Monitoring Urban Changes with Ensemble of Neural Networks and Deep-Temporal Remote Sensing Data

Georg Zitzlsberger, Michal Podhoranyi, Václav Svatoň, Milan Lazecký, and Jan Martinovič

VSB TECHNICAL UNIVERSITY OF OSTRAVA | IT4INNOVATIONS NATIONAL SUPERCOMPUTING CENTER

Objectives

Monitor urban changes with high resolution remote sensing data:

- Combining both **Synthetic Aperture Radar** (SAR) and **multi-spectral optical**
- Eras **1991-2011** (ERS-1 & ERS-2 and Landsat 5 TM) and **2017-now** (Sentinel 1 & 2)
- Train neural network model with synthetic labels to provide a **fully automated** process
- Use **publicly available level 1 data** for maximum applicability and spatiotemporal coverage
- Resilient handling of **partial and irregular observations**

Data Sources

Era	Site	SAR observations (ascending & descending)	Optical multispectral observations	Area (km^2)
ERS-1/2 & Landsat 5 TM 1991-2011	Rotterdam	974 (-118)	753 (-434)	712.9
	Liège	934 (-89)	888 (-620)	713.3
	Limassol	291 (-27)	380 (-61)	718.6
	Sources	SAR: 12.5m/pixel, ERS-1/2 *SAR_IMP_1P* Optical: 30m/pixel, Landsat 5 TM *L4-5 TM C1 L1*	esa / USGS	
Sentinel 1 & Sentinel 2 2017-2021	Rotterdam	1,603 (-4)	278 (-10)	523.6
	Liège	1,040 (-0)	332 (-35)	508.4
	Limassol	468 (-0)	407 (-35)	576.2
	Sources	SAR: 10m/pixel, Sentinel 1 *SENTINEL1_IW_[ASC|DSC]* Optical: 10m/pixel, Sentinel 2 *L1C*	sentinelhub	

Table 1. Used areas of interests with number total number of observations and removed ones in parenthesis. They are complemented by information of areas covered and used data sources with product designators (all level 1).

