Segmentation and sequential classification of a synthetized image composed of spatial environmental data for the compilation of a soil type map

Gábor Illés[1] illesg@erti.hu

Katalin Takács[2] takacs.katalin@rissac.hu

Annamária Laborczi[2] laborczi@rissac.hu

László Pásztor[2] pasztor@rissac.hu

[1]Forest Research Institute, National Agricultural and Innovation Centre
Várkerület 30/A, 9600 Sárvár, Hungary

[2]Institute for Soil Sciences and Agricultural Chemistry,
Centre for Agricultural Research, Hungarian Academy of Sciences
Herman Ottó út 15, 1022 Budapest, Hungary

Abstract

A unified, national soil type map with spatially consistent predictive capabilities was compiled applying traditional and newly tested Digital Soil Mapping classification methods: segmentation of a synthesized image consisting of predictor variables and multi-phase, sequential classification by Classification and Regression Trees, Random Forests and Artificial Neural Networks. Object based classification using spatial-thematic segments was applied to define mapping objects. Classifications were carried out on two levels to achieve better results. Performance of classifiers was continuously assessed and applied for the identification of best performing predictions, which were combined for the production of the final map.

Keywords: digital soil mapping, soil type map, segmentation, sequential classification

1 Introduction

In our work we aimed at compiling a unified, nationwide soil type map – which is consistent regardless of land use - with harmonized legend and spatially consistent predictive capabilities and accuracy for soils of croplands and forests. Tasks of national spatial planning and basement of agricultural adaptation strategies have increasingly required the availability of such a map product.

The traditional Hungarian classification system (based on the genetic approach of Dokuchaev, 1883) considers soil forming as a genetic process (pedogenesis), in which geographic conditions are substantial (Stefanovits, 1972; Szabolcs, 1966; Várallyay et al., 1979). The system sorts soils into main soil types such as skeletal soils, lithomorphic soils, brown forest soils, chernozems, salt-affected soils, meadow soils, alluvial and deluvial soils, and peat soils.

Traditionally in Hungary the soil cover under agricultural and forestry management is typically characterized independently and just approximately identically. Soil data collection is carried out and the databases of soil features are managed irrespectively. As a consequence, nationwide soil maps cannot be considered homogeneously predictive for soils of croplands and forests, plains and hilly/mountainous regions. In order to compile a national soil type map with harmonized legend as well as with spatially relatively homogeneous predictive power and accuracy, the authors unified the resources of forestry and agricultural areas.

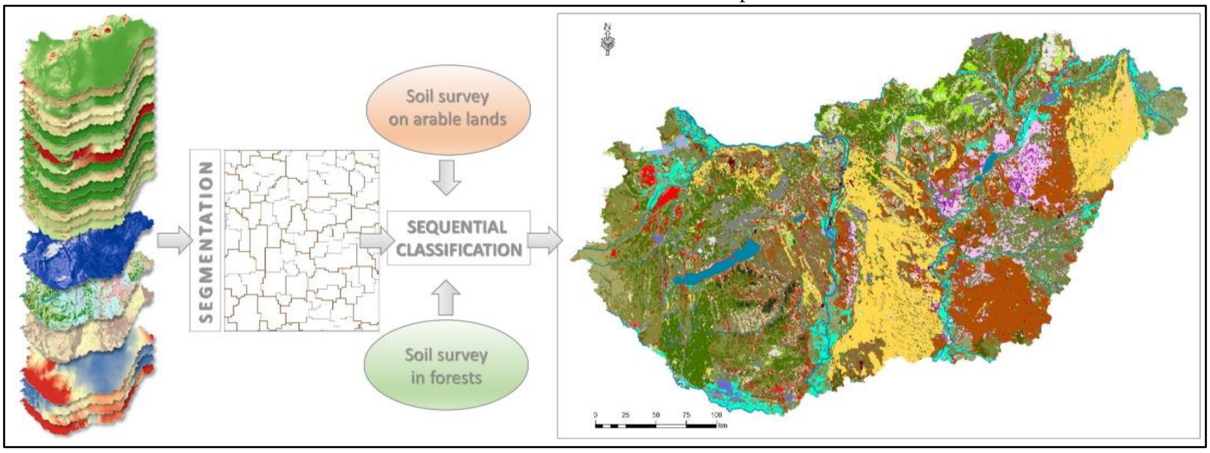
2 Materials and methods

Soil profile data originating from the two sources (agriculture and forestry) were cleaned up and harmonized according to a common soil type classification. For agricultural land, two independent datasets were involved in our study: the Hungarian Soil Information and Monitoring System (SIMS, 1995) and the Hungarian Detailed Soil Hydrophysical Database (MARTHA - Makó et al., 2010). The applied forestry database consists of data points of forest compartments that were subject to site and soil surveys assigned to forest management plans (State Forest Service - ÁESZ, 2004). For the mapping procedure a harmonised soil dataset was set up consisting of the detailed description of almost 60,000 soil profiles, describing 41 representative soil-types with spatial reference.

