi) wzrost lub spadek albo że sytuacja się nie zmieni. Różnice między odsetkiem odpowiedzi pozytywnych (wzrost, poprawa) i negatywnych (spadek, pogorszenie) stanowią cząstkowe wskaźniki ocen i prognoz, które służą do prezentacji tendencji (Golec, 2014; Idzik, 2015). Wskaźniki te przyjmują wartości z przedziału od -100 do $+100$ p.proc. Dodatnia wartość wskaźnika oznacza, że przeważają opinie pozytywne, a ujemna świadczy o przewadze opinii negatywnych. Wzrost wartości wskaźnika oznacza poprawę koniunktury, jego spadek zaś – pogorszenie koniunktury z punktu widzenia badanych przedsiębiorstw. Jak wskazuje większość ekonomistów zajmujących się badaniami na podstawie testów koniunktury, o koniunkturze gospodarczej nie można jednoznacznie powiedzieć, że jest dobra lub zła, a jedynie, że jest lepsza lub gorsza niż w okresie odniesienia. Stanowi to istotę testu koniunktury i wynika ze specyfiki funkcjonowania systemu gospodarczego (Bieć, 1996). Dodatkowo w przypadku testów koniunktury regułą są systematyczne różnice przeciętnych wartości wskaźników obserwowane w długim okresie w poszczególnych kategoriach ekonomicznych czy sektorach.

Metodologia badania koniunktury gospodarczej prowadzonego przez GUS metodą testu koniunktury jest szczegółowo opisana w zeszycie metodologicznym (GUS, 2018). W badaniu prowadzonym przez Kantar Polska uczestniczy 200 placówek bankowych z całej Polski. Dyrektorzy i kierownicy placówek przedstawiają oceny i prognozy dotyczące: zmian stanu depozytów ogółem i w podziale na gospodarstwa domowe oraz podmioty gospodarcze (z wyodrębnieniem depozytów złotych i walutowych i uwzględnieniem rachunków bieżących i terminowych), zmian stanu kredytów złotych ogólnie i w podziale na osoby prywatne i podmioty gospodarcze

oraz ogólnie kredytów walutowych, oprocentowania kredytów, relacji depozytów walutowych do złotych oraz kredytów dewizowych do złotych, a także ogólnej sytuacji ekonomicznej placówek. Respondenci formułują swoje oceny i prognozy poprzez stwierdzenie, że wystąpił (wystąpi) wzrost lub spadek albo że sytuacja się nie zmieni. Syntetycznym wskaźnikiem koniunktury bankowej jest Pengab (Golec, 2014; Idzik, 2015).

W zestawieniu 1 zamieszczono podstawowe informacje na temat wskaźników koniunktury: ośmiu pochodzących z badania GUS oraz ośmiu z badania koniunktury w bankowości Kantar Polska. Autorzy, korzystając z tych dwóch źródeł, mogli zachować tożsame podejście metodyczne w przypadku oceny koniunktury w bankowości, a także koniunktury konsumenckiej oraz koniunktury w przemyśle i usługach. Warto jednak dodać, że istnieje wiele alternatywnych wskaźników bazujących na testach koniunktury, opracowywanych przez różne instytucje i obrazujących ocenę koniunktury gospodarczej w różnych obszarach gospodarki. Jednym z najpopularniejszych jest PMI (Purchasing Managers' Index).

W omawianym badaniu wykorzystano jedynie wskaźniki diagnostyczne. Przebieg kryzysu spowodowanego przez pandemię COVID-19 oraz siłę stosowanych w tym czasie obostrzeń zobrazowano przy wykorzystaniu wartości wskaźnika siły restrykcji (ang. *stringency index*; Hale i in., 2021).

Zestawienie 1. Wskaźniki koniunktury wykorzystywane w badaniach GUS i Kantar Polska

Pełna nazwa	Skrócona nazwa
GUS	
Bieżący wskaźnik ufności konsumenckiej	BWUK
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: przetwórstwo przemysłowe	Przemysł
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: budownictwo	Budownictwo
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: handel; naprawa pojazdów samochodowych	Handel
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: transport i gospodarka magazynowa	Transport
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: zakwaterowanie i gastronomia	Zakwaterowanie i gastronomia
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: działalność finansowa i ubezpieczeniowa	Finanse
Wskaźnik ogólnego klimatu koniunktury gospodarczej w: obsługa rynku nieruchomości	Nieruchomości

Zestawienie 1. Wskaźniki koniunktury wykorzystywane w badaniach GUS i Kantar Polska (dok.)

Pełna nazwa	Skrócona nazwa
Kantar Polska	
Pengab ogółem	Pengab
Pengab rynku kredytowego ogółem	Pengab_K
Pengab rynku depozytowego ogółem	Pengab_D
Wskaźnik rynku kredytowego gospodarstw domowych	WRK_GD
Wskaźnik rynku kredytowego podmiotów gospodarczych	WRK_PG
Wskaźnik rynku depozytowego gospodarstw domowych	WRD_GD
Wskaźnik rynku depozytowego podmiotów gospodarczych	WRD_PG
Wskaźnik oceny sytuacji ekonomicznej banku	WOSEB

Źródło: opracowanie własne.

Szeregi czasowe otrzymane z danych z badania koniunktury gospodarczej z reguły ulegają dalszym przekształceniom, zanim dojdzie do sformułowania wniosków na ich podstawie. Punktem wyjścia do tych działań może być przyjęcie, że szereg wyjściowy (Y_t) składa się z długookresowego trendu-cyklu (TC_t), wahań sezonowych (S_t) oraz wahań losowych, przypadkowych (I_t). Z uwagi na to, że wahania sezonowe i przypadkowe zaburzają ocenę koniunktury gospodarczej, najczęściej uwzględnia się tylko składnik trendo-cykliczny lub wyłącznie cykliczny. W omawianym badaniu do oceny koniunktury i formułowania wniosków autorzy posłużyli się składową trendo-cykliczną.

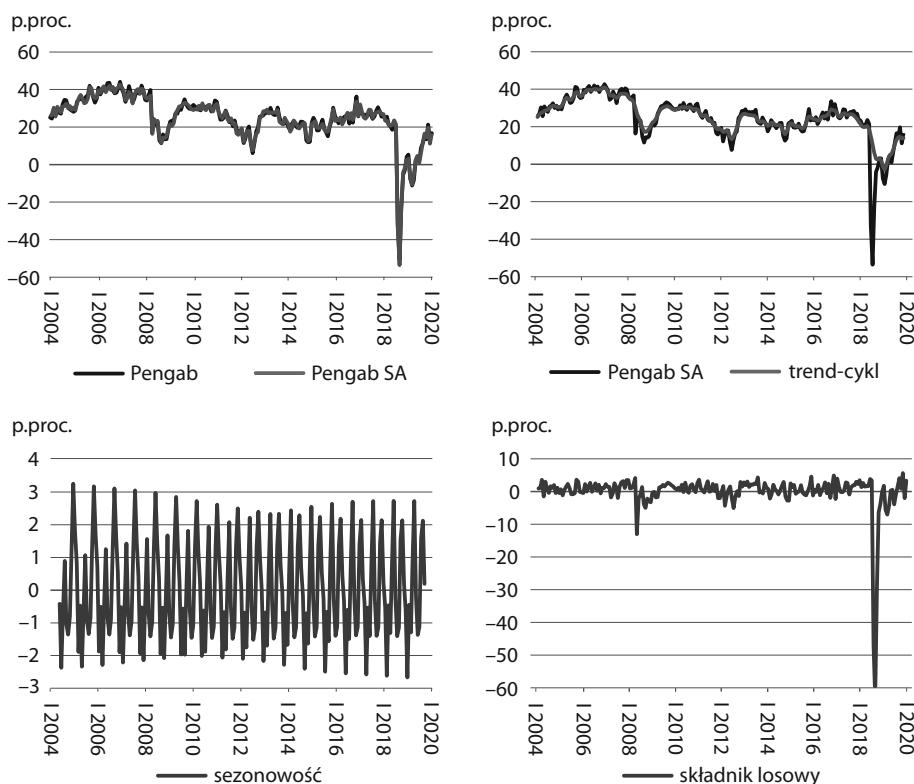
W celu otrzymania TC_t przeprowadzono dekompozycję szeregów czasowych za pomocą metody X-13-ARIMA. Dokonano tego w dwóch fazach. Pierwszą było wstępne modelowanie szeregu czasowego (regARIMA), drugą – właściwa dekompozycja szeregu czasowego, obliczenie prognoz i ocena jakości przyjętego modelu (szereż na ten temat: Center for Statistical Reseach and Methodology, 2020). Należy podkreślić, że w pierwszej fazie – i w modelu regARIMA – zrezygnowano z permanentnych zmian skokowych rozłożonych w czasie, które powodowały zaburzenia estymacji składnika długookresowego w późniejszych krokach.

W efekcie eliminacji wahań sezonowych otrzymano szeregi skorygowane sezonowo (dane SA), zawierające trend-cykl oraz wahania przypadkowe. Następnie, wykorzystując zawartą w metodzie X-13-ARIMA procedurę wygładzania X-11 oraz filtr Hendersona, dokonano rozdzielenia wyżej wymienionych komponentów. Finalny komponent trend-cykl wyliczono za pomocą algorytmu X-11 z zastosowaniem filtru Hendersona na danych wyrównanych sezonowo, którego rząd zależy od właściwości analizowanych danych.

Przykładowe wyniki obliczeń pokazano na wyk. 1. Pokazuje on oryginalny szereg czasowy wskaźnika Pengab, oszacowane wahania sezonowe, szereg czasowy skory-

gowany z wahań sezonowych (Pengab_SA), wygładzone wielkości składnika trendo-cyklicznego (trend-cykl) oraz wahania przypadkowe (wraz z obserwacjami nietypowymi). Należy tu zwrócić uwagę na zmienny w czasie charakter wahań sezonowych. Zauważalne jest również, że za gwałtowne i krótkotrwałe zmiany wartości wskaźników koniunktury gospodarczej odpowiada składnik przypadkowy; widać to zwłaszcza w 2020 r. Zachwianie rynku w czasie kryzysu COVID-19 doprowadziło do bezprecedensowego obniżenia wartości Pengab z poziomu 20,2 p.proc. w lutym 2020 r. do -50,85 p.proc. w kwietniu 2020 r. Miało ono charakter paniki trwającej 2–3 miesiące (marzec–maj 2020 r.). Z uwagi na krótkotrwałość tej zmiany została ona potraktowana w ramach procedury jako element przypadkowy, nietypowy. Niemniej jednak incydent ten miał swoje konsekwencje w postaci znaczącego spadku ogólnego klimatu koniunktury w późniejszych okresach.

Wykr. 1. Dekompozycja szeregu czasowego wskaźnika koniunktury w bankowości Pengab



Źródło: opracowanie własne na podstawie danych z Monitora Bankowego.

Analiza otrzymanych wyników, czyli składnika trendo-cyklicznego, miała w głównej mierze charakter jakościowy. Za pomocą metod graficznych porównano przebieg koniunktury w poszczególnych sektorach gospodarki oraz segmentach rynku bankowego. Skalę zmian percepcji koniunktury w czasie kryzysu COVID-19 oceniano w punktach procentowych.

3. Wyniki badania

3.1. Przebieg kryzysu

W grudniu 2019 r. w chińskim mieście Wuhan zdiagnozowano chorobę wywołaną nowym koronawirusem, która szybko zaczęła rozprzestrzeniać się na inne regiony Chin i inne części świata. Kalendarium rozwoju pandemii w Polsce przedstawiono w zestawieniu 2, a na wykř. 2 zestawiono Pengab ze wskaźnikiem odzwierciedlającym siłę obostrzeń towarzyszących pandemii w Polsce. Kluczowy dla gospodarki i sektora bankowego wydaje się okres od marca do czerwca 2020 r. 4 marca został odnotowany pierwszy przypadek zakażenia wirusem SARS-CoV-2 w Polsce. W tym okresie (tj. w marcu 2020 r.) wartość Pengab wynosiła 23,1 p.proc., a koniunktura w sektorze bankowym była oceniana jako dobra. Pomiar wykonany w kwietniu ujawnił spadek o 73,9 p.proc. w relacji do poprzedniego miesiąca, a Pengab osiągnęła najniższą w historii pomiarów (czyli od lutego 1993 r.) wartość –50,8 p.proc.

Zestawienie 2. Kalendarium rozwoju pandemii COVID-19

Miesiąc	Wydarzenia
2020	
Marzec	4.03 odnotowano pierwszy przypadek zakażenia wirusem SARS-CoV-2 w Polsce. 11.03 ogłoszono zamknięcie uczelni, szkół, przedszkoli, żłobków i ośrodków wychowawczych. 14.03 zamknięto granice Polski dla ruchu lotniczego oraz kolejowego, przywrócono kontrole paszportowe na granicach lądowych i wprowadzono zakaz wjazdu dla obcokrajowców. 27.03 we Włoszech odnotowano 919 przypadków śmiertelnych – największą dotychczas dzienną liczbę zmarłych na COVID-19 w jednym państwie. 31.03 ogłoszono dalsze restrykcje.
Kwiecień	8.04 rząd ogłosił nowy pakiet pomocowy dla polskich firm. 9.04 ogłoszono przedłużenie obowiązywania obostrzeń: zamknięcia punktów usługowych i ograniczeń w poruszaniu; zamknięcia szkół i ruchu lotniczego oraz kolejowego; zamknięcia granic. 15.04 ogłoszono wzmocnienie Funduszu Dróg Samorządowych w celu pobudzenia gospodarki. 16.04 był pierwszym dniem obowiązku noszenia maseczek. 29.04 premier RP ogłosił drugi etap odmrożenia gospodarki.
Maj	4.05 wszedł w życie drugi etap luzowania obostrzeń. 11.05 przedstawiono nowy rządowy program wsparcia turystyki krajowej (bon „1000 plus”). 13.05 ogłoszono trzeci etap odmrażania gospodarki. 16.05 przedsiębiorcy, niezadowoleni z działań rządu, protestowali w Warszawie. 21.05 GUS opublikował dane o produkcji przemysłowej w Polsce: w kwietniu spadek rok do roku wyniósł 24,6%. 27.05 ogłoszono zniesienie obowiązku noszenia maseczek i limitu osób w sklepach i kościołach od 30.05 oraz otwarcie klubów fitness, basenów, siłowni, sal zabaw i parków rozrywki, kin i teatrów, a także możliwość organizacji targów, wystaw i wesel do 150 osób od 6.06.

Zestawienie 2. Kalendarium rozwoju pandemii COVID-19 (dok.)

Miesiąc	Wydarzenia
2020 (dok.)	
Czerwiec	2.06 sejm przegłosował, a prezydent RP podpisał ustawę o wyborach prezydenckich w 2020 r. 10.06 rząd ogłosił, że w nocy z 12 na 13.06 Polska przywraca pełny ruch graniczny w ramach granic wewnętrznych UE.
Lipiec	2.07 Ministerstwo Zdrowia zmieniło zasady odbywania kwarantanny, znacznie ją skracając, aby ułatwić powrót potencjalnie chorych do codziennego funkcjonowania. 15.07 sejm przegłosował bon turystyczny. 21.07 rząd złagodził wiele obostrzeń dotyczących dystansu społecznego.
Wrzesień	10.09 minister zdrowia opublikował nową listę obszarów kraju z dodatkowymi obostrzeniami.
Październik	17.10 152 powiaty, czyli ok. połowy terytorium Polski, zostały objęte czerwoną strefą, w której zakazano organizacji imprez. 22.09 ogłoszono rozszerzenie czerwonej strefy na cały kraj od 24.10.
Listopad	17.11 ogłoszono kolejny program wsparcia przedsiębiorców działających w jednej z branż objętych obostrzeniami: program „2000 plus”. 21.11 premier zapowiedział otwarcie od 28.11 sklepów i punktów usługowych w ścisłym reżimie sanitarnym.
Grudzień	1.12 zwiększono wysokość mandatu za brak maseczki z 500 do 1000 zł. 19.12 rozpoczęto przyjmowanie wniosków o pomoc przedsiębiorcom w ramach „Tarczy 6.0”.
2021	
Styczeń	15.01 uruchomiono rejestrację na dobrowolne szczepienie przeciwko COVID-19 przez serwis internetowy, specjalną infolinię oraz w placówkach stacjonarnych.
Luty	5.02 premier ogłosił, że od 12.02 zostaną otwarte hotele, kina i teatry.
Czerwiec	7.06 ogłoszono, że od 16.06 zostaną poluzowane kolejne obostrzenia.
Wrzesień	22.09 rozpoczął się proces wystawiania skierowań do e-rejestracji na szczepienia dawką przypominającą dla wszystkich osób, które ukończyły 50. rok życia oraz dla pracowników ochrony zdrowia mających bezpośredni kontakt z pacjentami.
Październik	Gwałtownie wzrosła liczba zakażeń SARS-CoV-2. Zapowiedziano, że ewentualne obostrzenia będą wprowadzane w powiatach, w których będzie najwięcej nowych zakażeń i hospitalizacji.

Źródło: opracowanie własne na podstawie: Kantar (b.r.).

Aby zobrazować przebieg kryzysu COVID-19 w polskim sektorze bankowym, na wyk. 2 przedstawiono surowe wartości Pengab od stycznia 2020 r. do października 2021 r. na tle siły obowiązujących w tym czasie restrykcji związanych z pandemią. Wskaźnik siły restrykcji mierzy skalę obostrzeń w danym kraju i zawiera się w przedziale od 0 (zupełny brak obostrzeń) do 100 (pełne obostrzenia). Powstał on w ramach projektu Oxford Coronavirus Government Response Tracker (OxCGRT) i jest publikowany w ujęciu dziennym (Hale i in., 2021). Wartość tego wskaźnika w danym dniu jest obliczana jako średni wynik dziewięciu wskaźników, z których każdy przyjmuje wartość od 0 do 100. Dziewięć zmiennych używanych do obliczenia wskaźnika siły restrykcji to: zamknięcie szkół, zamknięcie miejsc pracy, odwoływanie imprez publicznych, ograniczenie zgromadzeń publicznych, ograniczenia dotyczące transportu publicznego, wymagania związane z pobytem w domu, publiczne kampanie informacyjne, ograniczenia dotyczące ruchów wewnętrznych oraz kontro-

le podróży międzynarodowych. Wskaźnik siły restrykcji odzwierciedla surowość polityki państw, która – jak wynika z wielu badań i obserwacji – przyczynia się do znacznych wahań aktywności gospodarczej.

Wykr. 2. Wartości Pengab na tle średniomiesięcznych wartości wskaźnika siły restrykcji w Polsce



Źródło: opracowanie własne na podstawie Hale'a i in. (2021) oraz danych z Monitora Bankowego.

Jak można zauważyć na wyk. 2, istnieje ujemny związek między siłą wprowadzanych obostrzeń a oceną koniunktury w sektorze bankowym dokonywaną przez kierowników placówek bankowych. Największym obostrzeniom – w marcu 2020 r. (wartość wskaźnika siły restrykcji 84,2 p.proc.) – towarzyszyła najniższa wartość Pengab (-50,8 p.proc.). Wraz z luzowaniem obostrzeń do września 2020 r. poprawiała się koniunktura. Reakcja na wzrost restrykcji obserwowana do końca 2020 r. nie była już tak gwałtowna. Z czasem również ujemny związek między siłą restrykcji a poziomem Pengab przestał być widoczny.

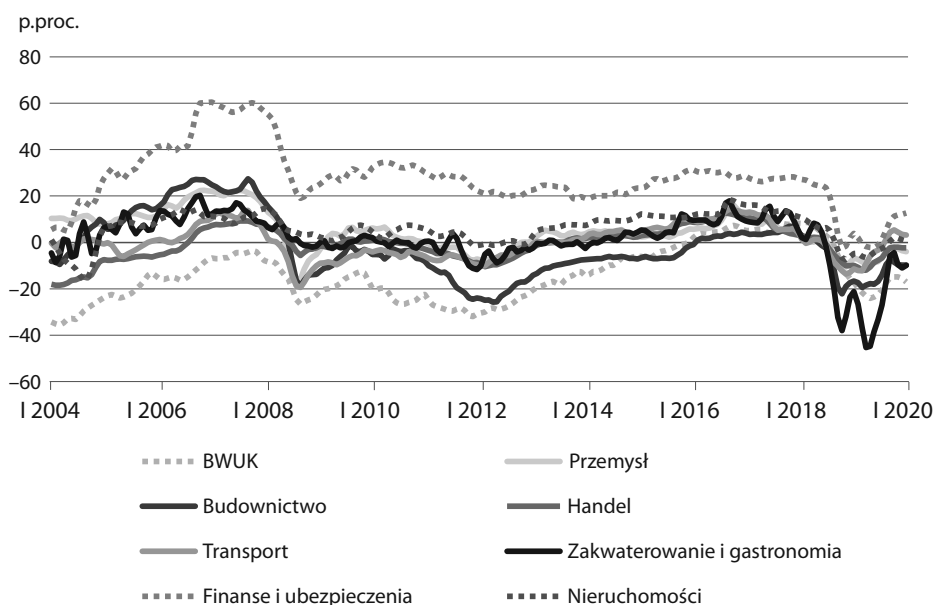
Reakcja wskaźnika Pengab podczas kryzysu w 2008 r. była znacznie wolniejsza – rynek potrzebował ponad półtora roku, aby gwałtownie zareagować na alarmujące informacje o rozwoju sytuacji na amerykańskim rynku finansowym. W krytycznym momencie w październiku 2008 r. jego wartość w stosunku do wartości z września 2008 r. spadła z 39,7 p.proc. do 11,4 p.proc., relatywnie niedużo w porównaniu z największym spadkiem (o 73,9 p.proc.) w czasie kryzysu COVID-19. Należy zaznaczyć, że w Polsce załamania koniunktury mierzonej testem koniunktury ani w przy-

padku pandemii, ani efektów kryzysu z 2008 r. nie znalazły w pełni potwierdzenia w twardych danych, czyli w zmianach wskaźników statystyki oficjalnej.

3.2. Koniunktura w bankowości na tle innych sektorów gospodarki

Ocena koniunktury gospodarczej na podstawie testów koniunktury jest różna w poszczególnych sektorach gospodarki (wykr. 3). W latach 2004–2021 najlepiej oceniana była koniunktura w działalności finansowej i ubezpieczeniowej. Średnia wartość wskaźnika ogólnego klimatu koniunktury gospodarczej w analizowanym okresie wynosiła tam 27,6 p.proc., podczas gdy w pozostałych sektorach zawierała się w przedziale od –0,7 do 6,1 p.proc. Poziom wartości wskaźnika bieżącego ufności konsumentckiej (WBUC) był jeszcze niższy i wynosił średnio –13,8 p.proc.

Wykr. 3. Koniunktura gospodarcza^a w wybranych sektorach gospodarki



a Wyrażona składową TC.

Źródło: obliczenia własne na podstawie: GUS (2021b).

Współzależność między poszczególnymi wskaźnikami koniunktury (należy pamiętać, że analizowany był składnik trendo-cykliczny – TC) w czasie mierzona współczynnikiem korelacji liniowej Pearsona jest zróżnicowana. Najmocniej powiązane są wskaźniki koniunktury w przemyśle i budownictwie (0,90), w handlu i na rynku nieruchomości (0,84) oraz w handlu i ufności konsumentckiej (0,82), a najsłabiej – w finansach i ubezpieczeniach oraz ocenie ufności konsumentckiej (0,29),

a także w przemyśle i ocenie ufności konsumenckiej (0,33). W analizowanym okresie wskaźnik koniunktury w sektorze finansów i ubezpieczeń był najsilniej skorelowany ze wskaźnikami koniunktury w budownictwie (0,77), zakwaterowaniu i gastronomii (0,75) oraz przemyśle (0,72). Wysokie i dodatnie wartości współczynnika korelacji świadczą o tym, że percepcja zmian koniunktury w wielu sekcjach PKD podlega podobnym zmianom. Wskazuje to na powiązanie koniunktury w poszczególnych branżach czy sektorach, a także na specyficzną wrażliwość wyników testów koniunktury na warunki prowadzenia działalności gospodarczej.

Analiza danych przedstawionych na wyk. 3 oraz w tabl. 1 pozwala porównać zmiany koniunktury w czasie kryzysu z lat 2008–2009 oraz kryzysu COVID-19. Można zauważyć, że spadki wartości wskaźników koniunktury w czasie obecnego kryzysu nie zawsze były większe niż w latach 2008–2009. W przypadku przemysłu, budownictwa, transportu oraz finansów i ubezpieczeń silniejsze spadki koniunktury nastąpiły w latach 2008–2009. Co ciekawe, w samej bankowości pogorszenie koniunktury (ocenianej na podstawie Pengab) w 2020 r. było nieco silniejsze niż w 2008 r. Kryzys COVID-19 najbardziej odczuły przedsiębiorstwa zajmujące się zakwaterowaniem i gastronomią – wartości wskaźników koniunktury w latach 2020–2021 spadły o blisko 59 p.proc., podczas gdy w latach 2008–2009 spadek wynosił niespełna 23 p.proc. Najmniej podatne na negatywne konsekwencje obecnego kryzysu okazały się rynek nieruchomości, przemysł i handel.

Tabl. 1. Wskaźniki koniunktury gospodarczej^a w czasie kryzysów

Wskaźniki	2008–2009			2020–2021		
	max 2007–2008	min 2008–2009	różnica (max – min)	max 2019–2020	min 2020–2021	różnica (max – min)
	w p.proc.					
Pengab	41,01	17,23	23,78	26,67	–2,94	29,61
BWUK	–3,45	–26,57	23,12	8,64	–24,11	32,74
Przemysł	22,51	–16,13	38,64	8,04	–10,54	18,58
Budownictwo	27,44	–18,83	46,27	7,08	–22,14	29,23
Handel	9,23	–5,61	14,84	6,27	–11,94	18,21
Transport	13,84	–19,46	33,30	7,06	–13,58	20,65
Zakwaterowanie i gastro- nomia	20,29	–2,60	22,89	13,66	–45,31	58,96
Finanse i ubezpieczenia	60,44	18,98	41,46	28,31	–2,39	30,70
Nieruchomości	14,60	0,65	13,95	13,66	–7,43	21,09

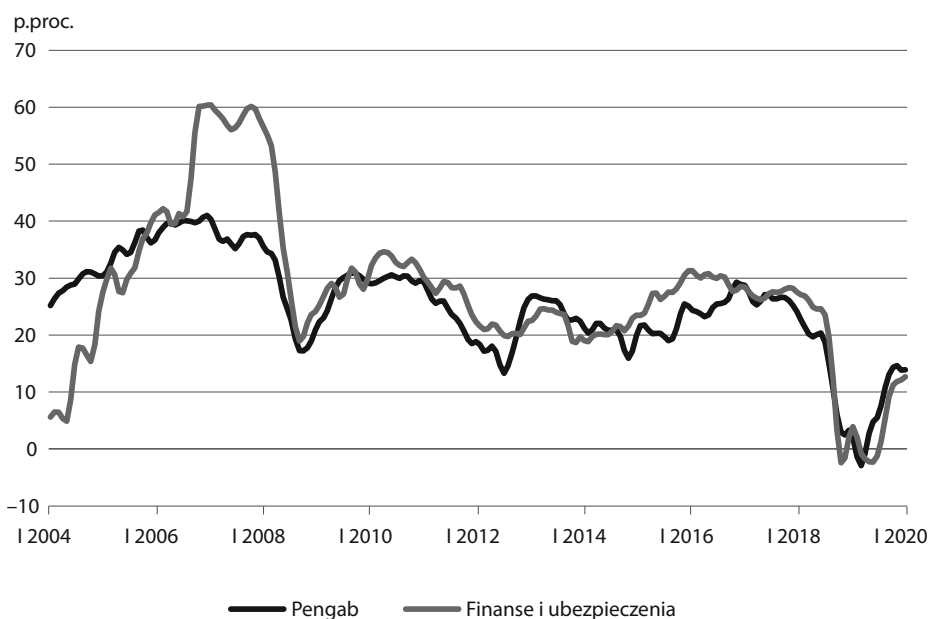
a Wyrażonej składową TC.

Uwaga. W celu przedstawienia skali zmian koniunktury, jakie nastąpiły w wyniku kryzysu, porównano różnice między szczytem koniunktury bezpośrednio przed kryzysem a jej dołkiem w czasie kryzysu.

Źródło: obliczenia własne na podstawie GUS (2021b) i danych z Monitora Bankowego.

Jednym z ważnych elementów działalności finansowej i ubezpieczeniowej jest sektor bankowy. Na wyk. 4 porównano wskaźnik ogólnej koniunktury w bankowości Pengab obliczany przez Kantar Polska ze wskaźnikiem koniunktury działalności finansowej i ubezpieczeniowej obliczanym przez GUS. Zauważalne jest wysokie skorelowanie tych wskaźników, szczególnie od 2009 r. W czasie kryzysu COVID-19 spadek wartości obu wskaźników jest porównywalny. Natomiast w latach 2008–2009 spadek koniunktury w działalności finansowej i ubezpieczeniowej był większy niż spadek koniunktury w samej działalności bankowej.

Wykr. 4. Koniunktura gospodarcza^a w działalności finansowej i ubezpieczeniowej oraz koniunktura w bankowości



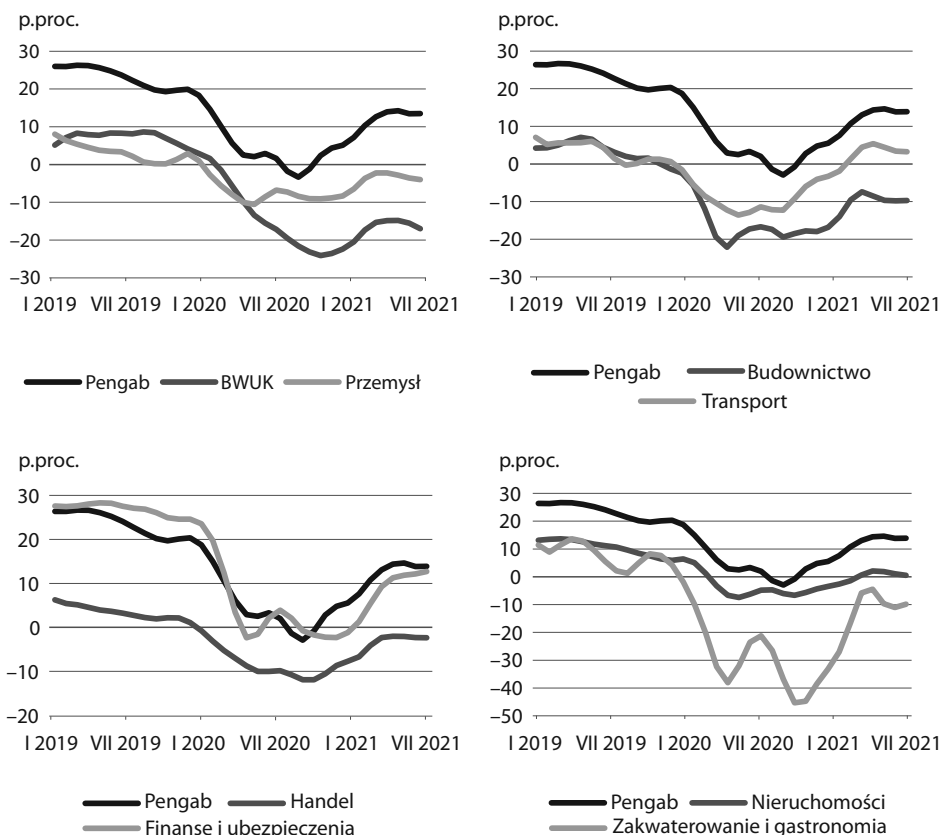
a Wyrażona składową TC.

