Statistics Poland

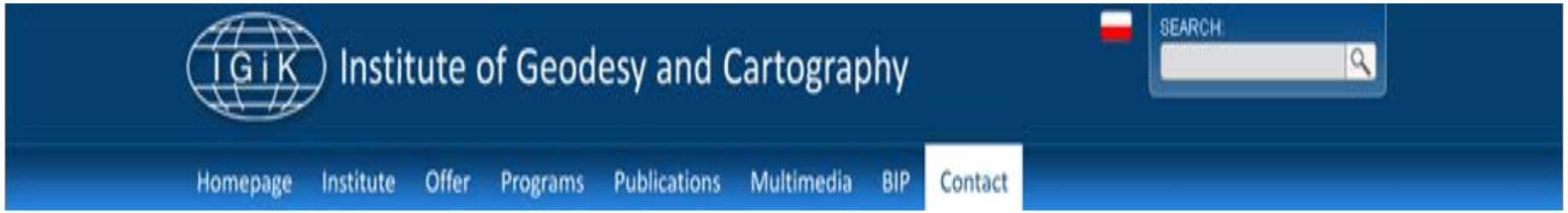
Case studies in the use of satellite data in agriculture statistics in Poland

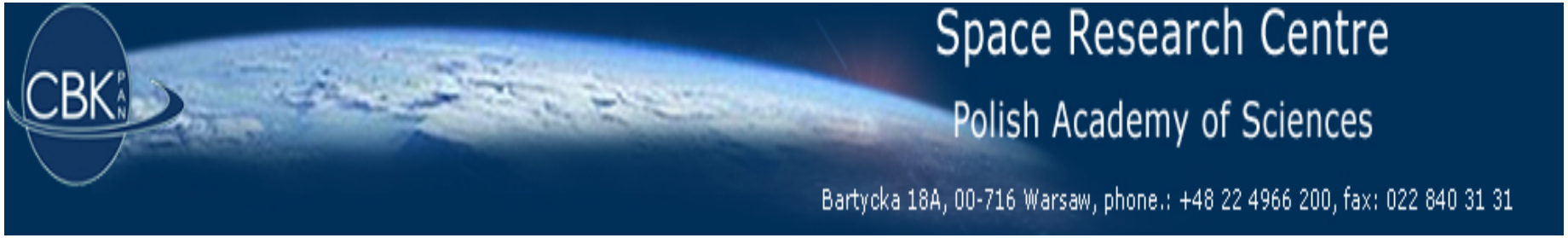
Tomasz Milewski

Statistics Poland

16 October 2019 Hangzhou

Cooperation:





Statistics supported by satellite data

- **Vegetation monitoring and yields by NOAA – for last 20 years**
- **Yield estimation by Sentinel– pilots since 2015**
- **Crop area recognition – pilots since 2015**
- **Monitoring of extremal conditions for plants i.e. ground frost, freezing, drought, floods – ad hoc projects**

I - NDVI

- **Normalized Difference Vegetation Index**

based on contrast between high reflectance values of emission for vegetation in near infrared band (channel 14) and low values in red emission band (maximum absorption) and is expressed with formula:

$$NDVI = (NIR - R) / (NIR + R)$$

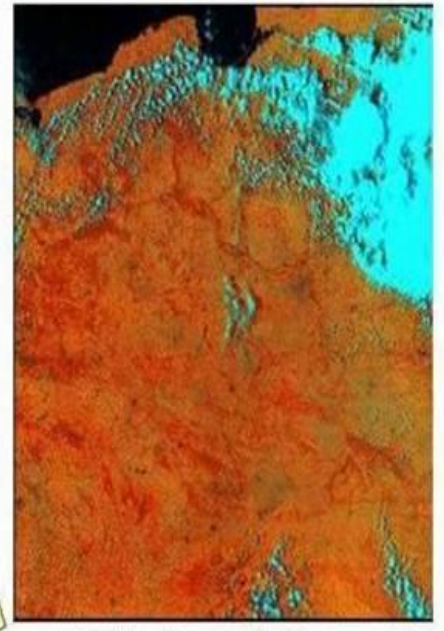
Together with values from range (-1, 1) we obtained information on the development state and condition of vegetation, and high values of index correspond to the areas overgrown with thick vegetation of good condition.

Work Station

Photo correction

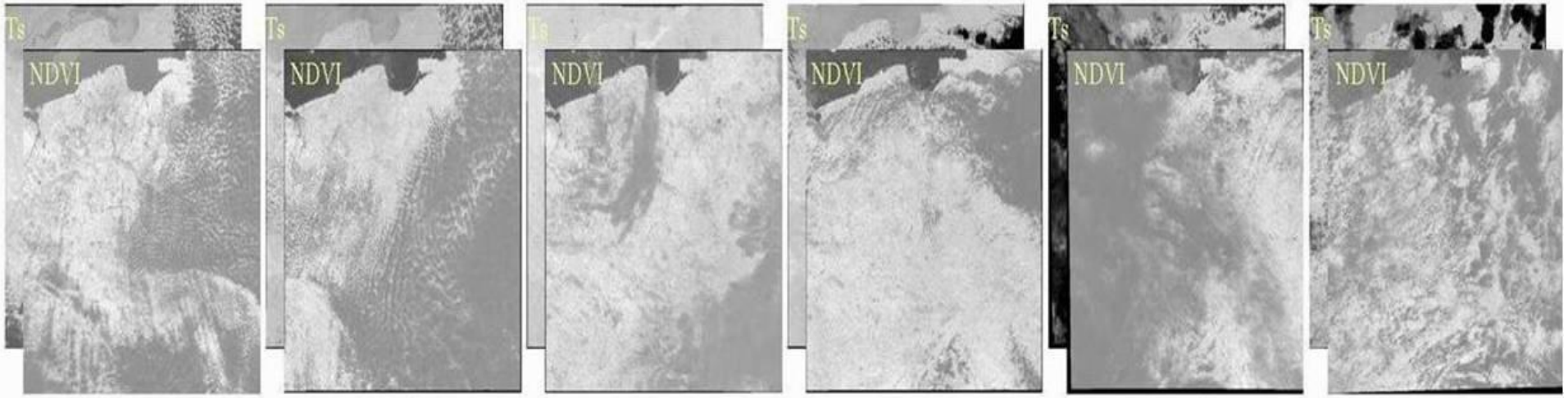
NOAA/AVHRR

Cut out country area, masking clouds

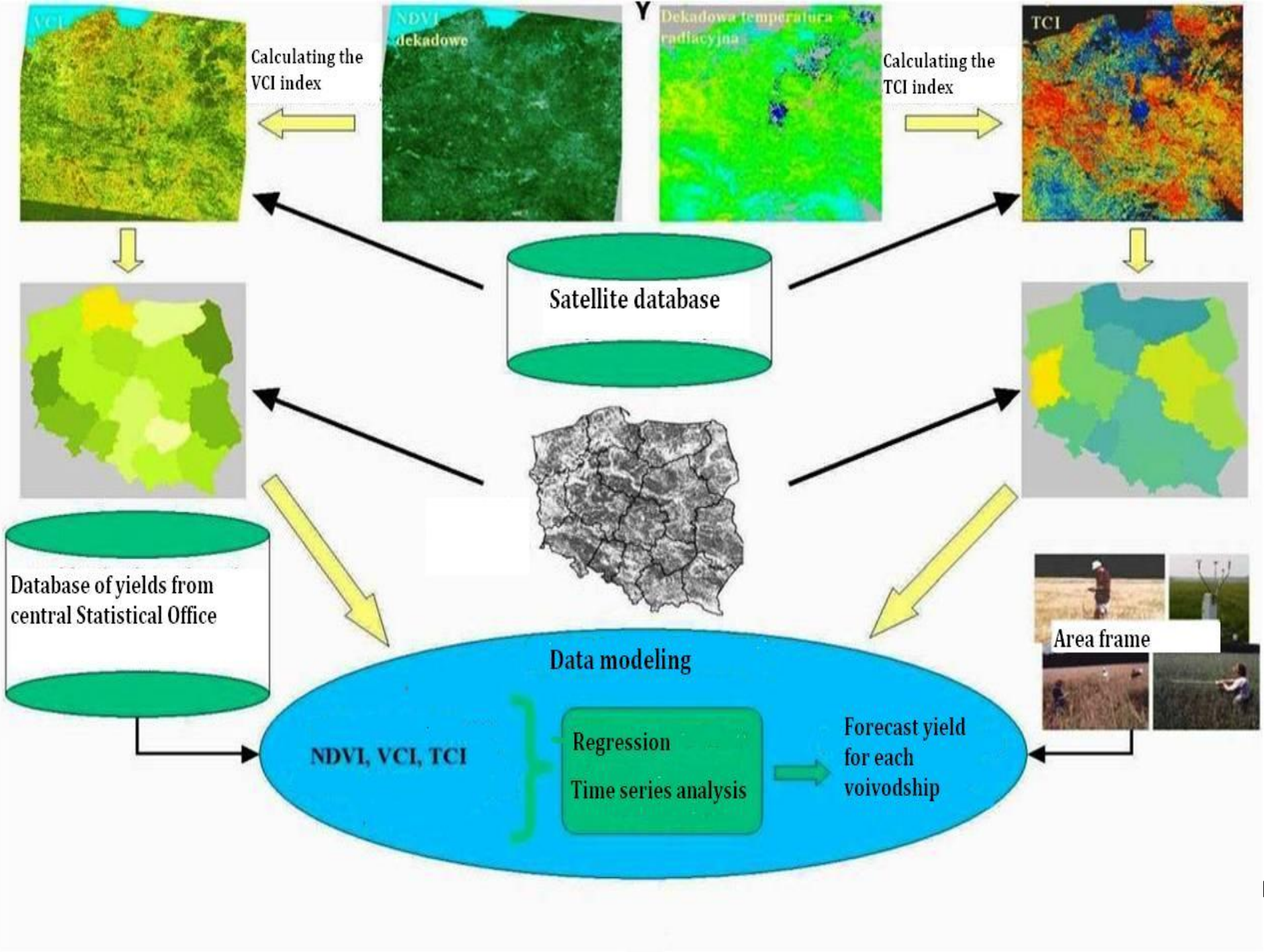


Poland

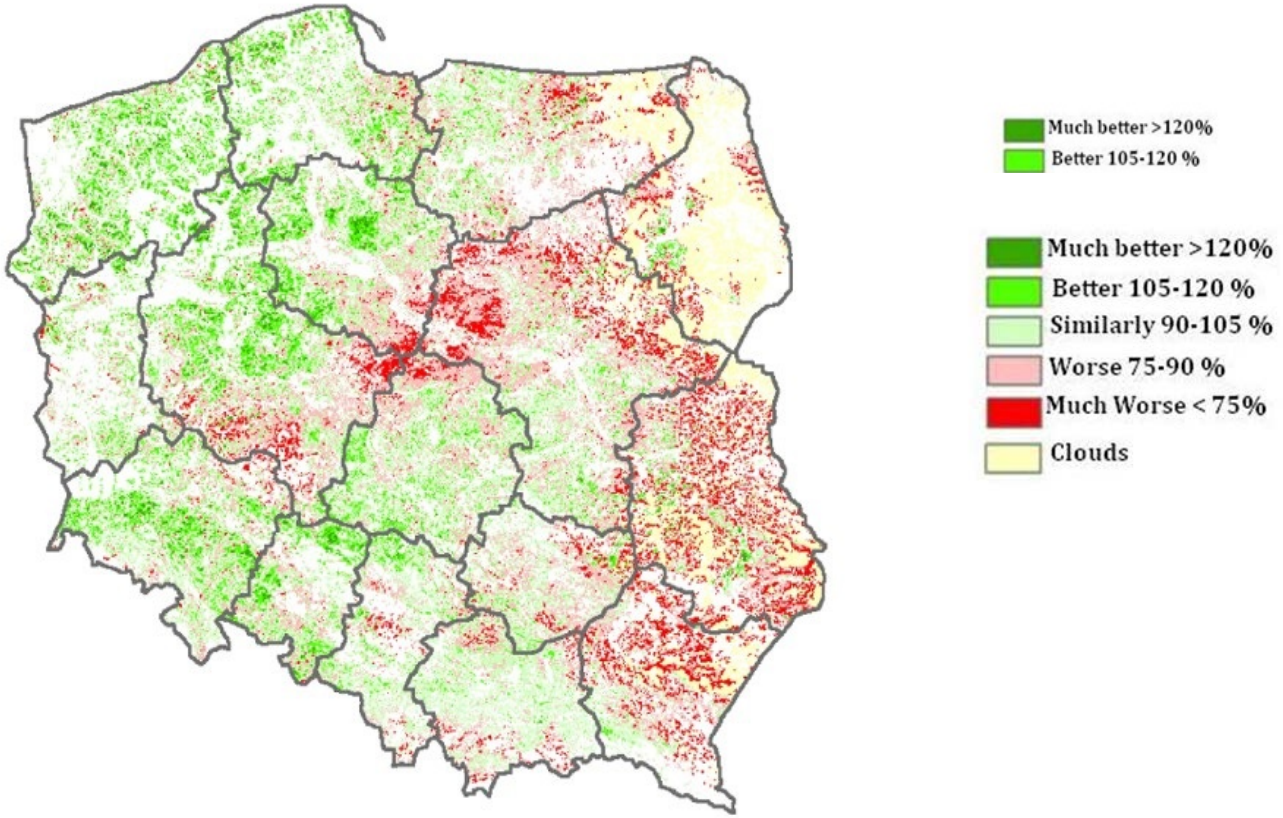
Calculation of radiation temperature and NDVI index



