

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 Geoenineering, 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 ogrodniczych. 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 ogrodniczych* (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):

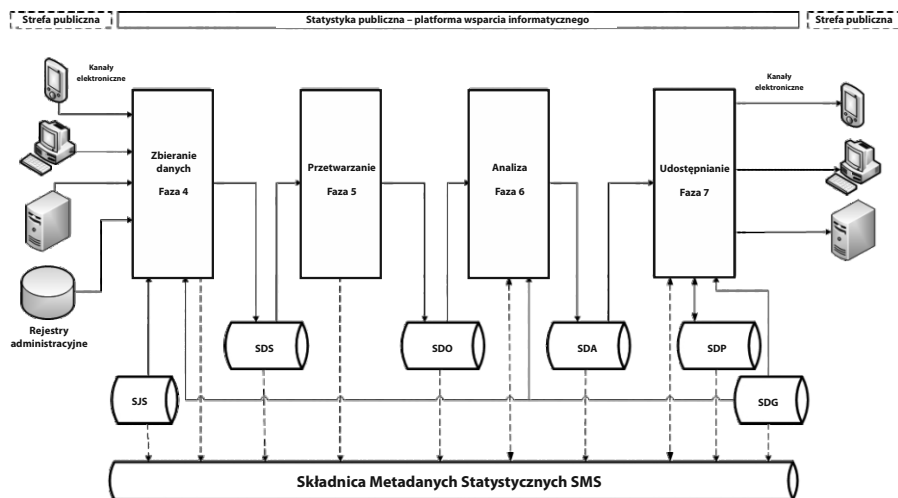
- opracowanie kompleksowej strategii wdrożenia koncepcji systemu satelitarnej identyfikacji i monitorowania upraw na potrzeby statystyki rolnictwa, zdefiniowanie szczegółów jego budowy oraz opracowanie rozwiązań systemowych, metodologicznych i funkcjonalnych;
- wypracowanie rozwiązań mających na celu 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).

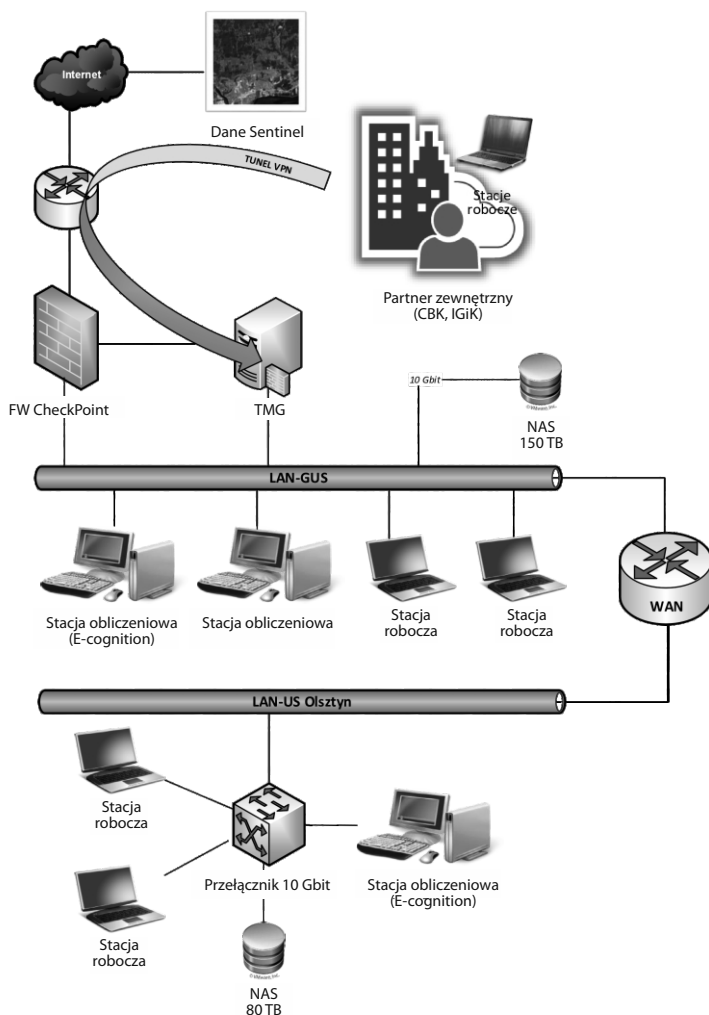
W systemie SATMIROL zostały stworzone moduły:

