Czech University of Life Sciences Prague

STEROPES

Semi-automatic supervised bare soil pixels retrieval: **Impact of different classification approaches on soil organic carbon prediction**

Vahid Khosravi, Asa Gholizadeh, Fabio Castaldi, Mohammadmehdi Saberioon, Prince Chapman Agyeman, Daniel Žížala, Radka Kodešová, Luboš Borůvka

Department of Soil Science and Soil Protection, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, Czech Republic

EJP SOIL
European Joint Programme

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 652615.

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

Introduction

Soil organic carbon (SOC)

- Important soil component
  - Soil fertility
  - Climate change mitigation

Remote sensing

- **Available data** for SOC estimation
  - Good spatial and temporal coverage
- Limitations of use – **limited bare soil exposure**
  - Snow cover
  - **Vegetation cover and crop residues**
    - No-tillage, conservation tillage, mulching, cover crops, strip-till ...

EJP SOIL
STEROPES

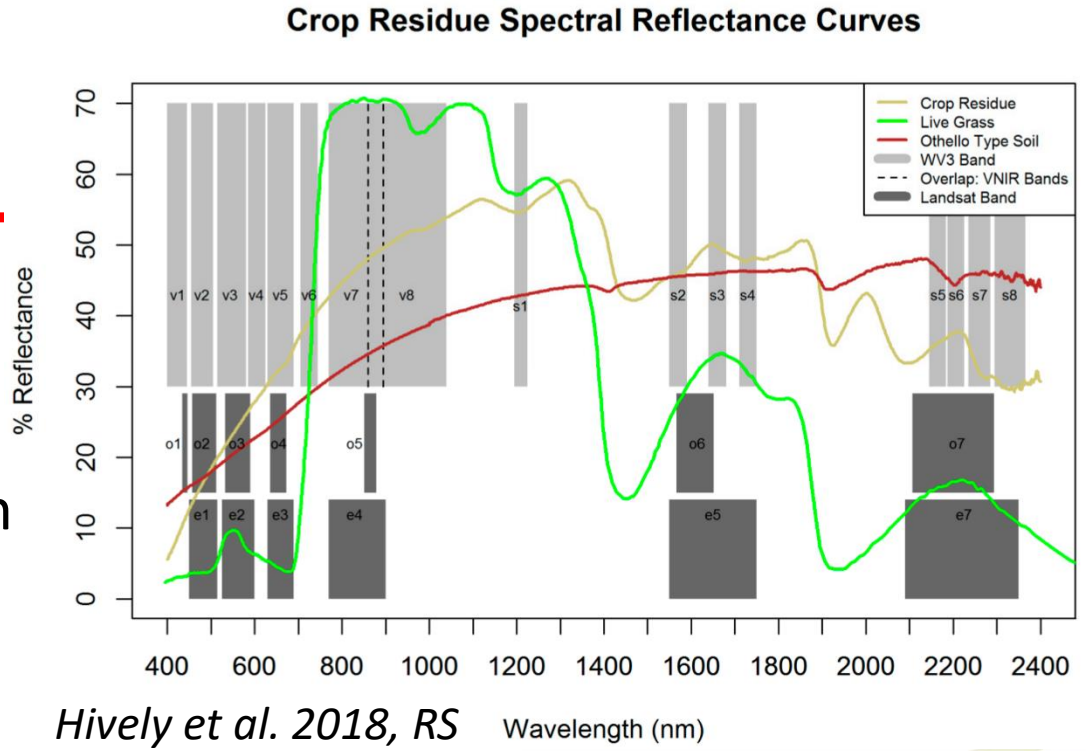
Objective

- To compare the performance of several **methods** (the index-based, unmixing-based, and integrated classification approaches on single-date multispectral satellite data) **to discriminate the bare soil pixels from dry and green vegetation**
- To evaluate the effect of classification on **SOC prediction** accuracy

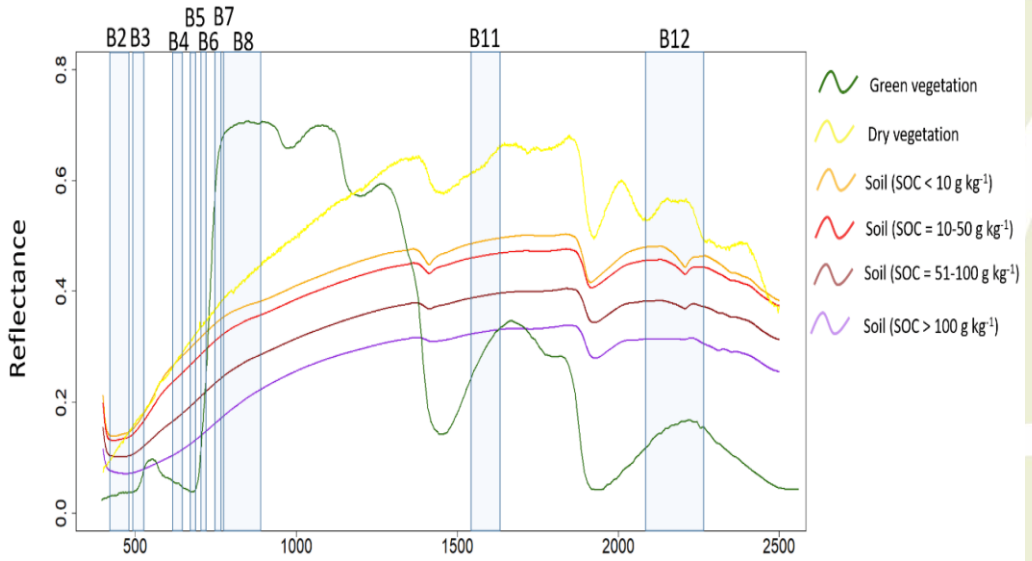
EJP SOIL
STEROPES

Spectral features

- **Green vegetation**
  - high reflectance in green (520 and 570 nm) and NIR (750 to 1000 nm) ranges
  - low reflectance in red (620 to 750 nm) region
- **Crop residues (dry vegetation)**
  - broad absorption band around 2100 nm (lignin, cellulose, and hemicellulose)
  - absorption feature at around 2300 nm (lignin)
  - descending reflectance trend within the range of 1600 to 2300 nm
- **Bare soil**
  - relatively constant reflectance trend between 1600 and 2300 nm