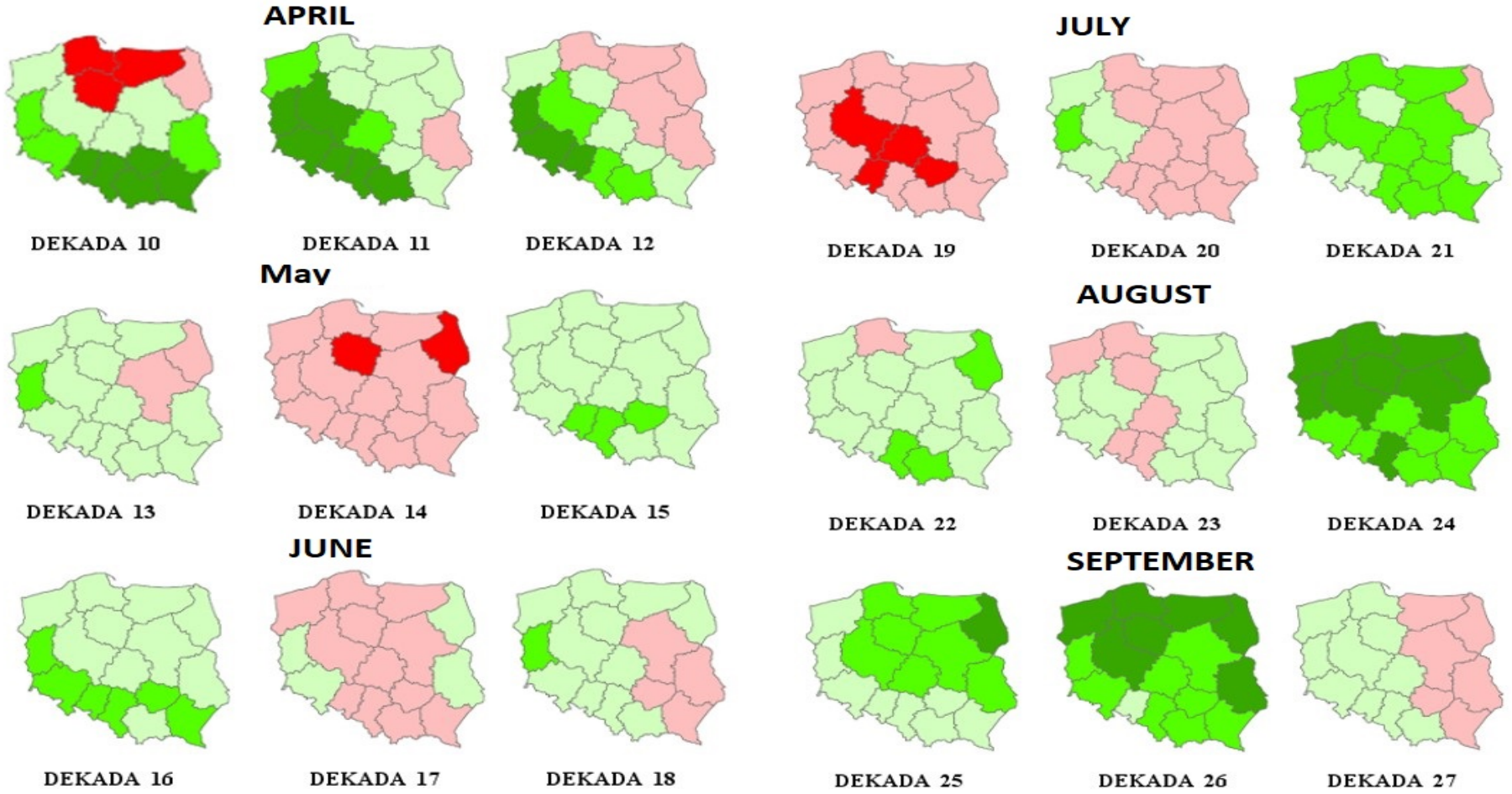
Create a decade


VCI
Calculating the VCI index
NDVI dekadowe
Dekadowa temperatura radiacyjna
Calculating the TCI index
TCI
Satellite database
Database of yields from central Statistical Office
Area frame
Data modeling
NDVI, VCI, TCI
Regression
Time series analysis
Forecast yield for each voivodship

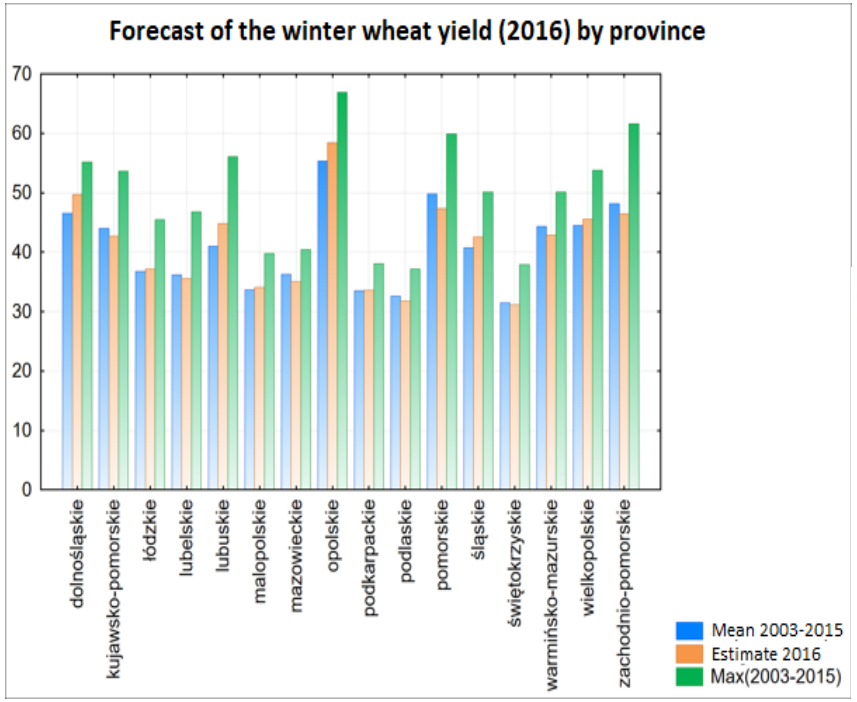
Comparison of NDVI in September 2016 to the average year

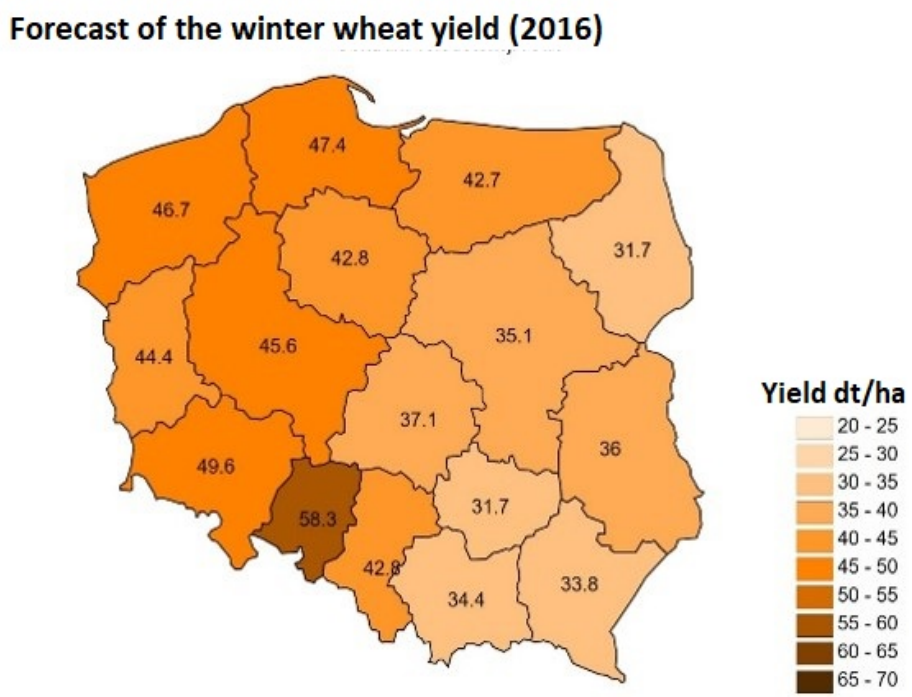


NDVI index in April – September 2016 for regions in comparison to the average year



Estimates of wheat yields for regions





Comparison of the yield estimates coming from satellites and current statistics

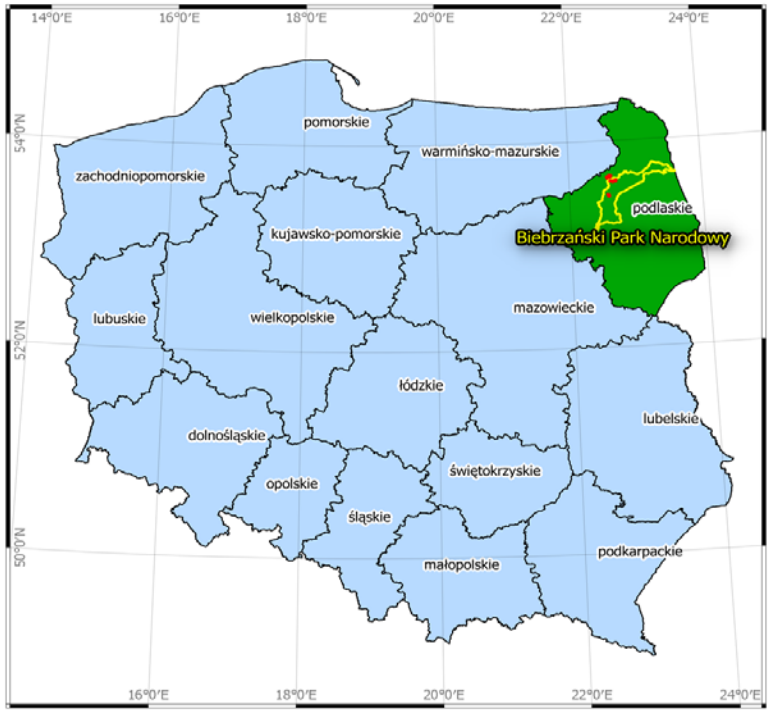
Voivodship (NTS 2)	Estimated yields (dt/ha)		Yield difference	
	Satellite	Statistical	dt/ha	%
DOLNOŚLĄSKIE	48,7	53,0	4,3	108,8
WARMIŃSKO-MAZURSKIE	43,3	50,9	7,6	117,6
WIELKOPOLSKIE	44,1	48,9	4,8	110,9
ZACHODNIOPOMORSKIE	46,1	59,2	13,1	128,4