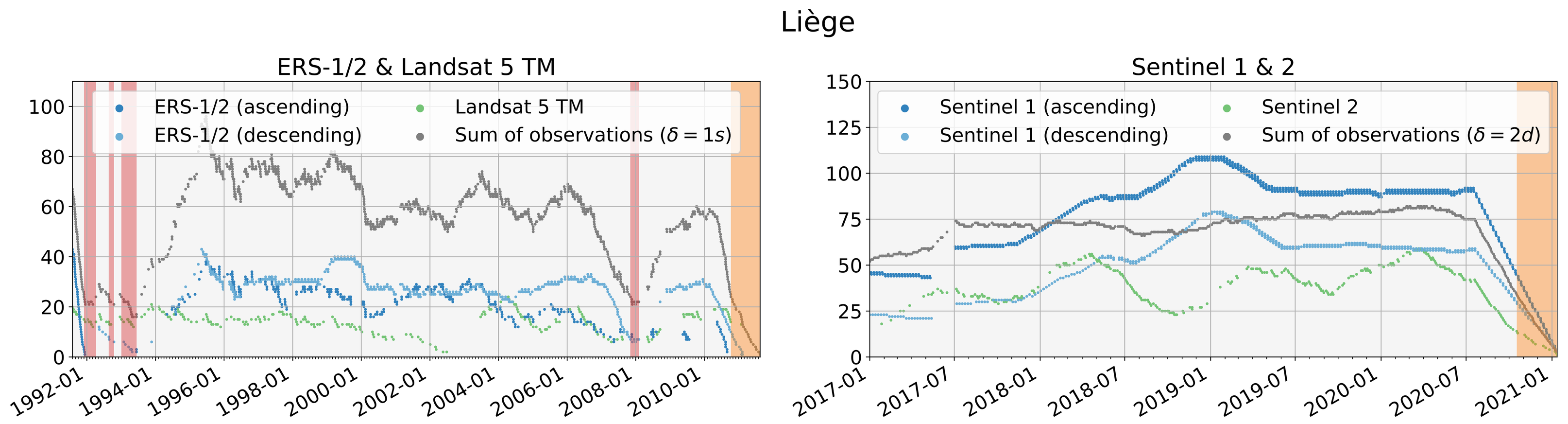


Figure 1. Number of remote sensing observations for the windows $\boldsymbol{w}^t_{i,j}$ for ERS-1/2 & Landsat 5 TM (left), and Sentinel 1 & 2 (right) era. The x-axis denotes the start times t of the according windows $\boldsymbol{w}^t_{i,j}$. In gray are accumulated observations ($\leq \Omega$) for the given step size δ. Areas in red/orange denote windows with observations $< \omega$; orange marks the end of mission/data.

