A, Preprocessing of Satellite images, cloud masking

According to the work plan the cloud mask and Quality Assessment bands of Landsat 8 images from 2014 to 2016 for Hungary. Because of the operational time of Landsat 8, starting at 11th February 2013 and it is still in operation, all Landsat 8 images were searched and ordered from USG Earth Explorer web site. Ten Landsat tile can cover the entire area of Hungary, and for each tile are captured 22-23 Landsat scenes per year. The number of images per year is more than 200, and it needed to be downloaded and organized into database. It does not require big data storage, because these layers are single band files, but for the later image processing a new back storage has to be implemented to our system. For computing the cloud free pixels of a given Landsat image the Cloud Mask band was used. which can be downloaded together with the optical bands containing surface reflectance (SR). In a Cloud Mask Band there three values: cloud, shadow of the cloud, and non-covered surface. During this research period all those Landsat 5 TM, Landsat 7 ETM+ and Landsat 8 OLI images we selected, which cover total area of Hungary in the shortest period in a year with the less cloud cover. 4 years could be selected according to the above mentioned condition from the database: 1986 (L5TM), 2003 (L7ETM+), 2015 and 2017 (both L8OLI). In the beginning the TOA, later Surface Reflectance Bands were ordered and downloaded. The 4 mosaics were constructed from 10 images using ERDAS 2013 and ERDAS 2020 Mosaic Pro tools.

These images were segmented and used to investigate by Péter Szilassi (senior researcher in this research project, and the result were published in Tájökológiai lapok 2017 15 (2) p.131-138 The title of the article is: Magyarországi kistájak felszínborítás változékonysága és felszínborítás mozaikosságuk változása (http://real.mtak.hu/71352/1/07_Szilassi.pdf)

B, Time series analysis of Landsat images for Csongrád County – Methodology development

Decision tree classification was used to produce the CLC00 database for Csongrád County, in Hungary. Three different Landsat image (30 April 2000, 3 July 2000 and 20 August 2000) as well as the Normalized Difference Vegetation Index (NDVI) and Tasseled Cap (TC) derived from the Landsat images and Shuttle Radar Topography Mission (SRTM) elevation data were used. Using all data and the entire data set as a training set, we achieved an overall accuracy of 78,6%. The classification process was able to delimit not only the pure land cover, but also most of the land use classes, so it is suitable for automatic production of the land use and some land use classes of the CLC map.

From the runs it was found that the highest accuracy (78.56%) is achieved when using all satellite images taken at different times, SRTM altitude data and Tasseled Cap transformation data. It can be concluded that only the NDVI index did not result in an improvement in accuracy (>0.02%), which may be due to the fact that the one-dimensional histogram of NDVI values between -1 and 1 does not clearly distinguish between CLC classes. The largest improvement in accuracy was due to the use of the July satellite image. Presumably, the information content of the reflectance differences due to the temporal variation resulted in the accuracy improvement. We tested the effect of increasing the study area on the accuracy of the classification and found that the size of the decision tree generated and the number of endings increased with the increase. As more data is available for model building, more endings can be generated by the model building procedure. It can be clearly seen that the accuracy varies logarithmically with the change in the size of the training area.

The result were published and presented in: Gudmann András – Mucsi László – Henits László: A CORINE felszínborítási térkép automatikus előállításának lehetősége döntésifa-osztályozó segítségével – Geodézia és Kartográfia 2019 71. évf. : 2 pp. 9-13.DOI: https://doi.org/10.30921/GK.71.2019.2.2 (Q4)

In part of the research project we produced land cover maps for the two sample areas for the year 2000 using two classification methods. For the classification, we used the spectral bands (7 bands) of Landsat 7 satellite imagery taken at six different dates in the given year, the vegetation, moisture and land cover indices (NDVI, EVI, SAVI, MSAVI, NDMI, NDBI, Urban index) calculated from them, Tasseled Cap transformation values (6 bands), and SRTM elevation data.

For both classifiers, the concentricity of the thematic layers created for the urban sample area is lower than that of the thematic strata created for the agricultural sample area. In the case of the decision tree, the map for the Szeged sample area has a accuracy of 81.47%, while the image for the agricultural area has a concentration of 89.75%. For the random forest classifier, the difference in concentrations is smaller, 95.13% and 97.05%, respectively. The difference in accuracy between the sample areas may be due to the complexity of the multiple classes in the urban area and the artificial surface classes

The result were published and presented in*: Gudmann András – Mucsi László: Döntési fa és véletlen erdő osztályozási módszerekkel készített felszínborítási térképek pontosságának összehasonlító elemzése - In: Molnár, Vanda Éva (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában X.: Theory meets practice in GIS Debrecen, Magyarország : Debreceni Egyetemi Kiadó (2019) 428 p. pp. 91-100. , 10 p.*