II – Pilot Estimation of grassland production using by drone and hyperspectral camera

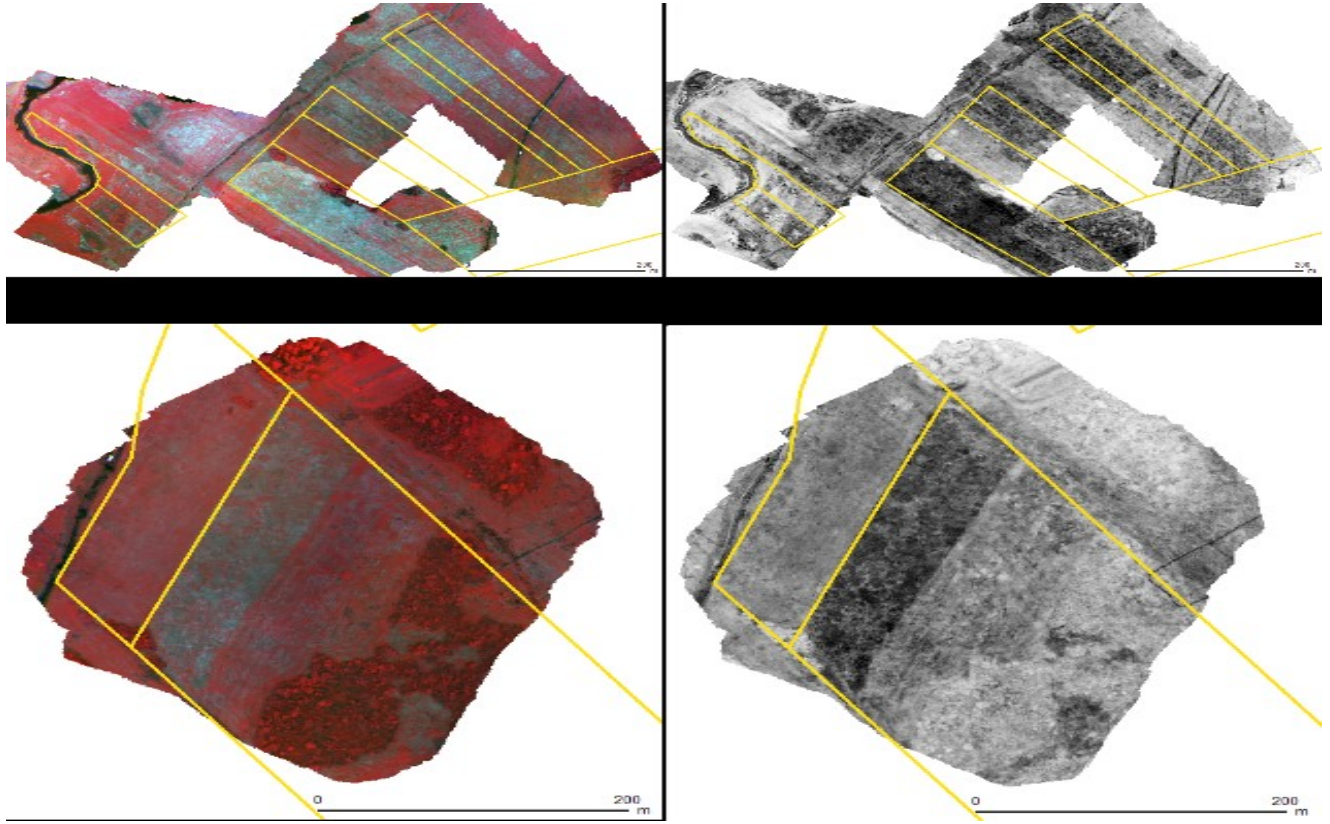
- Eurostat grant –*„Pilot studies to develop methodological improvements to agri-environmental statistics and statistics on grassland production”*
- CSO of Poland (the Employer) in Warsaw
- Institute of Geodesy and Cartography (IGiK) in Warsaw (the Contractor)
- The Institute of Technology of Sciences
- Statistical Office in Olsztyn
- Centre of New Technologies of the Warsaw University

Area Frame

- 21 plots indicated by the Employer
- 460.2 hectare
- Biebrza National Park
- 2 field campaigns
- 74 points (LAI)
- Landsat 8
- Drone



Two examples of data obtained from drone; on the left, the composition of channels 16-6-4 (bluish areas correspond to sections of meadows with larger share of dry vegetation) on the right established NDVIs.



(+)

- 397.65 ha; 4402 tons; 11.07 tons/ha
- Pictures at low hights
- 200m / 12 cm
- Mobility, flexibility and relatively small cost / single flight
- Good recognition of the imaged areas by drone operator

(-)

- ca.15 minutes/one flight
- high dependence on weather conditions
- eye contact
- camera (one lens vs many lens)

Simulation

- one voivodship - NUTS 2
- Total area = 20 180 km^2
- Agriculture area = 10 741 km^2
- daily area = 1 km^2
- working days = 2 685
- Daily costs = 210 euros
- Over **4 mln** euros (2014), **0.5** mln euros - 2018)

III - Pilot research to isolate particular groups of crops and land cover classes

- CSO - RSO Olsztyn (2015)
- CSO - RSO Olsztyn RSO & Lublin (2016)
- CSO - 16 RSO (2017)
- Space Research Centre - Polish Academy of Science

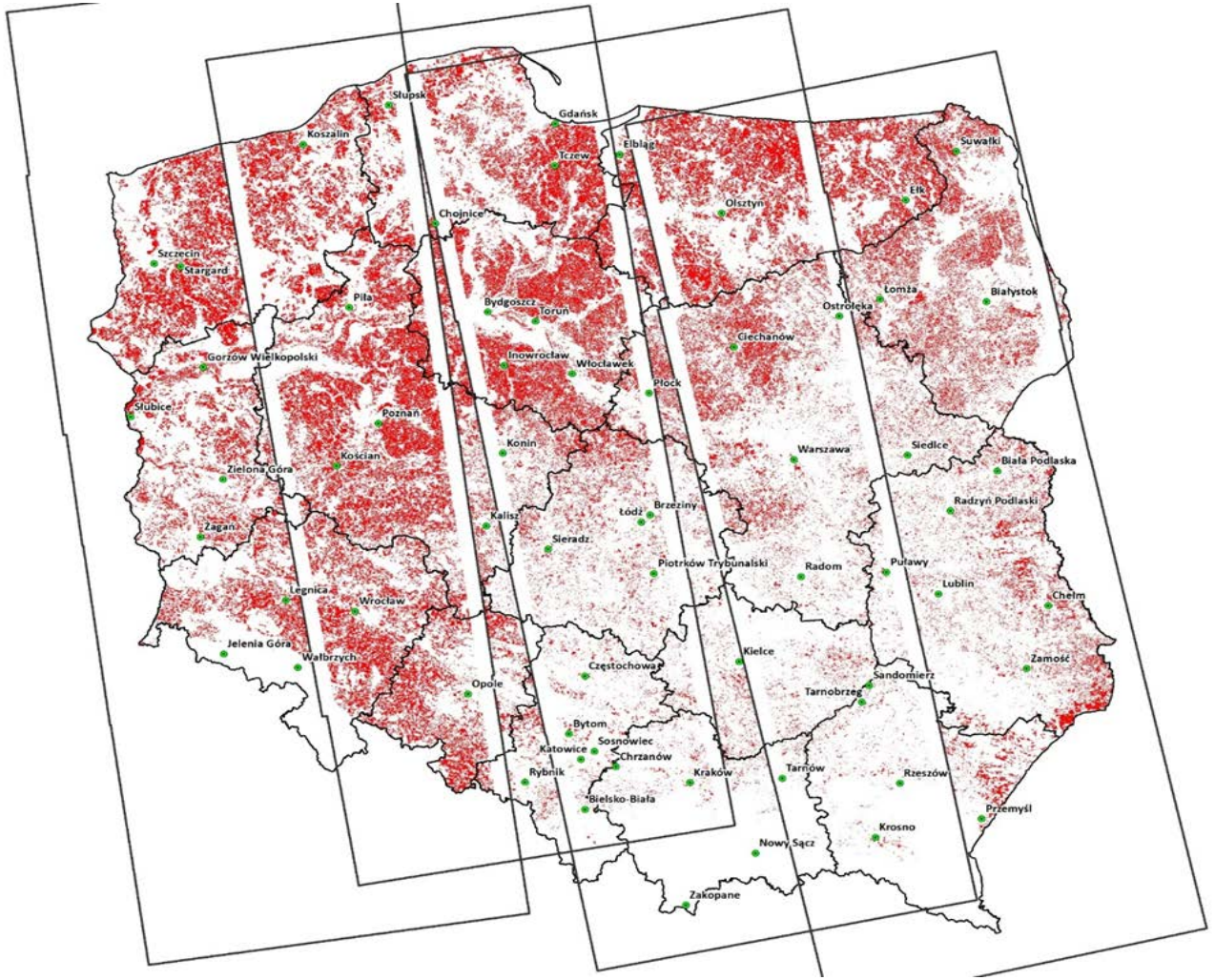
Methodology for selecting plots for in situ testing 2015-2017

- **Administrative sources**
  - cadastral vector data from the LPIS system (over 34 million records, 13GB),
  - vector layer of ARMA development fields (over 33 million records, 23GB),
  - general geographical database (GUGiK)
- **Other sources**
  - SRTM (Shuttle Radar Topography Mission) - areas with a slope of more than 3 °,
  - vector borders of images of SENTINEL-1

Method

- Access to EO data (SENTINEL 1 & 2 & 3, LANDSAT, MODIS, NOAA …)
- Segmentation – delimiting of uniform areas (admin data, EO data unsupervised classification)
- Training sample (*in situ*) – high quality data from fields (or admin data)
- Machine learning – random forest, support vector machine, neural networks (supervised classification)
- Validation (admin data, *in situ)*
- Generalization and presentation (maps, tabels …)

Sentinel 1 scenes and areas for *in situ* sampling



The network zones allocation to interviewers in relation to the transport system in the Warmińsko-Mazurskie Voivodship (Poland)

Example of point used for the test.

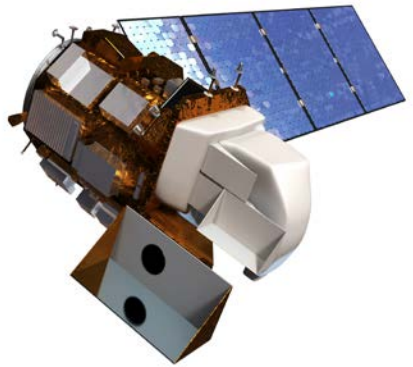
Visualization geodatabase file-based images.

Verification, standardization, Geographical reference = geodatabase.