Źródło: obliczenia własne na podstawie GUS (2021b) i danych z Monitora Bankowego.

Na wyk. 5 porównano wskaźnik koniunktury gospodarczej w bankowości ze wskaźnikami w innych sektorach w czasie kryzysu COVID-19. Z zestawienia tych danych wynika, że symptomy pogorszenia nastrojów wśród pracodawców i zarządzających przedsiębiorstwami obserwowano już przed wybuchem pandemii. Dodatkowo można zobaczyć, które sektory i obszary aktywności gospodarczej wcześniej, a które później zareagowały na kryzys. W przypadku budownictwa, przemysłu, transportu oraz nieruchomości wskaźnik koniunktury osiągnął wartości minimalne wcześniej, bo już w czerwcu i lipcu 2020 r. Równie wcześniej nastąpiło w tych sekto-

rach odbicie od poziomu minimum. Z opóźnieniem na wybuch pandemii zareagowały handel, bankowość (Pengab), konsumenci (BWUK) oraz przedsiębiorstwa z branży zakwaterowania i gastronomii. Co ciekawe, już w czerwcu 2021 r. nastąpiła poprawa koniunktury gospodarczej w kraju (zob. Cieślík i in., 2021; GUS, 2021a, 2020b). Generalnie wzrost koniunktury widoczny w analizowanych wskaźnikach waha się między 44% a 92% ich poprzedniego spadku. Największa poprawa koniunktury gospodarczej – po wcześniejszym pogorszeniu – ma miejsce w transporcie, gdzie 92% spadku zostało już zniwelowane. Dosyć silne, blisko 70-procentowe, odbicie odnotowano też w zakwaterowaniu i gastronomii.

Wykr. 5. Koniunktura^a w bankowości^b na tle koniunktury gospodarczej w innych sektorach



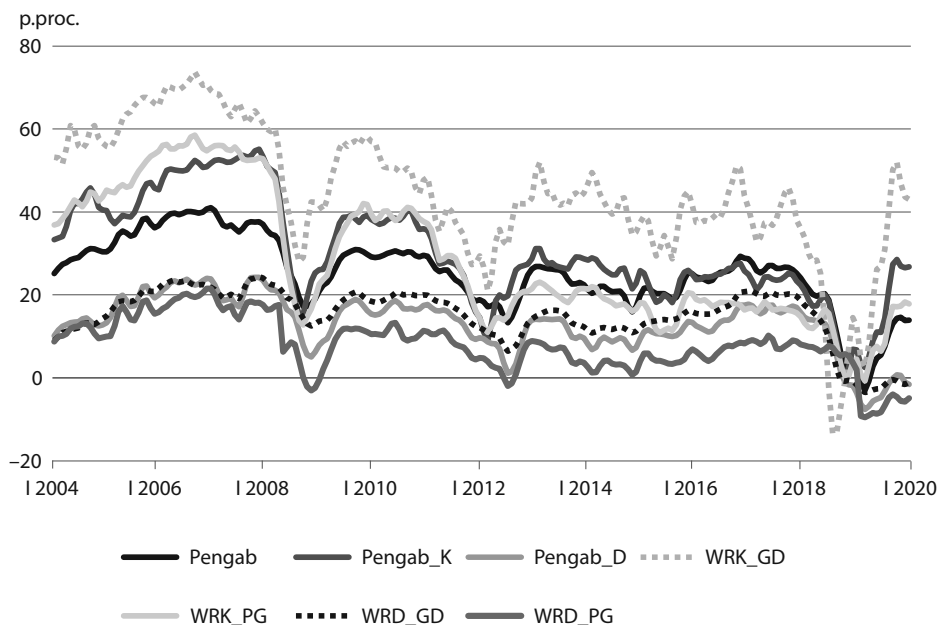
a Wyrażona składową TC. b Na podstawie Pengab.

Źródło: obliczenia własne na podstawie GUS (2021b) i danych z Monitora Bankowego.

3.3. Koniunktura w poszczególnych segmentach bankowości w czasie pandemii COVID-19

Pandemia COVID-19, wraz z jej konsekwencjami gospodarczymi, przejawia z perspektywy ekonomicznej wszystkie cechy kryzysu. W takich okolicznościach trudno przewidywać i oceniać zachowania uczestników rynku. Jednak mimo szczególnych uwarunkowań zachowania klientów na rynku finansowym są podatne na prawidłowości właściwe dla okresu kryzysu gospodarczego i powodują zmianę ocen koniunktury formułowanych przez bankowców. Wspólnym mianownikiem okresu kryzysowego, niezależnie od jego genezy, jest niepewność i utrata poczucia stabilności oraz trwałości praw i reguł, które wydawały się niezmiennie.

Wykr. 6. Koniunktura^a w bankowości



^a Wyrażona składową TC.

Źródło: obliczenia własne na podstawie danych z Monitora Bankowego.

Analizując zmiany koniunktury wyrażone składową trendo-cykliczną w poszczególnych segmentach bankowości w czasie dwóch kryzysów: w latach 2008–2009 oraz w latach 2020–2021, można zauważyć, że nie w każdym segmencie spadek koniunktury w czasie kryzysu COVID-19 był większy niż w czasie kryzysu finansowego (wykr. 6 oraz tabl. 2). Najsilniej negatywnie zareagował na pandemię indeks oceny sytuacji na rynku kredytów gospodarstw domowych (WRK_GD). Niepewna perspektywa zatrudnienia i ograniczenie dochodów gospodarstw domowych spowodowało

wały spadek jego wartości o 65 p.proc. w stosunku do okresu bezpośrednio sprzed pandemii. W tabl. 2 przedstawiono oszacowane spadki wskaźników (wyrażonych składową *TC*) w tych dwóch okresach. Na kryzys COVID-19 znacznie bardziej negatywnie niż na wcześniejszy kryzys zareagował także wskaźnik rynku depozytów gospodarstw domowych (*WRD_GD*), głównie z powodu niskich stóp procentowych. Rynku depozytowym (*Pengab_D*) w czasie obecnego kryzysu zanotowano większe spadki niż w czasie kryzysu z lat 2008–2009.

W pozostałych analizowanych przypadkach spadek wartości wskaźników koniunktury w czasie kryzysu COVID-19 nie przewyższał spadku notowanego w latach 2008–2009. Szczególnie wyróżnia się tu indeks oceny kredytów przedsiębiorstw (*WRK_PG*), którego spadek w okresie pandemii wyniósł 19,02 p.proc., a w kryzysie z lat 2008–2009 – aż 44,26 p.proc. Ogólnie rynek kredytowy (*Pengab_K*) w czasie pandemii ucierpiał mniej niż w czasie poprzedniego kryzysu.

Tabl. 2. Wskaźniki koniunktury w bankowości^a w czasie kryzysów

Wskaźniki	2008–2009			2020–2021		
	max	min	różnica	max	min	różnica
	2007–2008	2008–2009	(max – min)	2019–2020	2020–2021	(max – min)
	w p.proc.					
Pengab	41,01	17,23	23,78	26,67	-2,94	29,61
Pengab_K	55,09	15,26	39,83	28,56	0,38	28,18
Pengab_D	24,26	5,04	19,22	17,33	-7,61	24,94
WRK_GD	72,05	28,00	44,05	51,87	-13,46	65,32
WRK_PG	56,99	12,73	44,26	18,31	-0,71	19,02
WRD_GD	24,39	12,53	11,85	20,23	-3,52	23,75
WRD_PG	21,54	-3,04	24,58	9,02	-9,50	18,52

a Wyrażonej składową *TC*.

Uwaga. Jak przy tabl. 1.

Źródło: obliczenia własne na podstawie danych z Monitora Bankowego.

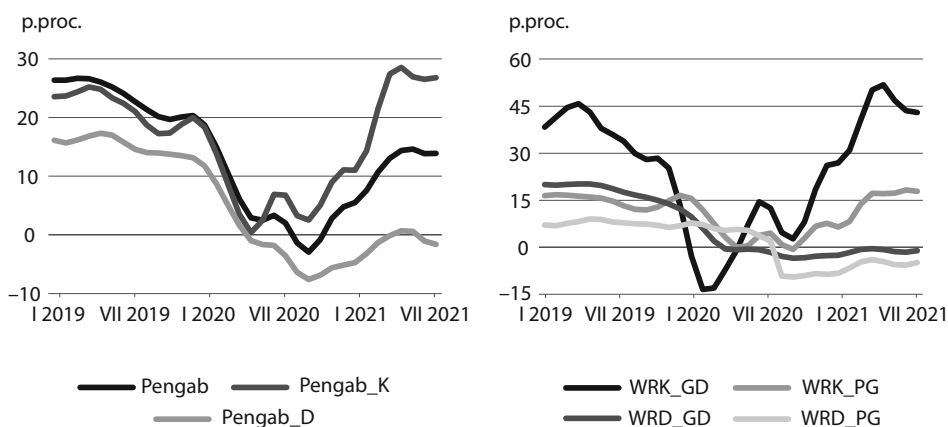
Skala spadku wartości wskaźnika koniunktury w bankowości w poszczególnych segmentach jest w pewnym stopniu skorelowana z ogólną wartością, jaką wskaźniki koniunktury przyjmowały przed wybuchem pandemii: im wyższe wartości przyjmował wskaźnik przed pandemią, tym większe spadki odnotowano później. Spadki koniunktury w sektorze bankowym w wyniku pandemii COVID-19 w największej mierze dotyczyły rynku kredytów gospodarstw domowych (spadek o 65,32 p.proc.), wskaźnika *Pengab* ogółem (spadek o 29,61 p.proc.), wskaźnika rynku kredytowego ogółem (spadek o 28,18 p.proc.) i wskaźnika rynku depozytowego ogółem (spadek o 24,94 p.proc.). Podobne wnioski w badaniach koniunktury w bankowości w tym czasie sformułowali Kluza i Krzesicki (2021).

Wchodzenie w fazę spadku wartości wskaźników koniunktury bankowej w okresie pandemii nastąpiło w podobnym czasie na rynku depozytowym (depozyty bieżą-

ce i terminowe gospodarstw domowych oraz depozyty bieżące i terminowe podmiotów gospodarczych) i kredytowym, ale faza ożywienia przebiegała na nich odmiennie. Rynek depozytowy wszedł w fazę wzrostu aktywności z półrocznym opóźnieniem w stosunku do rynku kredytowego, na którym już w maju 2021 r. aktywność powróciła do poziomu sprzed pandemii, podczas gdy aktywność na rynku depozytów nadal utrzymuje się na bardzo niskim poziomie (wykr. 7).

W czasie pandemii koniunktura na rynku kredytów podmiotów gospodarczych uległa osłabieniu. Wyrazem tego jest spadek wartości wskaźnika WRK_PG o ponad 15 p.proc. w okresie od lutego do czerwca 2020 r. Jednak w porównaniu z aktywnością klientów na rynku kredytów dla osób indywidualnych był on czterokrotnie mniejszy (wykr. 7). Co więcej, na rynku kredytów dla osób indywidualnych (gospodarstw domowych – WRK_GD) nastąpiła znaczna poprawa i pod koniec 2021 r. wskaźnik koniunktury (wyrażony składową TC) był na wyższym poziomie niż tuż przed wybuchem pandemii. Można to wiązać z polityką niskich stóp procentowych oraz wzrostem konsumpcji (Miniszewski, 2021).

Wykr. 7. Koniunktura^a w poszczególnych segmentach bankowości



^a Wyrażona składową TC.

Źródło: obliczenia własne na podstawie danych z Monitora Bankowego.

4. Podsumowanie

W omawianym badaniu dokonano oceny percepcji koniunktury gospodarczej w bankowości w czasie kryzysu COVID-19 oraz uwarunkowań zmian koniunktury w tym okresie. W analizie koniunktury szczególne wyzwanie stanowił dobór odpowiedniej metody pozwalającej na wyeliminowanie wpływu krótkotrwałych zmian

wskaźników koniunktury w czasie pierwszych dwóch, trzech miesięcy od wybuchu pandemii. Tak głębokich, krótkotrwałych i gwałtownych spadków wskaźników nie obserwowano bowiem dotąd w badaniach koniunktury gospodarczej. Szybkie zniwelowanie ok. 2/3 tych spadków uzasadnia słuszność potraktowania ich – przy wykorzystaniu metody X-13-ARIMA – jako zdarzeń nietypowych. Nie wyklucza to negatywnych średniookresowych skutków pandemii w postaci zmian koniunktury gospodarczej.

Przeprowadzony przegląd literatury przedmiotu wykazał, że wybuch pandemii COVID-19 wywarł ogromny wpływ na wszystkie obszary życia społeczno-gospodarczego, w tym na funkcjonowanie sektora bankowego w Polsce i na świecie. Kluczowe znaczenie dla koniunktury w bankowości w Polsce miała siła wprowadzanych restrykcji. Negatywny związek między siłą restrykcji a poziomem wskaźnika Pengab był szczególnie widoczny w 2020 r.

W całej historii badań koniunktury gospodarczej koniunktura w działalności finansowej i ubezpieczeniowej oraz w samej bankowości jest oceniana znacznie lepiej niż w przypadku innych rodzajów działalności gospodarczej. Niemniej jednak w czasie kryzysu spowodowanego przez pandemię COVID-19 i tutaj doszło do jej znacznego osłabienia. Nie było to tak duże pogorszenie, jak np. w branży hotelarsko-gastronomicznej, ale większe niż w przemyśle, handlu czy transporcie. Testy koniunktury pokazują, że wpływ kryzysu na sektor bankowy był nawet silniejszy w okresie pandemii niż podczas kryzysu finansowego z lat 2008–2009. Bankowość nie należy też do sektorów, które najszybciej zareagowały osłabieniem koniunktury w czasie pandemii. Wcześniej pogorszenie koniunktury uwidoczniło się w budownictwie, przemyśle, transporcie oraz na rynku nieruchomości.

Poszczególne segmenty rynku bankowego w czasie kryzysu COVID-19 zachowywały się bardzo różnie. Na rynku kredytowym można było zaobserwować zarówno dosyć gwałtowne spadki, jak i bardzo szybkie wzrosty wskaźników koniunktury, w efekcie czego w październiku 2021 r. ich wartości powróciły do poziomu sprzed wybuchu pandemii. Interesujące jest to, że skala spadku wskaźników koniunktury na rynku kredytowym gospodarstw domowych była ponadtrzykrotnie większa niż na rynku kredytowym przedsiębiorstw. Warto dodać, że podczas kryzysu finansowego 2008–2009 skala ta była porównywalna.

Na rynku depozytowym panuje odmienna sytuacja. Wskaźniki rynku depozytowego po wybuchu pandemii obniżyły się relatywnie nieznacznie, niemniej jednak skala odbicia do października 2021 r. wyniosła średnio ok. 1/3 spadku, co sugeruje, że postrzeganie rynku depozytowego w warunkach niskich stóp procentowych oraz rosnącej inflacji jest dosyć negatywne.

Bibliografia

- Adamowicz, E., Dudek, S., Konat, G., Majchrzak, K., Ratuszny, E., Walczyk, K. (2020). Koniunktura gospodarcza w Europie Środkowo-Wschodniej w dobie epidemii COVID-19. W: M. Strojny (red.), *Europa Środkowo-Wschodnia wobec globalnych trendów: gospodarka, społeczeństwo i biznes* (s. 61–127). Oficyna Wydawnicza SGH.
- Bieć, M. (1996). *Test koniunktury: metody, techniki, doświadczenia*. Oficyna Wydawnicza SGH.
- Center for Statistical Research and Methodology. (2020). *X-13ARIMA-SEATS Reference Manual*. <https://www2.census.gov/software/x-13arima-seats/x13as/windows/documentation/docx13as.pdf>.
- Cieślak, B., Piotrowski, M., Radzikowski, M. (2021). *Koniunktura w bankowości październik-grudzień 2021*. Instytut Rozwoju Gospodarczego, Szkoła Główna Handlowa w Warszawie. https://ssl-kolegia.sgh.waw.pl/pl/KAE/struktura/IRG/publikacje/Documents/biuletyn_banki_Q4_2021.pdf.
- Dayong, Z., Hu, M., Ji, Q. (2020). Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, 36, 1–6. <https://doi.org/10.1016/j.frl.2020.101528>.
- European Banking Authority. (2020). *The EU banking sector: first insights into the COVID-19 impacts* (Thematic note EBA/REP/2020/17). <https://service.betterregulation.com/document/442177>.
- Główny Urząd Statystyczny. (2011). *Badanie koniunktury gospodarczej*. https://stat.gov.pl/cps/rde/xbcr/gus/kon_badanie_koniunktury_gospodarczej_27022012.pdf.
- Główny Urząd Statystyczny. (2018). *Zeszyt metodologiczny. Badanie koniunktury gospodarczej*. <https://stat.gov.pl/obszary-tematyczne/koniunktura/koniunktura/badanie-koniunktury-gospodarczej-zeszyt-metodologiczny,5,9.html>.
- Główny Urząd Statystyczny. (2021a). *Koniunktura konsumencka – grudzień 2021 r. Informacja sygnałna*. <https://stat.gov.pl/obszary-tematyczne/koniunktura/koniunktura/koniunktura-konsumencka-grudzien-2021-roku,1,106.html>.
- Główny Urząd Statystyczny. (2021b). *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach (2000–2021) (grudzień 2021)*. <https://stat.gov.pl/obszary-tematyczne/koniunktura/koniunktura/koniunktura-w-przetworstwie-przemyslowym-budownictwie-handlu-i-uslugach-2000-2021-grudzien-2021,4,56.html>.
- Golec, M. M. (2014). Barometry koniunktury w sektorze bankowym. *Prace Naukowe Wyższej Szkoły Bankowej w Gdańsku*, 32, 215–229.
- Hale, T., Angrist, N., Goldszmidt, R., Kira, B., Petherick, A., Phillips, T., Webster, S., Cameron-Blake, E., Hallas, L., Majumdar, S., Tatlow, H. (2021). A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nature Human Behaviour*, 5(4), 529–538. <https://doi.org/10.1038/s41562-021-01079-8>.
- Hossain, M. (2021). The effect of the COVID-19 on sharing economy activities. *Journal of Cleaner Production*, 280 (part 1), 1–9. <https://doi.org/10.1016/j.jclepro.2020.124782>.
- Hurley, J., Walker, D. (2021). *Did the Covid-19 local lockdowns reduce business activity? Evidence from UK SMEs* (Staff Working Paper No 943). <https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2021/did-the-covid-19-local-lockdowns-reduce-business-activity-evidence-from-uk-smes.pdf>.

- Idzik, M. (2015). Aktywność klientów banków spółdzielczych na rynku usług bankowych w świetle wyników testu koniunktury. *Roczniki Naukowe Stowarzyszenia Ekonomistów Rolnictwa i Agrobiznesu*, 17(5), 65–70.
- International Monetary Fund. (2021). *World Economic Outlook Update October 2021*. <https://www.imf.org/en/Publications/WEO/Issues/2021/10/12/world-economic-outlook-october-2021>.
- Kantar Polska. (b.r.). *COVID-19 Barometr 2021*. Pobrane 10 stycznia 2022 r. z <https://tgismart.com/pl/dzienny-pomiar-nastrojow-ekonomicznych>.
- Kluza, S., Krzesicki, O. (2021). *Koniunktura w bankowości styczeń–marzec 2021*. Instytut Rozwoju Gospodarczego, Szkoła Główna Handlowa w Warszawie. https://ssl-kolegia.sgh.waw.pl/pl/KAE/struktura/IRG/publikacje/Documents/biuletyn_Q1_2021.pdf.
- Koronawirus. *Kolejny lockdown to nawet miliard zł strat dziennie*. (2021, 19 marca). <https://www.money.pl/gospodarka/koronawirus-kolejny-lockdown-to-nawet-miliard-zl-strat-dziennie-6619671119813280a.html>.
- Korzeb, Z., Niedziółka, P. (2020). Resistance of commercial banks to the crisis caused by the COVID-19 pandemic: the case of Poland. *Equilibrium. Quarterly Journal of Economics and Economic Policy*, 15(2), 205–234. <https://doi.org/10.24136/eq.2020.010>.
- Męcina, J., Potocki, P. (2020). *Wpływ COVID-19 na gospodarkę i rynek pracy w Polsce – wyciąg z raportu badawczego*. Fundacja Gospodarki i Administracji Polskiej.
- Miniszewski, M. (2021). *Konsumpcja w pandemii*. Polski Instytut Ekonomiczny.
- Narodowy Bank Polski. (2021). *Raport o stabilności systemu finansowego*. <https://www.nbp.pl/systemfinansowy/rsf122021.pdf>.
- Organisation for Economic Co-Operation and Development. (2020). *OECD Economic Outlook*. OECD Publishing. <https://doi.org/10.1787/0d1d1e2e-en>.
- Organisation for Economic Co-Operation and Development. (2021). *The COVID-19 crisis and banking system resilience. Simulation of losses on non-performing loans and policy implications*. OECD Publishing. <https://www.oecd.org/daf/fin/financial-markets/COVID-19-crisis-and-banking-system-resilience.pdf>.
- Rokicki, T. (2020). Zmiany w koniunkturze gospodarczej Polski w wyniku epidemii COVID-19. *Przegląd Prawno-Ekonomiczny*, (3), 105–126. <https://doi.org/10.31743/ppe.10029>.
- Rudke, M. (2020, 2 grudnia). *Zysk netto banków niższy o ponad połowę*. <https://www.rp.pl/banki/art391331-zysk-netto-bankow-nizszy-o-ponad-polowe>.
- Skikiewicz, R. (2020). Determinanty prowadzenia działalności w sektorze bankowym w świetle wyników badania metodą testu koniunktury. W: S. Kluza, K. Walczyk (red.), *20 lat koniunktury w sektorze bankowym – z badań Instytutu Rozwoju Gospodarczego SGH* (s. 73–88). Oficyna Wydawnicza SGH.
- S&P Global Ratings. (2020). *COVID-19 heat map: Post-crisis credit recovery could take to 2022 and beyond for some sectors*. <https://www.spglobal.com/ratings/en/research/articles/200624-covid-19-heat-map-post-crisis-credit-recovery-could-take-to-2022-and-beyond-for-some-sectors-11535796>.

Weersink, A., von Massow, M., Bannon, N., Ifft, J., Maples, J. (2021). COVID-19 and the agri-food system in the United States and Canada. *Agricultural Systems*, 188, 1–16. <https://doi.org/10.1016/j.agsy.2020.103039>.

Związek Banków Polskich. (2021). *Monitor Bankowy. Grudzień 2021*. https://www.zbp.pl/getmedia/03534aa0-0533-4a78-8862-d6b3d5f78f5c/mb_grudzien-2021.

Związek Przedsiębiorców i Pracodawców. (2020). *Raport z badania Finanse Polaków w czasie COVID-19. Jak pandemia wpłynęła na portfele i zwyczaje finansowe Polaków?* <https://zpp.net.pl/wp-content/uploads/2020/11/Raport-Finanse-Polakow-w-czasie-COVID-19.pdf>.

Związek Przedsiębiorstw Finansowych w Polsce. (2021, 26 lutego). *Polskie firmy w 2020. Zamknięcia i otwarcia biznesów*. <https://zpf.pl/polskie-firmy-w-2020/>.

Wpływ pandemii COVID-19 na stan zdrowia psychicznego społeczeństwa¹

Aneta Ptak-Chmielewska^a, Karolina Baszniak^b, Jarosław Kurpanik^c

Streszczenie. Pandemia COVID-19 odmieniła życie ludzi na całym świecie, m.in. wpłynęła na kondycję psychiczną i funkcjonowanie wielu rodzin. Głównym celem badania omawianego w artykule jest ocena wpływu pandemii COVID-19 na stan zdrowia psychicznego członków gospodarstw domowych. W badaniu posłużono się zbiorem danych pochodzących z ankiety *COVID Impact Survey*, przeprowadzonej w 2020 r. (w trakcie pierwszej fali pandemii) w Stanach Zjednoczonych wśród osób dorosłych przez organizację Data Foundation. Analizie poddano 6768 obserwacji. Oszacowano model regresji logistycznej oraz modele oparte na metodach data mining, takich jak: drzewa decyzyjne, wzmacnianie gradientowe, metoda *k*-najbliższych sąsiadów, sztuczne sieci neuronowe i metoda wektorów wspierających. Analiza skupień pozwoliła podzielić respondentów na grupy uwidaczniające cechy charakterystyczne i problemy członków gospodarstw domowych, a w utworzonym modelu uwzględniono kwestie zdrowia i zaburzeń psychicznych oraz ich związek z sytuacją finansową gospodarstw. Wyniki badania wskazują na to, że izolacja, zdalny tryb nauczania i pracy oraz mniejsza aktywność fizyczna przyczyniają się do pogarszania się stanu zdrowia psychicznego.

Słowa kluczowe: uczenie maszynowe, pandemia COVID-19, data mining, stan zdrowia psychicznego, gospodarstwo domowe, Stany Zjednoczone

JEL: C1, C45, I12

Influence of the COVID-19 pandemic on the mental health of society

Abstract. The COVID-19 pandemic changed the lives of people all over the world, by e.g. affecting the mental health and the functioning of many families. The main goal of the research presented in this paper is to assess the influence of the COVID-19 pandemic on the mental health of members of households. The research was performed on the basis of a data set from the *COVID Impact Survey* carried out by the Data Foundation think tank in 2020 (during the first wave of the COVID-19 pandemic) in the USA among adult respondents. The survey used 6,768

¹ Artykuł został opracowany na podstawie referatu wygłoszonego na konferencji Multivariate Statistical Analysis MSA 2021, która odbyła się w dniach 8–10 listopada 2021 r. w Łodzi. / The article is based on a paper delivered at the Multivariate Statistical Analysis MSA 2021 conference, held on 8–10 November 2021 in Łódź.

^a Szkoła Główna Handlowa w Warszawie, Instytut Statystyki i Demografii, Polska / SGH Warsaw School of Economics, Institute of Statistics and Demography, Poland.

ORCID: <https://orcid.org/0000-0002-9896-4240>. Autor korespondencyjny / Corresponding author, e-mail: aptak@sgh.waw.pl.

^b ING Tech Poland. ORCID: <https://orcid.org/0000-0003-3079-3526>. E-mail: karolina.baszniak@ing.com.

^c Uniwersytet Ekonomiczny w Katowicach, Wydział Informatyki i Komunikacji, Polska / University of Economics in Katowice, Faculty of Informatics and Communication, Poland.

ORCID: <https://orcid.org/0000-0002-4288-4647>. E-mail: jaroslaw.kurpanik@ue.katowice.pl.

observations. The authors estimated a model of logistic regression and models based on data mining methods, such as decision trees, XGBoost, the k -nearest neighbours method, artificial neural networks and a support vector machine. Cluster analysis allowed the division of the respondents into groups showing their characteristic features and problems, and the constructed model took into account their mental health issues and the relationship between those issues and the financial situation of households. The results demonstrate that isolation, remote education and work, and limited physical activity contribute to the worsening of the mental health of the population.

Keywords: machine learning, COVID-19 pandemic, data mining, mental health, household, USA

1. Wprowadzenie

Pandemia COVID-19, nazywana również pandemią koronawirusa, to globalna epidemia choroby zakaźnej wywołanej przez koronawirusa SARS-CoV-2. Pierwsze przypadki zakażenia odnotowano w grudniu 2019 r. w chińskim mieście Wuhan, a 20 stycznia 2020 r. Światowa Organizacja Zdrowia (World Health Organization – WHO) ogłosiła wybuch epidemii – stanu zagrożenia zdrowia publicznego o zasięgu międzynarodowym (WHO, b.r.).

Pandemia COVID-19 odmieniła życie ludzi na całym świecie. Nie tylko wycisnęła piętno na zdrowiu i życiu wielu osób, lecz także wpłynęła na gospodarkę, kulturę, politykę, edukację, środowisko czy styl życia. Wybuch pandemii stał się zagrożeniem dla stabilności światowych gospodarek – prognozuje się, że pozostaną niestabilne aż do momentu wyraźnego poprawienia się wyników finansowych podmiotów gospodarczych (Scope Ratings, 2020).

Wśród sektorów gospodarki najbardziej dotkniętych wybuchem pandemii należy wymienić:

- branżę turystyczną – ucierpiała z powodu wprowadzania zakazów podróżowania oraz zamykania miejsc publicznych i atrakcji turystycznych (Turner, 2020);
- sektor detaliczny – na jego kondycję ekonomiczną wpłynęło skrócenie godzin otwarcia sklepów, a nawet ich całkowite zamknięcie. Według raportu Yelp (2020) ok. 60% amerykańskich firm, które zostały zamknięte od początku pandemii, nie wznowi już działalności. Wybuch epidemii podawany był także jako przyczyna braków w zaopatrzeniu. Wynikało to m.in. z paniki zakupowej, prowadzącej do opróżniania półek sklepowych z artykułów spożywczych oraz do zakłóceń w operacjach fabrycznych czy logistycznych (Tyko i in., 2020);
- branżę kulturalną i sektor edukacyjny – muzea, kina, teatry i inne instytucje kulturalne na całym świecie zostały tymczasowo zamknięte, a wystawy i koncerty – odwołane lub przełożone. Wiele instytucji i placówek edukacyjnych podjęło próbę świadczenia usług za pośrednictwem platform cyfrowych.