C, Land cover mapping for Hungary using object based image analysis

The use of an object-based image analysis (OBIA) method has recently become quite common for classifying high-resolution remote-sensed images. However, despite OBIA's segmentation being equally useful for analysing medium-resolution images, it is not used for them as often. This part of the research project aimed to analyse the effect of landscape metrics that have not yet been used in image classification to provide additional information for land cover mapping to improve the thematic accuracy of satellite image-based land cover mapping. To this end, multispectral satellite images taken by Landsat 8 Operational Land Imager (OLI) and Sentinel-2 Multispectral Instrument (MSI) during three different seasons in 2017 were analysed. The images were segmented, and based on these segments, four patch-level landscape metrics (mean patch size, total edge, mean shape index and fractal dimension) were calculated. A random forest classifier was applied for classification, and the Coordination of Information on the Environment Land Cover (CLC) 2018 database was used as reference data. According to the results, landscape metrics both with and without segmentation can significantly improve the overall accuracy of the classification over classification based on spectral values. The highest overall accuracy was achieved using all data (i.e., spectral values, segmentation, and metrics).

Random points were generated (4000 points per class) as the training data. The training and test datasets were transformed to delimited text format, after which classification was performed using the WEKA software. The random forest classification method was chosen, which is an assembly classifier designed with decision trees. According to this study's findings, the aforementioned indices could increase the prediction accuracy. Moreover, these landscape metrics improved the classification process even without the segmentation layer upon which the calculations were based. Moreover, differences in minimal segment size were found to increase the classification accuracy in different fragmented areas to differing degrees. In both study areas, it was observed that the effect of the landscape metrics on classification accuracy was increased along with the minimal segment size. The best overall accuracies (Marosszög: 91.39% with a kappa of 0.81; Gödöllői-hills: 76.14% with a kappa of 0.7) were obtained using all data (spectral bands, segmentation layer and all landscape metrics) with the 10-ha (Marosszög) and 25-ha (Gödöllői-hills) minimal segment sizes.

Detailed results were published in*: Gudmann, András; Csikós, Nándor ; Szilassi, Péter ; Mucsi, László Improvement in Satellite Image-Based Land Cover Classification with Landscape Metrics, REMOTE SENSING 12 : 21 p. 3580 , 18 p. (2020) D1-Q1*

D, Image analysis of medium spatial resolution hyperspectral satellite images (DESIS) 2021 –first application in Hungary

DESIS images can be obtained free of charge by submitting a scientific and use intent to DLR. Thanks to the free availability of the data and the characteristics of the instrument, DESIS images can be used for a variety of purposes, such as: medium- and long-term environmental monitoring in mining areas, vegetation monitoring, soil degradation measurements, etc.. Our research sample project will run from May to the end of October 2021 and we have mainly adapted the acquisition window to the phenological phase of spring-sown maize and sunflower. No yield data are available for these crops at this stage of the project, so the available records were tested on plots sown with the Avenue variety of winter wheat. The farm in Mezőhegyes is located in the south-eastern part of Hungary, close to the Hungarian-Romanian border

The spectral reflectance curve of the different yield areas and patches, determined from DESIS images, clearly shows that the reflectance values of the different yield areas do not differ significantly within the visible light range. However, in the near infrared bands after the red edge (from 135 to 744 nm), the spectral properties of the different yield areas also differ. In these bands, higher yielding areas have higher spectral values, while lower yielding pixels have lower values, which is associated with higher chlorophyll content. The differences measured in the near infrared bands provide a reliable basis for yield estimation. The regression model used for yield estimation was successful in estimating expected yields at pixel and plot level. The average error values of the regression models for the 08 June and 21 June images are nearly identical (0.98; 0.94), with the best average error value for the 16 June image

By observing the plot-level statistical values (minimum, maximum, mean, median, standard deviation, range) of the values estimated by the regression models and the differences between the measured yields, it can be concluded that all three models gave good estimates. The difference between the mean yield values at all three dates is less than 0.2 t/ha and the difference based on the 16 June recording is only 0.01 t/ha (10 kg/ha).

Detailed results were published in *: Gudmann, András ; Nizom, Farmonov ; Bónus, Krisztián ; Mucsi, László DESIS hiperspektrális űrfelvételek első magyarországi alkalmazása terméshozam becslésre*
In: Molnár, Vanda Éva (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában XII.: Theory meets practice in GIS, Debrecen, Magyarország : Debreceni Egyetemi Kiadó (2021) 360 p. pp. 105-111. , 7 p.