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

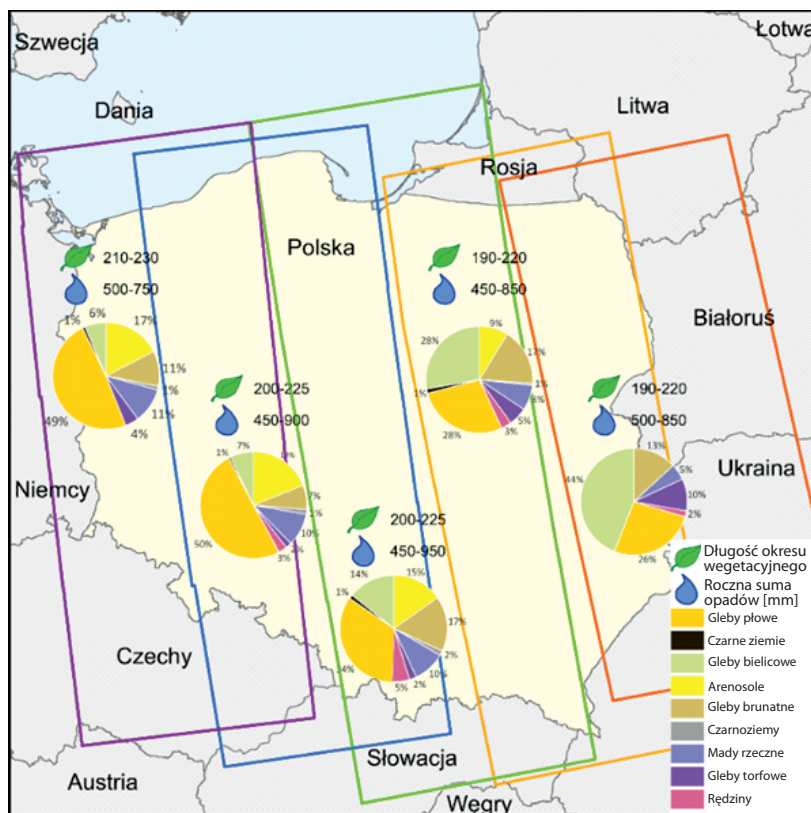
Wypracowywanie rozwiązań mających na celu rozpoznawanie upraw na podstawie danych satelitarnych odbywało się na terenach rolnych, które stanowią 60% powierzchni całego kraju. Większość terenów rolnych w Polsce – 68% – jest użytkowana jako grunty orne, 27% to łąki i pastwiska, a 5% stanowią uprawy trwałe. 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 jară, 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*

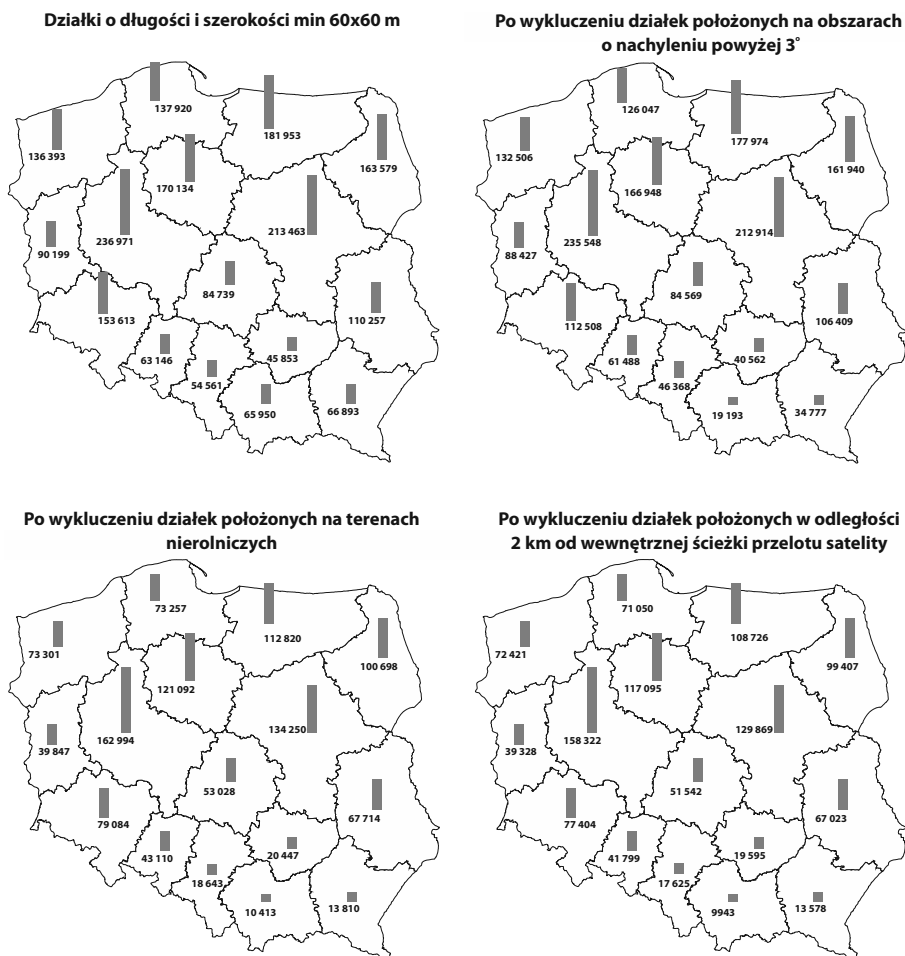
Badanie terenowe przeprowadzono we wszystkich województwach w okresie od czerwca do lipca 2019 r. i od czerwca do lipca 2020 r., a prace przygotowawcze trwały od stycznia do czerwca. Każdy z ankierów miał za zadanie dotrzeć do zaznaczonego pola, ocenić, czy uprawa jest jednorodna pod względem rozwoju 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>.