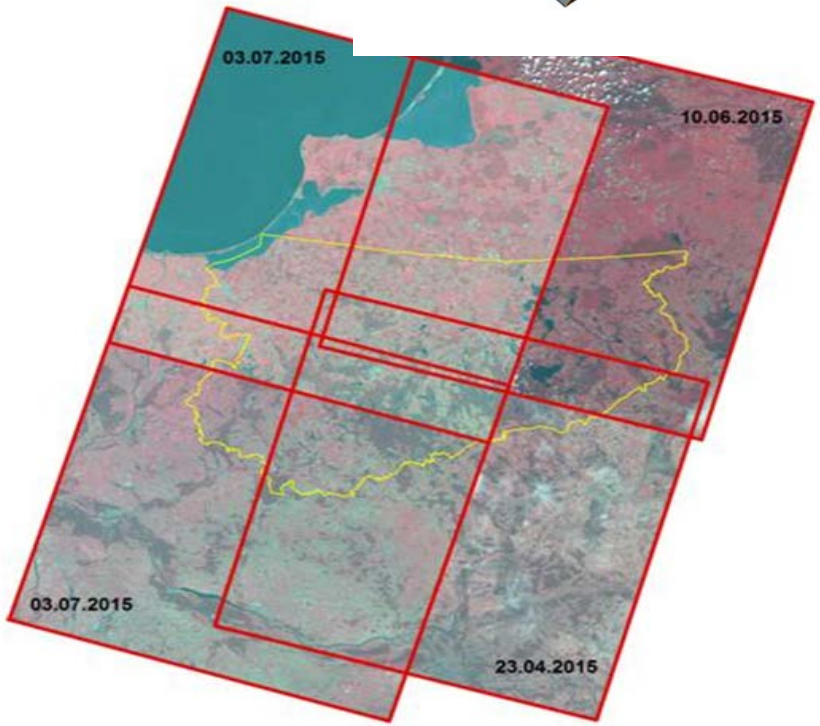
uprawa_jednorodna

OBJECTID *	Shape *	ID_punktu	Godzina	Data	Szerokosc	Dlugosc ge	Uprawa	Faza rozwoju roslin	Wysokosc uprawy w cm	Szerokosc miedzyrzedzi w cm
1	Punkt	1	09:08	2015-07-03	54,058137	19,266562	pszenica ozima	8 - dojrzewanie	96	12
2	Punkt	100	14:45	2015-06-25	53,919829	19,666468	pszenica ozima	7 - rozwoj owocow i nasion	70	14
3	Punkt	1000	08:25	2015-07-21	53,714661	20,42757	oleiste - inne	6 - kwitnienie	35	20
4	Punkt	1001	09:25	2015-07-21	53,291079	20,284482	gryka	6 - kwitnienie	64	12
5	Punkt	1002	11:05	2015-07-21	53,343572	19,711778	ziemniaki	4 - rozwoj wegetatywny	70	75
6	Punkt	1003	11:55	2015-07-21	53,431852	19,500038	mieszanki zbozowe jar	8 - dojrzewanie	56	12
7	Punkt	1004	12:05	2015-07-21	53,431811	19,490549	lubin	7 - rozwoj owocow i nasion	44	12
8	Punkt	1005	12:45	2015-07-21	53,490567	19,315142	buraki cukrowe	4 - rozwoj wegetatywny	60	45
9	Punkt	1006	14:15	2015-07-21	53,682421	19,300253	warzywa gruntowe	4 - rozwoj wegetatywny	28	50
10	Punkt	1007	14:25	2015-07-21	53,680559	19,299853	warzywa gruntowe	4 - rozwoj wegetatywny	20	150
11	Punkt	1008	14:45	2015-07-21	53,678235	19,292666	warzywa gruntowe	4 - rozwoj wegetatywny	60	80
12	Punkt	1009	14:50	2015-07-21	53,677609	19,29273	warzywa gruntowe	4 - rozwoj wegetatywny	52	45
13	Punkt	101	13:07	2015-07-02	54,025934	19,691007	rzepak ozimy	8 - dojrzewanie	164	25
14	Punkt	1010	15:00	2015-07-21	53,676681	19,292902	pszenica jara	7 - rozwoj owocow i nasion	73	14
15	Punkt	1011	15:10	2015-07-21	53,676703	19,294025	warzywa gruntowe	4 - rozwoj wegetatywny	35	55
16	Punkt	1012	15:20	2015-07-21	53,676519	19,296474	warzywa gruntowe	6 - kwitnienie	70	35
17	Punkt	1013	15:35	2015-07-21	53,6736	19,297324	warzywa gruntowe	4 - rozwoj wegetatywny	20	130
18	Punkt	1014	15:50	2015-07-21	53,681287	19,285646	warzywa gruntowe	4 - rozwoj wegetatywny	40	60
19	Punkt	1015	17:20	2015-07-21	53,694949	19,394514	jeczmien ozimy	9 - zamieranie	14	14
20	Punkt	1016	18:00	2015-07-21	53,797212	19,328756	motylkowe pastewne -	6 - kwitnienie	60	12
21	Punkt	1017	18:40	2015-07-21	53,855661	19,520475	motylkowe pastewne -	3 - wzrost pedu, powstawanie rozety	35	14
22	Punkt	1018	09:00	2015-07-22	53,818503	20,652017	plantacje drzew owoco	7 - rozwoj owocow i nasion	170	400

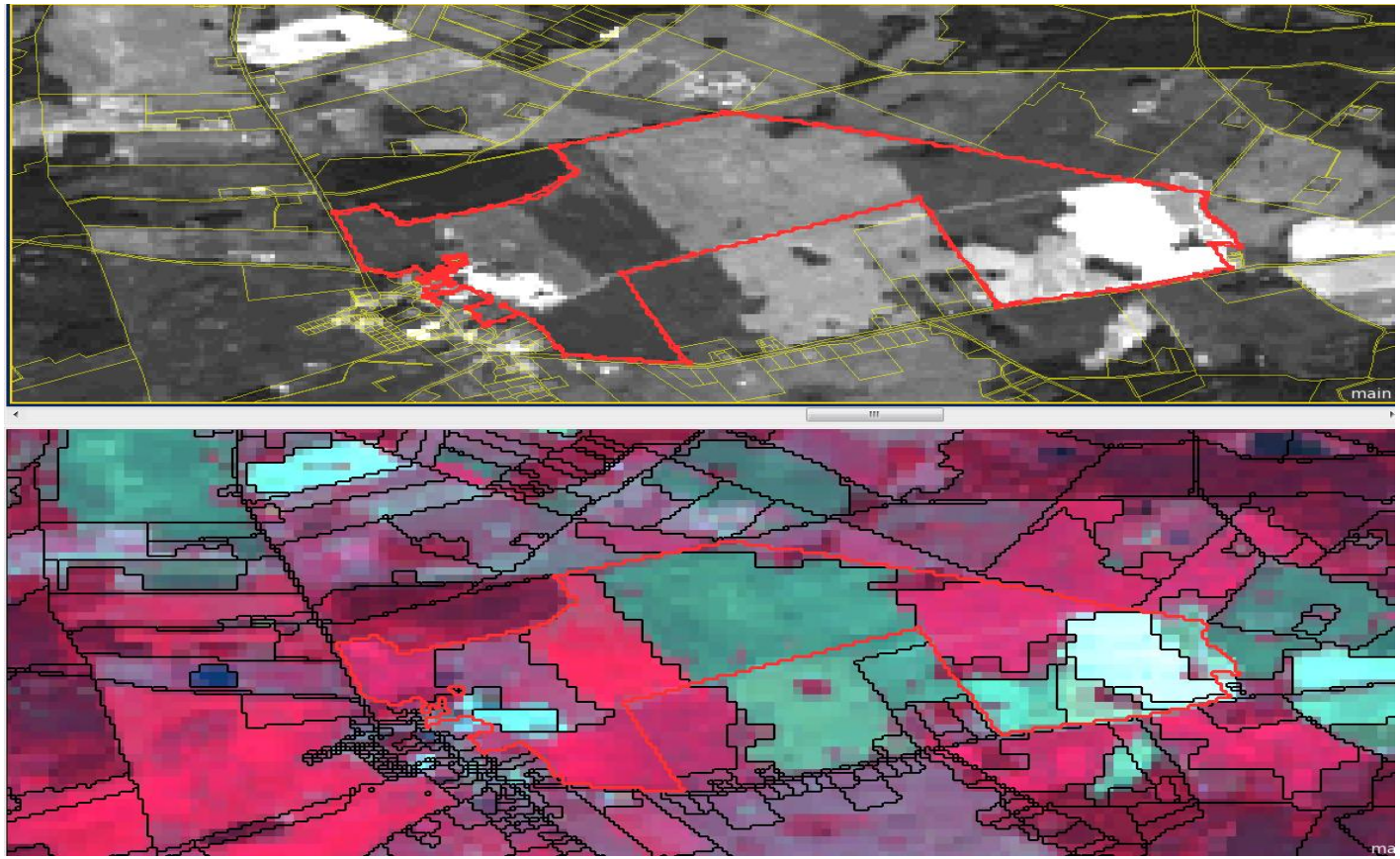
Data classification



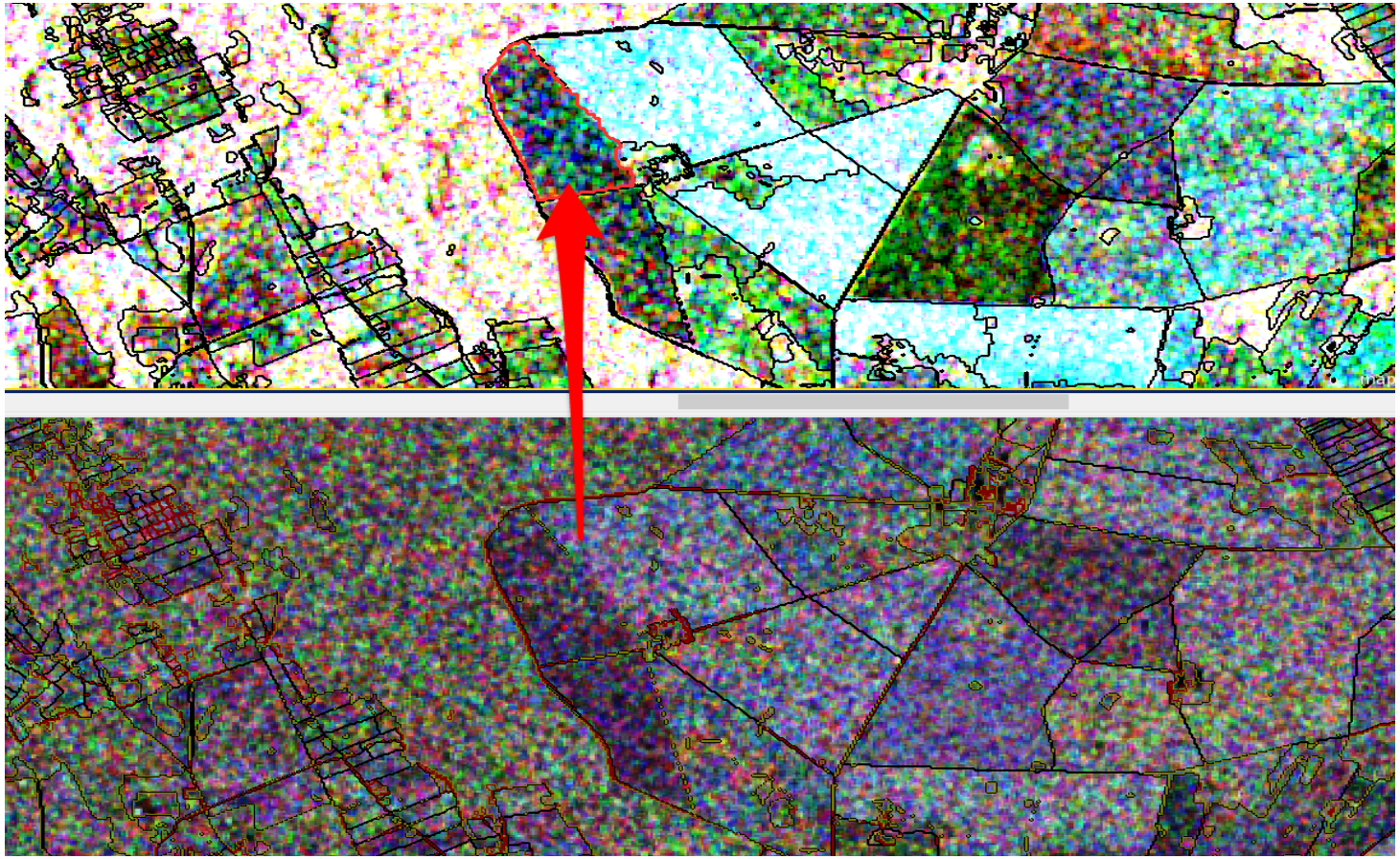
- Support Vector Machine (SVM),
- Decision Trees (DT),
- K-Nearest Neighbours (KNN)
  - Sigma,
  - Entropia,
  - Alfa,
  - Multi-temporal indicators,
  - Wishard distribution.