E, Multi-temporal image analysis on agrarian study are using random forest and support vector machine algorithms

In this part of the research project, the two most popular classification techniques the RF and SVM algorithms were used for supervised classification methods to map over agricultural fields during the summer of 2020. The averaged producer's accuracy (PA) and user's accuracy (UA) per class category, kappa and overall accuracy (OA) of SI, SB, and SI-SB-fAPAR which were obtained from both RF and SVM classification. According to the accuracy assessment, OA of the SB and SB-SI-fAPAR datasets values showed the highest accuracy from 97.8% to 98.7% to find affected corns for both RF and SVM classification methods. Furthermore, based on the input image lists, SVM and RF had not the highest differences between SB, SI and SB-SI-fAPAR. With the RF classification, Integration of SB-SI and fAPAR showed a slight improvement compared with SI-fAPAR feature datasets. Figure 5 demonstrates RF generated classification map by utilizing SB-SI-fAPAR datasets from Sentinel 2 images. For SVM classification, differences for all classification schemes percentage of overall accuracy were not big. However, a bit of an increase in the ratio of overall accuracy when SB and SB-SI-fAPAR were added to the classification and reported 95.96% and 96.37%. Figure 6 presents SVM generated map by using SB-SI-fAPAR datasets from S-2A (optical) images

The research evaluated the integration of multi-temporal S-2 images, spectral indices and fAPAR acquired in the maize growing season, which could detect affected corns from healthy corns by employing two robust RF and SVM machine learning algorithms in Mezőhegyes, South-Eastern Hungary. We identified affected hybrid corn which was caused by overwatering in Spring during the dry season and an extremely high precipitation rate in July. This information is crucial and lacks in our study area. Thereby, overall crop yield decreased to 10% as a result of these issues. Our result demonstrated that the highest OA of classification was obtained by RF using SB and SB-SI-fAPAR datasets. This approach achieved OA 97.35% (kappa 0.93) and 96.72% (kappa 0.93) respectively. However, SB-SI-fAPAR datasets showed a great AO rate for both SVM and RF classification to identify affected crops. Our research results showed that it is possible to identify affected corns and map crop fields based entirely on free and open-source satellite imagery at relative small agricultural fields since it is time-saving and cost-effective.

***Detailed results were published in** : Khilola, Amankulova ; Nizom, Farmonov ; Gudmann, András Viktor ; Krisztián, Bónus ; László, Mucsi Investigation the reason of affected Hybrid Corn in Agricultural Fields by Using Multi-Temporal Sentinel-2 Images in Mezőhegyes, South-Eastern Hungary In: Molnár, Vanda Éva (szerk.) Az elmélet és a gyakorlat találkozása a térinformatikában XII.: Theory meets practice in GIS Debrecen, Magyarország : Debreceni Egyetemi Kiadó (2021) 360 p. pp. 25-34. , 10 p.*

F, Methodology development for Land Cover Map and Land Use Mapping with a Combined Pixel-Based and Object-Based Approach Using Multi-Temporal Landsat Data, a Random Forest Classifier, and Decision Rules

This part of the research projectstudy was conducted to generate such maps in Binh Duong province, Vietnam, using a novel combination of pixel-based and object-based classification techniques and geographic information system (GIS) analysis on multi-temporal Landsat images. Firstly, the connection between land cover and land use was identified; thereafter, the land cover map and land use function regions were extracted with a random forest classifier. Finally, a land use map was generated by combining the land cover map and the land use function regions in a set of decision rules. The results showed that

land cover and land use were linked by spectral, spatial, and temporal characteristics, and this helped effectively convert the land cover map into a land use map. The final land cover map attained an overall accuracy (OA) = 93.86%, with producer's accuracy (PA) and user's accuracy (UA) of its classes ranging from 73.91% to 100%. Meanwhile, the final land use map achieved OA = 93.45%, and the UA and PA ranged from 84% to 100%. The study demonstrated that it is possible to create high-accuracy maps based entirely on free multi-temporal satellite imagery that promote the reproducibility and proactivity of the research as well as cost-efficiency and time savings.
With our proposed procedure, the maps achieved high overall accuracy. It shows the potential of using a combination of pixel-based and object-based classification techniques and GIS techniques on free multi-temporal satellite images to effectively extract and translate a land cover map into a land use map. Many studies have also shown effectiveness in using a combination of these methods in the area of land cover and land use mapping however, the difference in this study is the production of a land cover map and land use map separately, which these previous studies have not done.
As the final land use map in our research was generated by combining the land cover map and the land use function regions in a set of decision rules, its accuracy depended on that of the combined components. In land cover mapping, in our study, the higher the overall accuracy, the higher accuracy within classes and the low total disagreement of the final land cover map have shown a certain efficiency when using multi-temporal images in a pixel-based classification compared to using single-date images.