Pandemia miała negatywny wpływ na zdrowie ludzi w wielu aspektach, m.in. doprowadziła do ogólnego spadku liczby wizyt u lekarzy. Szacuje się, że w Stanach Zjednoczonych liczba konsultacji z powodu symptomów zawału serca zmniejszyła

się o 38%, a w Hiszpanii – o 40% (World Food Programme, 2021). Obostrzenia pandemiczne negatywnie oddziaływały również na zdrowie psychiczne ludzi na całym świecie. Dystans społeczny spotęgował poczucie samotności, przyczynił się do popadania w depresję, a w wyniku braku możliwości wyjścia z domu – nasilił przemoc domową (Surkova i in., 2020). Z badań ankietowych wynika, że w czerwcu 2020 r. 40% dorosłych Amerykanów doświadczało zaburzeń zdrowia psychicznego, a 11% poważnie rozważało podjęcie próby samobójstwa w ciągu ostatniego miesiąca (Czeisler i in., 2020).

Głównym celem badania omawianego w artykule jest ocena wpływu pandemii COVID-19 na stan zdrowia psychicznego członków gospodarstw domowych.

2. Badania społecznych i gospodarczych skutków pandemii COVID-19

Spoleczne i gospodarcze następstwa pandemii są przedmiotem wielu prac badawczych. W Europie m.in. European Foundation for the Improvement of Living and Working Conditions (Eurofound) prowadzi badanie ankietowe *Living, working and COVID-19*, dotyczące doświadczeń z życia i pracy w czasie pandemii, a jego celem jest wsparcie decydentów w wyprowadzeniu krajów z kryzysu. Zagregowane dane interaktywne z tego badania są publikowane na stronie Eurofound². Można je przeglądać w różnych zestawieniach i dla różnych krajów.

W lutym i marcu 2021 r. przeprowadzono trzecią rundę tego badania, a podsumowanie jej wyników opublikowano w raporcie *Living, working and COVID-19 (Update April 2021): Mental health and trust decline across EU as pandemic enters another year* (Eurofound, 2021). Przedstawiono w nim sytuację społeczną i ekonomiczną mieszkańców Europy podczas pierwszego roku życia zgodnie z restrykcjami wprowadzonymi po wybuchu pandemii COVID-19. Przeanalizowano podstawowe wnioski oraz prześledzono rozwój sytuacji w 27 krajach Unii Europejskiej od rozpoczęcia badania w kwietniu 2020 r. Skupiono się na problemach, z którymi zmagają się społeczeństwa w trakcie pandemii, jak np.: wzrost niepewności na rynku pracy wynikający z większego ryzyka jej utraty, spadek poziomu dobrobytu, pogorszenie w zakresie zapewnienia równości płci, spadek zaufania do instytucji publicznych, rosnące trudności z zachowaniem równowagi pomiędzy pracą a życiem prywatnym oraz wzrost nieufności wobec szczepień. Wyniki przedstawione w raporcie wskazują na potrzebę bardziej holistycznego podejścia do wsparcia tych grup ludności, które najbardziej ucierpiały w wyniku kryzysu.

Innym badaniem ankietowym, którego wyniki zostały wykorzystane w badaniu omawianym w artykule, jest *COVID Impact Survey*, przeprowadzone w Stanach

² Zob. <https://www.eurofound.europa.eu/data/covid-19>.

Zjednoczonych przez niezależną pozarządową organizację non profit Data Foundation. W związku z ograniczoną infrastrukturą monitorowania zdrowia Data Foundation zdecydowała się na przeprowadzenie statystycznie wiarygodnego badania ankietowego poruszającego tematykę zdrowia fizycznego i psychicznego, bezpieczeństwa ekonomicznego i żywnościowego oraz zatrudnienia. Zostało ono opracowane przy wsparciu czołowych krajowych ekspertów w dziedzinie zdrowia publicznego, ekonomii i nauk społecznych (Wozniak i in., 2020) w celu przekazania płynącej z niego wiedzy decydom i amerykańskiej opinii publicznej, jak również zachęcenia rządu do wykonywania podobnych analiz na większą skalę.

Ankiety przeprowadzono na reprezentatywnej próbie losowej³ z kilkunastu stanów i obszarów metropolitalnych. Wyniki dostarczają informacji na temat doświadczeń amerykańskiego społeczeństwa związanych z pandemią COVID-19.

Z kolei *Household Pulse Survey* to badanie przeprowadzone przez agencję rządową Stanów Zjednoczonych (United States Census Bureau) we współpracy z agencjami federalnymi, które dostarczyło danych na temat społecznego i ekonomicznego oddziaływania pandemii koronawirusa na amerykańskie gospodarstwa domowe. Próba biorąca udział w ankiecie została wybrana losowo, z zachowaniem reprezentatywności dla całego kraju. Kwestionariusz był krótki; respondenci otrzymywali go e-mailowo. Wyniki ankiety udostępniono zarówno w postaci danych jednostkowych, umożliwiającą dalszą analizę przez badaczy, jak i w formie interaktywnego pulpitu nawigującego, pozwalającego na zapoznanie się z zagregowanymi wynikami zaprezentowanymi na wykresach i w tabelach.

W czerwcu 2020 r. Chetty i in. opracowali, na podstawie danych z sektora prywatnego, raport o ekonomicznych skutkach pandemii COVID-19 w Stanach Zjednoczonych pt. *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data*. Badacze zamieścili w nim ogólnodostępną bazę danych, które w czasie rzeczywistym pokazują aktywność gospodarczą prywatnych firm. Publikowane statystyki dotyczą m.in. wydatków konsumentów, sytuacji biznesu i kluczowych wskaźników gospodarczych.

W grudniu 2020 r. ukazał się raport z badania wzdłużnego trzech holenderskich kohort kontrolnych na temat wpływu pandemii COVID-19 na zdrowie psychiczne ludzi z depresją, zaburzeniami lękowymi i obsesyjno-kompulsyjnymi oraz bez nich autorstwa badaczy związanych z międzynarodowym dziennikiem medycznym „The Lancet. Psychiatry”. W kwietniu i maju 2020 r. badacze rozsyłali kwestionariusze internetowe do osób z zaburzeniami psychicznymi i do ludzi zdrowych, zawierające pytania na temat postrzegania wpływu pandemii na zdrowie psychiczne oraz radzenia sobie z lękiem przed infekcją, a także cztery zwalidowane skale oceniające wystę-

³ Więcej informacji na stronie: <https://www.covid-impact.org/about-the-survey-questionnaire>.

powanie objawów depresji, zaburzeń lękowych, zmartwień i samotności używane we wcześniejszych badaniach. Liczbę i przewlekłość zaburzeń określono na podstawie diagnoz z poprzednich lat przy użyciu regresji logistycznej i modeli mieszanych.

Badanie wykazało, że w czasie pandemii COVID-19 u osób bez zaburzeń depresyjnych, lękowych lub obsesyjno-kompulsyjnych wystąpił większy wzrost częstości objawów depresji, zaburzeń lękowych, zmartwień i samotności niż u osób z zaburzeniami psychicznymi. W przypadku osób z największym obciążeniem psychicznym zauważono nawet niewielki spadek częstości występowania objawów (Pan i in., 2021).

Badania skutków pandemii COVID-19 są niezwykle istotne dla decydentów opracowujących strategię walki z jej negatywnymi skutkami oraz dla wszystkich, którym zrozumienie zmian, jakie przyniosły wydarzenia ostatnich dwóch lat, może ułatwić efektywne funkcjonowanie w życiu prywatnym i zawodowym.

3. Metody wykorzystane do analizy skutków pandemii COVID-19

W celu przestudiowania głównych problemów zdrowotnych, z jakimi zmagali się członkowie gospodarstw domowych w czasie pierwszej fali pandemii, wyniki ankiety *COVID Impact Survey* (COVID Impact Survey, b.r.) poddano analizie metodami statystycznymi (regresja logistyczna) oraz data mining.

Jednym z głównych skutków pandemii jest pogarszający się stan zdrowia psychicznego osób dotkniętych jej następstwami. Przy użyciu modelu predykcyjnego zidentyfikowano osoby, u których z dużym prawdopodobieństwem zostaną zdiagnozowane problemy psychiczne. Ma to bardzo duże znaczenie, ponieważ kondycja psychiczna jest nie mniej ważnym składnikiem ogólnego stanu zdrowia niż kondycja fizyczna. Choroby psychiczne, zwłaszcza depresja, zwiększają ryzyko wystąpienia wielu innych chorób, szczególnie przewlekłych, takich jak udar, cukrzyca typu II i choroby serca (National Institute of Mental Health, 2021). Szacunki sprzed pandemii COVID-19 pokazują, że z powodu samej depresji i lęków świat traci rocznie prawie bilion dolarów w obszarze produktywności ekonomicznej, a wyniki badań pozwalają stwierdzić, że każdy dolar wydany na rzetelną opiekę nad depresją i lękami zwróci się w przyszłości pięciokrotnie (WHO, 2020).

Badanie *COVID Impact Survey*, przeprowadzone wśród osób w wieku 18 lat i więcej, zostało oparte na danych z bazy National Survey Data wykorzystywanej do badań o zasięgu narodowym, reprezentacyjnych dla całej populacji gospodarstw domowych Stanów Zjednoczonych⁴. Podstawą konstrukcji próby reprezentatywnej stały się dane adresowe z bazy U.S. Postal Service (USPS) zawarte w Delivery

⁴ Więcej informacji na stronie covid-impact.org.

Sequence File (DSF), uwzględniającej ok. 97% gospodarstw domowych w Stanach. Wyniki były ważone zmiennymi demograficznymi z *Current Population Survey 2020*.

Jednym z najistotniejszych kroków podczas budowania modelu i analizowania wpływu określonych czynników na zmienną objaśnianą jest wybór zmiennych objaśniających. Aby zapewnić jak najlepszą jakość modelu, na podstawie literatury przedmiotu dokonano statystycznego doboru zmiennych z uwzględnieniem zależności zmiennych objaśniających od zmiennej objaśnianej oraz wyeliminowano zmienne objaśniające skorelowane ze sobą.

Początkowo zbiór składał się z 7505 obserwacji. Po usunięciu brakujących wartości (odpowieź „nie wiem”), przypadków pominięcia pytań lub odmowy odpowiedzi ostatecznie liczył 6655 obserwacji. Dodatkowo został podzielony na zbiory uczący (treningowy) i testowy w proporcji 75% do 25%⁵ (losowanie warstwowe z warstwowaniem po zmiennej zależnej). Z tego względu w zbiorze treningowym znalazło się 4991 obserwacji, a w testowym – 1664.

Zmienną objaśnianą jest zmienna binarna PHYS3H, określająca, czy dana osoba ma zdiagnozowaną chorobę o podłożu psychicznym. Zmienne objaśniające zostały przedstawione w zestawieniu 1. Przyjęcie zmiennej objaśnianej PHYS3H jest ograniczeniem przeprowadzonego badania, ponieważ dokładna data diagnozy jest nieznaną (mogła zostać postawiona jeszcze przed wybuchem pandemii), niemniej jednak wyniki analizy dostarczają cennych informacji, i to nie tylko z punktu widzenia polityki zdrowotnej.

Zestawienie 1. Zmienne objaśniające użyte do budowy modelu

Zmienne	Opis	Wartości	Typ
PHYS8	Czy swój stan zdrowia oceniłbyś jako: wyśmienity, bardzo dobry, dobry, umiarkowany czy zły? (American Psychiatric Association, 2019)	(1) Wyśmienity (2) Bardzo dobry (3) Dobry (4) Umiarkowany (5) Zły	porządkowa
PHYS1K	Czy w ciągu ostatnich siedmiu dni doświadczyłeś uczucia zmęczenia lub znużenia? (Kennedy, 2008)	(1) Tak (0) Nie	binarna
PHYS1Q	Czy w ciągu ostatnich siedmiu dni doświadczyłeś utraty apetytu? (American Psychiatric Association, 2019)	(1) Tak (0) Nie	binarna
SOC5A	Jak często w ciągu ostatnich siedmiu dni byłeś zdenerwowany, niespokojny? (Kanter i in., 2008)	(1) Wcale lub krócej niż dzień (2) 1–2 dni (3) 3–4 dni (4) 5–7 dni	porządkowa

⁵ Typowa proporcja przyjmowana w data miningu.

Zestawienie 1. Zmienne objaśniające użyte do budowy modelu (dok.)

Zmienne	Opis	Wartości	Typ
SOC5C	Jak często w ciągu ostatnich siedmiu dni czułeś się samotny? (Weeks i in., 1980)	(1) Wcale lub krócej niż dzień (2) 1–2 dni (3) 3–4 dni (4) 5–7 dni	porządkowa
SOC5D	Jak często w ciągu ostatnich siedmiu dni czułeś brak nadziei, myśląc o przyszłości? (Healthline, b.r.)	(1) Wcale lub krócej niż dzień (2) 1–2 dni (3) 3–4 dni (4) 5–7 dni	porządkowa
ECON6B	Czy w ciągu ostatnich siedmiu dni otrzymałeś dodatkową pomoc żywieniową, zawnioskowałeś o nią lub starałeś się zakwalifikować do programu dodatkowej pomocy żywieniowej (<i>Supplemental Nutrition Program</i>)? (Brown i Moran, 1997)	(1) Otrzymałem (2) Zawnioskowałem (3) Próbowałem zawnioskować (4) Nie otrzymałem lub nie wnioskowałem	nominalna
ECON7_1	Czy obecnie, aby pokryć dodatkowy wydatek, musiałbyś skorzystać ze swojej karty kredytowej i mógłbyś spłacić dług dopiero w następnym okresie rozliczeniowym? (Brown i Moran, 1997)	(1) Tak (0) Nie	binarna
PHYS2_18	Czy uważasz, że gdy zostajesz w domu z powodu złego samopoczucia, wynika to z panowania pandemii koronawirusa?	(1) Tak (0) Nie	binarna
PHYS3M	Czy stwierdzono u ciebie nadwagę lub otyłość? (Pereira-Miranda i in., 2017)	(1) Tak (0) Nie	binarna
ECON6A	Czy w ciągu ostatnich siedmiu dni otrzymałeś zasiłek dla bezrobotnych, zawnioskowałeś o taki zasiłek lub starałeś się go otrzymać? (Brown i Moran, 1997)	(1) Otrzymałem (2) Zawnioskowałem (3) Próbowałem zawnioskować (4) Nie otrzymałem lub nie wnioskowałem	nominalna
AGE7	Wiek w latach	(1) 18–24 (2) 25–34 (3) 35–44 (4) 45–54 (5) 55–64 (6) 65–74 (7) 75+	porządkowa
HHINCOME	Dochód gospodarstwa domowego (roczny) w tys. USD	(1) poniżej 10 (2) 10–20 (3) 20–30 (4) 30–40 (5) 40–50 (6) 50–75 (7) 75–100 (8) 100–150 (9) 150 i więcej	porządkowa

Źródło: opracowanie własne na podstawie COVID Impact Survey (b.r.).

Do budowy modeli predykcyjnych wykorzystano metody i modele statystyczne oraz data mining: regresję logistyczną (ang. *logistic regression* – LR) drzewa decyzyjne (ang. *decision trees* – DT), wzmacnianie gradientowe (Extreme Gradient Boosting – XGBoost), metodę k -najbliższych sąsiadów (ang. *k-nearest neighbours* – KNN), sztuczne sieci neuronowe (ang. *neural networks* – NN) i metodę wektorów wspierających (Support Vector Machines – SVM). Szczegółowy opis metod data mining można znaleźć m.in. w następujących publikacjach polskojęzycznych: Frątczak (2012), Larose (2008), Lasek i Pęczkowski (2013) oraz anglojęzycznych: Alpaydin (2004), Kim i Sohn (2010), Rokach i Maimon (2014), Sarma (2007), Hastie i in. (2001).

3.1. Regresja logistyczna (LR)

Model regresji logistycznej jest używany wtedy, gdy zmienna objaśniana ma charakter binarny. Regresja logistyczna pozwala na obliczenie prawdopodobieństwa przewidywanego zdarzenia i opiera się na funkcji logistycznej następującej postaci:

$$f(z) = \frac{1}{1+e^{-z}} = \frac{1}{1+e^{-(\alpha+\sum\beta_i x_i)}}, \quad (1)$$

gdzie:

α – stała regresji,

β_i – współczynnik regresji dla i -tej zmiennej objaśniającej,

x_1, x_2, \dots, x_k – zmienne niezależne.

S-kształtna funkcja logistyczna f , widoczna na wykresie, przyjmuje wartości z zakresu od 0 do 1.

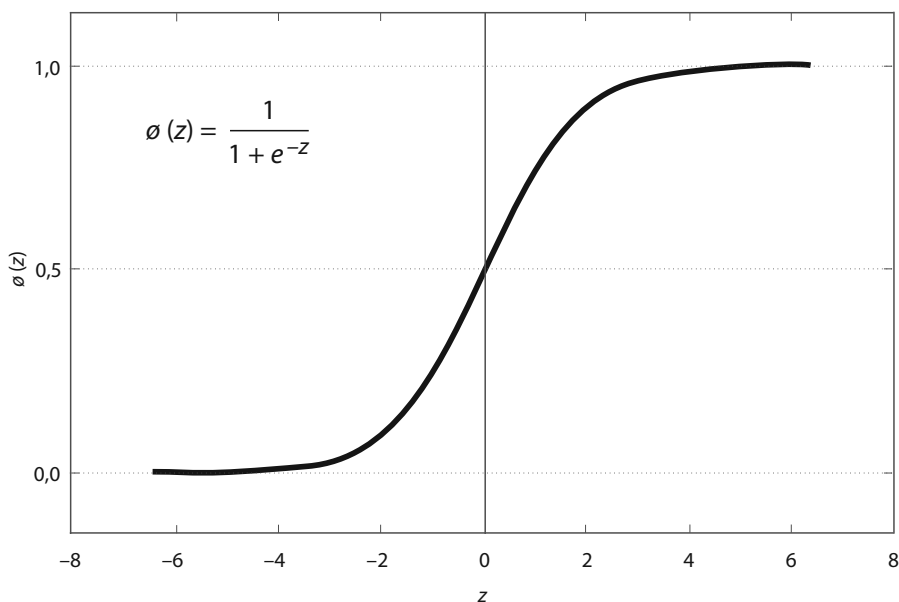
W regresji logistycznej wyniki są interpretowane w kategorii ilorazów szans. Szansę oblicza się, dzieląc prawdopodobieństwo sukcesu przez prawdopodobieństwo porażki, czyli:

$$odds = \frac{p}{1-p}. \quad (2)$$

Do selekcji zmiennych w regresji logistycznej stosuje się metody: wsteczną, postępującą lub krokową. W omawianym badaniu do selekcji zmiennych zastosowano metodę wsteczną (ang. *backward*), czyli wprowadzenie wszystkich zmiennych objaśniających, a następnie usuwanie po kolei najmniej istotnych statystycznie, przy czym istotność zazwyczaj mierzy się rezultatem odpowiedniego testu istotności (t -Studenta lub F -score) albo kryterium informacyjnym (np. kryterium informacyjnym Akaikego, Akaike information criterion – AIC). W tym przypadku wykorzysta-

no test t -Studenta. Ponadto częstą procedurą – uwzględnioną również tutaj – jest sprawdzenie występowania interakcji, czyli wpływu wybranych zmiennych niezależnych na inne zmienne.

Wykres funkcji logistycznej



Źródło: opracowanie własne.

3.2. Drzewa decyzyjne (DT)

Drzewa decyzyjne to narzędzie wykorzystywane do hierarchicznej segmentacji i podziału zbioru danych (schemat 1). Wyściowym elementem drzewa jest korzeń, który zawiera cały zbiór danych. Kolejno, za pomocą odpowiednich reguł, dokonuje się podziału zbioru danych na węzły (segmenty) i podziału segmentów na subsegmenty. Segment z subsegmentami tworzy gałąź drzewa. Ostateczny subsegment, który nie podlega dalszemu podziałowi, to tzw. liść. Konkretna obserwacja jest przypisana tylko do jednego liścia. Model drzewa pokazuje przyporządkowanie danych do liścia, służące do predykcji lub klasyfikacji.

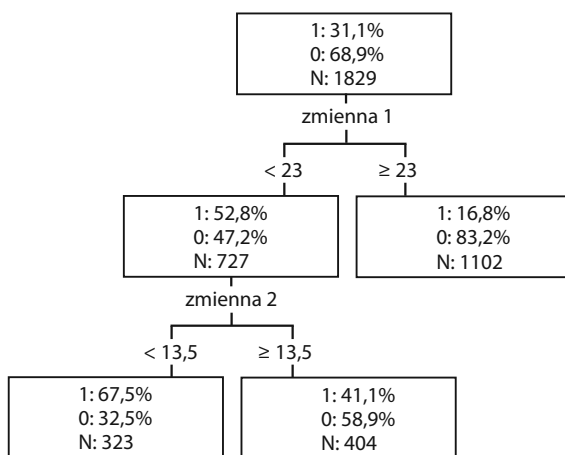
Do budowy drzew decyzyjnych niezbędne są duże zbiory danych z odpowiednio licznymi przypadkami zmiennej objaśnianej. Obserwacje odstające mogą zniekształcić wyniki modelu, ale największym zagrożeniem jest ryzyko przeuczenia modelu, czyli zbytowego dopasowania się modelu do danych uczących, które nie znajduje potwierdzenia w danych testowych. Drzewo decyzyjne nie zawiera oszacowania parametrów, ponieważ w przeciwieństwie do regresji opiera się na podziale zbioru

na oddzielne grupy. Reguły podziału mogą posłużyć do predykcji lub klasyfikacji nowych danych. Drzewo klasyfikacyjne służy do klasyfikacji danych w przypadku zmiennej objaśnianej binarnej lub porządkowej, natomiast drzewo regresyjne – do predykcji w przypadku zmiennej objaśnianej ilorazowej lub przedziałowej.

Podstawowe miary wykorzystywane w drzewie klasyfikacyjnym do pomiaru jakości podziału dla zmiennej zależnej binarnej lub porządkowej to:

- osiągnięty stopień separacji (test chi-kwadrat Pearsona);
- stopień redukcji zanieczyszczenia (entropia lub współczynnik Giniego).

Schemat 1. Przykładowe drzewo decyzyjne



Uwaga. N – liczebność.

Źródło: opracowanie własne.

Kryterium zatrzymania podziału może być osiągnięta minimalna liczba obserwacji w liściu, wielkość węzła lub liczba podziałów w ścieżce. Po zbudowaniu drzewo wymaga przycięcia do optymalnej wielkości.

Zaletami drzew decyzyjnych są przede wszystkim łatwa interpretowalność wyników oraz elastyczność modelu. Drzewa decyzyjne nie są wrażliwe na braki danych, nie wymagają założenia normalności rozkładów czy niezależności zmiennych objaśniających, którymi mogą być zmienne ciągłe, dyskretne, nominalne lub binarne, ani selekcji zmiennych, ponieważ automatycznie wybierane są tylko najistotniejsze zmienne, również z uwzględnieniem zależności nieliniowych.

Za największą wadę drzew decyzyjnych należy uznać ich niestabilność i skłonność do przetrenowania. Ponadto nie mają zastosowania w przypadku niedużych zbiorów danych. Ostateczne oszacowanie prawdopodobieństw jest zagregowane na poziomie subsegmentu końcowego.

3.3. Wzmacnianie gradientowe (XGBoost)

Algorytm XGBoost stanowi jedną z najpopularniejszych metod uczenia maszynowego. Bazuje na wielu klasyfikatorach, przy czym końcowa klasyfikacja poszczególnych obserwacji – na decyzjach podjętych przez pojedyncze drzewa decyzyjne. Podobnie jak w algorytmie AdaBoost wykorzystuje się strategię przyrostową, ponieważ jest mniej czasochłonna i skomplikowana niż trenowanie wszystkich drzew decyzyjnych równoległe (jak w przypadku lasów losowych).

Wzmacnianie gradientowe pozwala na agregację wielu stabilnych, ale niezbyt wydajnych klasyfikatorów. Główną koncepcją wzmacniania jest to, że w procesie iteracyjnym obserwacje otrzymują wagi wskazujące klasyfikatorowi, na których obserwacjach ma się koncentrować w kolejnej iteracji. Finalna decyzja klasyfikacyjna zależy od wyniku głosowania drzew klasyfikatorów.

Zaletę metody wzmacniania stanowi zdolność do redukcji błędu uczenia, który maleje w tempie wykładniczym. Od innych algorytmów XGBoost odróżnia zastosowanie kary, którą model otrzymuje za zbyt dużą liczbę liści w drzewie decyzyjnym. Takie podejście pozwala kontrolować złożoność modelu.

Algorytm składa się z dwóch części. W pierwszej minimalizuje się funkcję straty, czyli błąd (inaczej – funkcję kosztu), w drugiej kontrolowana jest złożoność modelu, co zapobiega przetrenowaniu.

3.4. Metoda k -najbliższych sąsiadów (KNN)

Model KNN jest jedną z najbardziej intuicyjnych i najprostszych metod uczenia z nadzorem. Nie uczy się do przodu, ale opóźnia uczenie i dokonuje klasyfikacji w momencie, kiedy otrzymuje żądanie zaklasyfikowania nowych danych. Z tego powodu jest określany jako uczenie oparte na przypadkach lub pamięci (ang. *instance-based learning, memory-based learning*). Polega na przypisaniu nowej niezaklasyfikowanej obserwacji do grupy, do której należy większość jej k -najbliższych sąsiadów. Metoda ta efektywnie redukuje błąd klasyfikacji, kiedy liczba obserwacji w próbie w zbiorze treningowym jest duża. Trafność klasyfikacji zależy od przyjętej wartości k , czyli optymalnej liczby sąsiadów, i w dużym stopniu od formuły odległości wykorzystanej do obliczenia dystansu między obiektami. W najprostszej wersji metody KNN dopasowanie zależy od liczebności kategorii najbliższej nowym danym.

3.5. Sztuczna sieć neuronowa (NN)

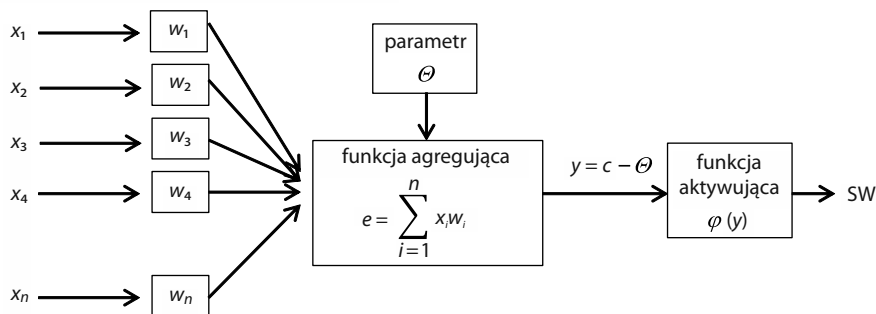
Konstrukcja modelu NN opiera się na podobieństwie do działania mózgu i systemu nerwowego organizmów żywych, czyli sieci neuronów i połączeń pomiędzy nimi. Wagi są modyfikowane podczas uczenia. Do neuronu dociera wiele sygnałów wej-

ściowych x_i , gdzie $i = 1, 2, \dots, n$, ale każdy neuron ma jedno wyjście. Każdej zmiennej wejściowej przypisana jest waga w_i . Po przypisaniu wag początkowych neuron jest aktywowany poprzez funkcję aktywacji (e), będącej sumą ilorazu zmiennych wejściowych i przypisanych wag. Wówczas zmienna y jest wyznaczana jako różnica pomiędzy wartością e a wartością progu Θ . Sygnał wyjściowy zależy od pobudzenia neuronu i funkcji aktywacji $\varphi(y)$. Forma tej funkcji określa typ neuronu (zob. schemat 2).

Sztuczne sieci neuronowe wykazują się elastycznością i szybko adaptują się do zmian. Są odporne na zaszumioną informację wejściową i nie wymagają żadnych założeń, takich jak np. normalność rozkładów. Zmienne objaśniające mogą być jakościowe i ilościowe. Możliwe jest modelowanie każdego rodzaju zależności nieliniowych i nieciągłych.

Model NN ma jednak ograniczenia, a największym jest brak możliwości bezpośredniej interpretacji otrzymanych wyników. Przy bardzo skomplikowanej architekturze sieci proces jej uczenia może być czasochłonny lub nie osiągnąć optymalnej redukcji błędu. Sieć nie selekcjonuje automatycznie zmiennych istotnych dla modelu i jest wrażliwa na przetrenowanie.

Schemat 2. Przykładowa sieć neuronowa



Uwaga. SW – sygnał wyjściowy.

Źródło: opracowanie własne.

3.6. Metoda wektorów wspierających (SVM)

Metoda SVM, nazywana też maszyną wektorów wspierających, opiera się na koncepcji płaszczyzn decyzyjnych, które definiują granice decyzji. Płaszczyzna decyzyjna oddziela obiekty należące do różnych klas. Najprostszym przykładem jest klasyfikator liniowy, który dzieli obiekty na klasy za pomocą linii prostej. Większość zagadnień wymaga jednak zastosowania bardziej skomplikowanych struktur do uzyskania

optymalnego podziału, czyli klasyfikacji obiektów ze zbioru testowego na podstawie zbioru uczącego. Klasyfikator bazujący na oddzielnych liniach do wydzielenia różnych klas obiektów jest nazywany *klasyfikatorem hiperpłaszczyznowym*.

Metoda SVM posługuje się hiperpłaszczyznami w przestrzeni wielowymiarowej do podziału obserwacji na klasy. Wspiera zarówno zadanie regresji, jak i klasyfikacji z wieloma zmiennymi ciągłymi i dyskretnymi – w przypadku tych drugich automatycznie tworzone są zmienne binarne odpowiadające kategoriom. Do skonstruowania optymalnej hiperpłaszczyzny metodą wektorów wspierających wykorzystuje się iteracyjny algorytm uczenia do minimalizacji funkcji błędu.

3.7. Ocena modeli

Skuteczność algorytmu uczenia maszynowego może być oceniona za pomocą macierzy pomyłek (ang. *confusion matrix*), która jest przedstawiana w postaci tabeli. Sprawdzając skuteczność algorytmu na zbiorze testowym, otrzymuje się informacje o błędach, które model popełnił w trakcie nauki na zbiorze treningowym. Jeśli klasyfikator jest binarny, macierz pomyłek ma postać macierzy 2×2 (zestawienie 2).

Zestawienie 2. Macierz pomyłek

Wartości przewidywane	Wartości faktyczne	
	pozytywne	negatywne
Pozytywne	prawdziwe pozytywne (TP)	fałszywe pozytywne (FP)
Negatywne	fałszywe negatywne (FN)	prawdziwe negatywne (TN)

Źródło: opracowanie własne na podstawie: Ghoneim (2019).