Segmentation

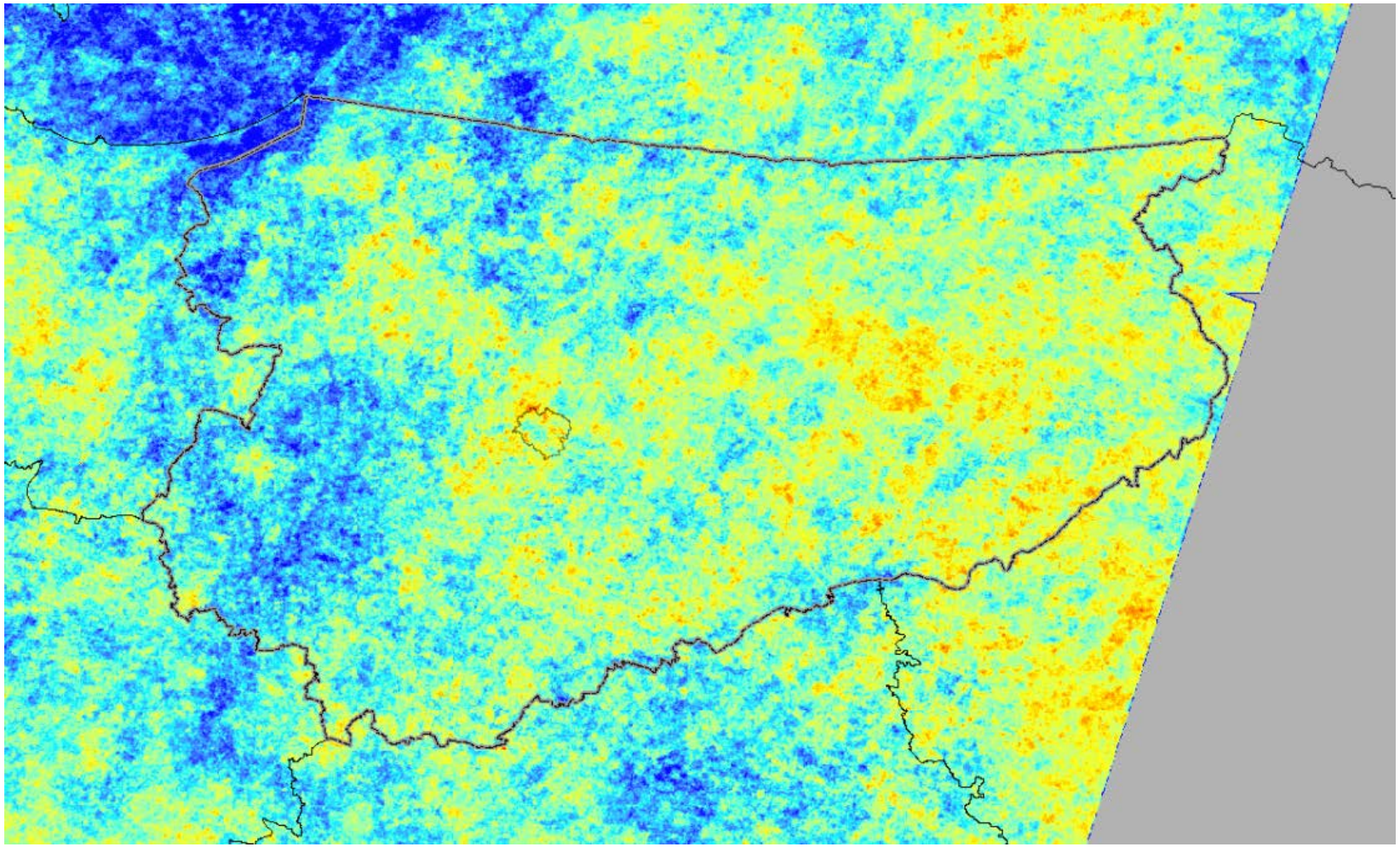


Segmentation



Results of the segmentation:
1 - correct
2 - incorrect

Claud ammont

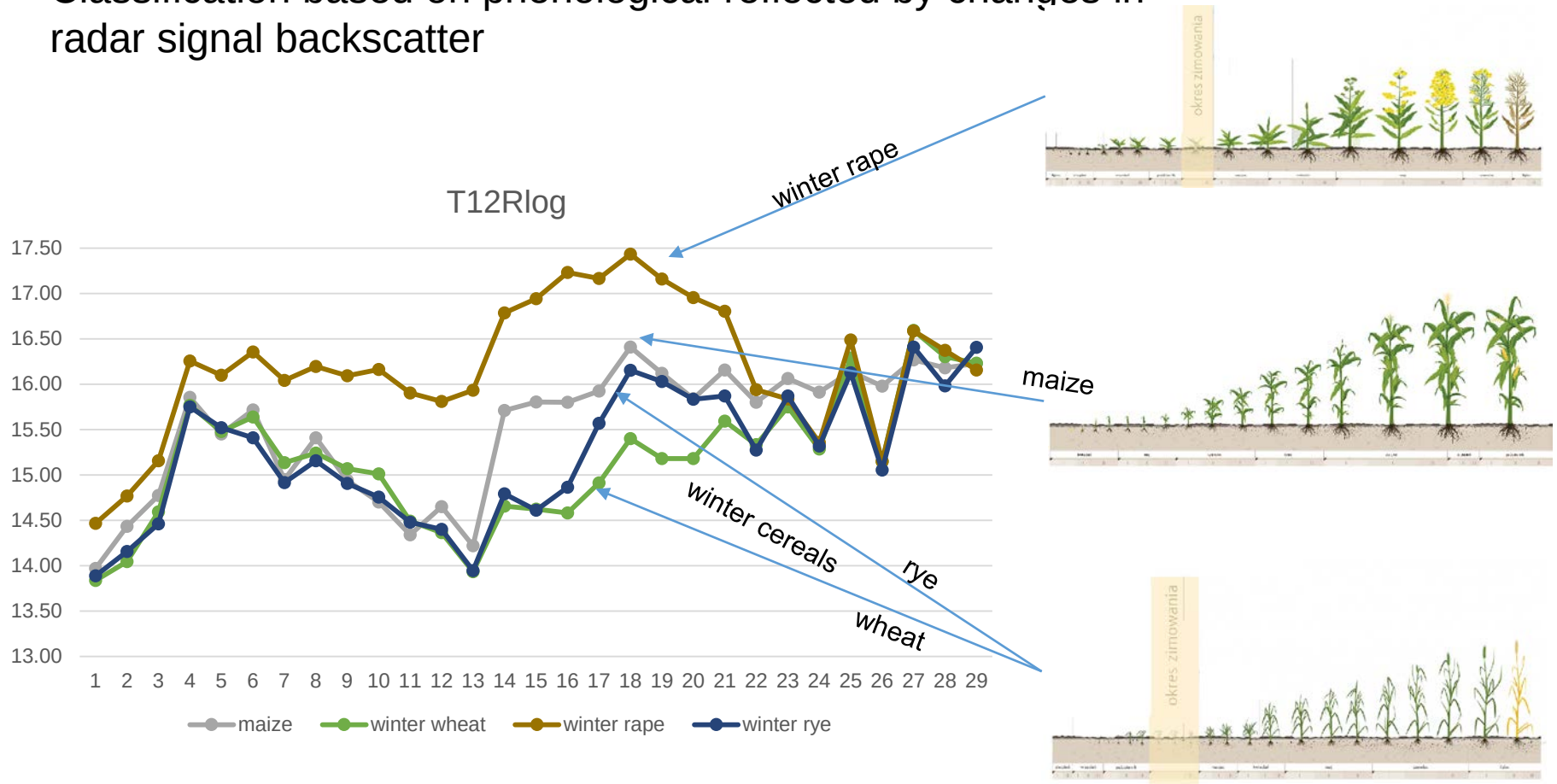


Ongoing projects:

- „EOStat – Agriculture Poland”, ESA
- „SATMIROL - Satellite identification and monitoring of crops for the needs of agricultural statistics” –GOSPOSTRATEG/NCBiR

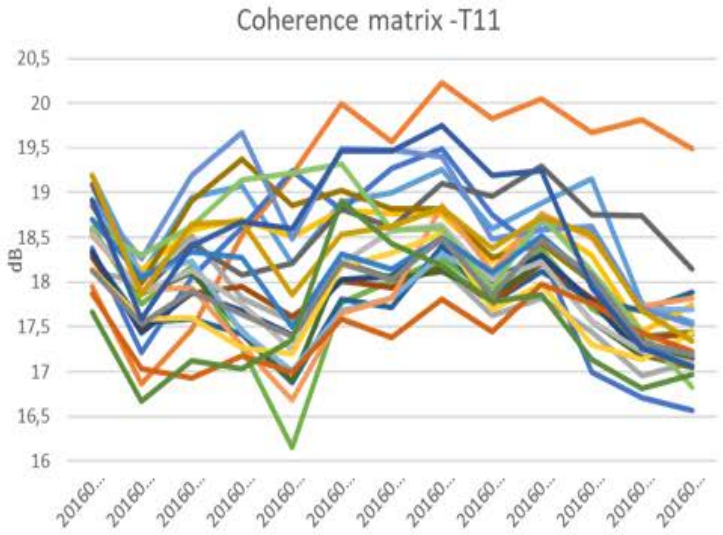
Crop mapping - CBK

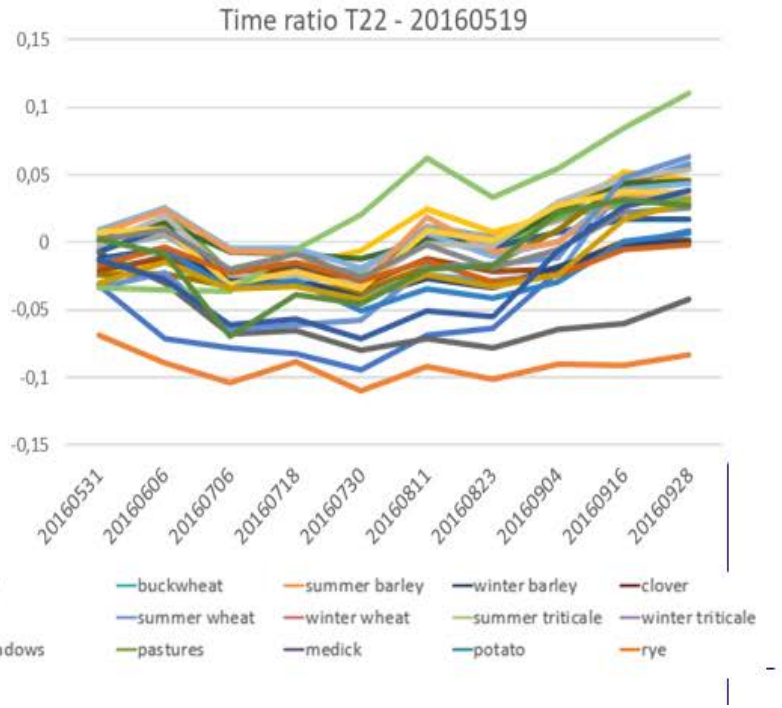
- Time series of radar data–Sentinel-1
- Classification based on phenological reflected by changes in radar signal backscatter