Pre-processing

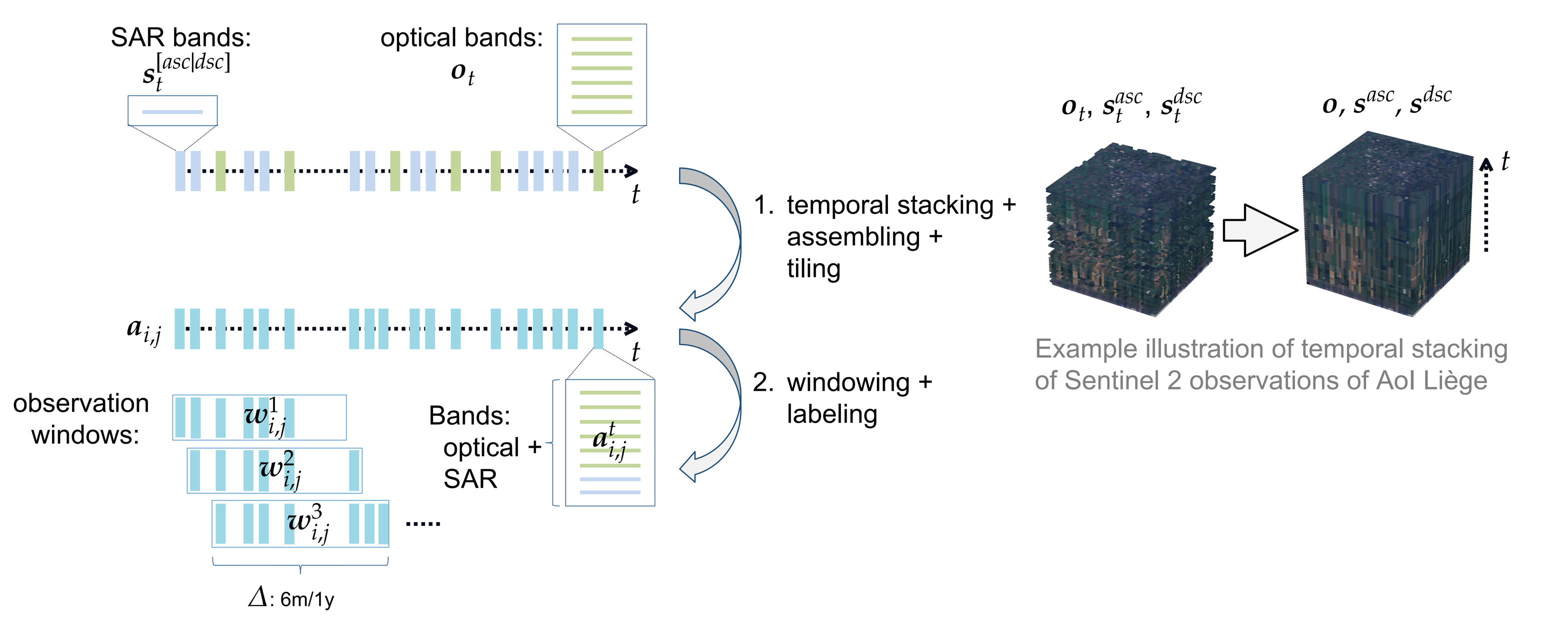
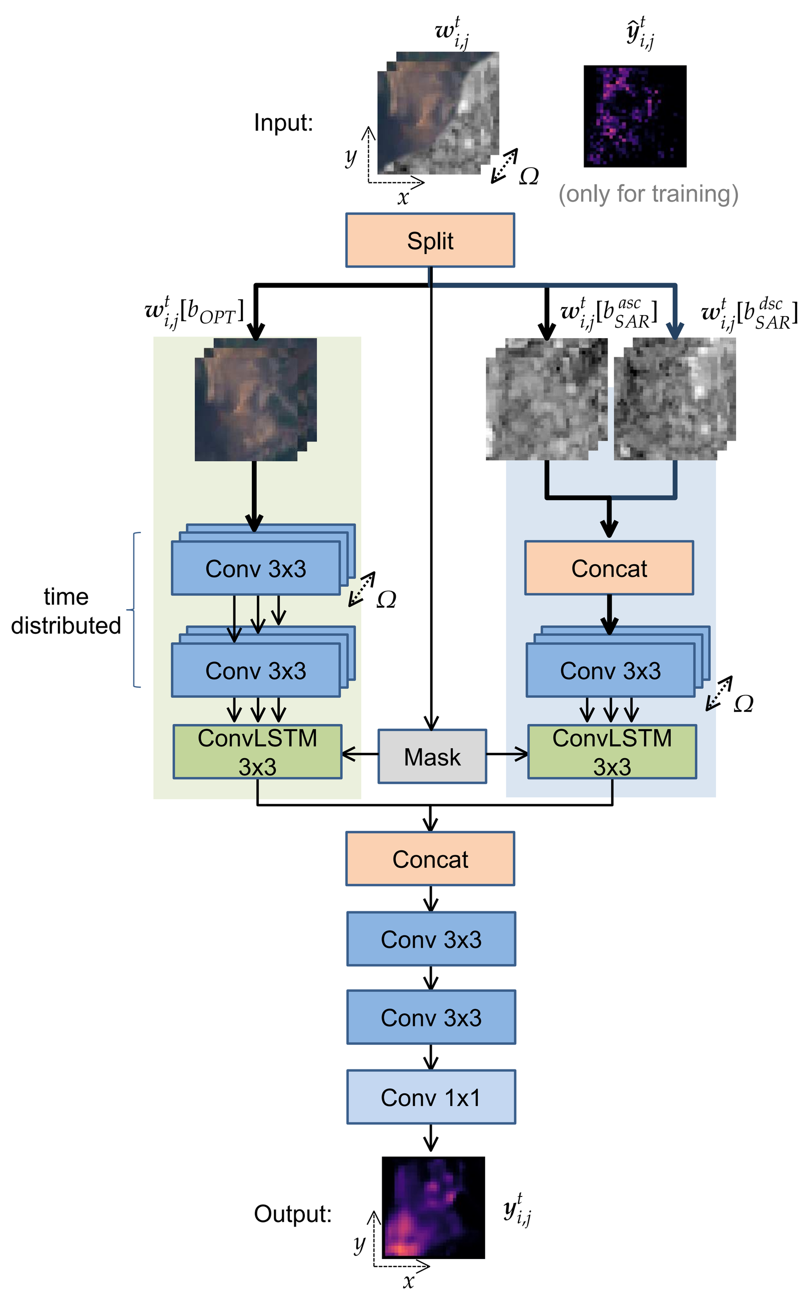


Figure 2. Two-step procedure to data preparation: From the series of available observations to the windows used for training/inference. One fourth and one eighth of the tiles are used for training and validation, respectively. Only a random one tenth of windows are considered each to minimize learning the sequence of observations.