Komórki (klasy) macierzy to:

- TP (ang. *true positive*) – liczba przypadków pozytywnych, które zostały poprawnie zaklasyfikowane jako pozytywne;
- FN (ang. *false negative*) – liczba przypadków pozytywnych, które zostały niepoprawnie zaklasyfikowane jako negatywne, czyli popełniono błąd drugiego rodzaju;
- FP (ang. *false positive*) – liczba przypadków negatywnych, które zostały niepoprawnie zaklasyfikowane jako pozytywne, czyli popełniono błąd pierwszego rodzaju;
- TN (ang. *true negative*) – liczba przypadków negatywnych, które zostały poprawnie zaklasyfikowane jako negatywne.

Na podstawie powyższych informacji można wprowadzić następujące miary (Ghoneim, 2019):

- trafność (ang. *accuracy*):

$$\text{trafność} = \frac{TP + TN}{TP + FN + FP + TN}; \quad (3)$$

- błąd klasyfikacji (ang. *error ratio*):

$$\text{błąd klasyfikacji} = \frac{FP + FN}{TP + FN + FP + TN}; \quad (4)$$

- precyzja (ang. *precision*) – udział poprawnych klasyfikacji przypadków pozytywnych w liczbie przewidywanych przypadków pozytywnych ogółem:

$$\text{precyzja} = \frac{TP}{TP + FP}; \quad (5)$$

- czułość (ang. *recall*) – udział poprawnych klasyfikacji przypadków pozytywnych w liczbie faktycznych przypadków pozytywnych ogółem:

$$\text{czułość} = \frac{TP}{TP + FN}; \quad (6)$$

- specyficzność (ang. *specificity*) – udział poprawnych klasyfikacji przypadków negatywnych w liczbie faktycznych przypadków negatywnych ogółem:

$$\text{specyficzność} = \frac{TN}{TN + FP}; \quad (7)$$

- *F-score* – średnia harmoniczna czułości oraz precyzji:

$$F\text{-score} = 2 \cdot \frac{\text{precyzja} \cdot \text{czułość}}{\text{precyzja} + \text{czułość}} \quad (8)$$

Miara ta, ze względu na uśrednienie, jest właściwa w przypadku, gdy wartość precyzji i czułości jest podobna. Wartość *F-score* zawiera się w przedziale od 0 do 1, przy czym 1 oznacza model idealny.

4. Wyniki analizy eksploracyjnej na podstawie *COVID Impact Survey*

Do eksploracji danych wykorzystano analizę skupień, a do budowy modeli predykcyjnych – regresję logistyczną, drzewa decyzyjne, wzmacnianie gradientowe, metodę k -najbliższych sąsiadów, sztuczne sieci neuronowe i metodę wektorów wspierających. Wybór metod był podyktowany częstością ich stosowania w przypadku zagadnień klasyfikacji i predykcji.

4.1. Analiza skupień

Analiza skupień to proces grupowania podobnych do siebie obserwacji w podzbiory wewnętrznie homogeniczne i jednocześnie odróżniające się od pozostałych (heterogeniczne), który zachodzi na podstawie określonych miar podobieństwa i przyjętego algorytmu. Podejście to ma szerokie zastosowanie w analityce biznesowej, a także w badaniach społecznych i demograficznych. Grupowanie ma na celu sortowanie różnych obiektów w taki sposób, aby stopień ich podobieństwa był maksymalny, jeśli należą do tej samej grupy (Panek i Zwierzchowski, 2013).

W przypadku danych pochodzących z ankiety badającej wpływ pandemii COVID-19 na zdrowie psychiczne i społeczne, a także sytuację ekonomiczną i finansową można dzięki segmentacji podzielić respondentów na mniejsze, jednorodnie grupy i określić ich cechy charakterystyczne. To ułatwia uzyskanie odpowiedzi na pytania, czy reakcje ludzi na skutki pandemii były podobne i jakie czynniki okazały się w tych przypadkach istotne.

W omawianym badaniu do analizy skupień wykorzystano wybrane zmienne, głównie quasi-ciągłe:

- SOC3A (Jak często w ciągu ostatniego miesiąca komunikowałeś się z bliskimi telefonicznie lub za pośrednictwem internetu (w tym mailowo)?);
- PHYS8 (Czy swój stan zdrowia ocenilibyś jako: wysmienity, bardzo dobry, dobry, umiarkowany czy zły?);
- SOC5A (Jak często w ciągu ostatnich siedmiu dni byłeś zdenerwowany, niespokojny?);
- SOC5B (Jak często w ciągu ostatnich siedmiu dni miałeś stany depresyjne?);
- SOC5C (Jak często w ciągu ostatnich siedmiu dni czułeś się samotny?);
- ECON4B (Pomyśl o tym, co wydarzy się za trzy miesiące – jak bardzo jest prawdopodobne, że będziesz wówczas zatrudniony?);
- ECON5AA (Wskaż, na ile prawdziwe dla ciebie jest stwierdzenie: „Martwimy się, że skończą się nam zapasy jedzenia, zanim zdobędziemy środki, aby kupić kolejne produkty”);

- ECON5AB (Wskaż, na ile prawdziwe dla ciebie jest stwierdzenie: „Jedzenie, które zakupiliśmy poprzednim razem, nie wystarczyło nam, ale nie mieliśmy środków, aby kupić kolejne produkty”);
- PHYS3H (Czy lekarz lub pracownik medyczny powiedział ci kiedykolwiek, że cierpisz z powodu złego stanu psychicznego?);
- AGE7 (wiek; zmienna ciągła);
- HHINCOME (dochód gospodarstwa domowego).

Po usunięciu obserwacji brakujących pozostałe 6768 poddano analizie. W pierwszym kroku oceniono korelację między zmiennymi. Quasi-ciągły charakter zmiennych pozwolił na zastosowanie współczynnika korelacji liniowej Pearsona. W tabl. 1 zamieszczono wyniki obliczeń dla poszczególnych par wybranych zmiennych.

Tabl. 1. Współczynniki korelacji Pearsona zmiennych quasi-ciągłych wykorzystanych w grupowaniu

Zmienne	SOC3A	PHYS8	SOC5A	SOC5B	SOC5C	ECON4B	ECON5AA	ECON5AB	HHINCOME	HHSIZE1
SOC3A	1,000	0,083	-0,006	-0,023	-0,007	0,095	-0,031	-0,036	-0,102	-0,016
PHYS8	0,083	1,000	0,179	0,167	0,165	0,205	-0,180	-0,179	-0,280	-0,018
SOC5A	-0,006	0,179	1,000	0,595	0,575	-0,011	-0,174	-0,164	-0,097	0,033
SOC5B	-0,023	0,167	0,595	1,000	0,590	-0,027	-0,183	-0,162	-0,102	0,019
SOC5C	-0,007	0,165	0,575	0,590	1,000	-0,033	-0,202	-0,172	-0,103	0,027
ECON4B	0,095	0,205	-0,011	-0,027	-0,033	1,000	-0,045	-0,050	-0,214	-0,186
ECON5AA ...	-0,031	-0,180	-0,174	-0,183	-0,202	-0,045	1,000	0,806	0,411	-0,144
ECON5AB ...	-0,036	-0,179	-0,164	-0,162	-0,172	-0,050	0,806	1,000	0,386	-0,124
HHINCOME	-0,102	-0,280	-0,097	-0,102	-0,103	-0,214	0,411	0,386	1,000	0,104
HHSIZE1	-0,016	-0,018	0,033	0,019	0,027	-0,186	-0,144	-0,124	0,104	1,000

Uwaga. Pogrubieniem zaznaczono zmienne silnie skorelowane.

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Analizując te wyniki, można dostrzec pewne związki między zmiennymi. Dla przykładu zmienna HHINCOME wykazuje ujemną korelację ze zmienną PHYS8 – wyższe dochody gospodarstwa domowego częściej wiążą się z ocenami stanu zdrowia jako wysmienity i bardzo dobry. Zbadanie korelacji zmiennych w zbiorze stanowiło ułatwienie przy wyborze zmiennych do segmentacji (pominięcie zmiennych silnie skorelowanych).

Skala wartości wybranych zmiennych nie była bardzo zróżnicowana, w związku z czym nie przeprowadzono ich standaryzacji, ponieważ nie poprawiłaby ona wyników analizy. Do grupowania wykorzystano natomiast metodę hierarchiczną Warda, która opiera się na łączeniu skupień o minimalnej sumie kwadratów odchyłeń między środkami ciężkości tych skupień. Wykorzystano kryterium CCC (Cubic Clustering Criterion) oraz miary pseudo- t^2 i pseudo- R . Kierując się tymi przesłankami,

zdecydowano się na utworzenie pięciu skupień. Ponadto – w celu pogłębienia analizy – zastosowano iteracyjną metodę *k*-średnich.

Aby lepiej określić profil osób znajdujących się w poszczególnych skupieniach, obliczono średnie wartości zmiennych w grupach (tabl. 2).

Tabl. 2. Średnie wartości zmiennych w poszczególnych skupieniach

Zmienne	Skupienia				
	1	2	3	4	5
SOC3A	1,424	1,387	1,430	1,788	1,324
PHYS8	2,317	1,965	2,555	1,788	2,446
SOC5A	1,510	1,419	1,831	1,788	2,026
SOC5B	1,473	1,416	1,891	1,292	2,044
SOC5C	1,445	1,418	1,902	1,260	2,022
ECON4B	4,668	1,396	2,352	4,718	1,989
ECON5AA	2,921	2,963	2,372	2,669	2,695
ECON5AB	2,955	2,985	2,487	2,725	2,777
PHYS3H	0,114	0,117	0,255	0,155	0,246
HHINCOME	6,850	7,542	2,791	2,891	6,377
HHSIZE1	1,995	2,341	2,620	1,465	3,756

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Po przeanalizowaniu wyników dotyczących poszczególnych skupień można zauważyć, że respondenci w dwóch licznych grupach (skupienie 1 i 2) rzadko deklaruwali problemy ze zdrowiem psychicznym czy fizycznym (zob. tabl. 3). Widoczna jest także zależność pomiędzy dochodami gospodarstwa domowego a oceną stanu zdrowia ankietowanych.

Tabl. 3. Liczebność grup i opis cech charakterystycznych wyróżnionych skupień

Skupienie	Liczebność	Opis
1	1336	osoby o stosunkowo wysokich dochodach i dobrym zdrowiu psychicznym, jednak często obawiające się o swoje zatrudnienie w najbliższym czasie
2	2187	osoby o najwyższych dochodach, najlepiej oceniające swój stan zdrowia, najmniej martwiące się o zatrudnienie w najbliższym czasie
3	1578	osoby o najniższych dochodach, często deklaruujące problemy finansowe, najgorzej oceniające stan swojego zdrowia psychicznego
4	709	osoby w wielu przypadkach samotne, martwiące się o zatrudnienie, nie najgorzej oceniające stan swojego zdrowia
5	958	osoby otoczone bliskimi, często deklaruujące uczucie zdenerwowania czy niepokojem, stosunkowo często sygnalizujące gorszy stan zdrowia psychicznego

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Interesująca jest zmienna ECON4B – oceniane przez respondentów prawdopodobieństwo bycia zatrudnionym w ciągu najbliższych trzech miesięcy. Dotyczące jej wyniki mogą mieć związek z aktualną sytuacją pandemiczną. Osoby najbardziej

dotknięte problemami z zatrudnieniem wypadały najgorzej w odpowiedziach na pytania dotyczące stanów depresyjnych, samotności i niepokoju. Oznacza to, że negatywne skutki pandemii miały wpływ na zdrowie psychiczne wielu osób.

4.2. Analiza i kategoryzacja zmiennych

Dla części zmiennych przeprowadzono kategoryzację. Dzięki łączeniu nieznacznie różniących się kategorii, także pod względem rozkładu zmiennej objaśnianej, usunięto grupy o niskiej liczebności (poniżej 5%). Zmienne AGE7 i HHINCOME skategoryzowano, aby uzyskać monotoniczność empirycznego logitu. Zmienne ostatecznie uwzględnione w modelu przedstawiono w zestawieniu 3.

Zestawienie 3. Zmienne objaśniające po kategoryzacji

Zmienne	Opis	Wartości	Typ
PHYS8	Czy swój stan zdrowia oceniałbyś jako: wyśmienity, bardzo dobry, dobry, umiarkowany czy zły?	(1) Wyśmienity (2) Bardzo dobry (3) Dobry (4) Umiarkowany (5) Zły	nominalna
PHYS1K	Czy w ciągu ostatnich siedmiu dni doświadczyłeś uczucia zmęczenia lub znużenia?	(1) Tak (0) Nie	binarna
PHYS1Q	Czy w ciągu ostatnich siedmiu dni doświadczyłeś utraty apetytu?	(1) Tak (0) Nie	binarna
SOC5A	Jak często w ciągu ostatnich siedmiu dni byłeś zdenerwowany, niespokojny?	(0) Prawie wcale (1) Co najmniej jeden dzień	binarna
SOC5C	Jak często w ciągu ostatnich siedmiu dni czułeś się samotny?	(0) Prawie wcale (1) Co najmniej jeden dzień	binarna
SOC5D	Jak często w ciągu ostatnich siedmiu dni czułeś brak nadziei, myśląc o przyszłości?	(0) Prawie wcale (1) Co najmniej jeden dzień	binarna
ECON6B	Czy w ciągu ostatnich siedmiu dni otrzymałeś dodatkową pomoc żywnościową, zawnioskowałeś o nią lub starałeś się zakwalifikować do programu dodatkowej pomocy żywnościowej (<i>Supplemental Nutrition Program</i>)?	(1) Zawnioskowałem lub planuję to zrobić (0) Nie wnioskowałem	binarna
ECON7_1	Czy obecnie, aby pokryć dodatkowe wydatki, musiałbyś skorzystać ze swojej karty kredytowej i mógłbyś spłacić dług dopiero w następnym okresie rozliczeniowym?	(1) Tak (0) Nie	binarna
PHYS2_18	Czy uważasz, że gdy zostajesz w domu z powodu złego samopoczucia, wynika to z panowania pandemii koronawirusa?	(1) Tak (0) Nie	binarna
PHYS3M	Czy stwierdzono u ciebie nadwagę lub otyłość?	(1) Tak (0) Nie	binarna

Zestawienie 3. Zmienne objaśniające po kategoryzacji (dok.)

Zmienne	Opis	Wartości	Typ
ECON6A	Czy w ciągu ostatnich siedmiu dni otrzymałeś zasiłek dla bezrobotnych, zawnioskowałeś o taki zasiłek lub starałeś się go otrzymać?	(1) Zawnioskowałem lub zamierzam zawnioskować (0) Nie zawnioskowałem	binarna
AGE7	Wiek w latach	(1) 18–24 (2) 25–34 (3) 35–44 (4) 45–54 (5) 55–64 (6) 65–74 (7) 75+	porządkowa
HHINCOME	Dochód gospodarstwa domowego (roczny) w tys. USD	(1) poniżej 10 (2) 10–20 (3) 20–30 (4) 30–40 (5) 40–50 (6) 50–75 (7) 75–100 (8) 100–150 (9) 150 i więcej	porządkowa

Źródło: opracowanie własne.

4.3. Model regresji logistycznej

Pierwszy etap polegał na zbudowaniu modelu z efektami głównymi, co pozwoliło na badanie istotności zmiennych objaśniających. Identyfikowano potencjalne zmienne zakłócające i zmienne do usunięcia z modelu. Przypomnijmy, że zmienną objaśnianą jest zmienna binarna PHYS3H. Przyjmuje ona wartość 1, gdy dana osoba ma zdiagnozowaną chorobę o podłożu psychicznym.

Za zmienną zakłócającą uznawano taką, której usunięcie wpływało na zmianę oszacowania parametru przy zmiennej o ponad 10%. W przypadku gdy usunięcie wybranej zmiennej nie oddziaływało wyraźnie na oszacowanie ocen parametrów przy pozostałych zmiennych, była ona kwalifikowana jako możliwa do usunięcia z modelu.

W kolejnych krokach z modelu usuwano zmienne nieistotne statystycznie, w pierwszej kolejności zmienne o wyższym p -value w teście istotności. W związku z tym, że jedną z takich zmiennych jest HHINCOME, istotna z punktu widzenia badania, postanowiono nie eliminować jej z modelu, a jedynie zbadać jej wpływ na zmienną ECON6A.

Następnie zbadano efekt usunięcia zmiennej PHYS1K. Jej usunięcie nie wpłynęło znacząco na oszacowania parametrów pozostałych zmiennych, co oznacza, że PHYS1K jest zmienną możliwą do usunięcia z modelu. Zdecydowano się również zbadać wpływ, jaki wywarłoby usunięcie zmiennych istotnych statystycznie na osza-

cowania parametrów pozostałych zmiennych. Redukując model kolejno o zmienne ECON6A, ECON7_1, PHYS2_18, PHYS8, ECON6B, PHYS3M i AGE7, zidentyfikowano zmiany oszacowania parametrów przy wielu zmiennych. Z tego powodu zbadano wprowadzenie do modelu 29 interakcji.

Dalszym etapem budowy modelu było wprowadzenie interakcji drugiego rzędu, wykrytych na podstawie różnic oszacowania parametrów wynikających z usunięcia wybranej zmiennej. Zastosowano metodę selekcji krokowej, która polega na wprowadzaniu lub usuwaniu zmiennych z modelu. Wymuszono dodanie zmiennej HHINCOME, aby zyskać szansę na zidentyfikowanie efektów zmiennych modyfikujących. Dla tego testu został przyjęty poziom istotności 0,05.

Wyniki analizy, przedstawione w tabl. 4, wskazują na istotność jednej interakcji – ECON6A i ECON6B – oraz dziesięciu innych zmiennych objaśniających. Jednak żadna ze zmiennych możliwych do usunięcia z modelu nie stworzyła interakcji istotnych statystycznie, przez co w modelu nie wystąpiły żadne zmienne modyfikujące efekty główne. Dołączenie interakcji do modelu sprawiło, że nieistotna statystycznie okazała się zmienna PHYS2_18, mimo że nie stworzyła z innymi zmiennymi żadnych interakcji.

Tabl. 4. Analiza istotności zmiennych dla modelu z interakcjami

Zmienne	Stopnie swobody	Chi-kwadrat Walda	p-value
Zasiłek dla bezrobotnych (ECON6A)	1	8,4	0,0038
Dochód gospodarstwa (HHINCOME)	1	0,7	0,3975
Stan zdrowia (PHYS8)	4	42,3	<0,0001
Uczucie zmęczenia (PHYS1K)	1	3,4	0,0637
Utrata apetytu (PHYS1Q)	1	14,5	0,0001
Zdenerwowanie lub niepokój (SOC5A)	1	19,9	<0,0001
Samotność (SOC5C)	1	37,7	<0,0001
Brak nadziei na przyszłość (SOC5D)	1	32,8	<0,0001
Program pomocy żywieniowej (ECON6B)	1	14,7	0,0001
Dług na karcie kredytowej (ECON7_1)	1	4,3	0,0383
Wpływ pandemii koronawirusa na złe samopoczucie (PHYS2_18)	1	3,8	0,0512
Nadwaga lub otyłość (PHYS3M)	1	31,1	<0,0001
Wiek (AGE7)	1	69,9	<0,0001
Zasiłek dla bezrobotnych wśród osób korzystających z programu pomocy żywieniowej (ECON6B · ECON6A)	1	5,5	0,0193

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Do oceny jakości modelu wykorzystano miary oparte na macierzy pomyłek na zbiorze testowym. Warto zwrócić uwagę na to, że przygotowanie modelu nie służyło uzyskaniu maksymalnej wartości predykcji. Jego celem była analiza wpływu pandemii COVID-19 na stan zdrowia psychicznego respondentów.

Model regresji logistycznej cechuje bardzo wysoka (na poziomie ponad 86%) trafność, czyli udział poprawnych klasyfikacji. Przy wysokiej precyzji (ponad 73%) uzyskano czułość na poziomie 31%, co oznacza, że 31% faktycznie chorych respondentów model prawidłowo zaklasyfikował jako chorych. Natomiast specyficzność modelu na poziomie 98% wskazuje, że dobór charakterystyk daje lepsze możliwości klasyfikacji respondentów zdrowych (tabl. 5).

Tabl. 5. Ocena modelu regresji logistycznej

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	89	33	.
zdrowi	197	1345	.
Trafność	0,86
Precyzja	0,73
Czułość	0,31
Specyficzność	0,98
<i>F-score</i>	0,43

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Ze względu na wyraźną dysproporcję pomiędzy grupą respondentów chorych a grupą respondentów zdrowych wyniki oceny modelu należy jednak interpretować ostrożnie. Wartość *F-score* jest w pełni interpretowalna tylko przy podobnej wartości precyzji i czułości.

Wyniki finalnego modelu zostały poddane analizie i interpretacji. W pierwszej kolejności zbadano oceny parametrów i ilorazów szans (tabl. 6). Jediną zmienną, która okazała się nieistotna statystycznie na poziomie istotności 0,1, jest HHINCOME.

Tabl. 6. Oceny parametrów i ilorazy szans w modelu regresji

Zmienne	Ocena parametru	Iloraz szans
Wyraz wolny	-1,77	.
Zasiłek dla bezrobotnych (ECON6A)	-1,00	0,37
Dochód gospodarstwa (HHINCOME)	-0,02	0,98
Stan zdrowia (PHYS8) ^a : (1) Wyśmienity	-0,52	0,59
(3) Dobry	0,25	1,29
(4) Umiarkowany	0,54	1,72
(5) Zły	0,83	2,30
Utrata apetytu (PHYS1Q)	0,48	1,62
Zdenerwowanie lub niepokój (SOC5A)	0,23	1,26
Samotność (SOC5C)	0,33	1,39
Brak nadziei na przyszłość (SOC5D)	0,30	1,35
Program pomocy żywieniowej (ECON6B)	-0,19	0,83
Dług na karcie kredytowej (ECON7_1)	-0,20	0,82
Wpływ pandemii koronawirusa na złe samopoczucie (PHYS2_18)	0,25	1,28

Tabl. 6. Oceny parametrów i ilorazy szans w modelu regresji (dok.)

Zmienne	Ocena parametru	Iloraz szans
Nadwaga lub otyłość (PHYS3M)	0,51	1,66
Wiek (AGE7)	-0,22	0,80
Zasiłek dla bezrobotnych wśród osób korzystających z programu pomocy żywieniowej (ECON6B · ECON6A)	0,23	.

a Kategoria referencyjna: Stan zdrowia (PHYS8) – (2) Bardzo dobry.

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Powyższe wyniki pozwalają wyciągnąć wiele wniosków. U osób oceniających swój stan zdrowia jako wysmienity szanse zdiagnozowania chorób o podłożu psychicznym są o ponad 40% mniejsze niż u osób określających swój stan zdrowia jako bardzo dobry (przy innych czynnikach niezmiennych). U osób oceniających swój stan zdrowia jako dobry szanse te są o prawie 30% wyższe, natomiast u osób oceniających swoje zdrowie najslabiej – blisko dwuipółkrotnie wyższe (przy założeniu *ceteris paribus*) niż u osób określających swój stan zdrowia jako bardzo dobry.

Zmienna HHINCOME okazała się statystycznie nieistotna. Zmienna ECON6A została włączona do modelu w interakcji ze zmienną ECON6B, co utrudnia interpretację tej drugiej.

W przypadku zmiennej PHYS1Q okazało się, że szanse zdiagnozowania choroby psychicznej są ponadpółtorakrotnie wyższe u osoby, która w ciągu ostatnich siedmiu dni doświadczyła utraty apetytu, niż u osoby, która nie miała takiego problemu (przy założeniu *ceteris paribus*).

Dla zmiennych SOC5A, SOC5C i SOC5D otrzymano następujące wyniki: u osób, które w ciągu ostatnich siedmiu dni były zdenerwowane lub zaniepokojone, szanse zdiagnozowania chorób o podłożu psychicznym są o ok. 26% wyższe, u osób, które w ostatnim czasie czuły się samotne – o blisko 40% wyższe, a u osób martwiących się o swoją przyszłość – o ponad 35% wyższe niż u respondentów, którzy takich problemów nie zgłosili (przy innych czynnikach niezmiennych).

W przypadku osób, u których stwierdzono nadwagę lub otyłość, szanse zdiagnozowania choroby psychicznej są o ponad 66% wyższe niż w przypadku osób z prawidłową wagą lub niedowagą. Natomiast wśród tych, którzy nieprzewidziane wydatki musieli opłacać z karty kredytowej, szanse te są o ok. 18% niższe.

Jeśli chodzi o zmienną AGE7, to prawdopodobieństwo zdiagnozowania choroby psychicznej spada z wiekiem o 20% rocznie (*ceteris paribus*).

Bardzo istotna z punktu widzenia analizy jest zmienna PHYS2_18. Szanse zdiagnozowania choroby o podłożu psychicznym u osób, które zauważają u siebie gorsze samopoczucie spowodowane pandemią, są o blisko 30% wyższe niż u osób, które tego nie dostrzegają (przy innych czynnikach niezmiennych).

4.4. Model drzewa decyzyjnego

W budowie drzewa decyzyjnego posłużono się metodą przycinania *prune costcomplexity*, która wykorzystuje analizę kosztów na podstawie walidacji krzyżowej. Miarą błędu jest współczynnik błędnych klasyfikacji. Liczba liści w modelu wyniosła 10. Statystyki dopasowania drzewa przedstawiono w tabl. 7.

Tabl. 7. Ocena modelu drzewa decyzyjnego

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	40	17	.
zdrowi	246	1361	.
Trafność	0,84
Precyzja	0,70
Czułość	0,14
Specyficzność	0,99
F-score	0,23

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Metoda przycinania drzewa niemal całkowicie niweluje problem nadmiernego dopasowania do zbioru treningowego, ponieważ statystyki dopasowania osiągają bardzo zbliżone wyniki zarówno w zbiorze treningowym, jak i testowym.

Model drzewa decyzyjnego cechuje wysoka trafność (ponad 84%), ale bardzo niska czułość, czyli zdolność do poprawnej klasyfikacji respondentów chorych. Specyficzność modelu bliska 1 wskazuje na wysoką precyzję poprawnej klasyfikacji respondentów zdrowych.

4.5. Model XGBoost

Użycie modelu XGBoost wiąże się w pierwszej kolejności z doбором odpowiedniego wzmacniacza, na podstawie którego dobierane są pozostałe parametry. W omawianym przypadku najlepsze wyniki uzyskano dla wzmacniacza *gbtree* wraz z maksymalną głębokością drzewa oraz minimalną sumą wag instancji ustawioną na 2.

Wprowadzenie losowości do drzew decyzyjnych i oparcie modelu na algorytmie XGBoost przyniosło wzrost trafności modelu i ponaddwukrotny wzrost czułości w porównaniu z drzewem decyzyjnym (tabl. 8).

Tabl. 8. Ocena modelu XGBoost

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	82	32	.
zdrowi	204	1346	.
Trafność	0,86
Precyzja	0,72
Czułość	0,29
Specyficzność	0,98
F-score	0,41

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

4.6. Model k -najbliższych sąsiadów

Zastosowanie algorytmu k -najbliższych sąsiadów jest związane głównie z odpowiednim doбором parametru k wyznaczającego liczbę sąsiadów, do których odległość od danego elementu zbioru jest najmniejsza. Dobór wartości parametru k należy przeprowadzić ostrożnie, ponieważ zbyt duża jego wartość może się przyczynić do rozmycia heterogeniczności obszarów podziału, a także skutkować występowaniem większej liczby błędów klasyfikacji rzadszych wzorców. W trakcie eksperymentów przetestowano wartości k z przedziału od 1 do 30, a najlepsze wyniki uzyskano dla $k = 24$.

Tabl. 9. Ocena modelu k -najbliższych sąsiadów

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	53	15	.
zdrowi	233	1363	.
Trafność	0,85
Precyzja	0,78
Czułość	0,19
Specyficzność	0,99
F-score	0,30

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Model k -najbliższych sąsiadów, podobnie jak model drzewa decyzyjnego, charakteryzuje się niską czułością i bardzo wysoką specyficznością (tabl. 9).

4.7. Model sztucznej sieci neuronowej

W zastosowanej w badaniu sztucznej sieci neuronowej każdy neuron z poprzedniej warstwy jest połączony z każdym neuronem z warstwy kolejnej. Występują tu trzy warstwy; pierwsza składa się z 26 neuronów, druga – z 13, a ostatnia (wyjściowa)

– z 1. Jako funkcję agregującą przyjęto funkcję liniową, a jako funkcję aktywacji – funkcję logitową.

Tabl. 10. Ocena modelu sztucznej sieci neuronowej

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	95	37	.
zdrowi	191	1341	.
Trafność	0,86
Precyzja	0,72
Czułość	0,33
Specyficzność	0,97
F-score	0,45

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Okazało się, że model sztucznej sieci neuronowej cechuje się wysoką trafnością, ale jednocześnie najwyższą czułością i wysoką specyficznością. Model poprawnie klasyfikuje chorych w 33% przypadków (tabl. 10).

4.8. Model oparty na metodzie wektorów wspierających

Przeprowadzono eksperymenty, w których przetestowano kilka rodzajów funkcji podziału przestrzeni decyzyjnej w celu wyznaczenia granicy przynależności klasowej poszczególnych punktów, m.in. liniową, sigmoidalną, wielowymiarową oraz radialną funkcję bazową, dla której uzyskano najlepsze wyniki klasyfikacji.

Tabl. 11. Ocena modelu opartego na metodzie wektorów wspierających

Wyszczególnienie	Klasa faktyczna		Wartość miary
	chorzy	zdrowi	
Klasa przewidywana: chorzy	35	8	.
zdrowi	251	1370	.
Trafność	0,84
Precyzja	0,81
Czułość	0,12
Specyficzność	0,99
F-score	0,21

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Czułość modelu opartego na metodzie wektorów wspierających okazała się najniższa. Model najlepiej rozpoznaje i klasyfikuje respondentów zdrowych, natomiast w przypadku respondentów chorych jego trafność klasyfikacji jest najniższa (tabl. 11).

4.9. Porównanie modeli

Do porównania modeli zastosowano miarę precyzji i czułości oraz uśrednioną miarę trafności predykcji, jaką jest *F1-score* na zbiorze testowym.

Tabl. 12. Porównanie modeli

Miary	RL	DT	XGBoost	KNN	NN	SVM
Trafność	0,86	0,84	0,86	0,85	0,86	0,84
Precyzja	0,73	0,70	0,72	0,78	0,72	0,81
Czułość	0,31	0,14	0,29	0,19	0,33	0,12
<i>F1-score</i>	0,43	0,23	0,41	0,30	0,45	0,21

Uwaga. Najwyższe wartości zaznaczono pogrubieniem.

Źródło: opracowanie własne z wykorzystaniem SAS Enterprise Guide.