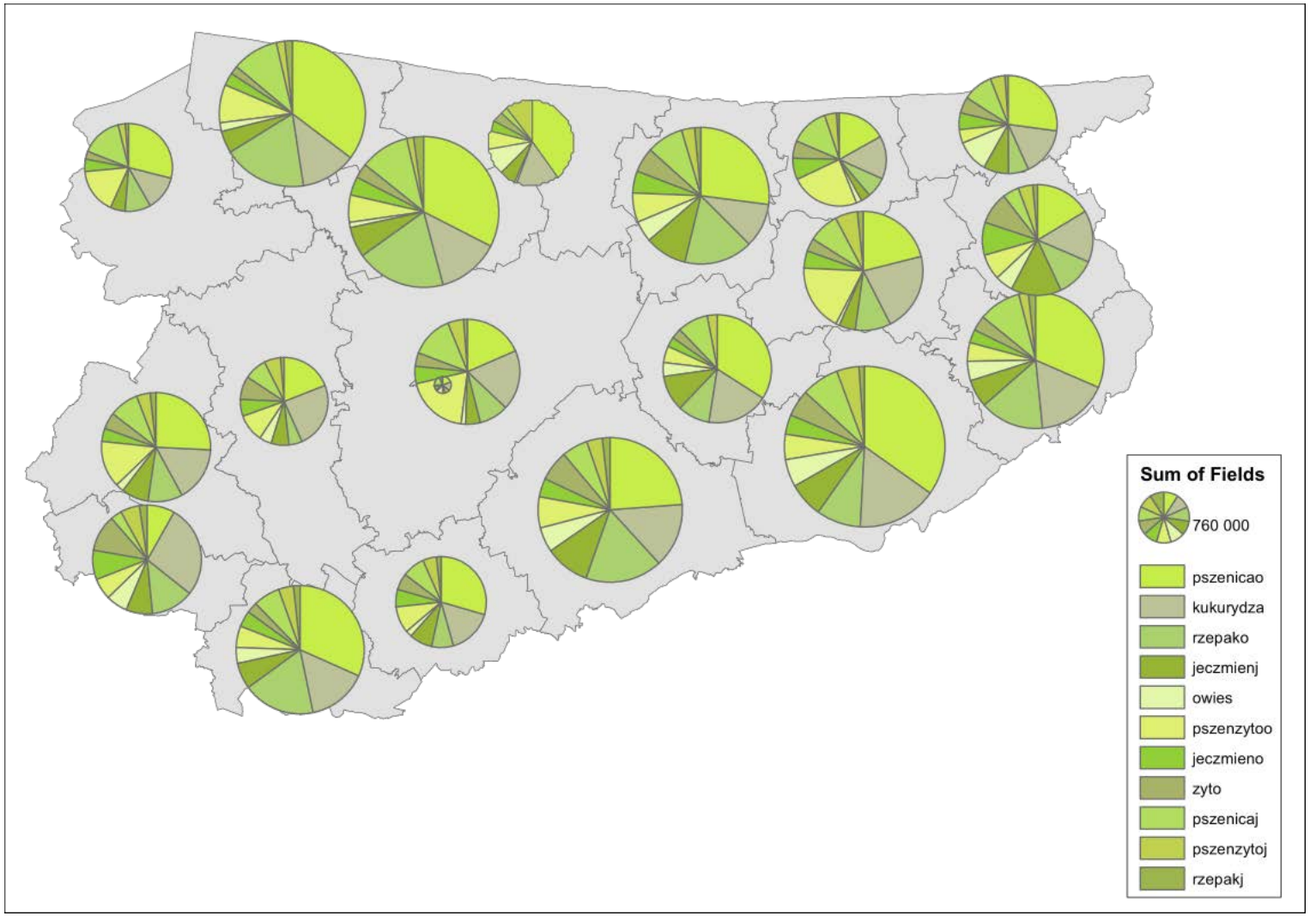
Enhancing of overall accuracy in classification - CBK

- Time ratios calculated for elements of coherence matices – 10%
- Analysis of increase/decrease functions describing crops' phenology – 13%
- Use the mixed datasets - 17% of overall accuracy

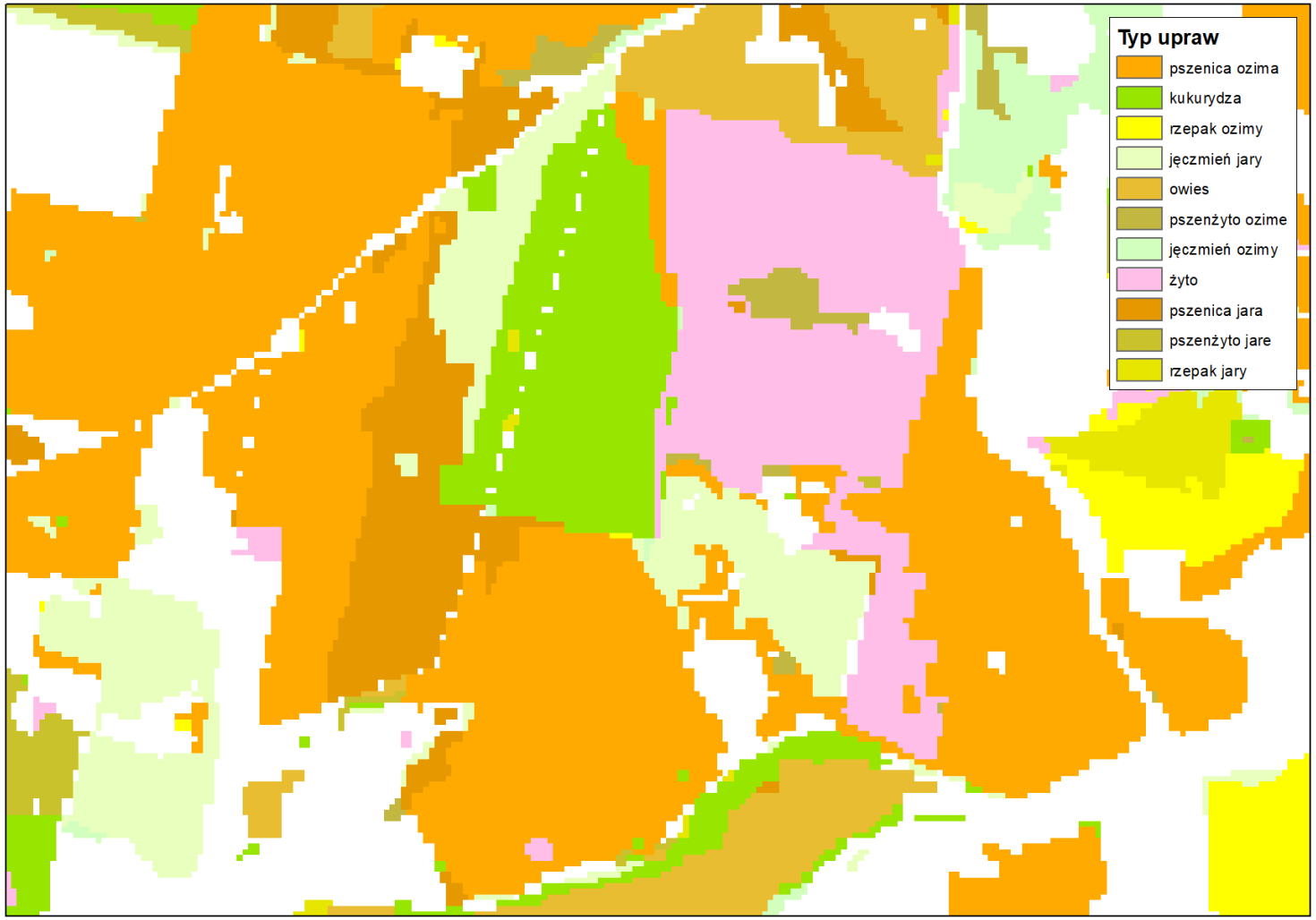




Crops in Warmińsko Mazurskie – CBK 2018

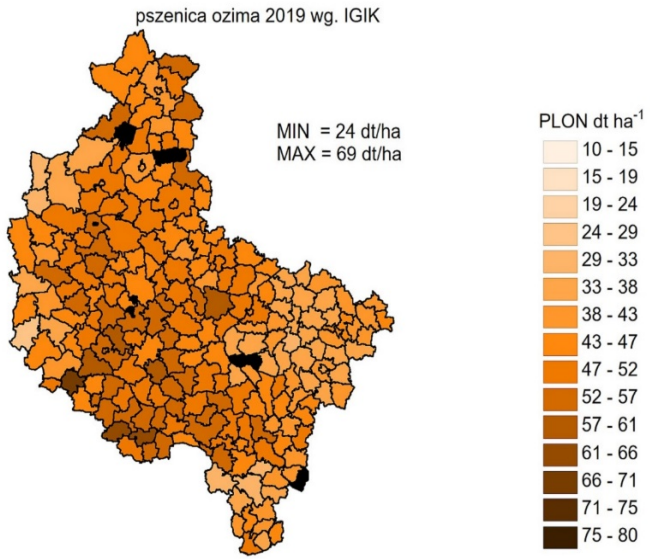


Crop mapping - CBK 2018



Crop estimation for wheat, maize and rape – IGIK 2019

voivodship	cereals dt ha^{-1}	wheat dt ha^{-1}	rape dt ha^{-1}	maize dt ha^{-1}
dolnośląskie	49.7	49.3	26.9	58.4
kujawsko-pomorskie	42.3	46.1	25.4	53.6
łódzkie	33.9	38.0	24.6	56.6
lubelskie	34.9	38.9	25.4	52.9
lubuskie	44.0	46.2	25.8	53.1
małopolskie	36.5	34.2	26.1	60.9
mazowieckie	30.1	35.4	21.2	52.3
opolskie	59.5	59.7	28.7	70.5
podkarpackie	33.4	33.5	20.8	69.9
podlaskie	27.3	32.6	22.3	49.4
pomorskie	41.7	54.0	27.1	32.8
śląskie	41.6	43.9	25.7	65.4
świętokrzyskie	30.5	31.7	23.6	49.6
warmińsko-mazurskie	37.4	45.3	22.3	42.6
wielkopolskie	43.7	46.3	24.4	56.1
zachodnio-pomorskie	47.7	51.6	26.4	46.7