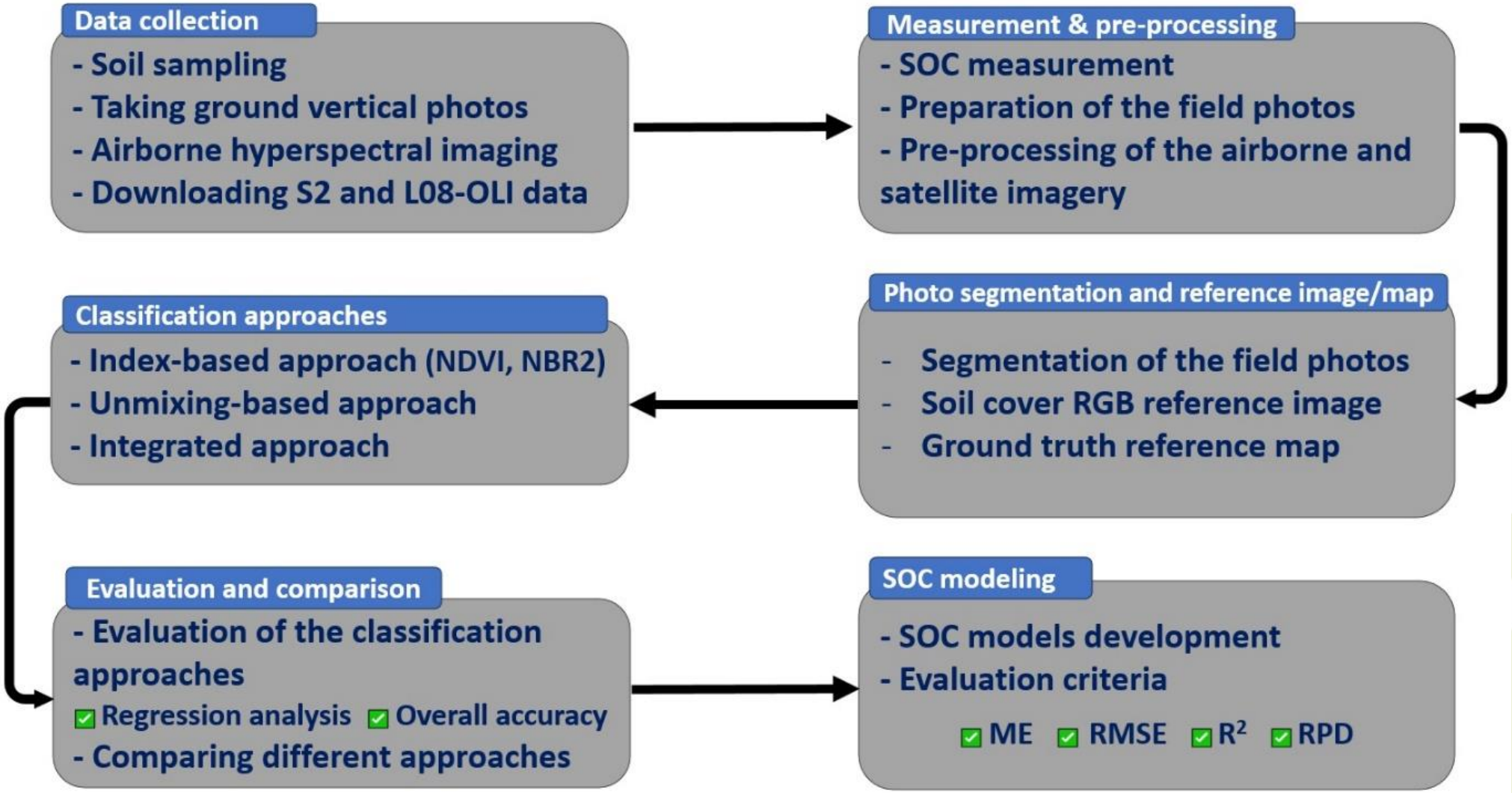


Hively et al. 2018, RS



Castaldi et al. 2019, RS

EJP SOIL
STEROPES



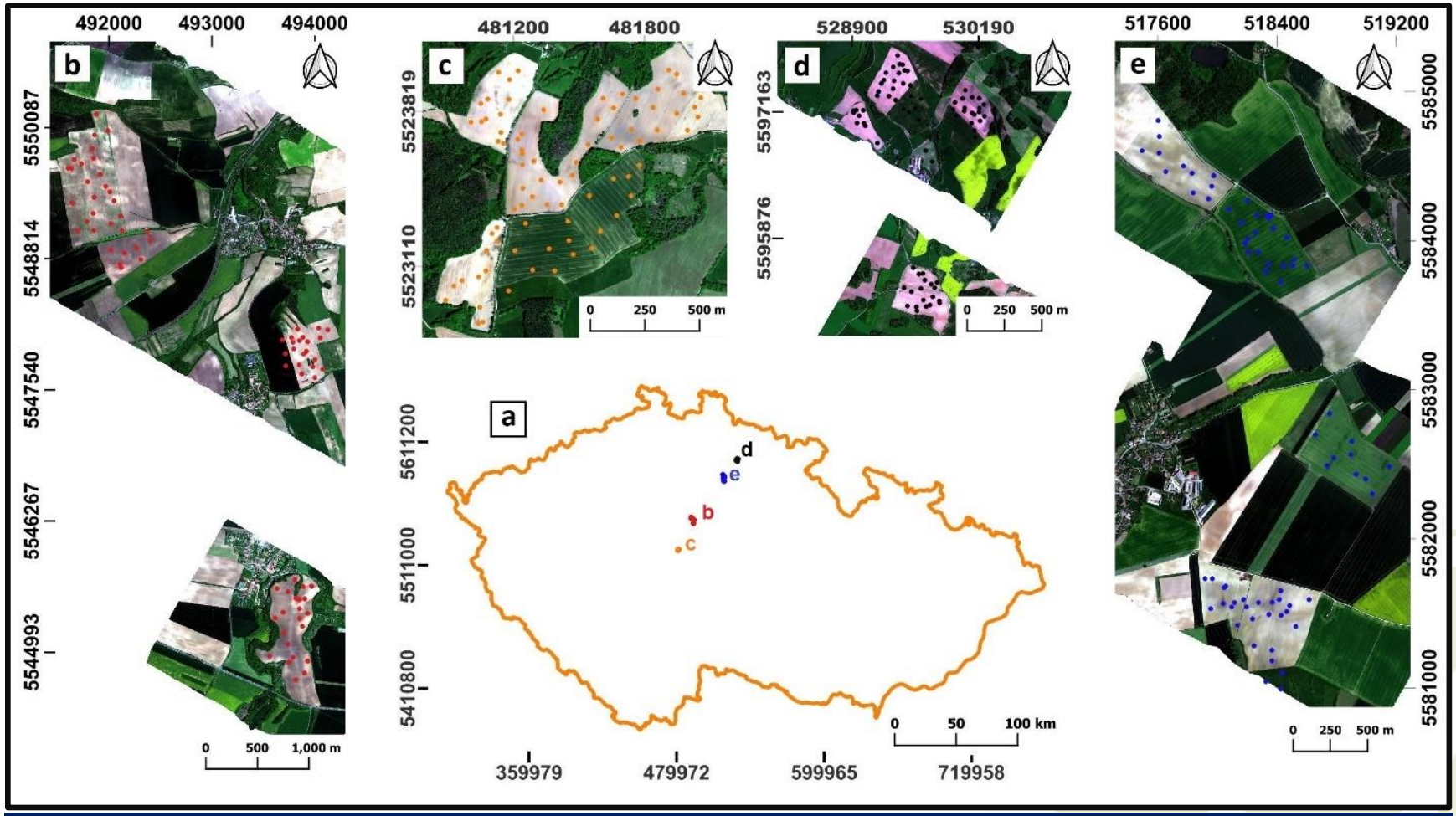
Construction of the study

EJP SOIL
STEROPES

Methodology

Field campaign

- 320 soil samples from four sites (Nová Ves, Jičín, Klučov and Přestavlky)
- SOC content of all samples was measured
- Digital ground vertical photographs were taken as soil cover ground truth



a) Study sites in the Czech Republic and sampling locations in b) Klučov, c) Přestavlky, d) Nová Ves nad Popelkou, and e) Jičín. Background is the airborne image acquired on June 3, 2021

EJP SOIL
STEROPES

Methodology

Multispectral satellite data

- **Landsat 8** (L08-OLI) – Operational land imager level 1T **(radiometrically and geometrically corr.)**
  - Fast line-of-sight atmospheric analysis of spectral hypercubes (FLAASH) algorithm **(for atmospheric correction)**
- **Sentinel 2** (S2) - Level 2A **(atmospherically and geometrically corrected)**
  - Resampling the 20m bands to 10m

Study sites, field campaign and image acquisition dates

Site	Area (ha)	Soil sampling date	Landsat-8 image acquisition date	Sentinel-2 image acquisition date	Airborne image acquisition date
Nová Ves	129	June, 2021	15.06.2021	04.06.2021	03.06.2021
Jičín	153	June, 2021	15.06.2021	04.06.2021	03.06.2021
Klučov	175	June, 2021	15.06.2021	19.06.2021	03.06.2021
Přestavlky	58	June, 2021	15.06.2021	19.06.2021	03.06.2021

Methodology

Airborne hyperspectral imaging

- **CASI/SASI (Itres Ltd., Calgary, Canada) sensors aboard Cessna 208B Grand Caravan photogrammetric aircraft**