Wyniki zastosowania metod uczenia maszynowego charakteryzują się tylko nieznacznie większą precyzją klasyfikacji niż wyniki regresji logistycznej. Gdy weźmie się pod uwagę wszystkie zastosowane modele, to okazuje się, że najlepsze (ale porównywalne z regresją) wyniki daje sztuczna sieć neuronowa (tabl. 12).

5. Podsumowanie

Głównym celem badania omawianego w artykule było ustalenie wpływu pandemii COVID-19 na stan zdrowia psychicznego członków gospodarstw domowych. Przeprowadzone analizy umożliwiły m.in. ocenę oddziaływania skutków rozprzestrzeniania się koronawirusa na gospodarstwa domowe. To ważne, ponieważ badacze mogą przekazać zdobytą wiedzę decydentom i opinii publicznej. Analizowano głównie wyniki badań i dane pochodzące ze Stanów Zjednoczonych, gdzie w ramach prac zleconych przez rząd i organizacje pozarządowe przeprowadzono ankiety dotyczące statusu zatrudnienia, bezpieczeństwa żywnościowego, warunków mieszkaniowych, zdrowia fizycznego, dostępu do opieki zdrowotnej czy problemów związanych z edukacją.

Jak twierdzą badacze publikujący w dzienniku medycznym „The Lancet”, u osób, które przed wybuchem pandemii nie miały zaburzeń depresyjnych, lękowych lub obsesyjno-kompulsyjnych, w czasie jej trwania stwierdzono dość intensywne występowanie ich objawów. Co więcej, w czerwcu 2020 r. 40% dorosłych Amerykanów doświadczało zaburzeń psychicznych, a 11% ujawniło, że w ostatnim miesiącu poważnie rozważało podjęcie próby popełnienia samobójstwa.

Warto podkreślić, że choroby psychiczne są jednym z bagatelizowanych skutków pandemii COVID-19. Izolacja, zdalna edukacja i praca, ograniczenie kontaktów

międzyludzkich czy zmniejszona aktywność fizyczna, a w czasie pierwszej fali pandemii także panika zakupowa i strach przed zachorowaniem, wycisnęły piętno na kondycji psychicznej ludzi na całym świecie. Skutkiem tych zjawisk będą ukryte koszty pandemii, które będziemy ponosić w przyszłości. Wyniki badań dotyczących omawianej tematyki mogą okazać się pomocne podczas opracowywania strategii łagodzących negatywne skutki pandemii.

Analiza skupień pozwoliła na stworzenie pięciu grup, z których dwie najbardziej liczne obejmowały osoby oceniające swój stan zdrowia jako wyśmienity lub bardzo dobry. W tych skupieniach średnia dochodów gospodarstw domowych była najwyższa, co skłoniło badaczy do przeprowadzenia dalszej analizy sytuacji finansowej respondentów i zbadania jej związku z możliwą chorobą o podłożu psychicznym. Z tego względu do modeli została wprowadzona zmienna HHINCOME, określająca dochody ankietowanych.

Kolejny interesujący wynik uzyskany na podstawie grupowania jest taki, że często stany niepokoju lub zdenerwowania identyfikowano u respondentów otoczonych bliskimi – mieszkających z kilkoma innymi osobami lub utrzymujących częsty kontakt z rodziną. Może to oznaczać, że nie tylko samotność, lecz także obawa o najbliższych wpływają na stan zdrowia psychicznego. Natomiast w skupieniu, w którym znalazło się wiele osób obawiających się o swoje zatrudnienie, samoocena stanu zdrowia dokonana przez respondentów była dobra.

W badaniu zastosowano także modele data mining oraz model regresji logistycznej. Uzyskane wyniki unaocniły, że wśród osób, które w ciągu ostatnich siedmiu dni doświadczyły uczucia zdenerwowania lub niepokoju, szanse zdiagnozowania chorób o podłożu psychicznym są o ok. 26% wyższe wśród tych, którzy w ostatnim czasie czuli się samotni (o blisko 40%) i martwili się o swoją przyszłość (o ponad 35%) niż wśród osób, które oceniły swój stan zdrowia jako bardzo dobry (*ceteris paribus*). Dodatkowo szanse te rosną znacząco w przypadku osób otyłych lub z nadwagą. Izolacja, zdalny tryb nauczania i pracy oraz zmniejszona aktywność fizyczna niewątpliwie przyczyniają się zatem do pogarszania się stanu zdrowia psychicznego.

Zastosowane w badaniu modele data mining dostarczyły wyników tylko nieznacznie lepszych, o wyższej precyzji klasyfikacji, niż model regresji logistycznej. Z porównania wszystkich zastosowanych modeli wynika, że najlepsze – ale porównywalne z rezultatami regresji – wyniki uzyskano za pomocą sztucznej sieci neuro-
nowej.

Bibliografia

- Alpaydin, E. (2004). *Introduction to Machine Learning*. MIT Press.
- American Psychiatric Association. (2019, April). *Warning Signs of Mental Illness*. <https://psychiatry.org/patients-families/warning-signs-of-mental-illness>.
- Brown, G. W., Moran, P. M. (1997). Single mothers, poverty and depression. *Psychological medicine*, 27(1), 21–33. <https://doi.org/10.1017/s0033291796004060>.
- Chetty, R., Friedman, J. N., Hendren, N., Stepner, M., The Opportunity Insights. (2020). *The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data* (NBER Working Paper 27431). <http://www.nber.org/papers/w27431>.
- COVID Impact Survey. (b.r.). *Reliable information about the impacts of the COVID-19 pandemic*. Pobrane 10 października 2021 r. z <https://www.covid-impact.org/>.
- Czeisler, M. É., Lane, R. I., Petrosky, E., Wiley, J. F., Christensen, A., Njai, R., Weaver, M. D., Robbins, R., Facer-Childs, E. R., Barger, L. K., Czeisler, C. A., Howard, M. E., Rajaratnam, S. M. W. (2020). *Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic*. Centers for Disease Control and Prevention.
- European Foundation for the Improvement of Living and Working Conditions. (2021). *Living, working and COVID-19 (Update April 2021): Mental health and trust decline across EU as pandemic enters another year*. Publications Office of the European Union. <https://doi.org/10.2800/76802>.
- Frątczak, E. (red.). (2012). *Zaawansowane metody analiz statystycznych*. Oficyna Wydawnicza Szkoły Głównej Handlowej w Warszawie.
- Ghoneim, S. (2019, April 2). *Accuracy, Recall, Precision, F-Score & Specificity, which to optimize on?*. <https://towardsdatascience.com/accuracy-recall-precision-f-score-specificity-which-to-optimize-on-867d3f11124>.
- Hastie, T., Tibshirani, R., Friedman, J. H. (2001). *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer-Verlag.
- Healthline. (b.r.). *Sings of Depression*. <https://www.healthline.com/health/depression/recognizing-symptoms#lostinterest>.
- Kanter, J. W., Busch, A. M., Weeks, C. E., Landes, S. J. (2008). The Nature of Clinical Depression: Symptoms, Syndromes, and Behavior Analysis. *The Behavior Analyst*, 31(1), 1–21. <https://doi.org/10.1007/BF03392158>.
- Kennedy, S. H. (2008). Core symptoms of major depressive disorder: relevance to diagnosis and treatment. *Dialogues in Clinical Neuroscience*, 10(3), 271–277. <https://doi.org/10.31887/DCNS.2008.10.3/shkennedy>.
- Kim, H. S., Sohn, S. Y. (2010). Support Vector Machines for Default Prediction of SMEs Based on Technology Credit. *European Journal of Operational Research*, 201(3), 838–846. <https://doi.org/10.1016/j.ejor.2009.03.036>.
- Larose, D. T. (2008). *Metody i modele eksploracji danych*. Wydawnictwo Naukowe PWN.
- Lasek, M., Pęczkowski, M. (2013). *Enterprise Miner. Wykorzystanie narzędzi Data Mining w systemie SAS*. Wydawnictwa Uniwersytetu Warszawskiego. <https://doi.org/10.31338/uw.9788323527701>.
- National Institute of Mental Health. (2021). *Chronic Illness and Mental Health Recognizing and Treating Depression*. https://www.nimh.nih.gov/sites/default/files/documents/health/publications/chronic-illness-mental-health/21-mh-8015-chronicillness-mentalhealth_1.pdf.

- Pan, K.-Y., Kok, A. A. L., Eikelenboom, M., Horsfall, M., Jörg, F., Luteijn, R. A., Rhebergen, D., van Oppen, P., Giltay, E. J., Penninx, B. W. J. H. (2021). The mental health impact of the COVID-19 pandemic on people with and without depressive, anxiety or obsessive-compulsive disorders: a longitudinal study of three Dutch case-control cohorts. *The Lancet. Psychiatry*, 8(2), 121–129. [https://doi.org/10.1016/S2215-0366\(20\)30491-0](https://doi.org/10.1016/S2215-0366(20)30491-0).
- Panek, T., Zwierzchowski, J. (2013). *Statystyczne metody wielowymiarowej analizy porównawczej. Teoria i zastosowania*. Oficyna Wydawnicza Szkoły Głównej Handlowej w Warszawie.
- Pereira-Miranda, E., Costa, P. R. F., Queiroz, V. A. O., Pereira-Santos, M., Santana, M. L. P. (2017). Overweight and Obesity Associated with Higher Depression Prevalence in Adults: A Systematic Review and Meta-Analysis. *Journal of the American Collage of Nutrition*, 36(3), 223–233. <https://doi.org/10.1080/07315724.2016.1261053>.
- Rokach, L., Maimon, O. (2014). *Data mining with decision trees. Theory and Applications* (2nd edition). World Scientific Publishing.
- Sarma, K. S. (2007). *Predictive Modeling with SAS Enterprise Miner. Practical Solutions for Business Applications*. SAS Institute.
- Scope Ratings. (2020, February 28). *Scope affirms China's sovereign rating at A+ and maintains the Outlook at Negative*. Pobrane 11 marca 2020 r. z <https://scoperatings.com/#!search/research/detail/162598EN>.
- Surkova, E., Nikolayevsky, V., Drobniewski, F. (2020). False-positive COVID-19 results: hidden problems and costs. *The Lancet. Respiratory Medicine*, 8(12), 1167–1168. [https://doi.org/10.1016/S2213-2600\(20\)30453-7](https://doi.org/10.1016/S2213-2600(20)30453-7).
- Turner, B. (2020, April 3). *'Most significant crisis in the history of travel': where to now for tourism?*. <https://www.smh.com.au/business/the-economy/most-significant-crisis-in-the-history-of-travel-where-to-now-for-tourism-20200227-p5450j.html>.
- Tyko, K., Guynn, J., Snider, M. (2020, February 28). *Coronavirus fears empty store shelves of toilet paper, bottled water, masks as shoppers stock up*. <https://eu.usatoday.com/story/money/2020/02/28/coronavirus-2020-preparation-more-supply-shortages-expected/4903322002/>.
- Weeks, D. G., Michela, J. L., Bragg, M. E. (1980). *Relation Between Loneliness and Depression: A Structural Equation Analysis*. American Psychological Association.
- World Food Programme. (2021, May 5). *Global Report on Food Crises – 2021*. <https://www.wfp.org/publications/global-report-food-crises-2021>.
- World Health Organization. (b.r.). *Coronavirus disease (COVID-19) pandemic*. Pobrane 10 października 2021 r. z <https://www.who.int/>.
- World Health Organization. (2020, October 5). *COVID-19 disrupting mental health services in most countries, WHO survey*. <https://www.who.int/news/item/05-10-2020-covid-19-disrupting-mental-health-services-in-most-countries-who-survey>.
- Wozniak, A., Willey, J., Benz, J., Hart, N. (2020). *COVID Impact Survey*. National Opinion Research Center.
- Yelp Economic Average. (2020, September). *Local Economic Impact Report*. <https://www.yelpeconomicaverage.com/business-closures-update-sep-2020>.
- Yuen, K. F., Wang, X., Ma, F., Li, K. X. (2020). The Psychological Causes of Panic Buying Following a Health Crisis. *International Journal of Environmental Research and Public Health*, 17(10), 1–14. <https://doi.org/10.3390/ijerph17103513>.

Satelitarne rozpoznawanie upraw i szacowanie ich powierzchni w ramach systemu SATMIROL

Magdalena Mleczek^a, Przemysław Slesiński^b, Tomasz Milewski^c,
Artur Łączyński^d, Dariusz Miziołek^e, Edyta Woźniak^f, Jędrzej Bojanowski^g

Streszczenie. Jednym z ważnych obszarów zastosowania satelitarnych obserwacji Ziemi jest statystyka rolnictwa. Wykorzystanie nowych technologii cyfrowych i danych satelitarnych przyczyniło się do jej unowocześnienia i pomaga rozwiązywać problemy, z jakimi mierzy się ona obecnie oraz z jakimi borykają się odbiorcy danych statystycznych z zakresu rolnictwa. Głównym celem niniejszego artykułu jest zaprezentowanie innowacyjnego systemu satelitarnej identyfikacji i monitorowania upraw rolnych – SATMIROL. System ten został opracowany przez Główny Urząd Statystyczny, Instytut Geodezji i Kartografii oraz Centrum Badań Kosmicznych Polskiej Akademii Nauk na potrzeby statystyki rolnictwa z wykorzystaniem danych satelitarnych z programu Copernicus i danych administracyjnych. W opracowaniu omówiono wybrane moduły tego systemu, w których metody teledetekcji satelitarnej znalazły zastosowanie do rozpoznawania upraw. Przedstawiono też wyniki szacunku powierzchni upraw w skali kraju w trzech okresach pomiarowych w latach: 2019, 2020 i 2021. Wprowadzenie systemu SATMIROL umożliwiło prezentację danych w ujęciu przestrzennym, na co nie pozwalają tradycyjne metody stosowane w produkcji statystycznej. Dzięki nowej metodzie dane mogą być przedstawiane na wszystkich poziomach podziału administracyjnego kraju (województw, powiatów i gmin).

Słowa kluczowe: teledetekcja satelitarna, obserwacje Ziemi, statystyka rolnictwa, identyfikacja upraw, szacunek powierzchni upraw rolnych

JEL: O31, Q15, R52

^a Uniwersytet Warmińsko-Mazurski w Olsztynie, Wydział Geoinżynierii, Instytut Geodezji; Urząd Statystyczny w Olsztynie, Ośrodek Informatyki Statystycznej, Polska / University of Warmia and Mazury in Olsztyn, Faculty of Geoen지니어ing, Institute of Geodesy; Statistical Office in Olsztyn, Statistical Computing Centre, Poland. ORCID: <https://orcid.org/0000-0002-9932-1810>. Autor korespondencyjny / Corresponding author, e-mail: magdalena.mleczek@uwm.edu.pl.

^b Urząd Statystyczny w Olsztynie, Ośrodek Informatyki Statystycznej, Polska / Statistical Office in Olsztyn, Statistical Computing Centre, Poland. ORCID: <https://orcid.org/0000-0001-5438-1281>.

^c Główny Urząd Statystyczny, Departament Rolnictwa, Polska / Statistics Poland, Agriculture Department, Poland. ORCID: <https://orcid.org/0000-0002-4927-2686>.

^d Główny Urząd Statystyczny, Departament Rolnictwa, Polska / Statistics Poland, Agriculture Department, Poland. ORCID: <https://orcid.org/0000-0001-8885-5860>.

^e Główny Urząd Statystyczny, Departament Rolnictwa, Polska / Statistics Poland, Agriculture Department, Poland. ORCID: <https://orcid.org/0000-0002-4733-816X>.

^f Centrum Badań Kosmicznych Polskiej Akademii Nauk, Polska / Space Research Centre of the Polish Academy of Sciences, Poland. ORCID: <https://orcid.org/0000-0002-7717-4848>.

^g Instytut Geodezji i Kartografii, Polska / Institute of Geodesy and Cartography, Poland. ORCID: <https://orcid.org/0000-0001-8460-4183>.

Satellite-based crop identification and area estimation under the SATMIROL system

Abstract. One important area where satellite-based Earth observation is applied in is agricultural statistics. The use of new digital technologies and satellite data have contributed to the modernisation of the aforementioned scientific area and is designed to overcome the problems that both agricultural statistics and the recipients of statistical data are faced with. The main aim of the article is to present an innovative system of satellite-based crop identification and crop growth monitoring – SATMIROL. The system was developed by Statistics Poland, the Institute of Geodesy and Cartography and the Space Research Centre of the Polish Academy of Sciences. SATMIROL provides agricultural statistics based on satellite data from the Copernicus programme and administrative data. The article focuses on selected modules of the system which involve the use of satellite remote sensing for crop identification. The study also presents a nationwide crop area estimation performed in three periods: 2019, 2020 and 2021. The implementation of SATMIROL allowed the presentation of statistical data on agriculture in spatial terms, which is not possible when applying traditional methods of data production. This new method of acquiring data enables their presentation at all levels of Poland's administrative division, i.e. by voivodship, powiat (district) and gmina (commune).

Keywords: satellite remote sensing, Earth observations, agricultural statistics, crops identification, crops area estimation

1. Wprowadzenie

Przydatność geoinformacji obrazowej do opisywania przestrzeni zarówno lokalnej, jak i globalnej oraz rosnące zapotrzebowanie na nią przyczyniły się do niebywałego rozwoju technik obrazowania satelitarnego. Wraz z nim powstały programy oferujące otwarty i bezpłatny dostęp do danych obserwacji Ziemi (ang. *Earth observation – EO*)¹. Łatwo dostępne, aktualne dane, rejestrowane z dużą częstotliwością, oraz dane archiwalne są cennym źródłem informacji dla usługodawców, organów publicznych i organizacji międzynarodowych, możliwym do wykorzystania w wielu dziedzinach, w tym w produkcji statystycznej. Ważnym obszarem zastosowania danych EO jest statystyka rolnictwa. Wykorzystanie nowych technologii cyfrowych i danych satelitarnych przyczyniło się do jej unowocześnienia (szybsze udostępnianie danych, obniżenie kosztów, zmniejszenie poziomu agregacji, uprzestrzennienie statystyk) oraz pomaga rozwiązywać problemy, z jakimi mierzy się statystyka w zakresie rolnictwa i z jakimi borykają się odbiorcy danych (Główny Urząd Statystyczny [GUS], 2020).

Aby dane EO mogły być efektywnie wykorzystane w statystyce publicznej, konieczne było – ze względu na specyfikę procesu przetwarzania danych oraz ich duży wolumen – zbudowanie odpowiedniego systemu. W odpowiedzi na to zapotrzebowanie GUS podjął współpracę z instytucjami naukowo-badawczymi. W latach 2019–

¹ Na przykład Copernicus, nadzorowany przez Europejską Agencję Kosmiczną (European Space Agency – ESA). Zob. <https://www.copernicus.eu/en>.

2021 w ramach programu „Społeczny i gospodarczy rozwój Polski w warunkach globalizujących się rynków – GOSPOSTRATEG” zrealizował razem z Instytutem Geodezji i Kartografii (IGiK) oraz Centrum Badań Kosmicznych Polskiej Akademii Nauk projekt badawczo-rozwojowy pn. „SATMIROL – Satelitarna identyfikacja i monitorowanie upraw na potrzeby statystyki rolnictwa” (GUS, b.r.). Zadaniem zespołu naukowo-badawczego były budowa oraz wdrożenie innowacyjnego systemu do identyfikacji i monitorowania upraw rolnych z wykorzystaniem danych satelitarnych, prowadzące do unowocześnienia obecnego systemu uzyskiwania danych z zakresu statystyki rolnictwa.

Realizacja projektu wpisała się w zasadniczy cel programu Ministerstwa Gospodarki (2012) dotyczący działań na rzecz rozwoju technologii kosmicznych i wykorzystywania systemów satelitarnych w Polsce – zwiększanie sprawności i efektywności działania administracji publicznej poprzez rozwój i wdrażanie rozwiązań opartych na technikach teledetekcji satelitarnej, umożliwiającej bardziej efektywne wykorzystywanie posiadanych zasobów i infrastruktury.

Głównym celem artykułu jest zaprezentowanie innowacyjnego systemu satelitarnej identyfikacji i monitorowania upraw rolnych – SATMIROL. W opracowaniu omówiono wybrane moduły tego systemu, w których metody teledetekcji satelitarnej znalazły zastosowanie do rozpoznawania upraw. Przedstawiono też wyniki szacunku powierzchni upraw w skali kraju w trzech okresach pomiarowych w latach: 2019, 2020 i 2021.

2. System satelitarnej identyfikacji i monitorowania upraw rolnych – SATMIROL

Badania statystyczne w zakresie rolnictwa, wymagane zarówno przez instytucje krajowe, jak i międzynarodowe, są prowadzone przez GUS w następujących obszarach: użytkowanie gruntów, powierzchnia zasiewów, okresowe oceny stanu upraw oraz produkcja ważniejszych upraw rolnych i ogrodnich. System SATMIROL, opracowany i wdrażany w ramach projektu „Satelitarna identyfikacja i monitorowanie upraw na potrzeby statystyki rolnictwa – SATMIROL” w latach 2019–2021, jest jednym z mechanizmów uzyskiwania danych z tego zakresu. Do podstawowych publikacji statystycznych GUS, w których dane z systemu SATMIROL znajdują zastosowanie, należą:

- opracowania:
 - *Rolnictwo* (miesiąc publikacji: lipiec),
 - *Rocznik Statystyczny Rolnictwa* (grudzień),
 - *Produkcja upraw rolnych i ogrodnich* (grudzień);

- informacje sygnałowe:
 - *Wstępna ocena przezimowania upraw* (kwiecień),
 - *Wiosenna ocena stanu upraw rolnych i ogrodnich* (maj),
 - *Wstępny szacunek głównych ziemiopłodów rolnych i ogrodnich* (lipiec),
 - *Wynikowy szacunek głównych ziemiopłodów rolnych i ogrodnich* (grudzień),
 - *Przedwynikowy szacunek głównych ziemiopłodów rolnych i ogrodnich* (wrzesień).

Dane te są także publikowane w czerwcu w Banku Danych Lokalnych (dziedzina Rolnictwo – Użytkowanie gruntów oraz Produkcja roślinna).

W ramach projektu SATMIROL zdefiniowano i zrealizowano siedem zadań (GUS, 2020):

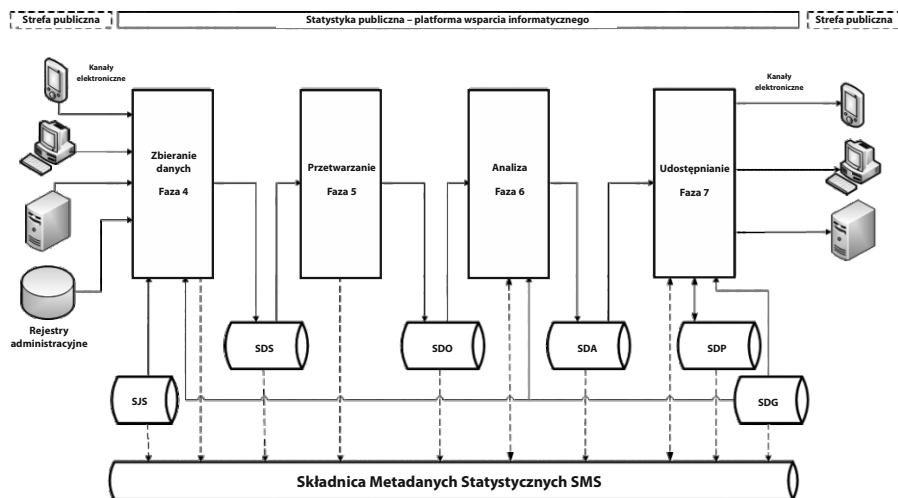
- opracowanie kompleksowej strategii wdrożenia koncepcji systemu satelitarnej identyfikacji i monitorowania upraw na potrzeby statystyki rolnictwa, zdefiniowanie szczegółów jego budowy oraz opracowanie rozwiązań systemowych, metodologicznych i funkcjonalnych;
- wypracowanie rozwiązań mających na celu rozpoznawanie upraw na podstawie danych satelitarnych na terenie Polski;
- wypracowanie rozwiązań mających na celu ocenę kondycji roślin uprawnych i monitorowanie sytuacji kryzysowych;
- kalibracja terenowa (*in situ*) danych satelitarnych;
- opracowanie procedur związanych z implementacją algorytmów dla obszaru rozpoznawania upraw;
- opracowanie procedur związanych z implementacją algorytmów dla obszaru oceny kondycji roślin uprawnych i monitorowania sytuacji kryzysowych;
- wdrożenie i uruchomienie systemu do identyfikacji i monitorowania upraw rolnych.

System SATMIROL został oparty na modelu usprawniającym procesy statystyczne w Polsce – Modelu Procesu Produkcji Statystycznej (MPPS; Dygaszewicz i Szafranski, 2018). Model ten opiera się z kolei na podstawowym modelu realizacji badań statystycznych (Generic Statistical Business Process Model – GSBPM)², wypracowanym przez Europejską Komisję Gospodarczą (United Nations Economic Commission for Europe – UNECE), który jest uznawany za branżowy standard i wdrażany przez urzędy statystyczne w ponad 80 krajach. MPPS uzupełniają pod-

² Model stworzono ok. 2006 r. w Nowej Zelandii z myślą o zastosowaniu go w badaniach gospodarczych. Stąd też jego nazwa, która w dosłownym tłumaczeniu brzmi: podstawowy model prowadzenia statystyki działalności gospodarczej. Jest to narzędzie do opisu działań statystycznych w kontekście odpowiednich procesów prowadzących do uzyskania oczekiwanych rezultatów i zarazem podstawa pracy statystycznej, z uwzględnieniem takich jej aspektów, jak jakość, efektywność, standaryzacja itp. W dużej mierze może on być traktowany także jako uniwersalny model statystycznego postępowania badawczego i za taki jest obecnie uważany.

procesy z komponentem geoprzestrzennym, weryfikacji i oceny jakości oraz planowania. Oba modele zostały zbudowane zgodnie z podejściem procesowym i obejmują całość działań organizacji, a nie tylko wybrane aspekty jej funkcjonowania. Na schemacie 1 przedstawiono fazy produkcji statystycznej i składnice w MPPS.

Schemat 1. Fazy produkcji statystycznej i składnice danych w MPPS



Uwaga. SJS – Składnica Jednostek Statystycznych, SDS – Składnica Danych Surowych, SDO – Składnica Danych Operacyjnych, SDA – Składnica Danych Analitycznych, SDP – Składnica Danych Publikacyjnych, SDG – Składnica Danych Geoprzestrzennych.
Źródło: Dygaszewicz i Szafrąński (2018, s. 18).

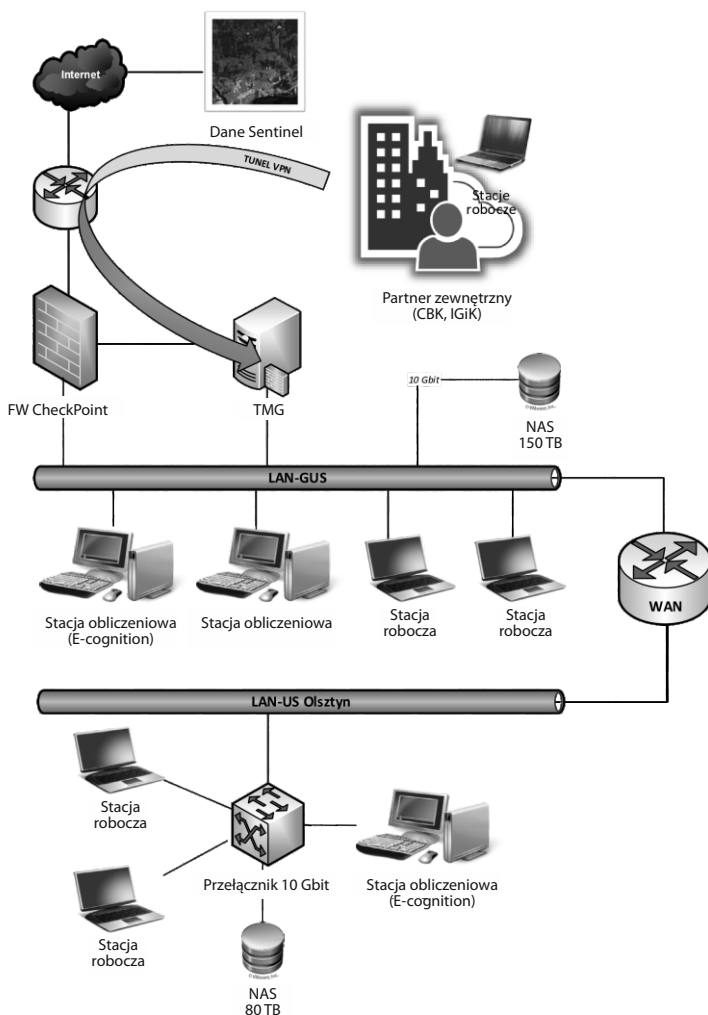
W systemie SATMIROL zostały stworzone moduły:

1. pobierania i wstępnego przetwarzania danych – zestaw narzędzi do automatycznego pobierania i przetwarzania zdjęć radarowych Sentinel-1 i optycznych Sentinel-2, w tym wieloczesowego przetworzenia polarymetrycznego zdjęć dwupolaryzacyjnych;
2. segmentacji i klasyfikacji upraw – zestaw narzędzi do automatycznej segmentacji danych obrazowych oraz klasyfikacji, w tym uczenia maszynowego oraz walidacji wyników;
3. automatycznej oceny warunków wzrostu i kondycji roślin uprawnych oraz detekcji sytuacji kryzysowych poprzez odniesienie chwilowych wartości satelitarnych wskaźników roślinnych uzyskanych z Sentinel-3 OLCI i SLSTR do wieloletnich danych referencyjnych (z sensora MODIS) oraz poprzez analizę warunków agrometeorologicznych na podstawie reanalizy klimatologicznej ERA-5;
4. raportowania i tworzenia statystyk na wszystkich poziomach podziału administracyjnego kraju (województw, powiatów i gmin).

Rozpoznawanie upraw i szacunek ich powierzchni opisywane w artykule są bezpośrednio związane z modułami 1, 2 i 4, dlatego opis modułu 3 pominięto.

Na schemacie 2 przedstawiono architekturę systemu SATMIROL.

Schemat 2. Infrastruktura systemu SATMIROL służąca do pobierania i przetwarzania danych



Źródło: GUS (2020, s. 4).

Poszczególne elementy systemu zostały zaimplementowane w dwóch miejscach w GUS oraz w Urzędzie Statystycznym w Olsztynie, w którym funkcjonuje Ośrodek

Badań Rolnictwa i pracuje zespół specjalistów ds. teledetekcji satelitarnej (zestawienie).