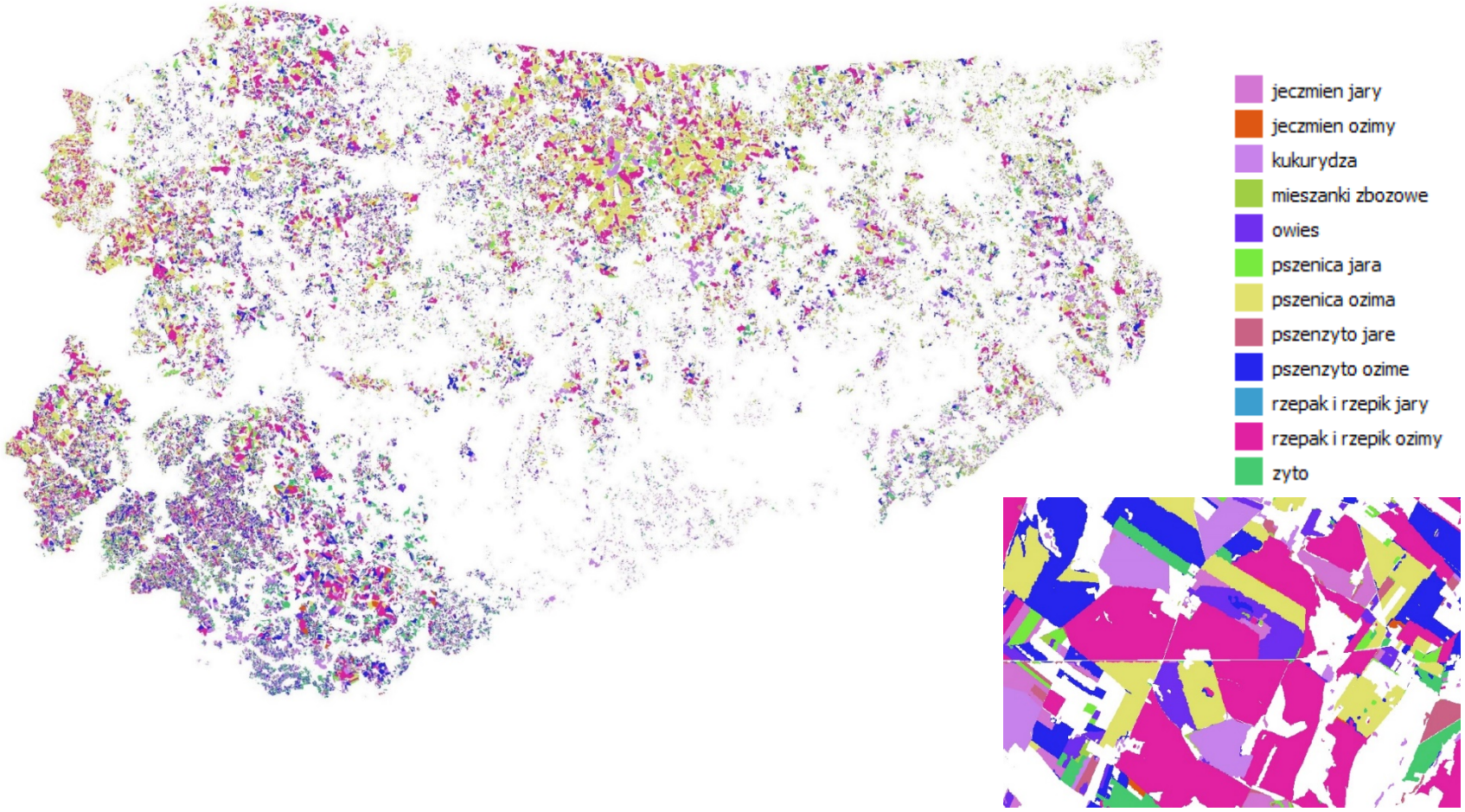
Results 2017 – RSO Olsztyn

	ARiMR_2017 [ha]	GUS_2017 [ha]	CLASSIFICATION AREA [ha]	Area % ARiMR	Area % GUS	Producer	User	KIA Per Class	OA	KIA	Producer (in situ)	User (in situ)	KIA Per Class (in situ)	OA	KIA
jeczmien jary	27562	39283	39367	143	100	0,7	0,75	0,67	0,76	0,73	0,92	0,69	0,91	0,7	0,67
jeczmien ozimy	2797	4187	2951	105	70	0,84	0,82	0,84			0,8	1	0,79		
kukurydza	51995	41758	73439	141	176	0,9	0,95	0,89			0,96	0,92	0,96		
mieszanki zbozowe	20377	39707	41537	204	105	0,4	0,48	0,35			0,17	0,5	0,12		
owies	25140	22470	41591	165	185	0,69	0,67	0,67			0,73	0,66	0,7		
pszenica jara	3368	28102	30524	906	109	0,53	0,62	0,49			0,69	0,64	0,66		
pszenica ozima	132034	116517	126174	96	108	0,84	0,83	0,82			0,77	0,71	0,75		
pszenzyto jare	6101	9660	9625	158	100	0,4	0,64	0,39			0,44	0,85	0,42		
pszenzyto ozime	58699	61789	73323	125	119	0,8	0,79	0,77			0,65	0,43	0,61		
rzepak i rzepik jary	3298	2591	1409	43	54	0,21	0,93	0,21			0,86	1	0,85		
rzepak i rzepik ozimy	92849	65894	104838	113	159	0,97	0,95	0,97			1	0,96	1		
zyto	29185	35783	38030	130	106	0,88	0,87	0,87			0,96	0,86	0,96		

	ARiMR_2017 [ha]	GUS_2017 [ha]	CLASSIFICATION AREA [ha]	Area % ARiMR	Area % GUS	Producer	User	KIA Per Class	OA	KIA	Producer (in situ)	User (in situ)	KIA Per Class (in situ)	OA	KIA
kukurydza	51995	41758	67381	130	161	0,91	0,95	0,89	0,91	0,87	0,96	0,96	0,96	0,9	0,9
zboza jare	82549	139223	177993	216	128	0,88	0,94	0,81			0,91	0,95	0,87		
zboza ozime	222716	218276	254080	114	116	0,93	0,92	0,89			0,93	0,96	0,88		
rzepak i rzepik jary	3298	2591	1627	49	63	0,18	0,86	0,18			1	1	1		
rzepak i rzepik ozimy	92849	65894	102305	110	155	0,97	0,95	0,96			1	1	1		

	ARiMR_2017 [ha]	GUS_2017 [ha]	CLASSIFICATION AREA [ha]	Area % ARiMR	Area % GUS	Producer	User	KIA Per Class	OA	KIA	Producer (in situ)	User (in situ)	KIA Per Class (in situ)	OA	KIA
jare	291798	262127	297645	102	114	0,88	0,92	0,82	0,92	0,88	0,86	0,95	0,74	0,9	0,8
ozime	314569	312839	353853	112	113	0,94	0,92	0,91			0,93	0,93	0,87		
TUZ i W	543437	387375	543504	100	140	0,94	0,93	0,91							

The use of satellite observations in agriculture - an example based on Sentinel-1 radar data



result of object Classification (map) – ANN/RF

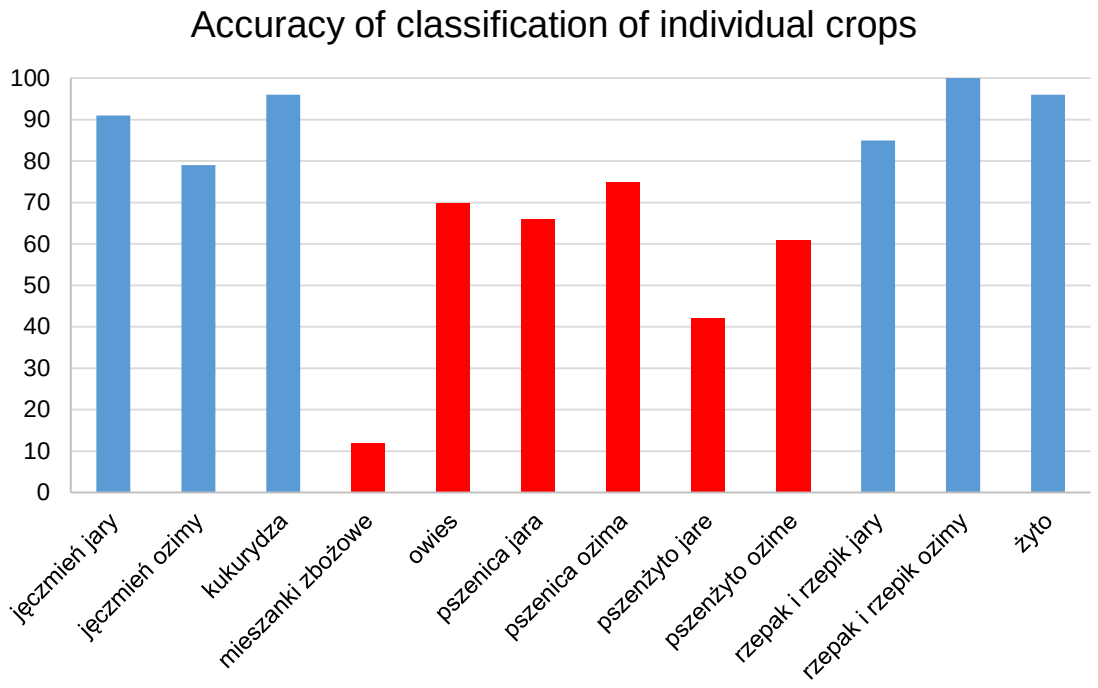
Results 2018 – RSO Olsztyn

id	uprawa_id	SG_RCzBR_2018	ARiMR_2018 [ha]	area [ha] ANN	Area % ARiMR	Producer	User	F-score	OA	KIA	Producer (in situ)	User (in situ)	F-score (in situ)	OA	KIA
1	buraki cukrowe	4633,06	3968	3793	96	0,92	0,98	0,95			1,00	0,99	0,99		
2	gryka	5195,61	8875	12007	135	0,77	0,81	0,79			0,74	1,00	0,85		
3	jęczmień jary	45407,98	41279	56360	137	0,78	0,70	0,74			0,98	0,85	0,91		
4	jęczmień ozimy	3808,92	4406	4556	103	0,88	0,94	0,91			0,65	1,00	0,79		
5	kukurydza	48973,49	62309	80485	129	0,96	0,90	0,93			1,00	0,86	0,93		
6	mieszanki zbożowe	40703,18	27382	21756	79	0,19	0,36	0,25			0,10	0,43	0,17		
7	owies	24834,77	31706	31326	99	0,63	0,65	0,64			0,70	0,68	0,69		
8	plantacje drzew owocowych	1109,96	1569	17684	1127	0,37	0,65	0,47			0,42	0,97	0,59		
9	plantacje krzewów owocowych	2185,58	1798	2752	153	0,23	0,40	0,29			0,26	1,00	0,42		
10	pszenica jara	63685,36	57526	58854	102	0,76	0,75	0,75			0,78	0,42	0,55		
11	pszenica ozima	93218,37	101500	102952	101	0,89	0,82	0,86	0,89	0,85	0,78	0,80	0,79	0,83	0,81
12	pszenżyto jare	7831,49	5368	6164	115	0,40	0,70	0,51			0,36	0,86	0,51		
13	pszenżyto ozime	55041,41	58498	68979	118	0,77	0,79	0,78			0,86	0,69	0,77		
14	rzepak i rzepik jary	3878,55	6355	6084	96	0,75	0,95	0,84			1,00	1,00	1,00		
15	rzepak i rzepik ozimy	42596,4	66416	73255	110	0,98	0,96	0,97			1,00	1,00	1,00		
16	TiUZ_W	392019,97	561947	576095	103	0,97	0,97	0,97			0,97	0,72	0,83		
17	ziemiaki	8792,9	4990	4405	88	0,57	0,65	0,61			0,91	1,00	0,95		
18	żyto	34100,63	28827	40032	139	0,87	0,80	0,83			0,94	0,78	0,85		
111	gorczyca	2284,42	2274	2340	103	0,82	0,93	0,87							
151	straczkowe	26761,15	47441	44775	94	0,73	0,90	0,81							
Suma		907063,2	1124434	1214651	108										

Comparison of crop areas for warmińsko-mazurskie voivodship from different methods

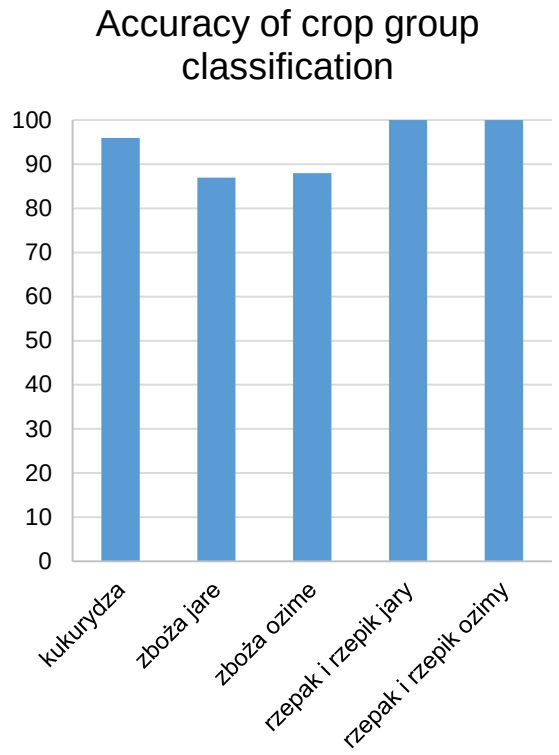
Crop plant	ARiMR_2018 [ha] farms > 10 ha, declarations	IGiK_2018 [ha] all parcels, Sen2Agri	US Olsztyn_2018 [ha] all parcels, ANN
wheat	159026	160795	161806
rape	72771	74406	79339
triticale	63866	85398	75143
maize	62309	62197	80485
leguminous	47441	61116	44775
barley	45685	58792	60916
oat	31706	47718	31326
rye	28827	37160	40032
buckwheat	8875	12674	12007
potatoes	4990	3443	4405
sugar beets	3968	4438	3793
mustard	2274	3158	2340
fruit plantations	1798	4908	2752
grasslands	561947	450998	576095