***Detailed results were published in**: Bui, Dang Hung; Mucsi, László From Land Cover Map to Land Use Map: A Combined Pixel-Based and Object-Based Approach Using Multi-Temporal Landsat Data, a Random Forest Classifier, and Decision Rules, REMOTE SENSING 13 : 9 Paper: 1700 , 24 p. (2021) Q1*

G, Data fusion of Sentinel 1 and Sentinel 2 data for land cover mapping –Methodology development

These study results demonstrate that fusion data from various sources at the three fusion levels can improve accuracy in land cover mapping. In these studies, various fusion techniques were used, ranging from simple to very complex methods. When fusion of S-1 and S-2 data based on Dempster-Shafer (D-S) theory at the decision level yielded better mapping results compared to others. It comes from the advantages of the D-S theory-based technique in reducing the impact of noise data and feature selection in land cover classification. The most obvious improvement was found in the classes of barren land and built up. As a result, the datasets fused at the decision level increased the OA by a range of 0.75% to 2.07% compared to the S-2 datasets. The fusion of S-1 and S-2 data with their derived textures and indices at the decision level using D-S theory brought the best results in this study, achieving an OA and Kappa coefficient of 92.67% and 0.91, respectively.

Moreover, the integration of SAR and optical products using the layer-stacking technique at the pixel level did not give more power to the classification process. It reduced the accuracy of the mapping result by 4.88% to 6.58% compared to that of the optical datasets. These findings may be influenced by the processing and selection of features, fusion technique, and classifier. Further studies on this issue are needed. Furthermore, the inclusion of GLCM textures and spectral indices in the datasets helped improve the mapping results. However, while the effectiveness of the textures is clear, the contribution of the indices needs to be studied further.

In general, the results of this study show that using the D-S fusion method for high-accuracy mapping in other urbanized areas holds great potential. This study represents an initial step, and it paves the way for further research on land cover mapping using additional available data from the active and passive sensors for performance improvement.

The detailed result were published in*: Bui, Dang Hung; Mucsi, László Comparison of layer-stacking and Dempster-Shafer theory-based methods using Sentinel-1 and Sentinel-2 data fusion in urban land cover mapping GEO-SPATIAL INFORMATION SCIENCE (2022) (Q1)*

H, Land use/land cover classification and change detection for Hungary with spatial information

This study aimed to test a possible classification method, to create large-scale LULC maps on 4 dates. Furthermore, we tested these predicted maps in a large-scale, pixel-based change analysis. In the presented method, we used HGBCT classifier to create maps, based on Landsat imagery with landscape metrics and texture data. By considering the predicted map, land use change analysis was performed, which matched the CLC change layers. The land use change mapping method presented meets the preliminary expectations.

The results of our presented method show a significant progress. In our classification scenarios, we used more 20 classes with low separability values and big mono-temporal datasets. Despite the use of this difficult scenario, the HGBCT classifier achieved satisfactory results (83.35%–92.63%). The other model-performance metrics, such as log loss and kappa, show corresponding results. Furthermore, the UAs and PAs of the models were high, only the class 211–Non-irrigated arable land, which has the largest extent, obtains a low PA. This low PA can occur for reasons such as training point randomness, class mixture caused by the CLC's MMU, and combination of the pixel-based mapping unit and an accurate model. As most of these reasons cannot be eliminated, the problem could occur in every setting of the method. According to these results, our predicted map was suitable for change mapping. Land use change mapping is difficult to evaluate because the relevant databases would have been created using different methodologies (nomenclatures, MMUs, and reference data). In this study, land use change maps were created at the pixel level (900 m^2) using the CLC nomenclature. The CLC change layers and the predicted change maps were comparable. However, the two datasets has different time-range, which slightly distorts the comparison. Besides, the MMUs of the two datasets were different (CLC change layers had 5-ha changes). Thus, the predicted maps had more details and larger values in certain categories than the CLC change layers, and their range of statistics was much larger. This difference could be attributed to the magnitudes of the changes, which were much bigger in the predicted maps (1.802%/year – 6.414%/year) than in the CLC change layers (0.259%/year – 0.516%/year). The MMU had a large impact on the analysis because the majority of the changes occurred in small and separate areas; thus, they are not shown in CLC change layers and may not be permanent, which is true especially in agriculture in which are many small parcels. These temporary, likely parcel-level, agricultural changes are responsible for approximately half of the changes. Overall, the predicted maps contain both permanent and temporary changes, and thus provide a detailed picture of the land use changes in the country. Thus, the maps can be used in detailed surveys.

Detailed results will be published in: Gudmann, András; Mucsi, László: Land use/land cover classification and change detection for Hungary with spatial information, in Geographica Pannonica (the paper was submitted to the journal at 10th May 2022, and after the first review it was accepted after minor revision)