- **Ground truth, digital ground vertical photos from the soil sampling sites**

Pre-processing

- Radiometric corrections by RadCorr Ver. 9.3.6.0
- Atmospheric corrections by MODTRAN in the ATCOR-4 Ver. 7.3
- Geometric corrections, orthorectification and georeferencing in GeoCor Ver. 3.7.2
- Removing noisy bands, transformation to absorbance, Savitzky-Golay smoothing and first derivative transformation

Methodology

Photo segmentation, reference image and reference map

- All photos taken from the sampling locations were segmented and the percentages of each cover class was determined

Representative photos of different soil cover classes at sampling points within the study sites:

a) bare soil,
b) green vegetation,
c) dry vegetation,
d) mixture of all cover classes

(a)

(b)

(c)

(d)

EJP SOIL
STEROPES

Methodology

Photo segmentation, reference image and reference map

- A **random forest** (RF) model was developed using the spectra of the **airborne image** at sampling points (as predictor variables) and the obtained **cover class** percentages (as response variables).
- This model was later applied to the whole airborne image to produce a 2.5m × 2.5m pixels reference soil cover **RGB** image (reference image) representing the percentages of the **three cover classes** in each band:
  - **R: bare soil**
  - **G: green vegetation**
  - **B: dry vegetation**
- Binary reference map was obtained by classification of the reference image (pixels with more than 80% bare soil labeled as bare soil, remaining as non-bare)