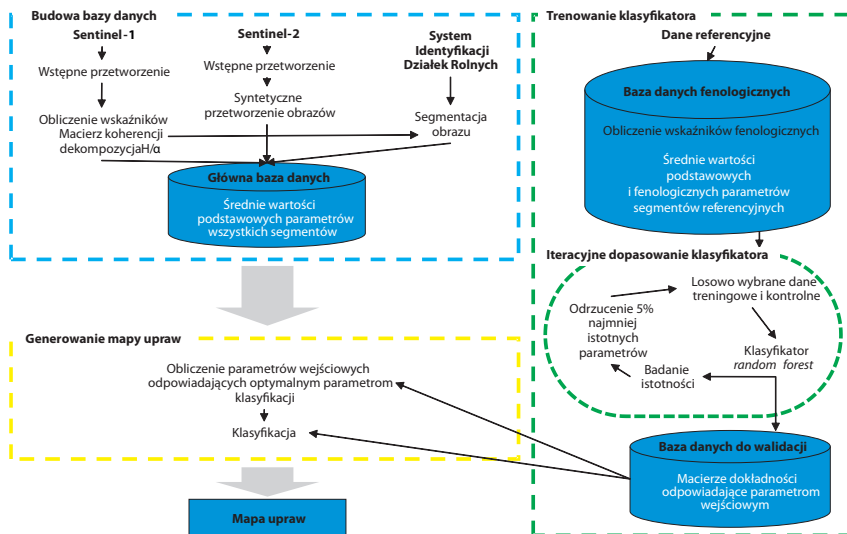
Mapa 2. Potencjalne działki reprezentatywne



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.

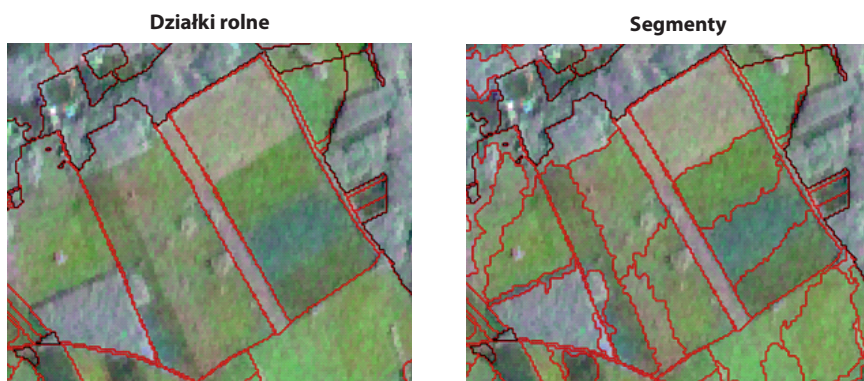
Obrazy satelitarne wymagają wstępnego przetworzenia, które w przypadku danych optycznych Sentinel-2 polega na zidentyfikowaniu obszarów zasłoniętych chmurami i ich usuwaniu. Braki są uzupełniane wartościami pikseli obrazów z sąsiadujących dat rejestracji. Z kolei wstępne przetworzenie obrazów radarowych 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