Zestawienie. Infrastruktura obliczeniowa

GUS	Urząd Statystyczny w Olsztynie
Zasoby dyskowe typu NAS o pojemności 150 TB (podłączone do sieci przepustowością 10Gb/s)	Zasoby dyskowe typu NAS o pojemności 80 TB (podłączone do sieci przepustowością 10Gb/s)
Dwie stacje obliczeniowe (10/20 rdzeni, 3,0 GHz, 128 GB RAM)	Stacja obliczeniowa (10/20 rdzeni, 3,0 GHz, 128 GB RAM)
Oprogramowanie do analizy obiektowej obrazów	Oprogramowanie do analizy obiektowej obrazów Oprogramowanie do analiz i przetwarzania danych geoprzestrzennych
Dwie stacje robocze (4/8 rdzeni, 3,5 GHz, 64 GB RAM)	Dwie stacje robocze (4/8 rdzeni, 3,5 GHz, 64/128 GB RAM)

Źródło: opracowanie na podstawie: GUS (2020).

3. Dane wykorzystywane w systemie SATMIROL

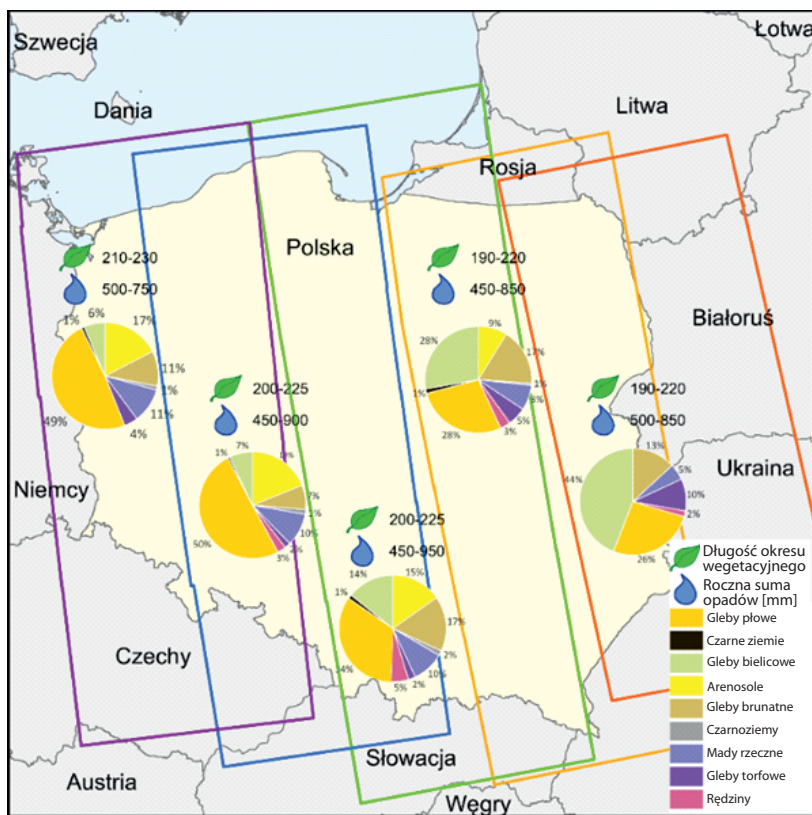
Wypracowywanie rozwiązań mających na celu rozpoznawanie upraw na podstawie danych satelitarnych odbywało się na terenach rolnych, które stanowią 60% powierzchni całego kraju. Większość terenów rolnych w Polsce – 68% – jest użytkowana jako grunty orne, 27% to łąki i pastwiska, a 5% stanowią uprawy trwałe. Według danych katastralnych istnieje prawie 11 mln działek rolnych o zróżnicowanym rozkładzie powierzchni: 4% działek ma powierzchnię mniejszą niż 1 ha, 39% – 1–5 ha, 23% – 5–10 ha, 19% – 10–20 ha, 12% – 20–50 ha i tylko 4% – ponad 50 ha. Wschodnia (zwłaszcza południowo-wschodnia) część Polski charakteryzuje się dużą liczbą bardzo małych działek. Do rozpoznawania upraw za pomocą metody SATMIROL wytypowano 25 upraw: buraki cukrowe, gorczycę, grykę, jęczmień jary, jęczmień ozimy, kukurydzę, mieszanki zbożowe, owies, plantacje drzew owocowych, plantacje krzewów owocowych, proso, pszenicę jara, pszenicę ozimą, pszenżyto jare, pszenżyto ozime, rzepak jary, rzepak ozimy, strączkowe, trawy i użytki zielone, truskawkę, tytoń, warzywa, ziemniaki, zioła i przyprawy oraz żyto ozime.

3.1. Dane satelitarne

Podstawą opracowania metodologii rozpoznawania upraw były dane satelitarne pochodzące z programu Copernicus. Posłużono się dwoma typami obrazów: radarowymi z satelitów Sentinel-1A i Sentinel-1B oraz optycznymi z satelitów Sentinel-2A i Sentinel-2B. Wszystkie dane zostały pobrane automatycznie ze strony ESA lub platformy CreoDIAS za pomocą opracowanych procedur.

Sentinel-1 to aktywny system obrazujący w pasmie C (o długości fali wynoszącej w przybliżeniu 5,3 cm), który rejestruje obrazy niezależnie od występującego zachmurzenia (ESA, 2013). Wykorzystane dane obejmowały produkty SLC (ang. *single look complex*) rejestrowane w trybie IW (ang. *interferometric wide swath*) z orbity zstępującej w układzie polaryzacyjnym VV+VH. Rozdzielczość użytych obrazów wynosi w przybliżeniu $2,3 \times 13,9$ m, a szerokość pasa obrazowania – 280 km. Obszar Polski jest objęty pięcioma pasami obrazowania, których granice stanowiły główne kryterium podziału opracowania na niezależne części (mapa 1).

Mapa 1. Obszar opracowania z granicami zasięgu pasów obrazowania satelitów Sentinel-1 oraz warunki środowiskowe charakteryzujące poszczególne pasy obrazowania



Źródło: opracowanie własne na podstawie danych satelitarnych Sentinel-1 oraz danych IGIK.

Obrazy Sentinel-1 zostały pobrane dla trzech okresów pomiarowych w latach: 2019, 2020 i 2021. Przed sezonem wegetacyjnym upraw (tj. od połowy lutego do końca marca) obrazy były pobierane co 12 dni, a w czasie sezonu wegetacyjnego (od początku kwietnia do końca września) co 6 dni. W każdym roku pomiarowym zebrano ok. 1050 scen satelitarnych, każda o wielkości 6,5–7 GB.

Drugim źródłem danych satelitarnych był system Sentinel-2, który jest pasywny i rejestruje 13 kanałów spektralnych (ESA, 2015). Wymiar pojedynczej sceny wynosi 100×100 km, a wymiar piksela jest uzależniony od kanału spektralnego i dla kanałów B2–B4 i B8 wynosi 10 m, dla kanałów B5–B7, B8A i B11–B12 – 20 m i dla kanałów B1 i B9–B10 – 60 m (ESA, 2015). Wykorzystane dane należały do produktów skorygowanych atmosferycznie oznaczonych symbolem L2A – drugi poziom przetworzenia. W każdym roku pomiarowym zgromadzono blisko 1500 scen satelitarnych, każda o wielkości ok. 1 GB.

3.2. Dane administracyjne

Do budowy systemu satelitarnej identyfikacji i monitorowania upraw wykorzystano dane administracyjne obejmujące obszar całego kraju. Ich źródłem były następujące instytucje:

- GUS – dane *in situ* za lata 2019–2021 (zob. podrozdział 3.3);
- Agencja Restrukturyzacji i Modernizacji Rolnictwa (ARiMR) – dane z Systemu Identyfikacji Działek Rolnych (Land-parcel Identification System – LPIS) oraz Zintegrowanego Systemu Zarządzania i Kontroli (Integrated Administration and Control System – IACS) za lata 2019–2021, tj.:
 - warstwa wektorowa MKO JPO (dawniej PEG) – maksymalny kwalifikowalny obszar – wyznaczony dla każdej działki ewidencyjnej obszar gruntów rolnych kwalifikujących się do przyznania płatności,
 - warstwa wektorowa MKO ORNY – obszar w ramach MKO JPO obejmujący grunt orny,
 - warstwa wektorowa pól zagospodarowania,
 - zinformatywowana baza wniosków o przyznanie płatności – baza danych w formie pliku tekstowego została poddana geokodowaniu, harmonizacji i standaryzacji;
- Główny Urząd Geodezji i Kartografii (GUGiK) – dane z bazy danych obiektów topograficznych (BDOT) za 2019 r. – warstwa wektorowa danych obiektów topograficznych.

Na podstawie danych *in situ* (GUS) oraz zinformatywowanej bazy wniosków o przyznanie płatności (ARiMR) opracowano bazę próbek uczących i kontrolnych zasilających algorytmy uczenia maszynowego, w tym na potrzeby walidacji wyników

klasyfikacji obiektowej. Na podstawie warstwy wektorowej MKO JPO, MKO ORNY, pól zagospodarowania (ARiMR) oraz bazy danych obiektów topograficznych (GUGiK) wyznaczono tereny wyłącznie rolne podlegające procesowi klasyfikacji obrazu.

3.3. Dane z badania *in situ*

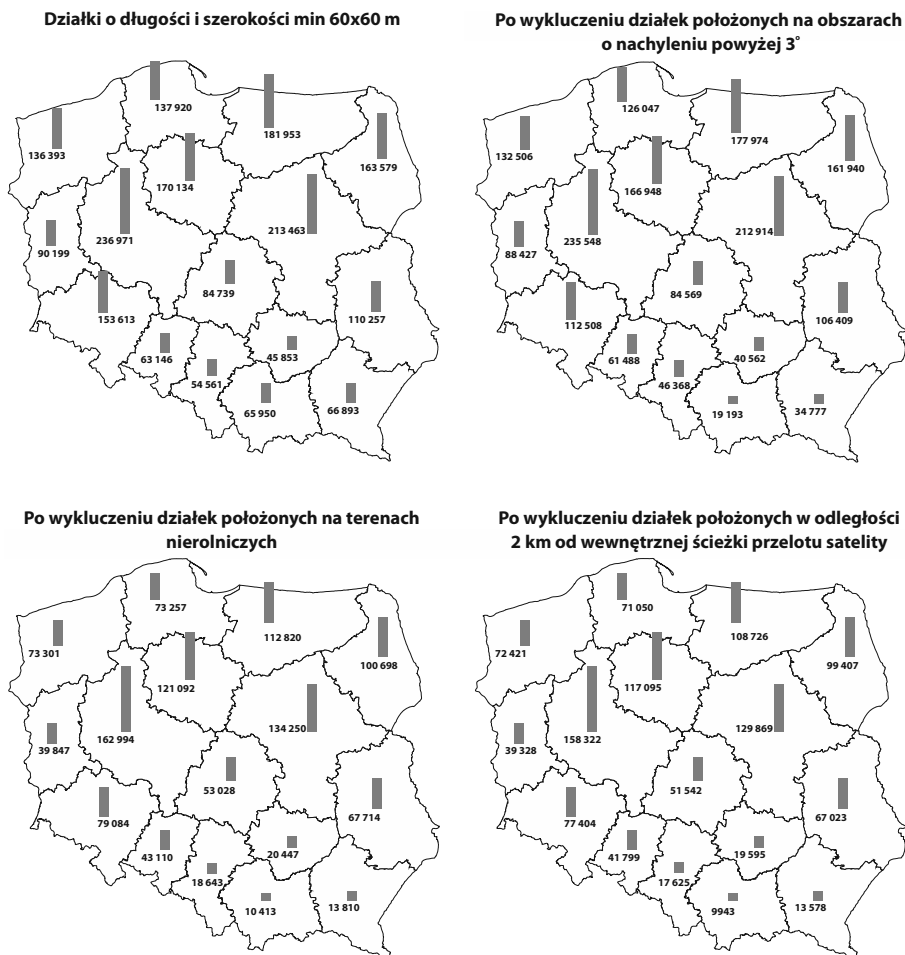
Badanie terenowe przeprowadzono we wszystkich województwach w okresie od czerwca do lipca 2019 r. i od czerwca do lipca 2020 r., a prace przygotowawcze trwały od stycznia do czerwca. Każdy z ankierów miał za zadanie dotrzeć do zaznaczonego pola, ocenić, czy uprawa jest jednorodna pod względem rozwoju na obszarze co najmniej 60×60 m, oraz zebrać dane wyszczególnione w formularzu badania *in situ*. Ankierzy byli wyposażeni w tablety z odbiornikiem GPS i aplikacją do nawigacji umożliwiającą dotarcie do badanego pola, rejestrację współrzędnych oraz wykonanie geotagowanych zdjęć. Dla każdej obserwacji na polu należało wykonać opis zgodnie ze standardowym formularzem. Do opisanego wybrano 22 uprawy w każdym województwie. W przypadku każdej z upraw należało opisać co najmniej 20 plantacji reprezentatywnych (minimum 50 opisów dla każdej uprawy w danym pasie przelotu satelity). Do wstępnego wyboru pól reprezentatywnych, spełniających kryteria: przynależności do terenów rolnych, minimalnej długości i szerokości 60 m, nachylenia obszaru poniżej 3° oraz minimalnej odległości od wewnętrznej ścieżki przelotu satelity, wykonano analizy przestrzenne na podstawie następujących danych:

- administracyjnych:
 - danych wektorowych o charakterze katastralnym z LPIS (ponad 34 mln rekordów, 13 GB),
 - wektorowej warstwy pól zagospodarowania ARiMR (ponad 33 mln rekordów, 23 GB),
 - danych z BDOO;
- przestrzennych:
 - SRTM (Shuttle Radar Topography Mission)³,
 - wektorowych granic pasów zobrazowań satelitów teledetekcyjnych Sentinel-1.

Na mapie 2 przedstawiono liczbę potencjalnych działek reprezentatywnych w obrębie województw.

³ Zob. <https://www2.jpl.nasa.gov/srtm/dataprod.htm>.

Mapa 2. Potencjalne działki reprezentatywne

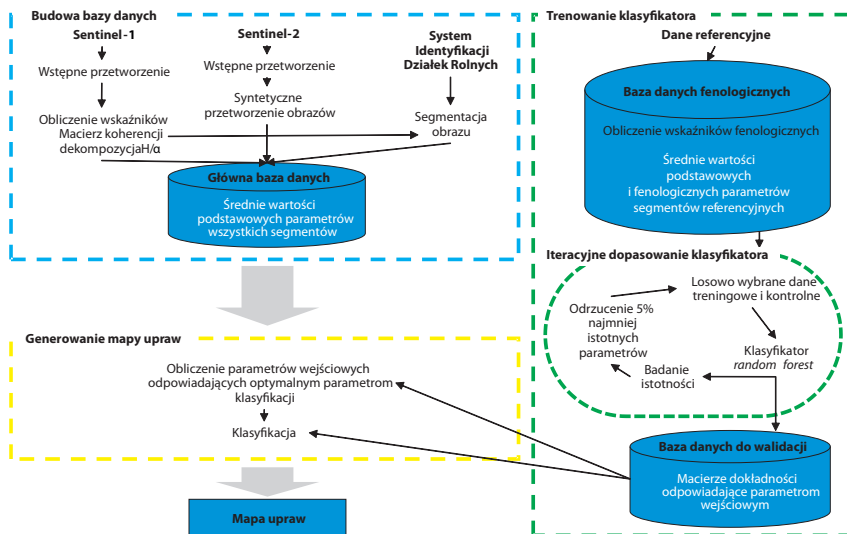


Źródło: GUS (2019).

4. Metoda rozpoznawania upraw

W ramach systemu SATMIROL opracowano metodę rozpoznawania upraw, którą przedstawiono na schemacie 3. Opiera się ona na trzyetapowej procedurze klasyfikacyjnej. Pierwszy etap polega na zbieraniu statystyk i tworzeniu bazy danych (obszar ograniczony niebieską linią przerywaną), drugi – na trenowaniu i uczeniu systemu klasyfikacyjnego (obszar ograniczony zieloną linią przerywaną), a trzeci – na zastosowaniu tego algorytmu i generowaniu mapy upraw (obszar ograniczony pomarańczową linią przerywaną).

Schemat 3. Metoda rozpoznawania upraw



Źródło: GUS (2020, s. 9).

4.1. Zbieranie statystyk i tworzenie bazy danych

Baza danych budowana jest na podstawie obrazów satelitarnych z programu Copernicus (Sentinel-1 i Sentinel-2) oraz danych administracyjnych z LPIS i BDOO dla analizowanego okresu.

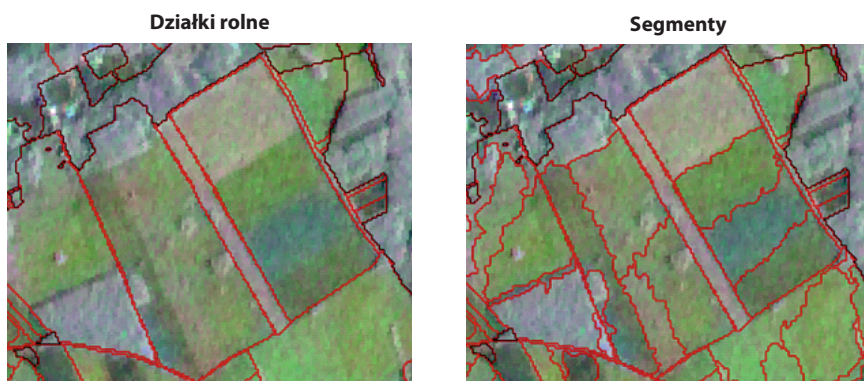
Obrazy satelitarne wymagają wstępnego przetworzenia, które w przypadku danych optycznych Sentinel-2 polega na zidentyfikowaniu obszarów zasłoniętych chmurami i ich usuwaniu. Braki są uzupełniane wartościami pikseli obrazów z sąsiadujących dat rejestracji. Z kolei wstępne przetworzenie obrazów radarowych Sentinel-1 obejmuje trzy kroki: scalenie, ortorektyfikację i filtrację przestrzenną obrazów. Scalanie obrazu wynika ze specyfiki obrazowania za pomocą progresywnego skanera TOPSAR (Terrain Observation by Progressive Scans), który generuje niezależne fragmenty obrazów (ang. *bursts*; De Zan i Monti Guarnieri, 2006). Ortorektyfikacja polega na transformacji obrazów z geometrii obrazowania do układu współrzędnych. Wszystkie obrazy są ortorektyfikowane do układu współrzędnych geograficznych opartego na elipsoidzie WGS84. W procesie filtracji redukuje się szum (efekt plamkowania) przy użyciu filtra Enhanced LEE w oknie 3×3 piksele (Lang i in., 2014; Zhu i in., 2013).

Wstępnie przetworzone dane radarowe są poddawane wieloczasowemu przetworzeniu polarymetrycznemu, tj. przetworzeniu informacji z każdego z obrazów

(w pewnym przedziale czasu) związanej z odpowiedzią obiektów na powierzchni ziemi na kontakt z falami spolaryzowanymi. Wynikiem przetworzenia polarymetrycznego jest uzyskanie macierzy koherencji T_2 , a następnie parametrów dekompozycji polarymetrycznej H/α (Cloude, 2007; Cloude i Pottier, 1997). Dla każdej daty rejestracji z serii czasowej i dla każdego pasa przelotu otrzymuje się sześć warstw obrazów wejściowych do tworzenia wskaźników.

Tak przygotowane dane satelitarne stanowią podstawę do utworzenia bazy danych. Składa się ona z obiektów przestrzennych – segmentów – utworzonych w procesie segmentacji wstępnie przetworzonych obrazów w granicach obszarów rolnych. Segmentacja jest wykonywana za pomocą programu eCognition w dwóch etapach: na pierwszym tworzy się obiekty odpowiadające zasięgom działek LPIS, a na drugim sprawdzana jest homogeniczność informacji obrazowej tych działek przy użyciu zdjęć satelitarnych. Działki niejednorodne, np. obsiane wieloma uprawami, są dzielone na mniejsze obiekty (mapa 3). W trakcie segmentacji, na podstawie bazy LPIS i BDOO, eliminuje się również wszystkie obszary nierolne, które na kolejnych etapach klasyfikacji nie są brane pod uwagę.

Mapa 3. Przykładowa wizualizacja różnic granic działek rolnych i granic segmentów



Źródło: opracowanie własne na podstawie danych ESA i ARIMR.

4.2. Trenowanie i uczenie systemu klasyfikacyjnego

Drugi etap rozpoczyna się od obliczenia wieloczasowych wskaźników fenologicznych dla segmentów, dla których istnieją dane referencyjne, i poddania ich klasyfikacji obiektowej. Wskaźniki opisujące zmiany fenologiczne upraw wykorzystują różnicę pomiędzy parametrami polarymetrycznymi oraz zmiany w czasie poszczególnych

parametrów. Celem opracowania radarowych wskaźników wieloczasowych jest wychwycenie różnic pomiędzy poszczególnymi uprawami, możliwe dzięki wykorzystaniu zmian w typie i natężeniu odbicia fali radarowej w zależności od charakterystycznego tempa wzrostu poszczególnych upraw.

W ramach projektu opracowano znormalizowany wskaźnik wieloczasowy, obliczany na podstawie poszczególnych elementów macierzy koherencji lub elementów dekompozycji H/a dla dwóch terminów. Przy obliczaniu wskaźnika dla wszystkich kombinacji dostępnych terminów przyjmuje się wcześniejszy termin jako początkowy i późniejszy jako końcowy. Obliczono ok. 3800 tego typu parametrów. Innym typem opracowanych parametrów pozwalających określić tempo wzrostu upraw są parametry a i b funkcji liniowej opisującej przyrost natężenia odbicia sygnału w poszczególnych okresach. Obliczono ok. 7500 takich parametrów. Do obliczeń związanych ze wzrostem upraw wykorzystywane są również znormalizowane wskaźniki elementów macierzy koherencji, które opisują zmianę w sposobie odbicia fali w zależności od kształtu uprawy i są liczone dla poszczególnych terminów. Obliczono ok. 110 parametrów tego typu.

Procedura wieloczasowego przetwarzania zdjęć optycznych jest analogiczna do procedury zaprojektowanej dla zdjęć radarowych. Dla poszczególnych kanałów spektralnych generowane są: znormalizowany wskaźnik wieloczasowy, parametry a i b funkcji liniowej opisującej przyrost odbiciowości w poszczególnych pasmach spektralnych w poszczególnych okresach oraz znormalizowane wskaźniki pomiędzy kanałami spektralnymi.

Obliczone obrazy wskaźników stanowią warstwy wejściowe, które klasyfikuje się wielokrotnie metodą *random forest* na podstawie pól treningowych z danych referencyjnych (Belgiu i Drąg, 2016; Pelletier i in., 2016; Rodriguez-Galiano i in., 2012; Song i in., 2017; Thanh Noi i Kappas, 2018). Dokładność klasyfikacji określana jest na niezależnych od pól treningowych polach kontrolnych pochodzących również z danych referencyjnych. Analiza dokładności pozwala określić przydatność poszczególnych warstw wejściowych. Po każdej klasyfikacji odrzucanych jest 5% najmniej znaczących warstw i proces klasyfikacji jest powtarzany. W ten sposób otrzymuje się zestaw optymalnych danych wejściowych do klasyfikacji oraz klasyfikator, który należy zastosować na kolejnym etapie procedury.

4.3. Zastosowanie algorytmu i generowanie mapy upraw

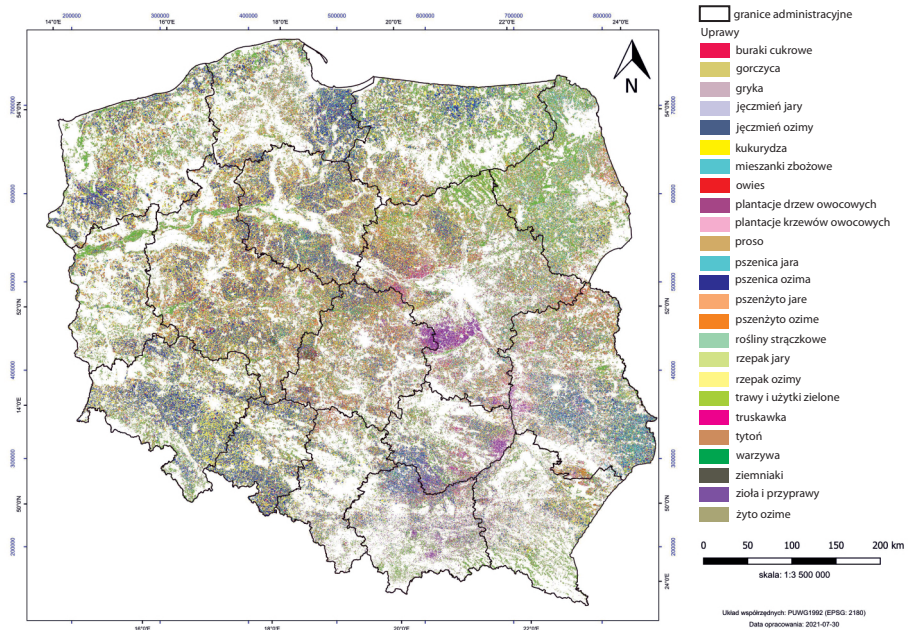
Trzeci etap omawianej metody rozpoczyna się od wybrania wskaźników wyznaczonych na poprzednim etapie jako najistotniejsze dla wszystkich segmentów w granicach gruntów rolnych i wykonania klasyfikacji, w wyniku której uzyskuje się obraz przedstawiający rozkład przestrzenny klasyfikowanych upraw.

Kolejny krok stanowi wykonanie walidacji wyników finalnej klasyfikacji, która odbywa się przy użyciu danych *in situ* (podobnie jak na drugim etapie na niezależnych od pól treningowych polach kontrolnych pochodzących z danych referencyjnych), danych z deklaracji rolników złożonych do systemu ARiMR oraz danych z kontroli na miejscu ARiMR. Na podstawie obliczonej macierzy błędów wyznaczane są miary niedoszacowania PA (ang. *producer's accuracy*), przeszacowania UA (ang. *user's accuracy*), dokładności poszczególnych upraw *F1-score* (średnia harmoniczna pomiędzy miarą niedoszacowania a miarą przeszacowania) i dokładności całkowitej klasyfikacji OA (ang. *overall accuracy*; Congalton, 1991; Congalton i Green, 2019). Im wartości są bliższe 1 (lub 100 w przypadku wyrażenia wartości procentowo), tym wyższa jest jakość wyników uzyskanych w procesie klasyfikacji. W najlepszym przypadku dokładność przyjmuje wartość 1. Dane walidacyjne powinny stanowić przynajmniej 20% prób danych referencyjnych.

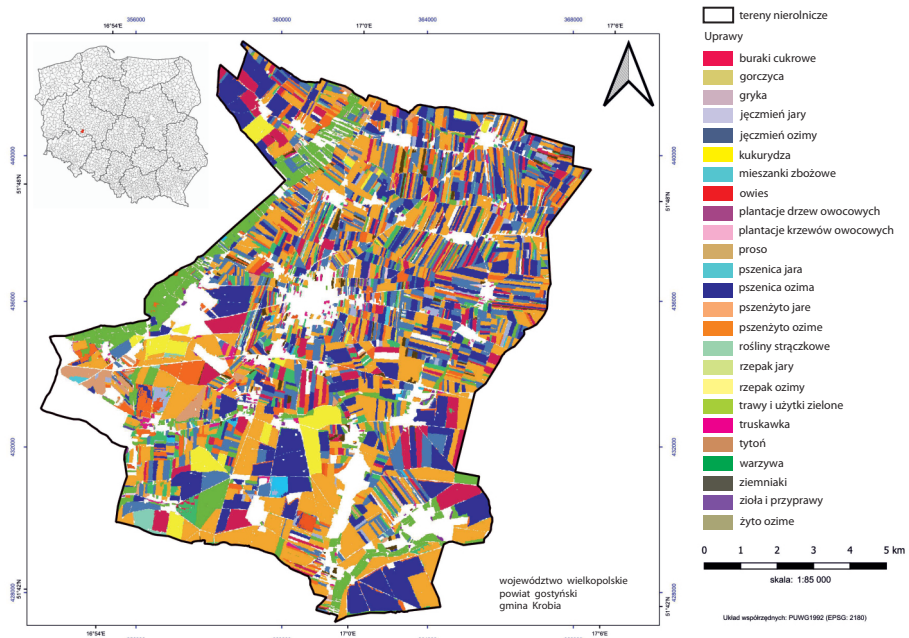
Na podstawie obrazu klasyfikacji oblicza się powierzchnie upraw. Są one sumowane na wszystkich poziomach administracyjnych kraju (województwa, powiaty i gminy). Numer identyfikacyjny poszczególnych jednostek pozwala na integrację tych danych z systemem raportowania GUS.

5. Wyniki

W przeprowadzonym badaniu klasyfikacji poddano 23 klasy upraw (bez mieszanek zbożowych i tytoniu) w latach 2019–2020 oraz 25 klas upraw w 2021 r. Dokładność całkowita klasyfikacji wyniosła odpowiednio 85% dla 2019 r. i 2020 r. oraz 78% dla 2021 r. Na mapie 4 przedstawiono wyniki z 2021 r. w ujęciu przestrzennym w postaci mapy upraw dla Polski, a na mapie 5 – fragment upraw w większej skali.

Mapa 4. Uprawy w 2021 r. – Polska

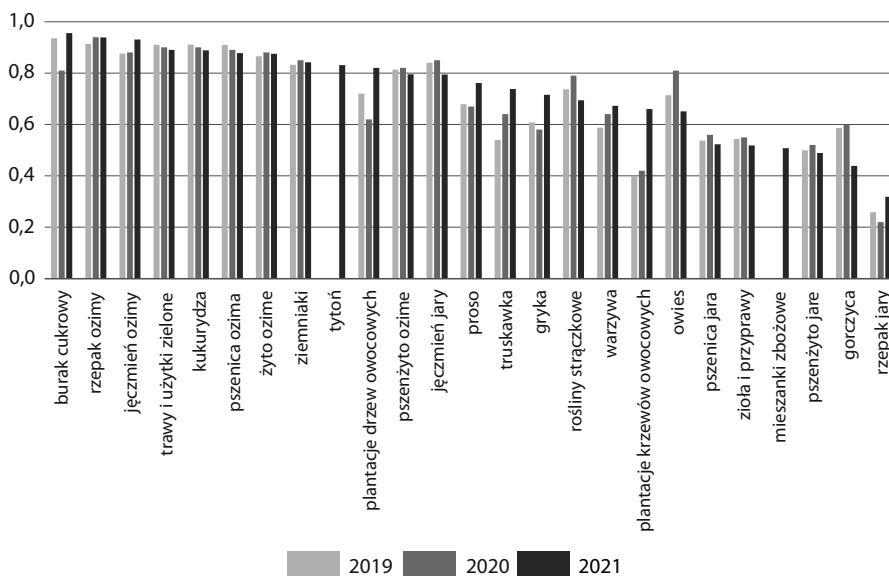
Źródło: opracowanie własne na podstawie danych ESA, ARiMR i GUS.