EJP SOIL
STEROPES

NDVI=((R_NIR−R_Red)/(R_NIR+R_Red))

Methodology – image classification

1. Index-based approach

- **NDVI (Normalized difference vegetation index)** to mask the green vegetation:
  - threshold values from 0.05 to 0.25, with 0.01 increment

$$NDVI = (NIR1 - Red)/(NIR1 + Red)$$

- **NBR2 (Normalized Burn Ratio 2)** to mask the dry vegetation:
  - threshold values from 0.05 to 0.25, with 0.01 increment

$$NBR2 = (SWIR1 - SWIR2)/(SWIR1 + SWIR2)$$

- The reference map was used for defining the optimal threshold values

EJP SOIL
STEROPES

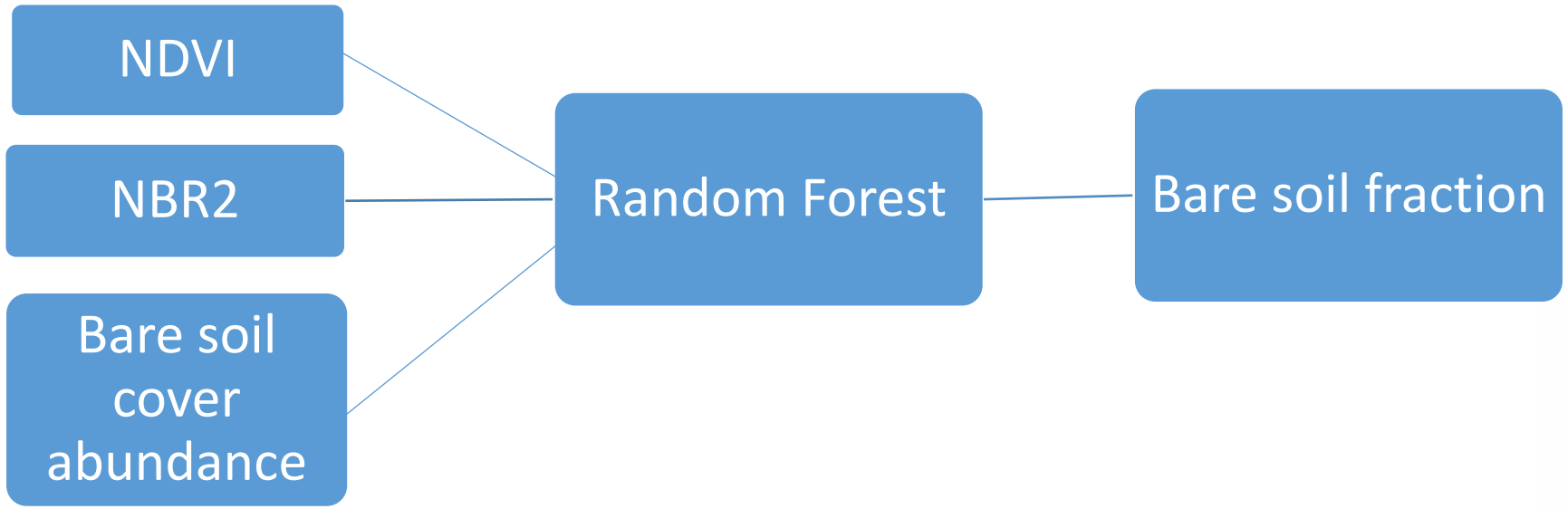
Methodology – image classification

2. Unmixing-based approach

- The **linear spectral unmixing** (LSU) was implemented to disintegrate the pixels fractions of bare soil, green vegetation, and dry vegetation
- The pure pixels were manually extracted from the airborne imagery
- Fifteen pure pixels of each cover class were extracted as endmembers
- Abundance maps of the satellite images pixels were produced
- Different thresholds (from 50% to 95% with the 5% increment) were considered to label the bare soil pixels

Methodology - image classification

3. Integrated approach



- Different thresholds (from 50% to 95% with increment of 5%) to label the mixed pixels

EJP SOIL
STEROPES

Methodology – model development for SOC prediction

Model development procedure

- Kennard–Stone (KS) for **splitting data**
  - training (75%)
  - testing (25%)
- 5-fold cross-validation for training
- **Random Forest** (RF) for modeling

Evaluation criteria

- Index of determination (R^2)
- Root mean square error (RMSE)
- Mean error (ME)
- Residual prediction deviation (RPD)

EJP SOIL
STEROPES

Results

Segmented photos

- All digital photos were segmented and percentages of every soil cover class were calculated.
- Visual comparison indicated that the segmentation was satisfactory and soil cover fractions were successfully extracted.
- Samples-pixels with each soil cover class percentage greater than 80% were labeled as that cover type

EJP SOIL
STEROPES

Results

Samples statistics

Descriptive statistics of samples' SOC content (%) in different cover classes and sites

	Bare soil				Green vegetation				Dry vegetation			
	n	Min	Max	Mean	n	Min	Max	Mean	n	Min	Max	Mean
Nova Ves	63	0.88	3.97	1.42	12	1.06	1.85	1.37	5	1.18	1.57	1.36
Jičín	22	0.81	1.78	1.06	40	0.69	1.61	1.03	17	0.5	1.87	1.03
Klučov	61	0.78	1.54	1.09	3	0.98	1.54	1.17	15	0.77	1.61	1.18
Přestavlky	65	0.63	2.13	1.20	14	0.31	1.51	1.05	3	0.94	1.29	1.12
All	211	0.63	3.97	1.22	69	0.31	1.85	1.1	40	0.5	1.87	1.13