Papers & Github

MDPI Remote Sensing [5] | BLENDED [4] | GitHub

Model Architecture and Synthetic Label Generation



Supervised learning with synthetic labels as

$$\hat{\boldsymbol{y}}^t_{i,j} := s_{cm}(\boldsymbol{w}^t_{i,j}[b^{asc}_{SAR}], \boldsymbol{w}^t_{i,j}[b^{dsc}_{SAR}]) \cdot o_{cm}(\boldsymbol{w}^{t-\Delta}_{i,j}[b_{OPT}], \boldsymbol{w}^{t+\Delta}_{i,j}[b_{OPT}])$$

with

$$s_{cm}(\boldsymbol{p}^{asc}, \boldsymbol{p}^{dsc}) = \frac{O(\boldsymbol{p}^{asc}, \sigma, \eta) + O(\boldsymbol{p}^{dsc}, \sigma, \eta)}{2}$$

and

$$o_{cm}(\boldsymbol{p}^-, \boldsymbol{p}^+) = abs\left(E\Big((\boldsymbol{p}^-)_{mean^t}, \alpha, \gamma\Big) - E\Big((\boldsymbol{p}^+)_{mean^t}, \alpha, \gamma\Big) \right).$$

Methods:

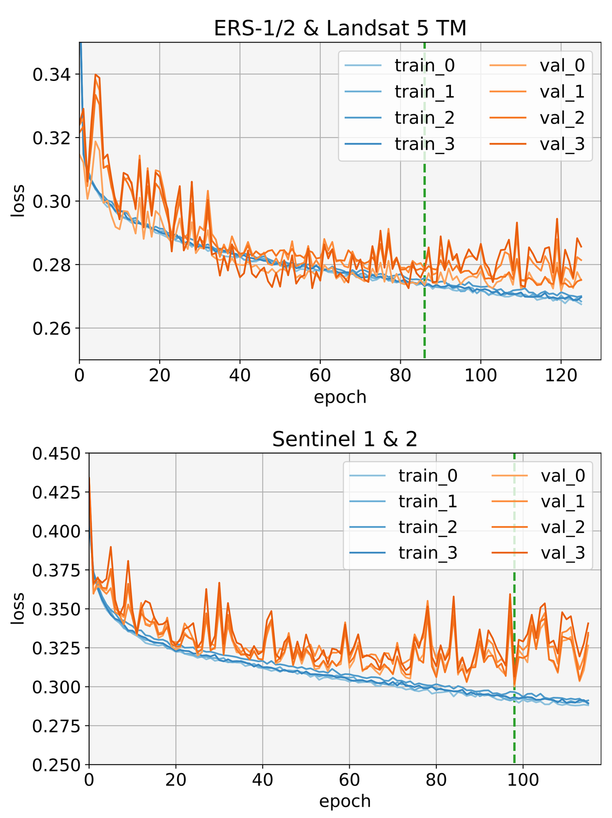
- $O(\ldots)$ is the Omnibus test statistic to identify changes in SAR data (Conradsen et al. [2])
- $E(\ldots)$ is our extended version [5] using the enhanced normalized difference impervious surfaces index (ENDISI, Chen et al. [1])

Figure 3. Ensemble neural network model architecture. Arrows indicate data flow of windows $\boldsymbol{w}^t_{i,j}$ for different data types (background green: multispectral optical, blue: SAR).

Parameter	Mnemonic	ERS-1/2 & Landsat 5 TM Rotterdam/Liège	ERS-1/2 & Landsat 5 TM Limassol	Sentinel 1 & 2 Rotterdam/Liège	Sentinel 1 & 2 Limassol
$b^{[asc\|dsc]}_{SAR}$	SAR bands	1 (VV)	1 (VV)	2 (VV+VH)	2 (VV+VH)
b_{OPT}	optical bands	7	7	13	13
α	shift	0.25	0.5	0.25	0.5
γ	scale	30.0	30.0	10.0	10.0
σ	significance	0.1	0.1	0.001	0.001
η	ENL	3	3	4	4
δ	step $(\frac{\cdot}{\text{observation}})$	1 second	1 second	2 days	2 days
ω	min. window size	25	25	35	35
Ω	max. window size	110	110	92	92

Table 2. Parameters used for training/inference for the different areas of interest and eras.