Mapa 5. Uprawy w 2021 r. – gmina Krobia

Źródło: opracowanie własne na podstawie wyników z systemu SATMIROL.

Wyniki dokładności klasyfikacji poszczególnych upraw dla trzech lat pomiarowych: 2019, 2020 i 2021 zostały przedstawione na wykr. 1. Im wartości bliższe 1, tym wyższa jest jakość wyników uzyskanych w procesie klasyfikacji.

Wykr. 1. Wyniki dokładności klasyfikacji poszczególnych upraw wyrażone miarą *F1-score*



Źródło: opracowanie własne na podstawie wyników z systemu SATMIROL.

Dziesięć klas, tj. burak cukrowy, jęczmień jary, jęczmień ozimy, kukurydza, pszenica ozima, pszenżyto ozime, rzepak ozimy, trawy i użytki zielone, ziemniaki oraz żyto ozime, zostało rozpoznanych z najwyższą dokładnością – powyżej 0,8 – we wszystkich trzech latach pomiarowych. Tytoń, klasyfikowany jedynie w 2021 r., również uzyskał dokładność powyżej 0,8. Najniższą dokładność klasyfikacji – ok. 0,5 lub mniej – we wszystkich latach pomiarowych uzyskano dla takich upraw, jak: pszenica jara, pszenżyto jare, rzepak jary oraz ziola i przyprawy. Mieszanki zbożowe były klasyfikowane tylko w 2021 r. i dokładność ich identyfikacji wyniosła 0,51. Zauważa się znaczny wzrost dokładności klasyfikacji w 2021 r. w porównaniu z latami 2019 i 2020 w przypadku plantacji drzew owocowych – do 0,82, plantacji krzewów owocowych – do 0,66, a także truskawek – do 0,74.

Na podstawie wykonanej klasyfikacji obliczono powierzchnię upraw. Zestawienie powierzchni upraw w skali kraju pokazane jest w tabelicy oraz graficznie zaprezentowane na wykr. 2.

Tablica. Powierzchnia upraw i dokładność klasyfikacji poszczególnych upraw

Uprawy	Powierzchnia w ha			Dokładność klasyfikacji (F1-score)		
	2019	2020	2021	2019	2020	2021
Burak cukrowy	244 045	273 843	237 304	0,94	0,81	0,96
Gorczyca	159 522	252 752	60 570	0,59	0,60	0,44
Gryka	244 044	341 461	116 305	0,61	0,58	0,72
Jęczmień jary	456 418	395 793	300 163	0,84	0,85	0,79
Jęczmień ozimy	243 231	327 738	310 264	0,88	0,88	0,93
Kukurydza	1 278 561	1 288 319	1 583 815	0,91	0,90	0,89
Mieszanki zbożowe	-	-	354 798	-	-	0,51
Owies	706 883	726 687	427 516	0,71	0,81	0,65
Plantacje drzew owocowych	551 052	611 540	318 730	0,72	0,62	0,82
Plantacje krzewów owocowych	854 070	786 572	511 710	0,40	0,42	0,66
Proso	104 519	147 649	118 251	0,68	0,67	0,76
Pszemica jara	388 722	391 783	323 236	0,54	0,56	0,52
Pszemica ozima	1 617 127	1 477 333	1 735 239	0,91	0,89	0,88
Pszemczyto jare	339 615	570 224	402 000	0,50	0,52	0,49
Pszemczyto ozime	1 002 628	1 090 168	939 308	0,81	0,82	0,79
Rośliny strączkowe	242 669	257 483	306 249	0,74	0,79	0,69
Rzepak jary	369 827	255 760	135 912	0,26	0,22	0,32
Rzepak ozimy	643 604	753 921	888 104	0,91	0,94	0,94
Trawy i użytki zielone	2 492 199	2 839 917	3 288 459	0,91	0,90	0,89
Truskawka	161 558	177 729	510 869	0,54	0,64	0,74
Tytoń	-	-	68 935	-	-	0,83
Warzywa	182 902	157 001	186 618	0,59	0,64	0,67
Ziemniaki	157 719	178 484	172 746	0,83	0,85	0,84
Zioła i przyprawy	197 762	191 539	98 978	0,54	0,55	0,52
Żyto ozime	749 147	782 753	691 409	0,86	0,88	0,88

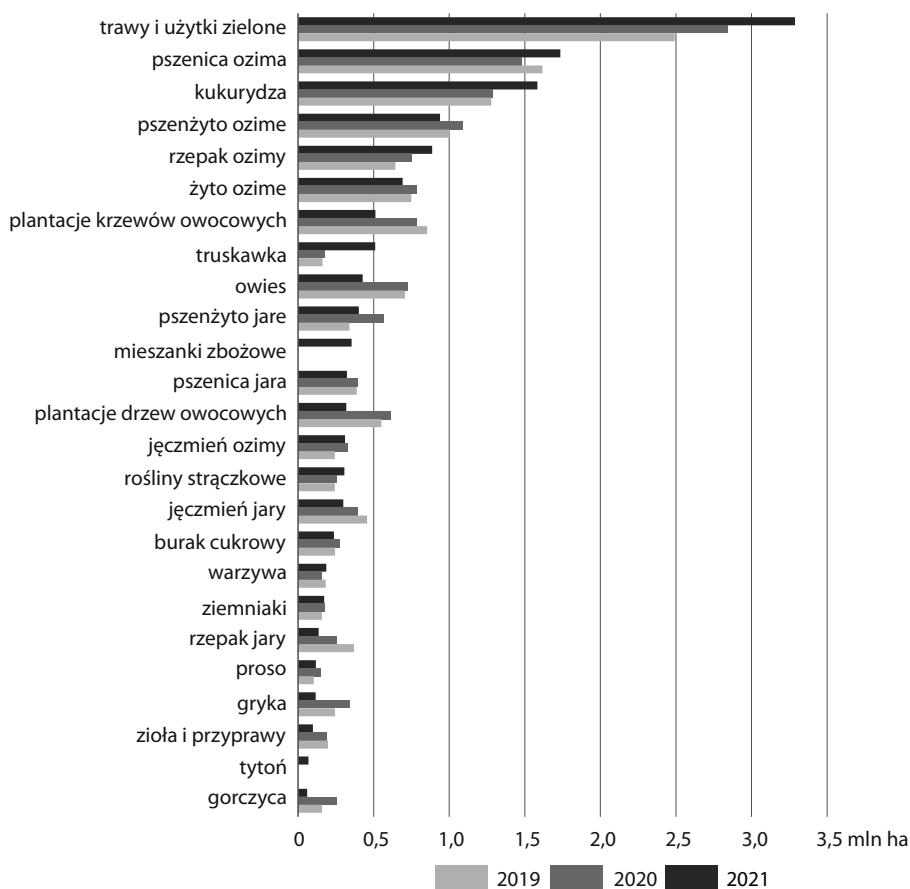
Źródło: opracowanie własne na podstawie wyników z systemu SATMIROL.

W latach 2019 i 2020 widoczne jest duże podobieństwo wielkości powierzchni w przypadku grupy zbóż zbliżonych pokrojowo.⁴ Wśród nich dominują zboża ozime, preferowane przez rolników z uwagi na występujące w Polsce wiosenne susze. Uprawy jare pojawiają się na ogół po stratach zimowych lub związane są z brakiem czasu na siew ozimin albo trudnością w przygotowaniu pól. W 2021 r. z powodu niekorzystnych warunków atmosferycznych, a tym samym opóźnienia wiosennych prac polowych, zaobserwowano zmniejszenie ogólnej powierzchni uprawy zbóż podstawowych (zbliżonych pokrojowo). Dotyczy to zwłaszcza form jarych wszystkich gatunków. Ich miejsce zajęła kukurydza, którą wysiewa się w późniejszym terminie agrotechnicznym. W latach 2019 i 2020 ogólna powierzchnia kukurydzy jest porównywalna. Powiększanie się powierzchni uprawy kukurydzy w 2021 r. wynika

⁴ Czyli zbliżonych wyglądem i zewnętrznym ukształtowaniem, do których zalicza się takie cechy, jak: rozmiar pędów, kierunek wzrostu, stopień zdrewniania, ulistnienie, liczbę oraz rodzaj odgałęzień czy miejsce narządów płciowych.

również z zapotrzebowania na paszę oraz rosnącego zużycia tego surowca w biogazowniach.

Wykr. 2. Powierzchnia upraw



Źródło: opracowanie własne na podstawie wyników z systemu SATMIROL.

W skali kraju obserwuje się systematyczny wzrost powierzchni uprawy rzepaku ozimego, co wiąże się głównie ze zwiększonym zapotrzebowaniem surowcowym przemysłu olejarskiego. Poza tym stanowi on doskonały przedplon dla uprawy pszenicy ozimej. Powierzchnia uprawy ziemniaków w skali kraju jest ustabilizowana. Obserwuje się systematyczny wzrost areału uprawy roślin strączkowych, korzystnie wpływających na warunki fizyczno-chemiczne gleb.

Zaobserwowano zwiększenie zgrupowanej powierzchni trwałych użytków zielonych, traw i motylkowych drobnonasiennych, stanowiących zaplecze paszowe dla przeżuwaczy. Uzyskana duża różnica areału traw i użytków zielonych w Polsce wynika również ze zmiany zastosowanych danych. Jak wspomniano w poprzednim

rozdziale, w procesie segmentacji opracowywane są tylko fragmenty obrazów w granicach obszarów rolnych zdefiniowanych przez bazę zewnętrzną, w tym wypadku LPIS. W badaniu w 2019 r. uwzględniono jedynie trawy na gruntach ornych, ponadto w badaniu w 2021 r. zastosowano zaktualizowaną bazę gruntów ornych zakwalifikowanych do jednolitej płatności obszarowej przez ARiMR.

Pozostałe rozpoznane klasy upraw, z uwagi na mniejszą powierzchnię, cechują się dużą zmiennością, która nie zawsze uwidacznia się na poziomie Polski, ale jest znacząca na poziomie województw. Wymaga to dalszego doskonalenia procesu rozpoznawania tych upraw.

6. Podsumowanie

W artykule zaprezentowano najważniejsze elementy systemu satelitarnej identyfikacji i monitorowania upraw rolnych – SATMIROL. Przedstawiono również wyniki uzyskane w latach 2019–2021 za pomocą jednego z jego modułów, służącego do identyfikacji upraw i szacowania ich powierzchni w skali całego kraju metodami teledetekcji satelitarnej.

Wprowadzenie systemu SATMIROL umożliwiło prezentację danych w ujęciu przestrzennym, na co nie pozwalają tradycyjne metody zbierania danych statystycznych. Dzięki nowej metodzie dane mogą być przedstawiane na wszystkich poziomach podziału administracyjnego kraju – według województw, powiatów i gmin. Zbieranie w sposób tradycyjny nie tylko danych o powierzchni upraw, lecz także pozostałych danych wynikowych, takich jak wstępna ocena przezimowania upraw, wiosenna ocena stanu upraw oraz ocena warunków wzrostu i kondycji roślin, na poziomie agregacji niższym niż województwa jest możliwe tylko w wyniku pełnych spisów rolnych (czyli co 10 lat). Jednak dane te dotyczą powierzchni upraw w gospodarstwach rolnych i nie pozwalają na prezentowanie wyników dla określonych obszarów. A przecież dane dla całego obszaru gminy, powiatu czy województwa są niezbędne do właściwego kształtowania polityk rozwojowych na poziomie krajowym i regionalnym.

System SATMIROL jest jednym z mechanizmów uzyskiwania strategicznych danych z zakresu rolnictwa, pozwalającym odbiorcom na aktualną ocenę sytuacji i bieżące wspieranie działań w ramach polityk rolnych, w tym rolnośrodowiskowych. Przyczynia się tym samym do poprawy realizacji działań w obszarze ochrony i ograniczenia degradacji środowiska przyrodniczego, również w obliczu klęsk żywiołowych.

System ten jest także istotnym elementem wprowadzania innowacyjności w obszarze statystyki rolnictwa. Tradycyjnie stosowane metody są oparte na zbieraniu

danych poprzez ankietowanie. Biorąc pod uwagę duży zakres informacji do zebrania, stanowi to obciążenie dla respondentów i wiąże się z kosztami ponoszonymi na przygotowanie i przeprowadzenie badania ankietowego. System wykorzystujący dane administracyjne, satelitarne i z badania terenowego *in situ* przyczynia się do zmniejszenia obciążenia respondentów i ankieterów oraz obniżenia kosztów badań prowadzonych metodą ankietową, a także wpływa na poprawę jakości danych statystycznych.

Bibliografia

- Belgiu, M., Drąg, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>.
- Cloude, S. R. (2007, 22–26 stycznia). *The Dual Polarisation Entropy/Alpha Decomposition: A PALSAR Case Study* [referat]. POLinSAR 2007, the 3rd International Workshop on Science and Applications of SAR Polarimetry and Polarimetric Interferometry, Frascati.
- Cloude, S. R., Pottier, E. (1997). An entropy based classification scheme for land application of polarimetric SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 35(1), 68–78. <http://doi.org/10.1109/36.551935>.
- Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37(1), 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B).
- Congalton, R. G., Green, K. (2019). *Assessing the Accuracy of Remotely Sensed Data. Principles and Practices*. CRC Press Taylor & Francis Group.
- De Zan, F., Monti Guarnieri, A. (2006). TOPSAR: Terrain Observation by Progressive Scans. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9), 2352–2360.
- Dygaszewicz, J., Szafranski, B. (2018). Badania statystyczne – ujęcie modelowe. *Wiomości Statystyczne*, 63(12), 5–22. <https://doi.org/10.5604/01.3001.0014.0726>.
- European Space Agency. (2013). *Sentinel-1 User Handbook*. https://sedas.satapps.org/wp-content/uploads/2015/07/Sentinel-1_User_Handbook.pdf.
- European Space Agency. (2015). *Sentinel-2 User Handbook*. https://sentinels.copernicus.eu/web/sentinel/user-guides/document-library/-/asset_publisher/xslst4309D5h/content/sentinel-2-user-handbook.
- Główny Urząd Statystyczny. (b.r.). *Satelitarna identyfikacja i monitorowanie upraw na potrzeby statystyki rolnictwa – SATMIROL*. <https://stat.gov.pl/statystyki-eksperymentalne/gospostrateg/satmirol-satelitarna-identyfikacja-i-monitorowanie-upraw-na-potrzeby-statystyki-rolnictwa/>.
- Główny Urząd Statystyczny. (2019). *Raport okresowy nr 1 z realizacji projektu w ramach Programu „Społeczny i gospodarczy rozwój Polski w warunkach globalizujących się rynków” GOSPOSTRATEG*.

- Główny Urząd Statystyczny. (2020). *Raport okresowy nr 3 z realizacji projektu w ramach Programu „Społeczny i gospodarczy rozwój Polski w warunkach globalizujących się rynków” GOSPOSTRATEG.*
- Lang, F., Yang, J., Li, D. (2014). An adaptive enhanced Lee speckle filter for polarimetric SAR image. *Cehui Xuebao / Acta Geodaetica et Cartographica Sinica*, 43(7), 690–697. <https://doi.org/10.13485/j.cnki.11-2089.2014.0112>.
- Ministerstwo Gospodarki. (2012). *Program działań na rzecz rozwoju technologii kosmicznych i wykorzystywania systemów satelitarnych w Polsce.* <https://www.gov.pl/attachment/b4c8fb77-becc-4919-9796-d49cbf67ad0f>.
- Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G. (2016). Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment*, 187, 156–168. <https://doi.org/10.1016/j.rse.2016.10.010>.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>.
- Song, Q., Hu, Q., Zhou, Q., Hovis, C., Xiang, M., Tang, H., Wu, W. (2017). In-Season Crop Mapping with GF-1/WFV Data by Combining Object-Based Image Analysis and Random Forest. *Remote Sensing*, 9(11), 1–19. <https://doi.org/10.3390/rs9111184>.
- Thanh Noi, P., Kappas, M. (2018). Comparison of Random Forest, k -nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification using Sentinel-2 Imagery. *Sensors*, 18(1), 1–20. <https://doi.org/10.3390/s18010018>.
- Zhu, J., Wen, J., Zhang, Y. (2013). A new algorithm for SAR image despeckling using an enhanced Lee filter and median filter. *2013 6th International Congress on Image and Signal Processing (CISP)*, (s. 224–228). <https://doi.org/10.1109/CISP.2013.6743991>.

Anatol Kula (1938–2022)



Fot. Archiwum „WS”

Anatol Kula urodził się 3 lipca 1938 r. w Dąbrowie Górniczej w rodzinie urzędnika miejscowej kopalni. Studiował na Wydziale Technologii Rolno-Spożywczej Szkoły Głównej Gospodarstwa Wiejskiego w Warszawie. Ukończył również Studium Dziennikarskie Uniwersytetu Warszawskiego. W latach 1971–1999 (z kilkuletnią przerwą) był dziennikarzem Polskiego Radia. Początkowo zajmował się tematyką rolniczą, a później – głównie społeczną. Z mikrofonem i magnetofonem zjeździł całą Polskę. To zajęcie znakomicie łączyło się z jego największą życiową pasją – podróżowaniem.

Pracę w Głównym Urzędzie Statystycznym rozpoczął w 1986 r. Najpierw był zatrudniony w Departamencie Informacji i Analiz, następnie w Gabinetzie Prezesa, a od 1992 r. kolejno w Departamencie Analiz i Opracowań Zbiorczych, Departamencie Informacji oraz Biurze Informacji i Wydawnictw. Po przejściu na emeryturę w 2003 r. kontynuował pracę w GUS.

W 1994 r. związał się z „Wiadomościami Statystycznymi” – został sekretarzem redakcji i członkiem Kolegium Redakcyjnego. Zbiegło się to w czasie z objęciem stanowiska redaktora naczelnego „WS” przez prof. dr. hab. Tadeusza Walczaka, pod którego kierunkiem czasopismo coraz wyraźniej otwierało się na środowisko naukowe. Bardzo dobra, bliska współpraca sekretarza redakcji z redaktorem naczelnym niewątpliwie przyczyniła się do wzmocnienia pozycji „WS” jako periodyku naukowego, zyskującego uznanie również poza statystyką publiczną. Anatol Kula był sekretarzem redakcji do 2006 r., a w posiedzeniach Kolegium Redakcyjnego brał udział do 2014 r. Później wspomagał „WS” jako redaktor – aż do 2018 r., kiedy w wieku 80 lat zrezygnował z aktywności zawodowej. Nie można pominąć jego pracy poza czasopiśmem: opracowywał redakcyjnie monografie z serii Biblioteka Wiadomości Statystycznych, a w latach 1989–2000 redagował comiesięczny dodatek GUS do dziennika „Rzeczpospolita” pt. *Statystyka Polski*.

Był postacią nieszablonową. Sam siebie określał mianem obserwatora świata i taki właśnie był: ciekawy ludzi, nowych miejsc, pełen zainteresowań i niesłabnącej potrzeby odkrywania. Niektórzy znali go przede wszystkim jako niestrudzonego turystę i zapalonego kibica sportowego, inni – jako miłośnika kina i teatru, a jeszcze inni – jako smakosza. Długo potrafił opowiadać o swoich wyprawach do Maroka, Francji,

Włoch czy Hiszpanii, które fascynowały go zarówno przyrodniczo, jak i kulturowo. Był erudytą, z którym prowadziło się inspirujące rozmowy, oraz dżentelmenem, jakich dzisiaj już nie ma wielu. Wraz z przyjaciółmi regularnie uczestniczył w spotkaniach kółka brydżowego. Cenił sobie warszawskie życie kulturalne, a w ostatnim czasie, kiedy pandemia wymusiła wiele ograniczeń, odkrył też atuty Legionowa, w którym mieszkał.

Zmarł 21 sierpnia 2022 r. w Nowym Dworze Mazowieckim. Będzie go nam bardzo brakowało.

Redakcja „Wiadomości Statystycznych. The Polish Statistician”

WYDAWNICTWA GUS. SIERPIEŃ 2022 PUBLICATIONS OF STATISTICS POLAND. AUGUST 2022

W ofercie wydawniczej Głównego Urzędu Statystycznego z ubiegłego miesiąca warto zwrócić uwagę na następującą publikację:

Among Statistics Poland's last month's publications, we would like to recommend:



Tytuł: *Regiony Polski 2022*

Title: *Regions of Poland 2022*

Język: polski, angielski

Language: Polish, English

Dodatkowe informacje: opracowanie dostępne w wersji drukowanej i elektronicznej

Additional information: publication available both in the printed and in the electronic form

W najnowszym z folderów *Regiony Polski*, wydawanych z częstotliwością roczną od 2007 r., przedstawiono charakterystykę sytuacji społeczno-gospodarczej makroregionów (NUTS 1) i regionów (NUTS 2) w podziale

statystycznym obowiązującym od 1 stycznia 2021 r. Prezentowane dane dotyczą 2021 r., z uwzględnieniem wstępnych wyników Narodowego Spisu Powszechnego Ludności i Mieszkań 2021, lub ostatniego dostępnego roku. W celu zilustrowania zmian w czasie wybrane kategorie zostały przedstawione w ujęciu dynamicznym. Dodatkowo w publikacji zamieszczono informacje na temat obszarów realizacji zintegrowanych inwestycji terytorialnych miast wojewódzkich.

W sierpniu br. ukazały się ponadto:

- „Biuletyn statystyczny” nr 7/2022;
- *Ceny robót budowlano-montażowych i obiektów budowlanych (czerwiec 2022 r.);*
- *Koniunktura w przetwórstwie przemysłowym, budownictwie, handlu i usługach 2000–2022 (sierpień 2022 r.);*
- *Produkcja ważniejszych wyrobów przemysłowych w lipcu 2022 r.;*
- *Sytuacja społeczno-gospodarcza kraju w lipcu 2022 r.;*
- „Wiadomości Statystyczne. The Polish Statistician” nr 7/2022.

Wszystkie publikacje GUS w formie elektronicznej są dostępne na stronie stat.gov.pl/publikacje/publikacje-a-z.

Electronic forms of all the publications by Statistics Poland are available at stat.gov.pl/en/publications.

Joanna Sadowy

Główny Urząd Statystyczny, Departament Opracowań Statystycznych
Statistics Poland, Statistical Products Department

DLA AUTORÓW FOR THE AUTHORS

(for the English translation of the information given below, please visit ws.stat.gov.pl/ForAuthors)

W „Wiadomościach Statystycznych. The Polish Statistician” („WS”) zamieszczane są artykuły o charakterze naukowym poświęcone teorii i praktyce statystycznej, które prezentują wyniki oryginalnych badań teoretycznych lub analitycznych wykorzystujących metody statystyki matematycznej, opisowej bądź ekonometrii. Ukazują się również artykuły przeglądowe, recenzje publikacji naukowych oraz inne opracowania informacyjne. W czasopiśmie publikowane są prace w języku polskim i angielskim.

Od 2007 r. „WS” znajdują się na liście polskich punktowanych czasopism naukowych MEiN. Zgodnie z komunikatem Ministra Edukacji i Nauki z dnia 1 grudnia 2021 r. w sprawie wykazu czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych „WS” otrzymały 70 punktów.

„Wiadomości Statystyczne. The Polish Statistician” są udostępniane w następujących bazach, repozytoriach, katalogach i wyszukiwarkach: Agro, BazEkon, Central and Eastern European Academic Source (CEEAS), Central and Eastern European Online Library (CEEOL), Central European Journal of Social Sciences and Humanities (CEJSH), Directory of Open Access Journals (DOAJ), EBSCO Discovery Service, European Reference Index for the Humanities and Social Sciences (ERIH Plus), Exlibris Primo, Google Scholar, ICI Journals Master List, ICI World of Journals, Norwegian Register for Scientific Journals and Publishers (The Nordic List) oraz Summon.

Za publikację artykułów na łamach „WS” autorzy nie otrzymują honorariów ani nie wnoszą opłat.

1. Zgłaszanie artykułów

Prace należy przysyłać na adres: redakcja.ws@stat.gov.pl.

Artykuł powinien być utrzymany w formie bezosobowej i zawierać streszczenie, słowa kluczowe, kod/kody JEL oraz ORCID, adres e-mail i afiliację autora. Tytuł, streszczenie, słowa kluczowe i afiliacja powinny być podane w języku polskim i angielskim.

Jeżeli w pracy występują tablice, wykresy lub mapy, powinny być umieszczone w treści artykułu. Ponadto należy je przesłać osobno (najlepiej w formacie Excel); szczegółowe informacje są zawarte w pkt 4.2.14.

Autor jest zobowiązany do podania w artykule wszelkich źródeł finansowania badań będących podstawą pracy. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub sympozjum naukowego albo jeżeli artykuł został przez autora umieszczony w repozytorium internetowym lub zgłoszony w innym wydawnictwie do opublikowania w innej wersji językowej, to podczas składania tekstu do publikacji w „WS” należy poinformować o tym redakcję.

Prosimy o niestosowanie stylów i ograniczenie formatowania do wymogów redakcyjnych. Więcej informacji w rozdziale *Wymogi redakcyjne*.

Razem z artykułem należy przesłać skan oświadczenia (do pobrania ze strony internetowej czasopisma) o oryginalności pracy, niezłożeniu jej w innym wydawnictwie i udzieleniu wydawcy „WS” licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0), zawierającego numer ORCID, afiliację lub afiliacje oraz dane kontaktowe autora, wraz ze wskazaniem proponowanego działu czasopisma.

Załączenie skanu oświadczenia jest warunkiem poddania pracy ocenie wstępnej i skierowania do recenzji.

2. Przebieg prac redakcyjnych

Zgłoszony artykuł jest oceniany i opracowywany w czteroetapowym procesie:

1. **Ocena wstępna**, dokonywana przez redakcję. Polega na weryfikacji naukowego charakteru artykułu oraz jego struktury i zawartości pod kątem wymogów redakcyjnych, a także zgodności tematyki z profilem czasopisma. Autor uzupełnia i poprawia artykuł stosownie do uwag redakcji, a w przypadku nieuwzględnienia danej uwagi uzasadnia swoje stanowisko. **Razem z poprawionym artykułem autor przesyła w osobnym pliku zanonimizowaną wersję pracy, przeznaczoną do recenzji.** Anonimizacja polega na utajnieniu nazwiska autora (także we właściwościach pliku), usunięciu podziękowań i informacji o źródłach finansowania, a także innych informacji wskazujących na afiliację lub umożliwiających zidentyfikowanie autora. Warunkiem skierowania pracy do recenzji jest potwierdzenie oryginalności tekstu uzyskane za pomocą systemu antyplagiatowego Similarity Check. W przypadku wykrycia znacznego podobieństwa do innych prac artykuł zostanie odrzucony.
2. **Ocena recenzentów**, dokonywana przez specjalistów w danej dziedzinie. Artykuł oceniają dwaj recenzenci spoza jednostki naukowej, przy której afiliowany jest autor; w przypadku pracy w języku angielskim co najmniej jeden recenzent jest afiliowany przy jednostce zagranicznej. W razie sprzecznych opinii dwóch recenzentów powoływany jest trzeci recenzent. Recenzenci kierują się kryteriami oryginalności i jakości opracowania zarówno w odniesieniu do treści, jak i formy artykułu.

Autorzy artykułów, które otrzymały pozytywne oceny, wprowadzają poprawki zalecane przez recenzentów i przesyłają do redakcji zmodyfikowaną wersję pracy. Jeśli pojawi się różnica zdań dotycząca zasadności proponowanych zmian, autorzy są zobligowani do uzasadnienia swojego stanowiska.

3. **Ocena Kolegium Redakcyjnego (KR)**, decydująca o przyjęciu pracy do publikacji. Jest dokonywana na podstawie recenzji, z uwzględnieniem opinii redaktorów tematycznego i merytorycznego. Polega m.in. na weryfikacji dokonania przez autora zmian w artykule stosownie do uwag recenzentów. KR ocenia artykuł pod względem poprawności i spójności merytorycznej oraz zaleca autorowi wprowadzenie poprawek, jeśli są one konieczne, aby praca spełniała wymogi czasopisma. Autorowi przysługuje prawo do odwołania od decyzji o niepublikowaniu artykułu. W takim przypadku powinien on skontaktować się z redakcją „WS” i przedstawić uzasadnienie. Ostateczna decyzja w tej sprawie należy do redaktora naczelnego.

W „WS” publikowane są wyłącznie te artykuły, które otrzymają pozytywną ocenę na każdym z wymienionych etapów i zostaną poprawione przez autora zgodnie z otrzymanymi uwagami (chyba że autor przedstawi argumenty uzasadniające nieuwzględnienie danej uwagi).

Artykuły przyjęte przez KR do publikacji są zamieszczane na stronie internetowej czasopisma w zakładce Early View, gdzie znajdują się do czasu opublikowania w konkretnym wydaniu „WS”.

4. **Opracowanie redakcyjne, autoryzacja i korekta.** Artykuł zakwalifikowany do druku jest poddawany opracowaniu merytorycznemu i językowemu. Redakcja zastrzega sobie prawo

do zmiany tytułu i śródtytułów, modyfikowania tablic, wykresów i innych elementów graficznych oraz przerezegowania treści bez naruszenia zasadniczej myśli autora.

Po opracowaniu redakcyjnym artykuł jest przesyłany do autoryzacji. Tekst zatwierdzony przez autora, po składzie i łamaniu, jest poddawany korekcie i rewizji (II korekcie). Autor dokonuje korekty autorskiej tekstu na etapie rewizji. Wykresy i inne materiały graficzne są opracowywane na podstawie plików i danych przekazanych przez autora i podawane korekcie i rewizji. Autor dokonuje ich akceptacji na etapie rewizji.

W przypadku odkrycia błędów w opublikowanym artykule zamieszcza się na łamach „WS” sprostowanie, a artykuł w wersji elektronicznej jest poprawiany i umieszczany na stronie internetowej „WS” ze stosownym wyjaśnieniem.

3. Zasady etyki publikacyjnej COPE

Redakcja „WS” podejmuje wszelkie starania w celu utrzymania najwyższych standardów etycznych zgodnie z wytycznymi Komitetu ds. Etyki Publikacyjnej (COPE), dostępnymi na stronie internetowej www.publicationethics.org, oraz wykorzystuje wszystkie możliwe środki mające na celu zapobieżenie nadużyciom i nierzetelności autorskiej. Przyjęte zasady postępowania obowiązują autorów, Radę Naukową, Kolegium Redakcyjne, redakcję, pracowników Wydziału Czasopism Naukowych GUS, recenzentów i wydawcę.