n: number of samples, Min: Minimum, Max: Maximum

EJP SOIL
STEROPES

Results

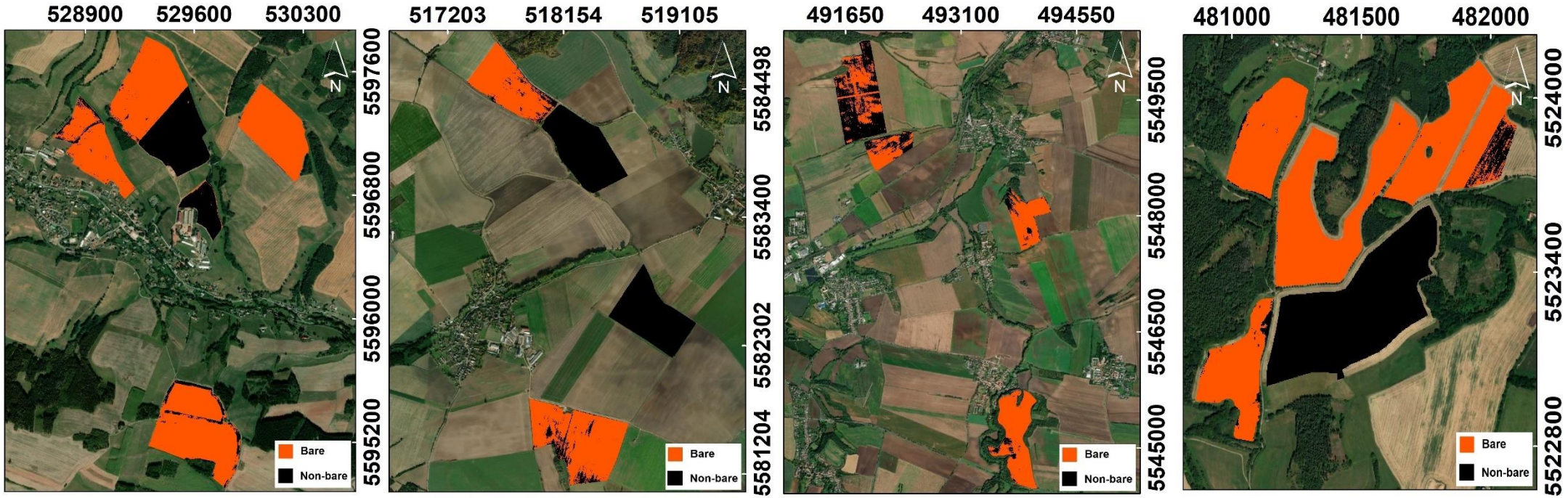
Soil cover reference image

Reference image (R: bare soil, G: green vegetation, and B: dry vegetation) for a) Nová Ves nad Popelkou, b) Jičín, c) Klučov, and d) Přestavlky sites

EJP SOIL
STEROPES

Results

Soil cover reference map



Reference map prepared after classification of the reference image for a) Nová Ves nad Popelkou, b) Jičín, c) Klučov, and d) Přestavlky sites

EJP SOIL
STEROPES

Results

Satellite images classification results

Linear regression parameters and OA (%) (all for bare soil) obtained by different classification approaches

	Index-based		Unmixing-based		Integrated	
	L08-OLI	S2	L08-OLI	S2	L08-OLI	S2
R^2	0.75	0.78	0.77	0.81	0.81	0.87
RMSE	0.05	0.03	0.03	0.02	0.02	0.01
RPD	2.11	2.13	2.12	2.16	2.15	2.21
OT	NDVI: 0.22 NBR2: 0.13		0.90	0.90	0.85	0.85
OA (%)	68	74	70	78	76	85

OT: optimal threshold, OA: overall accuracy

EJP SOIL
STEROPES

Results

Satellite images classification results

Linear regression parameters and OA (%) (all for bare soil) obtained by different classification approaches

	Index-based		Unmixing-based		Integrated	
	L08-OLI	S2	L08-OLI	S2	L08-OLI	S2
R^2	0.75	0.78	0.77	0.81	0.81	0.87
RMSE	0.05	0.03	0.03	0.02	0.02	0.01
RPD	2.11	2.13	2.12	2.16	2.15	2.21
OT	NDVI: 0.22 NBR2: 0.13		0.90	0.90	0.85	0.85
OA (%)	68	74	70	78	76	85