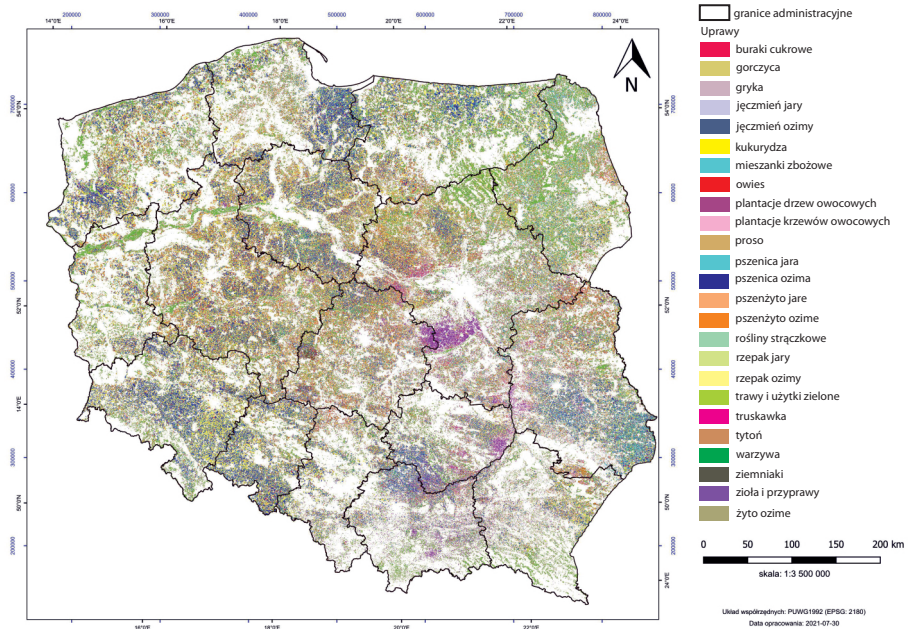
Trzeci etap omawianej metody rozpoczyna się od wybrania wskaźników wyznaczonych na poprzednim etapie jako najistotniejsze dla wszystkich segmentów w granicach gruntów rolnych i wykonania klasyfikacji, w wyniku której uzyskuje się obraz przedstawiający rozkład przestrzenny klasyfikowanych upraw.

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.

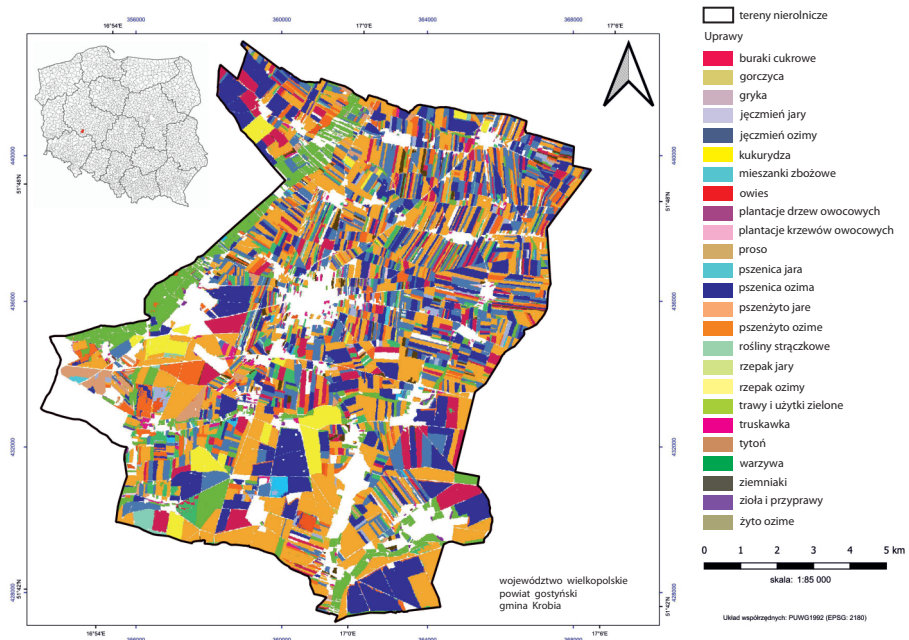
Na podstawie obrazu klasyfikacji oblicza się powierzchnie upraw. Są one sumowane na wszystkich poziomach administracyjnych kraju (województwa, powiaty i gminy). Numer identyfikacyjny poszczególnych jednostek pozwala na integrację tych danych z systemem raportowania GUS.