The use of satellite observation in agriculture - an example



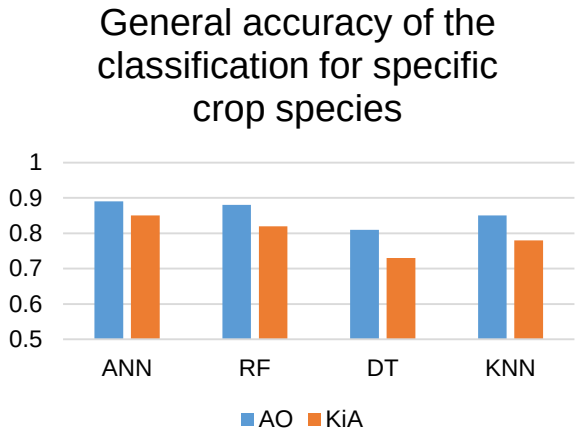
Average accuracy = 83 %
Kappa index = 81 %

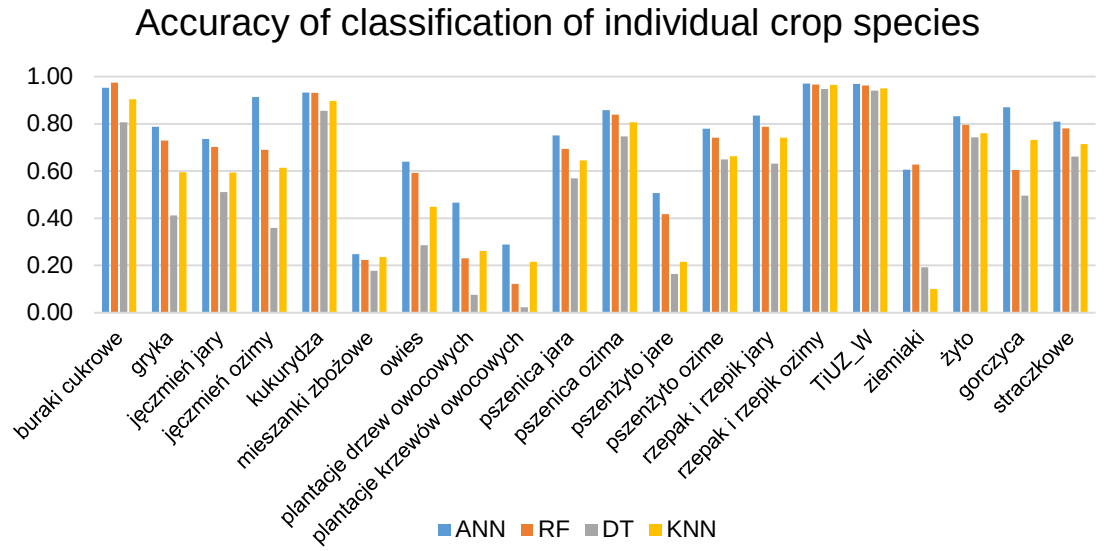
Average accuracy = 93 %
Kappa index = 90 %

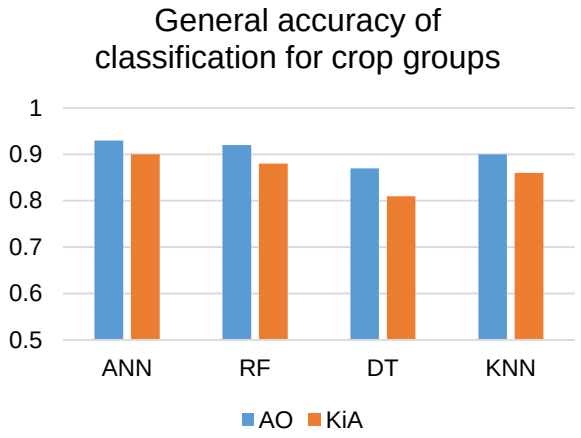


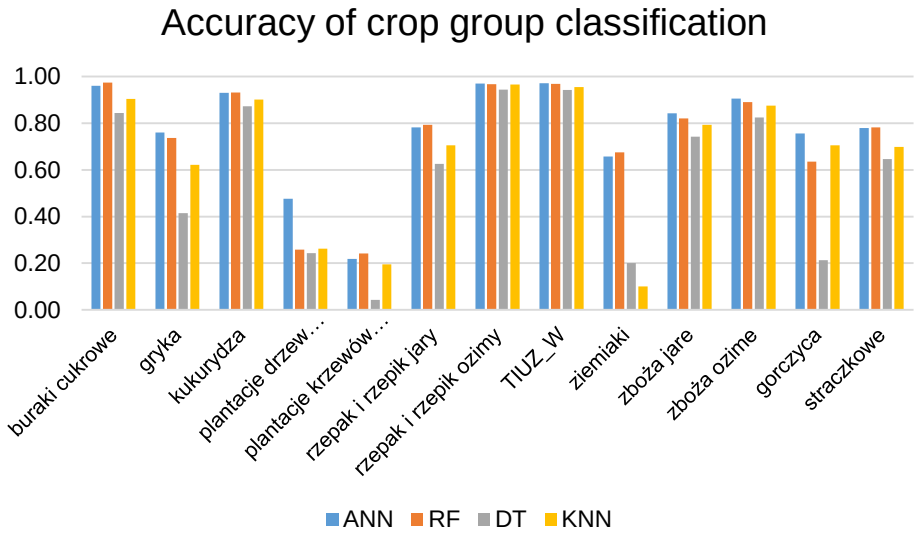
Statistical Office
in Olsztyn

The use of satellite observations in agriculture - an example based on Sentinel-1 radar data



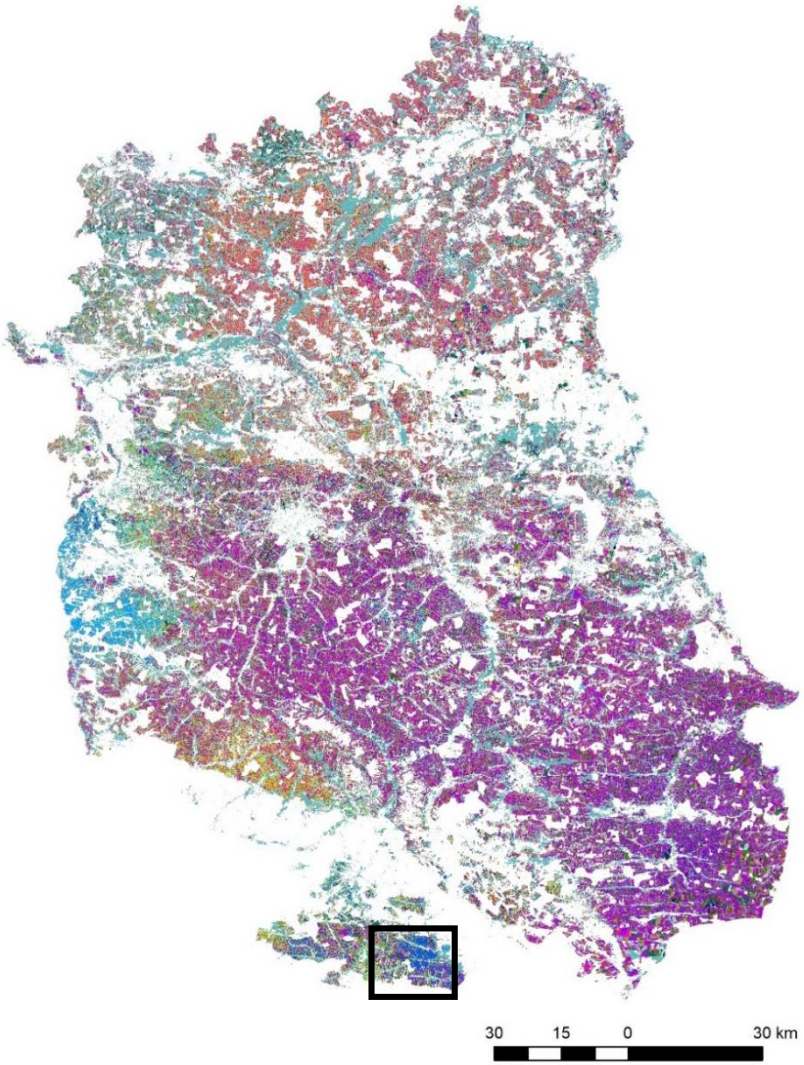


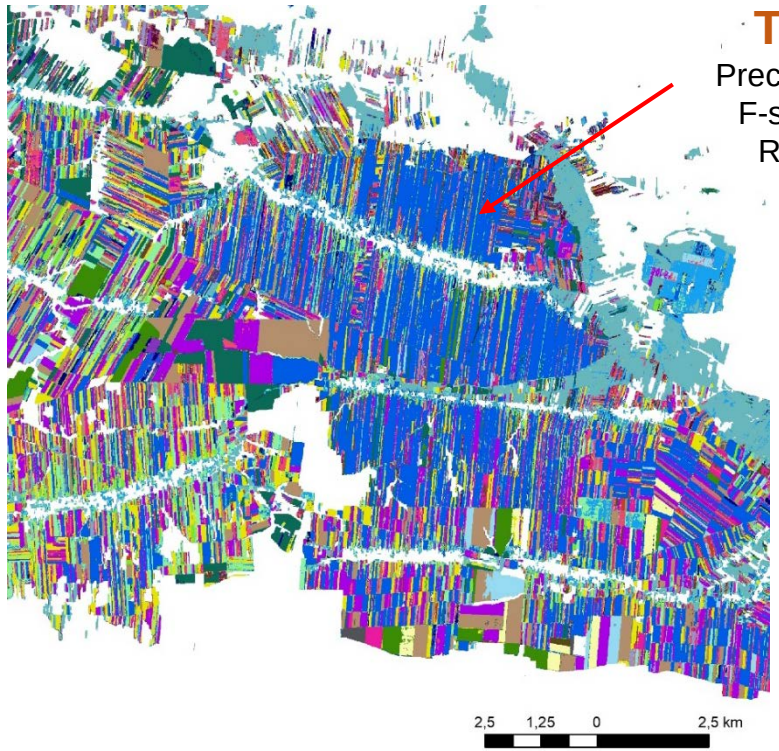




Statistical Office in Olsztyn

The use of satellite observations in agriculture - an example based on optical data Sentinel-2





Tobaco:
Precision = 0.86
F-score = 0.90
Recall = 0.95

- season 2018
- Lubelskie province
- Sentinel-2 data (854 scenes)
- Sen2Agri
- 24 classes

OAA = 70 % KIA = 0.68

Future plans for use of satellite data

- Developing the system of the crop area assessment and replacement of existing surveys and estimates
- Switching to the satellite technics and harvest modelling and replacing the existing system
- Delivering information on lower levels of aggregation (a parcel ?)
- Earlier information providing (July or earlier)
- More applications in land cover and land use statistics
- Enhanced use in forestry and grassland statistics
- Absorption of scientific solutions by the statistics (agriculture, environment, climate changes, land management …)

Conclutions

- Benefits: response burden, cost, low agregation, etc..
- Big Data: dealing with the huge amount of data
- Machine learning, training sample, validation
- Combining Sentinel 1 & Sentinel 2 &.. & Admin Data
- Common Agricultural Policy after 2020
- Extremal conditions
- ….

Statistics Poland

Thank you for your attention☺

Tomasz Milewski
Expert, Agriculture Department,
Statistics Poland
Tel. +48 22 608 34 12
t.milewski@stat.gov.pl