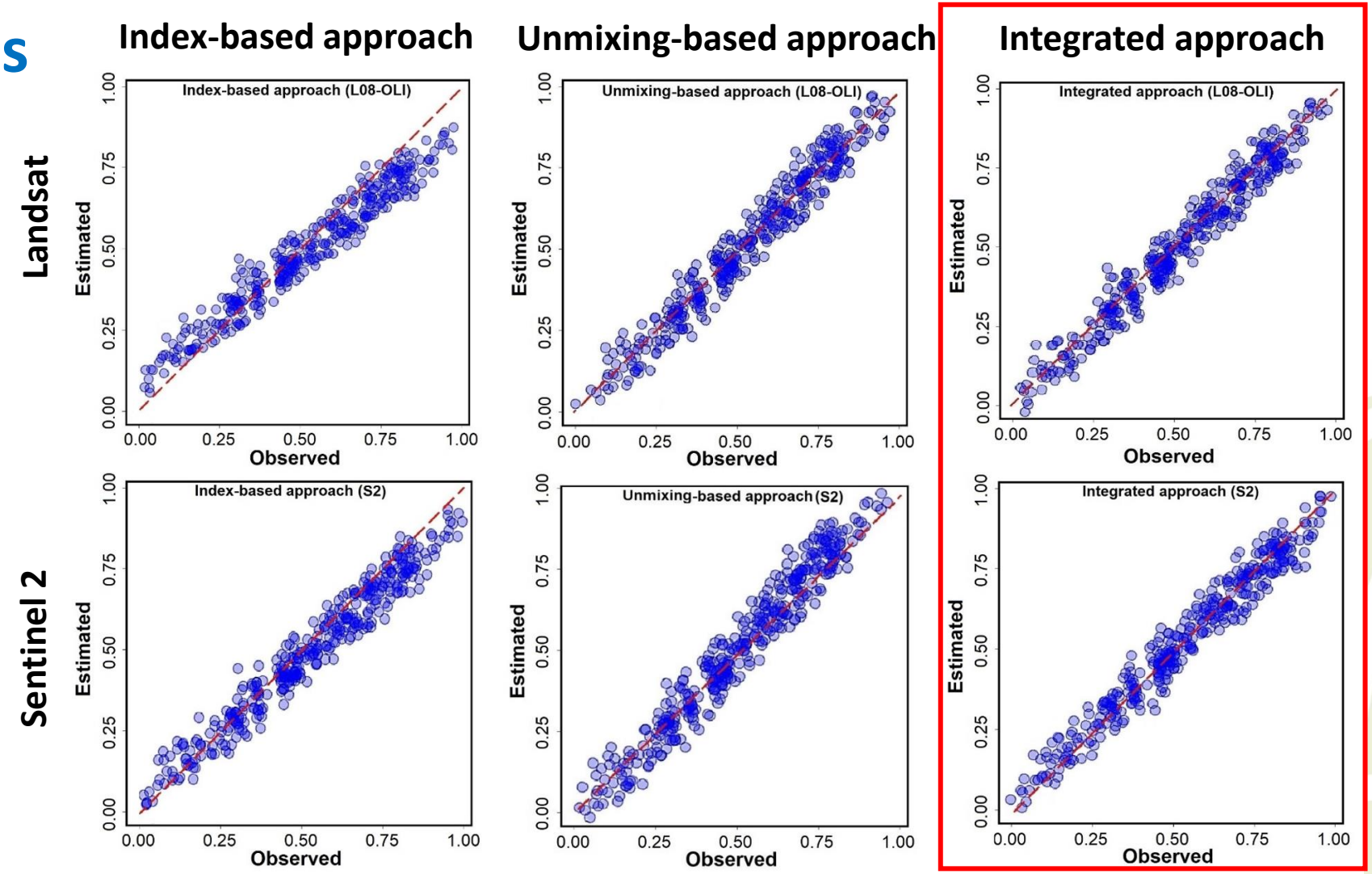
OT: optimal threshold, OA: overall accuracy

EJP SOIL
STEROPES

Results

Classification results

Scatter diagrams comparing the observed and estimated bare soil fractions using different approaches
(at 400 selected pixels distributed evenly throughout the whole study area)