Training



- Four NVIDIA Tesla V100-SXM2 GPUs
- Barbora cluster at IT4Innovations
- Tensorflow 2.4 with Keras and Horovod (data parallel)
- Batch size of 32 per GPU (128 effective), fixed learning rate of 0.004 and Sync-SGD with momentum of 0.8
- Tanimoto with complement loss (Diakogiannis et al. [3])
- Data sizes (TFRecord files with GZIP compression):
  - 58 GB training, 7 GB validation } ERS-1/2 & Landsat 5 TM (40:23h)
  - 187 GB training, 24 GB validation } Sentinel 1 & 2 (23:30h)
- Only training data is shuffled; every GPU contains a fixed $\frac{1}{4}^{th}$ of the overall validation data (*val_[0-3]*)

Figure 4. Loss values over epochs for both eras for the ensemble with both multispectral optical and SAR data. The green vertical indicates the best ensemble parameters for each era.

Results

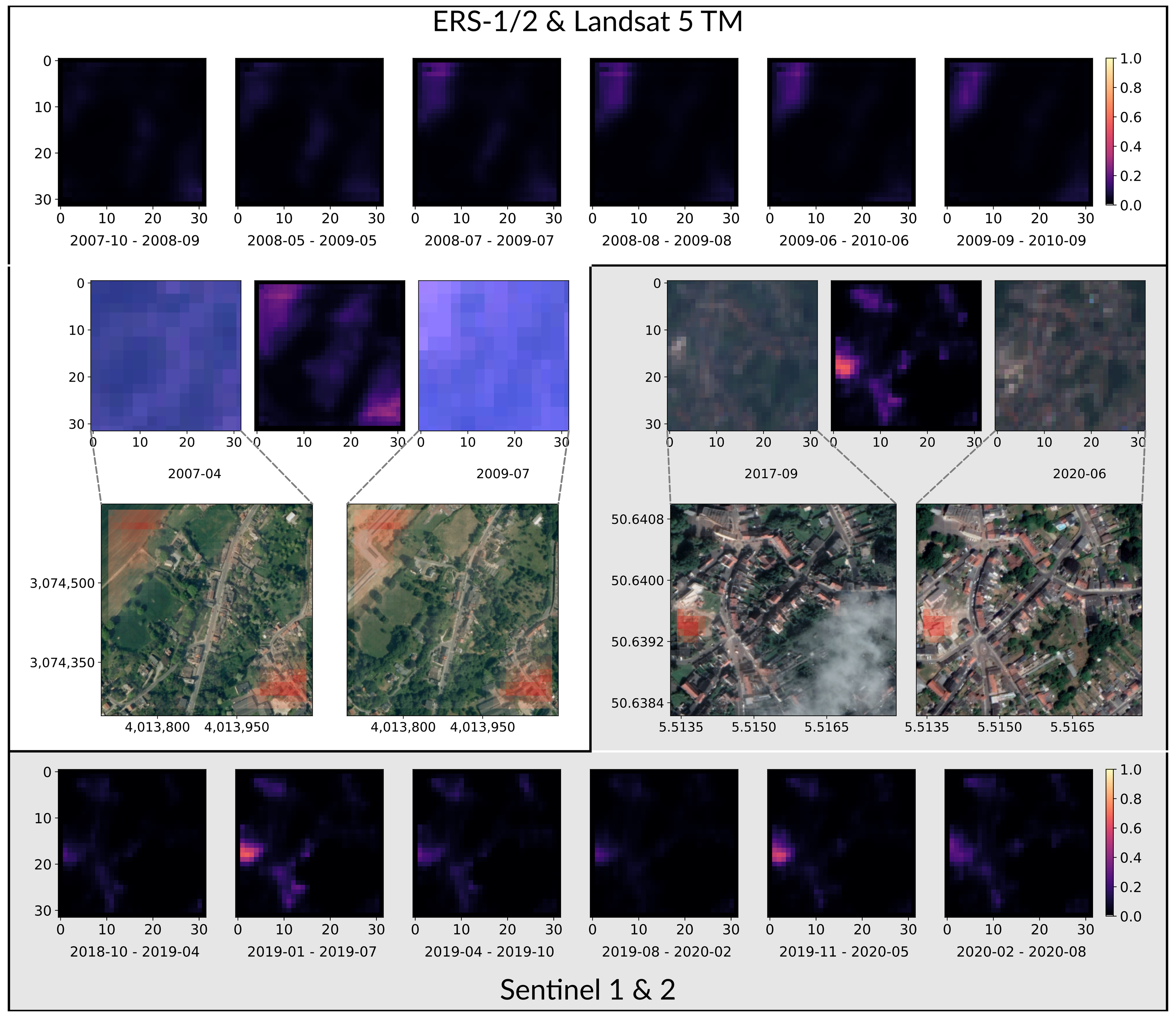


Figure 5. Detection examples for ERS-1/2 & Landsat 5 TM (top) and Sentinel 1 & 2 eras (bottom; gray background). For each, a detection series of six windows are shown on top and bottom, respectively. The center shows for both eras the true color optical observations at two points in time (left, right) with the maximum of all detected changes over the entire era (middle). Zoomed in are corresponding very-high-resolution imagery from Google Earth, © 2021 Maxar Technologies with detections superimposed in red. Higher values in detection outputs $\boldsymbol{y}^t_{i,j}$ indicate a higher intensity of change.

Advantages:

- Supports "real-time" monitoring, which requires no averaging or lead time
- Fully automated but parameterizable training and inference processes
- Combination of SAR and optical observations provides higher temporal resolution to analyze fine-grained changes
- A network ensemble enables better control and use of just one remote sensing type

Current Limitations (cf. [5]):

- High reflective surfaces (e.g., white roofs, metal objects) induce large gradients
- Artifacts due to incomplete cloud removal
- Mobile objects cause false positives
- Farming land can be wrongly detected as impervious surface in case of droughts
- Detections show a slight bias towards optical observations