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

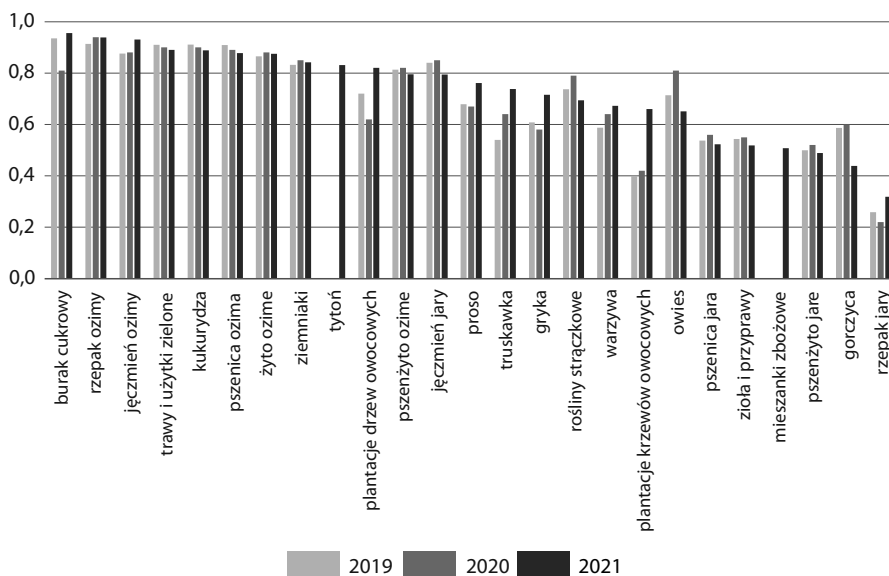
Źródło: opracowanie własne na podstawie danych ESA, ARiMR i GUS.

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 tablicy 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

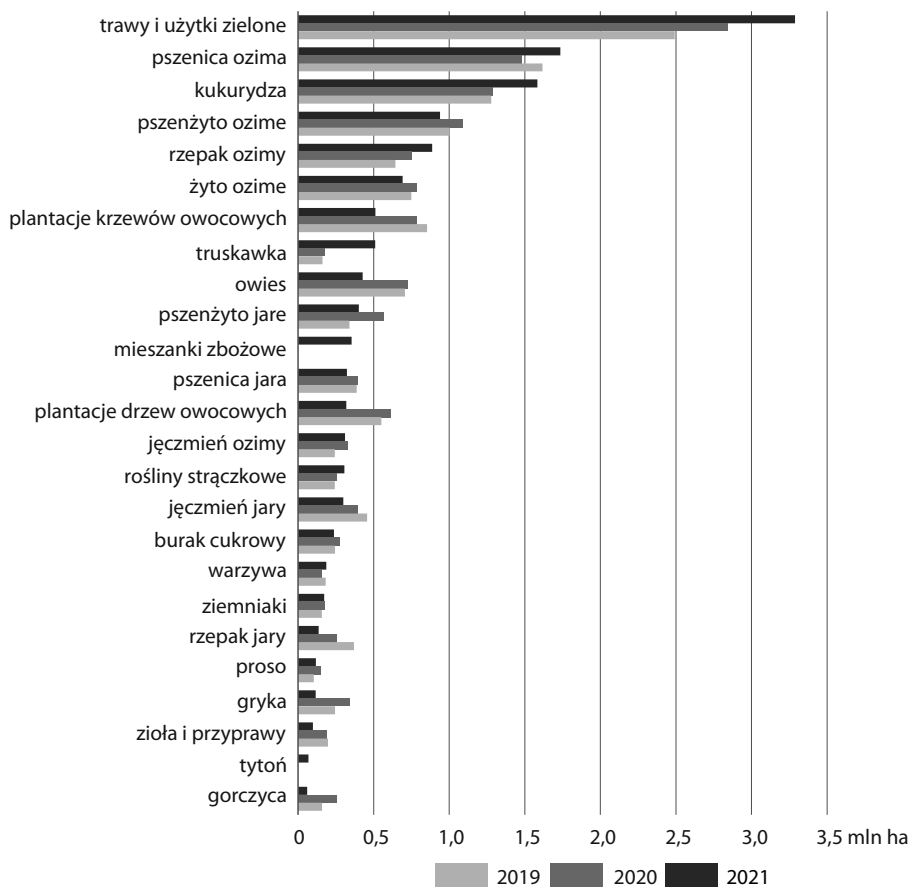
Źródło: opracowanie własne na podstawie wyników z systemu SATMIROL.

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>.