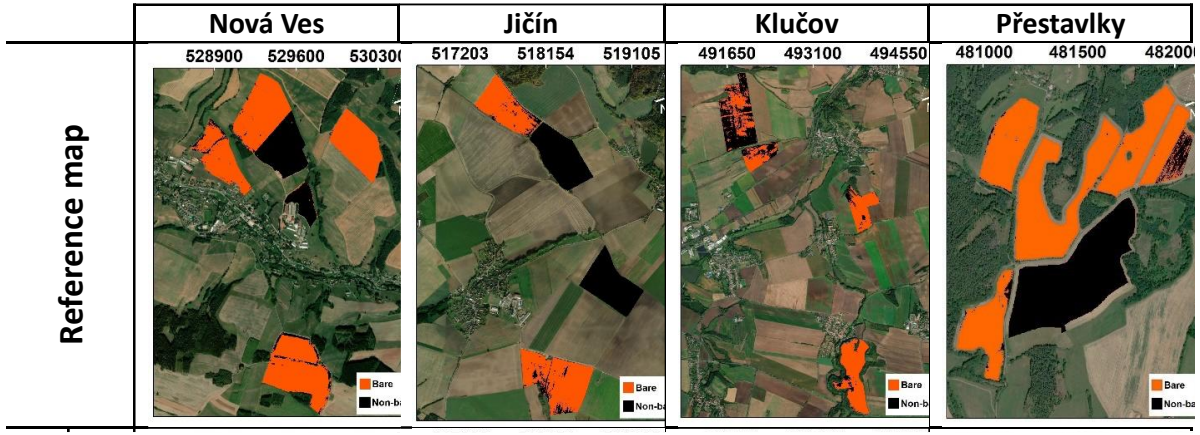


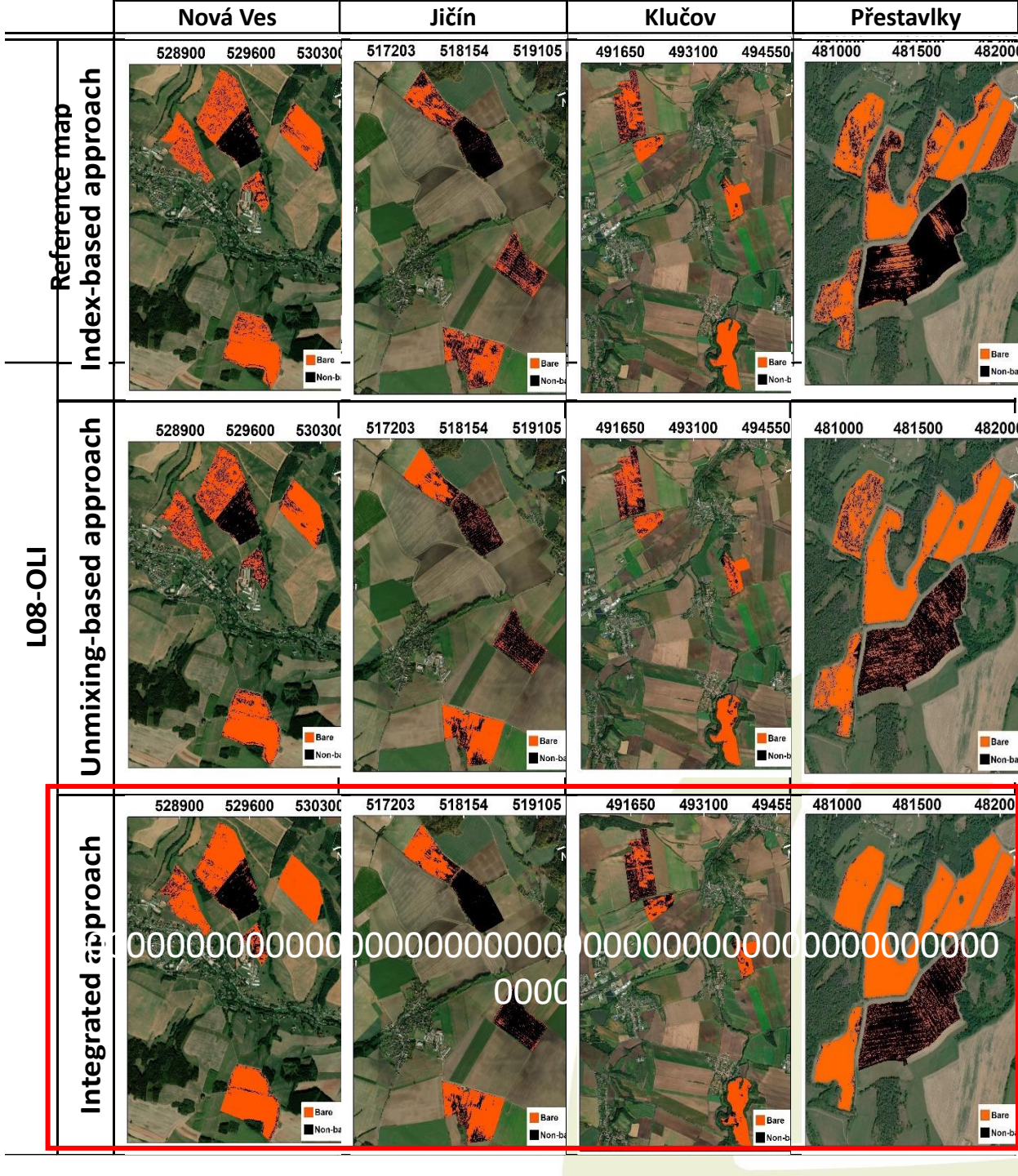
EJP SOIL
STEROPES

Results

Spatial distribution maps

Bare soil pixels retrieved using different classification approaches on Landsat data