Acknowledgements

This research was funded by ESA via the *Blockchain ENabled DEep Learning for Space Data* (BLENDED) project [4] (SpaceApps Subcontract No. 4000129481/19/I-IT4I), and by the Ministry of Education, Youth and Sports from the National Programme of Sustainability (NPS II) project "IT4Innovations excellence in science - LQ1602" and by the IT4Innovations infrastructure which is supported from the Ministry of Education, Youth and Sports of the Czech Republic through the e-INFRA CZ (ID:90140) via Open Access Grant Competition (OPEN-21-31).

References

[1] Junyi Chen, Kun Yang, Suozhong Chen, Chao Yang, Shaohua Zhang, and Liang He. Enhanced normalized difference index for impervious surface area estimation at the plateau basin scale. *Journal of Applied Remote Sensing*, 13(1):1 – 19, 2019.

[2] Knut Conradsen, Allan Aasbjerg Nielsen, and Henning Skriver. Determining the points of change in time series of polarimetric sar data. *IEEE Transactions on Geoscience and Remote Sensing*, 54(5):3007–3024, 2016.

[3] Foivos I. Diakogiannis, François Waldner, Peter Caccetta, and Chen Wu. Resunet-a: A deep learning framework for semantic segmentation of remotely sensed data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 162:94–114, April 2020.

[4] Bernard Valentin, Leslie Gale, Hakim Boulahya, Betty Charalampopoulou, Charalampopoulou Christos K., Dimitris Poursanidis, Nektarios Chrysoulakis, Václav Svatoň, Georg Zitzlsberger, Michal Podhoranyi, Dušan Kolář, Vladimír Veselý, Ondrej Lichtner, Michal Koutenský, Dominika Regéciová, and Matúš Múčka. Blended - using blockchain and deep learning for space data processing. In Pierre Soille, Sveinung Loekken, and Sergio Albani, editors, *Proceedings of the 2021 conference on Big Data from Space*, JRC125131, pages 97–100. Publications Office of the European Union, 18–20 May 2021.

[5] Georg Zitzlsberger, Michal Podhorányi, Václav Svatoň, Milan Lazecký, and Jan Martinovič. Neural network-based urban change monitoring with deep-temporal multispectral and sar remote sensing data. *Remote Sensing*, 13(15), 2021.

georg.zitzlsberger@vsb.cz