3.1. Odpowiedzialność autorów

1. Artykuły naukowe kierowane do opublikowania w „WS” powinny zawierać precyzyjny opis badanych zjawisk i stosowanych metod oraz autorskie wnioski i sugestie dotyczące rozwoju badań i analiz statystycznych. Autorzy powinni wyraźnie określić cel artykułu oraz jasno przedstawić wyniki przeprowadzonej analizy. Prezentacja efektów badań statystycznych zaprojektowanych i przeprowadzonych przez autorów wymaga opisanie zastosowanej w nich metodologii. W przypadku nowatorskich metod analizy pożądane jest podanie przykładu ilustrującego ich zastosowanie w praktyce statystycznej. Autorzy ponoszą odpowiedzialność za treści prezentowane w artykułach. W razie zgłaszania przez czytelników zastrzeżeń odnoszących się do tych treści autorzy są zobligowani do udzielenia odpowiedzi za pośrednictwem redakcji.
2. Na autorach spoczywa obowiązek zapewnienia pełnej oryginalności przedłożonych prac. Redakcja nie toleruje przejawów nierzetelności naukowej autorów, takich jak:
 - duplikowanie publikacji – ponowne publikowanie własnego utworu lub jego części;
 - plagiat – przywłaszczenie cudzego utworu lub jego fragmentu bez podania informacji o źródle;
 - fabrykowanie danych – oparcie pracy naukowej na nieprawdziwych wynikach badań;
 - autorstwo widmo (*ghost authorship*) – nieujawnianie współautorów, mimo że wnieśli oni istotny wkład w powstanie artykułu;
 - autorstwo gościnne (*guest authorship*) – podawanie jako współautorów osób o znikomym udziale lub niebiorących udziału w opracowywaniu artykułu;
 - autorstwo grzecznościowe (*gift authorship*) – podawanie jako współautorów osób, których wkład jest oparty jedynie na słabym powiązaniu z badaniem.

Autorzy deklarują w stosownym oświadczeniu, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i nie został złożony w innym wydawnictwie oraz że jest ich oryginalnym dziełem, i określają swój wkład w opracowanie artykułu. Jeżeli doszło do zaprezentowania podobnych materiałów podczas konferencji lub sympozjum naukowego, to podczas składania tekstu do publikacji w „WS” autorzy są zobowiązani poinformować o tym redakcję.

3. Autorzy są zobowiązani do podania w treści artykułu wszelkich źródeł finansowania badań będących podstawą pracy.
4. Główną odpowiedzialność za rzetelność przekazanych informacji, łącznie z informacją na temat wkładu poszczególnych współautorów w powstanie artykułu, ponosi zgłaszający artykuł.
5. Autorzy zgłaszający artykuły do publikacji w „WS” biorą udział w procesie recenzji double-blind peer review, dokonywanej przez co najmniej dwóch niezależnych ekspertów z danej dziedziny. Po otrzymaniu pozytywnych recenzji autorzy wprowadzają zalecane przez recenzentów poprawki i dostarczają redakcji zaktualizowaną wersję opracowania wraz z pisemnym poświadczeniem uwzględnienia poprawek. Jeśli pojawi się różnica zdań co do zasadności proponowanych zmian, należy wyjaśnić, które poprawki zostały uwzględnione, a w przypadku ich nieuwzględnienia – uzasadnić swoje stanowisko.
6. Jeżeli autorzy odkryją w swoim maszynopisie lub tekście już opublikowanym błędy, nieścisłości bądź niewłaściwe dane, powinni niezwłocznie poinformować o tym redakcję w celu dokonania korekty, wycofania tekstu lub zamieszczenia sprostowania. W przypadku korekty artykułu już opublikowanego jego nowa wersja jest zamieszczana na stronie internetowej „WS” wraz ze stosownym wyjaśnieniem.

3.2. Odpowiedzialność Rady Naukowej, Kolegium Redakcyjnego i Wydziału Czasopism Naukowych GUS

1. Rada Naukowa (RN) kształtuje profil programowy czasopisma, określa kierunki jego rozwoju i konsultuje jego zakres merytoryczny.
2. Kolegium Redakcyjne (KR) podejmuje decyzję o publikacji danego artykułu z uwzględnieniem ocen recenzentów oraz opinii zespołu redakcyjnego. W swoich rozstrzygnięciach członkowie KR kierują się kryteriami merytorycznej oceny wartości artykułu, jego oryginalności i jasności przekazu, a także ścisłego związku z celem i zakresem tematycznym „WS”. Oceniają artykuły niezależnie od płci, rasy, pochodzenia etnicznego, narodowości, religii, wyznania, światopoglądu, niepełnosprawności, wieku lub orientacji seksualnej ich autorów.
3. Zespół redakcyjny, wyodrębniony z KR, tworzą redaktor naczelny i jego zastępca, redaktorzy tematyczni i redaktor merytoryczny. Członkowie zespołu redakcyjnego weryfikują nadsyłane artykuły pod względem merytorycznym, oceniają ich zgodność z celem i zakresem tematycznym „WS” oraz sprawdzają spełnienie wymogów redakcyjnych i przestrzeganie zasad rzetelności naukowej. Ponadto wybierają recenzentów w taki sposób, aby nie wystąpił konflikt interesów, i dbają o zapewnienie uczciwego, bezstronnego i terminowego procesu recenzowania.
4. Za sprawny przebieg procesu wydawniczego, poinformowanie wszystkich jego uczestników o konieczności przestrzegania obowiązujących zasad i przygotowanie artykułów do

publikacji odpowiadają pracownicy Wydziału Czasopism Naukowych (WCN) GUS. W celu uzyskania obiektywnej oceny oryginalności nadsyłanych artykułów przed skierowaniem ich do recenzji WCN wykorzystuje system antyplagiacyjny. Informacje dotyczące artykułu mogą być przekazywane przez WCN wyłącznie autorom, recenzentom, członkom RN i KR oraz wydawcy.

5. Zmiany dokonane w tekście na etapie przygotowania artykułu do publikacji nie mogą naruszać zasadniczej myśli autorów. Wszelkie modyfikacje o charakterze merytorycznym są z nimi konsultowane.
6. W przypadku podjęcia decyzji o niepublikowaniu artykułu nie może on zostać w żaden sposób wykorzystany przez wydawcę lub uczestników procesu wydawniczego bez pisemnej zgody autorów. Autorzy mogą się odwołać od decyzji o niepublikowaniu artykułu. W tym celu powinni się skontaktować z redaktorem naczelnym lub sekretarzem redakcji „WS” i przedstawić stosowną argumentację. Odwołania autorów są rozpatrywane przez redaktora naczelnego.
7. Członkowie RN i KR ani pracownicy WCN nie mogą pozostawać w jakimkolwiek konflikcie interesów w odniesieniu do artykułów zgłaszanych do publikacji. Przez konflikt interesów należy rozumieć sytuację, w której jakiegokolwiek interesy lub zależności (służbowe, finansowe lub inne) mogą mieć wpływ na ocenę artykułu lub decyzję o jego publikacji.
8. W celu przeciwdziałania nierzetelności naukowej wymagane jest złożenie przez autorów oświadczenia, w którym deklarują, że zgłaszany artykuł nie narusza praw autorskich osób trzecich, nie był dotychczas publikowany i jest ich oryginalnym dziełem, a także określają swój wkład w opracowanie artykułu.
9. W celu zapewnienia wysokiej jakości recenzji wymagane jest złożenie przez recenzentów oświadczenia o przestrzeganiu zasad etyki recenzowania COPE i niewystępowaniu konfliktu interesów.
10. W przypadku uzasadnionego podejrzenia na jakimkolwiek etapie procesu wydawniczego, że autorzy dopuścili się nierzetelności naukowej (zob. pkt 3.1. Odpowiedzialność autorów), zespół redakcyjny skrupulatnie zbada sprawę ewentualnego nadużycia. Jeśli nierzetelność autorów zostanie udowodniona, to zgłoszony przez nich artykuł zostanie odrzucony przez KR, a autorzy otrzymają informację o podjętej decyzji wraz z jej uzasadnieniem.
11. Czytelnicy, którzy mają wobec autorów opublikowanego artykułu uzasadnione podejrzenia o nierzetelność naukową, powinni powiadomić o tym redaktora naczelnego lub sekretarza redakcji. Po zbadaniu sprawy ewentualnego nadużycia czytelnicy zostaną poinformowani o rezultacie przeprowadzonego postępowania. W przypadku potwierdzenia nadużycia, na łamach czasopisma zostanie zamieszczona stosowna informacja.

3.3. Odpowiedzialność recenzentów

1. Recenzenci przyjmują artykuł do recenzji tylko wtedy, gdy uznają, że:
 - posiadają odpowiednią wiedzę w określonej dziedzinie, aby rzetelnie ocenić pracę;
 - zgodnie z ich stanem wiedzy nie istnieje konflikt interesów w odniesieniu do autorów, przedstawionych w artykule badań i instytucji je finansujących, co potwierdzają w oświadczeniu;
 - mogą wywiązać się z terminu ustalonego przez redakcję, aby nie opóźnić publikacji.

2. Recenzenci są zobligowani do zachowania obiektywności i poufności oraz powstrzymania się od osobistej krytyki. Zawsze powinni uzasadnić swoją ocenę, przedstawiając stosowną argumentację.
3. Recenzenci powinni wskazać ważne dla wyników badań opublikowane prace, które w ich ocenie powinny zostać przywołane w ocenianym artykule.
4. W razie stwierdzenia wysokiego poziomu zbieżności treści recenzowanej pracy z innymi opublikowanymi materiałami lub podejrzenia innych przejawów nierzetelności naukowej recenzenci są zobowiązani poinformować o tym redakcję.
5. Po ukończeniu recenzji przechowywanie przesłanych przez redakcję materiałów (w jakiegokolwiek formie) oraz posługiwanie się nimi przez recenzentów jest niedozwolone.

3.4. Odpowiedzialność wydawcy

1. Materiały opublikowane w „WS” są chronione prawem autorskim.
2. Wydawca udostępnia pełną treść wszystkich artykułów w internecie w trybie otwartego dostępu, tj. bezpłatnie i bez technicznych ograniczeń, od 1 stycznia 2022 r. na licencji Creative Commons Uznanie autorstwa – Na tych samych warunkach 4.0 (CC BY-SA 4.0). W przypadku artykułów zgłoszonych do „WS” od 2022 r. dozwolone jest dzielenie się artykułem (kopiowanie i rozpowszechnianie go w dowolnym medium i formie) oraz adaptowanie go (w dowolnym celu, także komercyjnym) na warunkach określonych w tej licencji. Z pozostałych artykułów zamieszczonych w czasopiśmie można korzystać w ramach otwartego dostępu, zgodnie z ustawą o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego.
3. Wydawca deklaruje gotowość do opublikowania poprawek, wyjaśnień oraz przeprosin.

4. Wymogi redakcyjne

Zgodnie z wymogami czasopisma omawiany w artykule problem badawczy powinien być jednoznacznie zdefiniowany oraz istotny dla oceny zjawisk społecznych lub gospodarczych. Artykuł powinien zawierać wyraźnie określony cel badania, precyzyjny opis badanych zjawisk i stosowanych metod, uzyskane wyniki przeprowadzonej analizy oraz autorskie wnioski.

4.1. Struktura i zawartość artykułu

Wymagane elementy artykułu:

1. Tytuł.
2. Dane autora: imię i nazwisko, afiliacja w języku polskim i angielskim, ORCID, adres e-mail. Wśród autorów artykułu wieloautorskiego należy wskazać autora korespondencyjnego.
3. Streszczenie (zalecana objętość – do 1200 znaków ze spacjami, forma bezosobowa). W przypadku artykułu opisującego badanie empiryczne powinno zawierać: cel, przedmiot, okres i metodę badania, źródła danych i najważniejsze wnioski z badania. W przypadku artykułów o innym charakterze należy podać co najmniej cel pracy, jej przedmiot i najważniejsze wnioski.

Streszczenie to podstawowe źródło informacji o artykule, warunkujące też decyzję czytelnika o zapoznaniu się z całą pracą. Dlatego powinno być przygotowane ze szczególną starannością i dbałością o umieszczenie w nim wszystkich wymaganych elementów.

4. Słowa kluczowe – najistotniejsze pojęcia lub wyrażenia użyte w pracy (nie mniej niż trzy). Powinny być zawarte w streszczeniu i/lub tytule.
5. Kod/kody z klasyfikacji Journal of Economic Literature (JEL).
6. Tłumaczenie tytułu, streszczenia i słów kluczowych (na język angielski w przypadku artykułu napisanego w języku polskim, a na język polski w przypadku artykułu napisanego w języku angielskim).
7. W artykule opisującym badanie empiryczne wymagane są następujące części:
 - wprowadzenie, zawierające syntetyczne przedstawienie zagadnień teoretycznych, uzasadnienie podjęcia danego problemu badawczego, cel badania i krytyczne odniesienie do literatury przedmiotu. W wyjątkowych przypadkach, kiedy istotne dla podjętego tematu jest obszerniejsze przedstawienie dyskusji toczącej się w literaturze, przegląd literatury może stanowić odrębną część artykułu;
 - metoda badania, uwzględniająca przedmiot i okres badania, źródła danych i zastosowane metody badawcze, w tym uzasadnienie ich wyboru;
 - wyniki badania – analiza danych oraz interpretacja wyników i odniesienie ich do rezultatów wcześniejszych badań (dyskusja). W uzasadnionych przypadkach dyskusja może stanowić odrębną część artykułu;
 - podsumowanie, które powinno być związane i odzwierciedlać istotę problemu badawczego przedstawionego w artykule, bez podawania danych liczbowych; końcowe wnioski powinny odnosić się do treści artykułu, a w szczególności do celu badania.Wszystkie części powinny być opatrzone numerami.
8. Bibliografia, zawierająca pełny wykaz prac i materiałów przywołanych w artykule, przygotowana zgodnie z wymogami czasopisma.

4.2. Przygotowanie artykułu

1. Artykuł powinien być utrzymany w formie bezosobowej.
2. Tekst należy zapisać alfabetem łacińskim. Nazwy własne, tytuły itp. oryginalnie zapisane innym alfabetem powinny być poddane transliteracji.
3. Nie należy stosować stylów; formatowanie należy ograniczyć do wymogów redakcyjnych.
4. Objętość artykułu łącznie ze streszczeniem, słowami kluczowymi, bibliografią, tablicami, wykresami i innymi materiałami graficznymi nie powinna być mniejsza niż 10 stron maszynopisu ani przekraczać 20 stron.
5. Edytor tekstu: Microsoft Word, format *.doc lub *.docx.
6. Krój czcionki:
 - Arial – tytuł, autor, streszczenia, słowa kluczowe, kody JEL, śródtytuły, elementy graficzne (tablice, zestawienia, wykresy, schematy), przypisy;
 - Times New Roman – tekst główny, bibliografia.
7. Wielkość czcionki:
 - 14 pkt – tytuł, autor, tytuły rozdziałów;
 - 12 pkt – tekst główny, tytuły podrozdziałów;
 - 10 pkt – pozostałe elementy.
8. Marginesy – 2,5 cm z każdej strony.

9. Interlinia – 1,5 wiersza; tablice i przypisy – 1 wiersz; przed tytułami rozdziałów i podrozdziałów oraz po nich – pusty wiersz.
10. Wcięcie akapitowe – 0,4 cm; bibliografia – bez wcięcia, wysunięcie 0,4 cm.
11. Przy wycienieniach należy posłużyć się listą punktowaną z punktarami w postaci kropek (wysunięcie 0,4 cm, wcięcie 0 cm); wiersze (oprócz ostatniego) zakończone średnikiem.
12. Strony ponumerowane automatycznie.
13. Tablice i elementy graficzne (wykresy, mapy, schematy) muszą być przywołane w tekście.
14. Wykresy, mapy i schematy należy zamieścić w tekście głównym. Wykresy powinny być edytowalne (optymalnie wykonane w programie Excel; w przypadku wykonania w programie graficznym powinny mieć postać wektorową). Wykresy i inne materiały graficzne należy przekazać osobno w pliku programu Excel lub innym edytowalnym w pakiecie Microsoft Office. W przypadku wykonania ich w innym programie odpowiedni format pliku będzie ustalany indywidualnie.
15. Tablice muszą być edytowalne. Nie należy stosować rastrów, cieniowania, pogrubiania czy też podwójnych linii itp.
16. Wskazówki dotyczące opracowywania map znajdują się w publikacji *Mapy statystyczne. Opracowanie i prezentacja danych*, dostępnej na stronie internetowej GUS.
17. Pod tablicami i każdym elementem graficznym należy podać źródło opracowania, a także objaśnić użyte w nich skróty i symbole.
18. Literowe symbole liczb i innych wielkości niezłożonych należy zapisywać małą lub dużą literą i pismem pochyłym (np. a , A , $y(x)$, a_i); wektorów – pismem pochyłym i pogrubionym (np. \mathbf{a} , \mathbf{A} , \mathbf{w} , $\mathbf{y}(x)$, \mathbf{w}_i); macierzy – pismem prostym i pogrubionym (np. \mathbf{A} , \mathbf{a} , \mathbf{M} , $\mathbf{Y}(x)$, \mathbf{M}_i).
19. Objasnienia znaków umownych w tablicach: kreska (–) – zjawisko nie wystąpiło; zero (0) – zjawisko istniało w wielkości mniejszej od 0,5; (0,0) – zjawisko istniało w wielkości mniejszej od 0,05; kropka (.) – brak informacji, konieczność zachowania tajemnicy statystycznej, wypełnienie pozycji jest niemożliwe lub niecelowe; „w tym” – oznacza, że nie podaje się wszystkich składników sumy.
20. Stosowane są następujące skróty: tablica – tabl., wykres – wyk.
21. Wszystkie zawarte w artykule informacje, dane i stwierdzenia wykraczające poza wiedzę powszechną – np. wyniki badań innych autorów, zarówno o charakterze empirycznym, jak i koncepcyjnym – muszą być opatrzone przypisem bibliograficznym. Przez wiedzę powszechną należy rozumieć informacje ogólnie znane i niebudzące wątpliwości ani kontrowersji w danej grupie społecznej, np. utworzenie GUS w 1918 r. lub powstanie UE w 1993 r. na podstawie traktatu z Maastricht. Natomiast dane statystyczne udostępniane lub publikowane np. przez GUS lub Eurostat nie należą do takich informacji. Charakteru wiedzy powszechnej nie mają również stwierdzenia odnoszące się do idei, zjawisk i procesów społecznych, politycznych czy gospodarczych. Nawet pozornie zdroworozsądkowe idee zmieniają bowiem swój sens w zależności od kultury, języka lub dyscypliny naukowej, a także bywają w rozmaity sposób konceptualizowane, jak np. pojęcie poznania w naukach społecznych.

Podanie źródła jest konieczne niezależnie od tego, czy informacje lub stwierdzenia są ujęte w ramy cytatu, czy przedstawione bez dosłownego przytoczenia, np. w formie parafrazy. Jeżeli stwierdzenie może budzić jakiegokolwiek wątpliwości odbiorców, autor powinien wskazać stosowne źródło podawanej informacji.

22. Przypisy rzeczowe, słownikowe lub informacyjne należy umieszczać na dole strony. Przypisy bibliograficzne, zgodnie ze standardem APA (American Psychological Association), należy podawać w tekście głównym.
23. Bibliografię należy przygotować zgodnie ze standardem APA.

4.3. Zasady przywoływania publikacji w treści artykułu

Wyszczególnienie	Przykład przywołania	
	w odsyłaczu	w treści zdania
Autor indywidualny		
Jeden autor	(Iksiński, 2001)	Iksiński (2001)
Dwóch autorów	(Iksiński i Nowak, 1999)	Iksiński i Nowak (1999)
Trzech autorów lub więcej	(Jankiewicz i in., 2003)	Jankiewicz i in. (2003)
Autor instytucjonalny		
Nazwa funkcjonuje jako powszechnie znany skrótowiec: pierwsze przywołanie w tekście	(International Labour Organization [ILO], 2020)	International Labour Organization (ILO, 2020)
kolejne przywołanie	(ILO, 2020)	ILO (2020)
Pełna nazwa	(Stanford University, 1995)	Stanford University (1995)
Typ publikacji		
Publikacja bez ustalonego autorstwa	(<i>Skrócony tytuł ...</i> , 2015)	<i>Pełny tytuł</i> (2015)
Publikacja bez roku wydania	(Iksiński, b.r.)	Iksiński (b.r.)
Akt prawny	(Pełny tytuł)	Pełny tytuł
Strona internetowa / Zbiór danych: znana data publikacji	(Iksiński, 2020) / (Nazwa instytucji, 2020)	Iksiński (2020) / Nazwa instytucji (2020)
nieznana data publikacji	(Iksiński, b.r.) / (Nazwa instytucji, b.r.)	Iksiński (b.r.) / Nazwa instytucji (b.r.)
Rodzaj przywołania		
Przywoływanie kilku prac (porządek prac – chronologiczny, porządek autorów – alfabetyczny)	(Iksiński, 1997, 1999, 2004a, 2004b; Nowak, 2002)	Iksiński (1997, 1999, 2004a, 2004b) i Nowak (2002)
Przywoływanie publikacji za innym autorem (uwaga: w bibliografii należy wymienić tylko pracę czytaną)	(Nowakowski, 1990, za: Zienniecka, 2007)	Nowakowski (1990, za: Zienniecka, 2007)

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

4.4. Przykłady opisu bibliograficznego

Bibliografia powinna być zamieszczona na końcu opracowania. Prace należy uszeregować alfabetycznie według nazwiska pierwszego autora. W przypadku dwóch lub więcej prac tego samego autora / tych samych autorów trzeba je uporządkować chronologicznie według roku publikacji. Jeśli kilka prac tego samego autora / tych samych autorów zostało opublikowanych w tym samym roku, należy podać je w kolejności alfabetycznej według tytułu i odpowiednio oznaczyć literami a, b, c itd.

Typ publikacji	Przykład opisu bibliograficznego
Artykuł w czasopiśmie	
W wersji drukowanej	Nazwisko, X. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik (zeszyt)</i> , strona początku–strona końca.
Dostępny w internecie, z DOI	Nazwisko, X., Nazwisko 2, Y. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik(zeszyt)</i> , strona początku–strona końca. https://doi.org/xxx .
Dostępny w internecie, bez DOI	Nazwisko, X., Nazwisko 2, Y., Nazwisko 3, Z. (rok). Tytuł artykułu. <i>Tytuł czasopisma, rocznik(zeszyt)</i> , strona początku–strona końca. https://xxx .
Maszynopis	
Niepublikowany / przygotowywany przez autora / zgłoszony do publikacji, ale jeszcze niezaakceptowany	Nazwisko, X. (rok). <i>Tytuł</i> [maszynopis niepublikowany / w przygotowaniu / zgłoszony do publikacji].
Zaakceptowany do publikacji	Nazwisko, X. (w druku). Tytuł artykułu. <i>Tytuł czasopisma</i> .
Opublikowany nieformalnie (np. na stronie internetowej autora)	Nazwisko, X., Nazwisko 2, Y. (rok). <i>Tytuł artykułu</i> . https://xxx .
Opublikowany w trybie early view / ahead of print / online first	Nazwisko, X. (rok). Tytuł artykułu. <i>Tytuł czasopisma</i> . Early view / Ahead of print / Online first. https://xxx .
Książka	
W wersji drukowanej	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo.
Dostępna w internecie, z DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo. https://doi.org/xxx .
Dostępna w internecie, bez DOI	Nazwisko, X. (rok). <i>Tytuł książki</i> . Wydawnictwo. https://xxx .
W przekładzie	Nazwisko, X. (rok). <i>Tytuł książki</i> (tłum. Y. Nazwisko). Wydawnictwo.
Wydanie wielotomowe: tom zatytułowany	Nazwisko, X. (rok). <i>Tytuł książki: nr tomu. Tytuł tomu</i> . Wydawnictwo.
tom niezatytułowany	Nazwisko, X. (rok). <i>Tytuł książki (nr tomu)</i> . Wydawnictwo.
Kolejne wydanie	Nazwisko, X. (rok). <i>Tytuł książki (nr wydania)</i> . Wydawnictwo.
Pod redakcją: w języku polskim	Nazwisko, X. (red.). (rok). <i>Tytuł książki</i> . Wydawnictwo.
w języku angielskim	Nazwisko, X. (Ed.). (rok). <i>Tytuł książki</i> . Wydawnictwo.
Rozdział w pracy zbiorowej	Nazwisko, X. (rok). Tytuł rozdziału. W: Y. Nazwisko, Z. Nazwisko 2 (red.), <i>Tytuł książki</i> (s. strona początku–strona końca). Wydawnictwo. https://doi.org/xxx lub https://xxx .
Inne prace	
Raport: autor indywidualny	Nazwisko, X. (rok). <i>Tytuł raportu</i> . Wydawnictwo.
autor instytucjonalny	Nazwa instytucji. (rok). <i>Tytuł raportu</i> . Wydawnictwo.
Working Papers	Nazwisko, X. (rok). <i>Tytuł pracy</i> (nazwa serii i numer). https://doi.org/xxx lub https://xxx .
Sesja konferencyjna / prezentacja / referat	Nazwisko, X. (rok, dzień i miesiąc). <i>Tytuł pracy</i> [typ wystąpienia, np. referat]. Nazwa konferencji, miejsce konferencji.
Rozprawa doktorska: nieopublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [nieopublikowana rozprawa doktorska]. Nazwa instytucji nadającej tytuł doktorski.
opublikowana	Nazwisko, X. (rok). <i>Tytuł pracy</i> [rozprawa doktorska, nazwa instytucji nadającej tytuł doktorski]. https://xxx .
Akt prawny	Pełny tytuł aktu prawnego wraz z datą publikacji w dzienniku urzędowym.

Typ publikacji	Przykład opisu bibliograficznego
Strona internetowa	
Znana data publikacji, zawartość strony się nie zmienia	Nazwisko, X. (rok, dzień i miesiąc). <i>Tytuł</i> . https://xxx .
Nieznana data publikacji, zawartość strony się zmienia	Nazwa instytucji. (b.r.). <i>Tytuł</i> . Pobrane dzień, miesiąc i rok pobrania z https://xxx .
Zbiór danych	
Surowe dane nieopublikowane	Nazwisko, X. (rok wydania pracy, w której dane są wykorzystywane) [opis danych, np. surowe dane nieopublikowane dotyczące...]. Źródło danych (np. nazwa uniwersytetu).
Dane opublikowane: znana data publikacji, zawartość zbioru się nie zmienia	Nazwisko, X. (rok). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. https://xxx .
nieznana data publikacji, zawartość zbioru się zmienia	Nazwa instytucji. (b.r.). <i>Nazwa zbioru danych</i> [zbiór danych]. Wydawca. Pobrane dzień, miesiąc i rok pobrania z https://xxx .

Źródło: opracowanie na podstawie: American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th edition). <https://doi.org/10.1037/0000165-000>.

Praca przygotowana w sposób niezgodny z powyższymi wskazówkami będzie odesłana do autora z prośbą o dostosowanie formy artykułu do wymogów redakcyjnych.

DZIAŁY „WS” – TEMATYKA ARTYKUŁÓW WS SECTIONS – TOPICS OF THE ARTICLES

Pełny opis zakresu tematycznego działów: ws.stat.gov.pl/AimScope

Description of the topics covered in each section: ws.stat.gov.pl/AimScope

Studia metodologiczne / Methodological studies

- Oryginalne teoretyczne rozwiązania metodologiczne ze wskazaniem ich praktycznej użyteczności
- Prace przeglądowe i porównawcze oraz dotyczące etyki w statystyce, które wnoszą pionierski wkład poznawczy do obecnego stanu wiedzy

Statystyka w praktyce / Statistics in practice

- Nowatorskie zastosowania narzędzi i modeli statystycznych oraz analiza i ocena statystyczna zjawisk społeczno-ekonomicznych i innych, prowadzona w szczególności na danych z zasobów statystyki publicznej
- Wykorzystanie narzędzi informatycznych do uzyskiwania i przetwarzania informacji statystycznych, naliczania danych wynikowych, ich prezentacji i rozpowszechniania
- Projektowanie badań statystycznych, uzyskiwanie, integracja i przetwarzanie danych oraz generowanie wynikowych informacji statystycznych i kontrola ich ujawniania

Studia interdyscyplinarne. Wyzwania badawcze / Interdisciplinary studies. Research challenges

- Wyzwania badawcze wynikające z rosnących potrzeb użytkowników danych statystycznych i wymagające zaangażowania znacznych środków oraz rozwiązań z różnych dziedzin nauki i techniki
- Wykorzystanie technologii informacyjnych i komunikacyjnych, innowacyjność, przetwarzanie i analiza zagadnień związanych z data science i big data
- Wyniki badań prowadzonych przez przedstawicieli dyscyplin innych niż statystyka z wykorzystaniem metod statystycznych

Spisy powszechne – problemy i wyzwania / Issues and challenges in census taking

- Propozycje rozwiązań – zarówno organizacyjnych, jak i metodologicznych – możliwych do zastosowania w spisach oraz rezultaty analiz danych spisowych
- Praktyczne aspekty związane z gromadzeniem i udostępnianiem danych ze spisów, w tym dotyczące obciążenia odpowiedzi i ochrony tajemnicy statystycznej

Edukacja statystyczna / Statistical education

- Metody i efekty nauczania statystyki oraz popularyzacja myślenia statystycznego i rzetelnego posługiwania się informacjami statystycznymi
- Problemy związane z kształceniem w zakresie umiejętności stosowania statystyki na wszystkich poziomach edukacji, a także dotyczące wykorzystywania nowoczesnych koncepcji i metod dydaktycznych oraz pomocy naukowych w nauczaniu statystyki

Z dziejów statystyki / From the history of statistics

- Historia prowadzenia obserwacji statystycznych oraz rozwoju ich metodologii i narzędzi
- Życie i osiągnięcia zawodowe wybitnych statystyków, jak również działalność najważniejszych instytucji i organizacji statystycznych w Polsce i za granicą

In memoriam

- Nekrologi i artykuły wspomnieniowe

Informacje. Recenzje. Dyskusje / Discussions. Reviews. Information

- Teksty nierecenzowane i niemające charakteru artykułów naukowych: sprawozdania z konferencji naukowych i innych wydarzeń dotyczących statystyki, recenzje książek, omówienia nowości wydawniczych GUS, polemiki i dyskusje