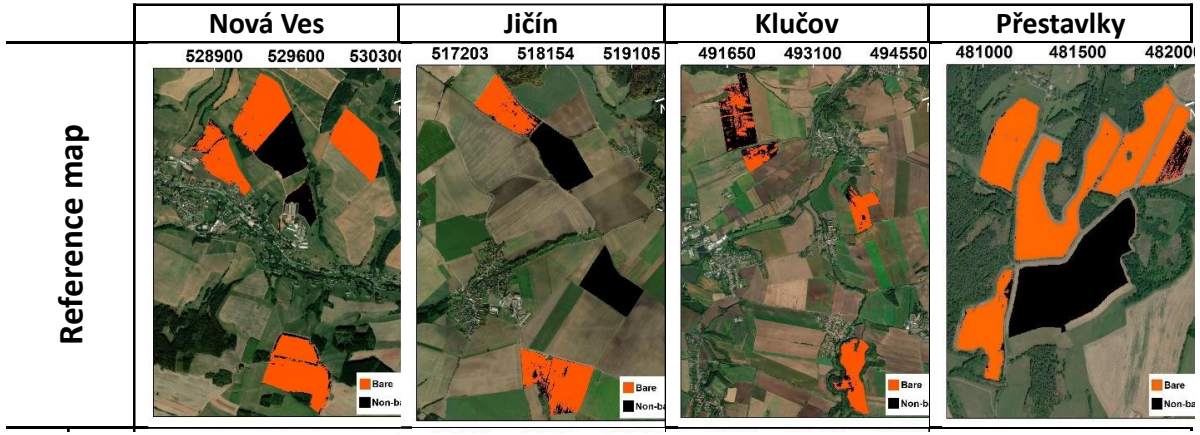


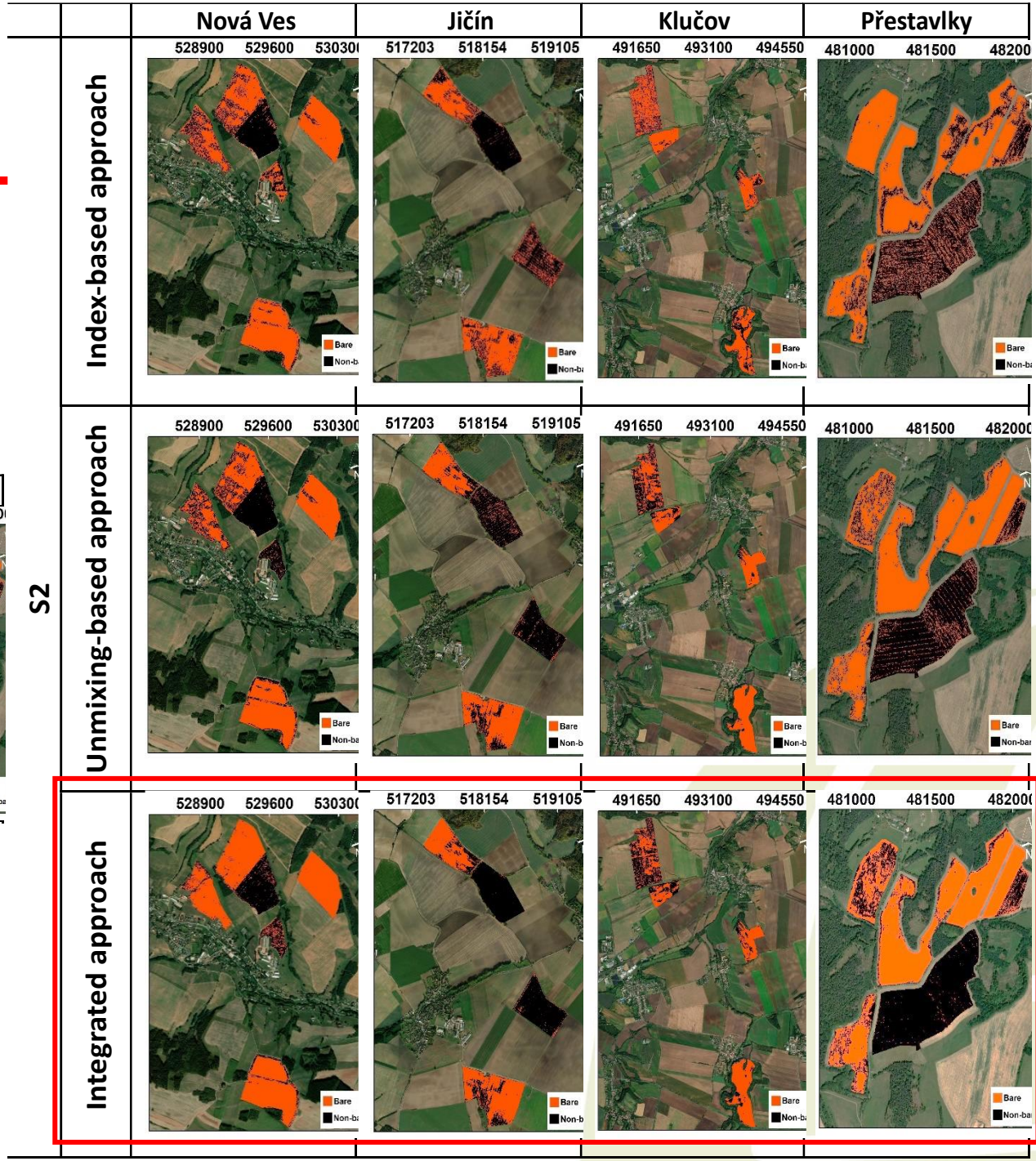
EJP SOIL
STEROPES

Results

Spatial distribution maps

Bare soil pixels retrieved using different classification approaches on Sentinel 2 data





EJP SOIL
STEROPES

Results

SOC modeling on retrieved bare soil

SOC prediction performance of models developed on bare soil sample-pixels retrieved via different classification approaches (validation data)

Site	Approach	Landsat 8-OLI						Sentinel-2					
		n	ME	RMSE	R^2	CCC	RPD	n	ME	RMSE	R^2	CCC	RPD
Nová Ves	**Index-based**	58	0.15	0.31	0.29	0.32	1.47	61	0.13	0.28	0.32	0.38	1.62
	Unmixing	59	-0.10	0.28	0.27	0.30	1.61	63	0.11	0.25	0.31	0.42	1.83
	Integrated	50	-0.12	0.25	0.32	0.34	1.82	52	-0.08	0.22	0.37	0.39	2.06
Jičín	**Index-based**	17	-0.12	0.22	0.37	0.41	1.16	19	-0.1	0.19	0.43	0.49	1.35
	Unmixing	18	0.14	0.19	0.46	0.53	1.35	20	0.12	0.21	0.49	0.51	1.22
	Integrated	15	-0.09	0.17	0.51	0.62	1.51	17	0.06	0.14	0.52	0.55	1.83
Klučov	**Index-based**	56	0.09	0.20	0.50	0.56	0.90	59	-0.06	0.16	0.55	0.62	1.12
	Unmixing	56	0.10	0.17	0.53	0.55	1.05	58	-0.09	0.14	0.58	0.60	1.28
	Integrated	44	-0.07	0.13	0.58	0.62	1.38	50	0.07	0.11	0.64	0.67	1.63
Přestavlky	**Index-based**	60	0.10	0.23	0.46	0.52	1.35	63	0.08	0.17	0.55	0.61	1.82
	Unmixing	61	0.08	0.16	0.57	0.59	1.94	65	-0.06	0.18	0.59	0.65	1.72
	Integrated	52	-0.04	0.13	0.60	0.63	2.38	55	0.05	0.10	0.63	0.69	3.10
All areas	**Index-based**	191	-0.16	0.25	0.33	0.37	1.42	202	0.14	0.27	0.24	0.41	1.32
	Unmixing	194	0.09	0.30	0.32	0.32	1.18	206	0.12	0.29	0.33	0.37	1.22
	Integrated	161	0.11	0.23	0.36	0.40	1.54	174	-0.08	0.20	0.39	0.42	1.78

n: number of retrieved bare soil sample-pixels

EJP SOIL
STEROPES

Conclusions

- Separate implementation of the approaches yielded lower accuracy results
- Integration of the two approaches showed the best performance in estimation of the bare soil fractions, especially on S2 data
- The OA of binary classification was the highest for integrated approach reaching 85%, followed by the unmixing-based approach with OA value of 78% (both on S2 data)
- Models developed on the sample-pixels retrieved using the integrated approach showed superior performance followed by those constructed on unmixing-based approach related data
- Nevertheless, the effect of locality on the prediction accuracy was bigger than the effect of the method used

EJP SOIL
STEROPES


Czech University
of Life Sciences Prague
Thank you for your attention
EJP SOIL
STEROPES
E-mail: boruvka@af.czu.cz