A corresponding dataset of spatially exhaustive, ancillary, environmental variables – including legacy soil data– was established covering the whole area of the country. Topographic features were taken into consideration as Digital Elevation Model (EU-DEM, 2015) and its numerous derivatives. Lithology categories (Gyalog & Síkhegyi, 2005), level of groundwater, climatic properties, land use (CORINE Land Cover - Büttner et al., 2004), vegetation (remotely sensed MODIS images) as well as chemical and physical property maps of legacy soil data (Pásztor et al., 2012) are also represented as different variables.

Various methods were tested for the compilation of the target map. First of all, a segmentation process of a synthesized image

Figure 1: Compilation process of the nationwide soil type map: segmentation of environmental co-variables, sequential classifications and the result map.



consisting of the predictor variables was carried out to delineate homogeneous spatial entities that were used later as objects for classifications. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse (Benz et al., 2004). The purpose of the introduction of object based classification was to delineate areas composed of a set of similar locations featured by the applied environmental co-variables, which are different from surrounding areas. The assumption is that these areas are exposed to similar soil forming processes and consequently can be considered as individual soil bodies.

Then we elaborated a two-level, multi-step, sequential classification method (Figure 1.). On the first level main soil type groups were classified and predicted in multiple steps. Soil types were targeted only in a second phase within the areas formerly attributed with their respective (containing) main soil type groups. On both levels in each step multiple classification models were applied. Six models use CART (Classification and Regression Trees), five are based on RF (Random Forests) and one is based on ANN (Artificial Neural Networks) classification. The models with identical classification tools differ either in the inherent parameters of the method or on the segmentation level, on which they are applied.

Two types of validation were carried out. On the one hand, profiles were split into learning and test sets, 20 % of the profiles were left out for validation. The validation provided by the test sets was used for the estimation of classification accuracy, which was carried out on both levels and in all steps. The best performing classifier was identified in each phase for each soil type. Besides the former data driven validation a trial was done for a certain external validation. A set of digitally processed, large scale legacy soil type maps were also available for a non-systematic comparison. The predicted raster map was compared to the legacy soil type maps on a pixel by pixel level.

3 Results and discussion

The main final product of our work is a newly compiled nationwide soil type map with harmonized legend and spatially consistent predictivity. Both the thematic and spatial representation of hilly/mountainous areas is much more detailed than on former national soil maps. Nevertheless, the mosaic-like pattern of lowlands is retained and the large scale geographical structural elements are very well reflected. The newly compiled product was compared to the two earlier nationwide soil type maps on a pixel by pixel level. For this purpose, they were rasterized with 1 ha resolution to the applied grid system. The similarity was measured by overall accuracy and overall kappa (Cohen, 1960; Rossiter, 2014). Both measures showed that the maps are rather dissimilar, that is in spite of the overall resemblance, locally they contain divergent local predictions.

Evaluation of the results showed that the object based, multi-level mapping approach performs significantly better than the simple classification techniques. A combination of best performing classifiers, when each classifier's vote on the same object is weighted according to its confidence in the voted class, led to the final product: a unified, national, soil type map with spatially consistent predictive capabilities.

The importance of the newly prepared map could actually be evaluated from the practical point of view. This is the first countrywide soil type map that unifies expert inputs and databases from both the agricultural farmlands and forested areas. As a consequence, this map can be equally used for agricultural or forestry oriented purposes providing interoperability between the sectors. Because of the robustness and huge data background, the map is suitable to be involved in nationwide spatial and land use management planning.

We do not consider our map as an ultimate product, since it could and should be refined and improved in a number of ways. The workflow inherently makes it possible to keep the map easily updated or refined if new qualified data becomes available.

4 Acknowledgement

Our work was supported by the Hungarian National Scientific Research Foundation (OTKA) under Grant K105167 and AGRARKLÍMA.2 VKSZ_12-1-2013-0034.

References

Büttner, G. et al., 2004. National land cover database at scale 1:50000 in Hungary. EARSeL eProceedings, 3(3), pp.323–330.

Dokuchaev, V. V., 1883. The Russian Chernozem [In Russian], St. Petersburg, Russia.

EU-DEM, 2015. Digital Elevation Model over Europe. http://www.eea.europa.eu/data-and-maps/data/eu-dem.

Gyalog, L. & Síkhegyi, F., 2005. Geological Map of Hungary, 1:100.000. Geological Institute of Hungary, Budapest [In Hungarian]. http://loczy.mfgi.hu/fdt100/.

Makó, A. et al., 2010. Introduction of the Hungarian Detailed Soil Hydrophysical Database (MARTHA) and its use to test external pedotransfer functions. Agrokémia és Talajtan, 59(1), pp. 29–38.

Pásztor, L. et al., 2012. Compilation of 1 : 50,000 scale digital soil maps for Hungary based on the digital Kreybig soil information system. Journal of Maps, 8(3), pp.215–219.

SIMS, 1995. *Hungarian Soil Information and Monitoring System. Methodology.*, Budapest: Ministry of Agriculture, Plant Protecting and Agro-ecological Department.

State Forest Service - ÁESZ, 2004. Guidelines for forest management planning. pp.21–55.

Stefanovits, P., 1972. *Talajtan*, Budapest: Mezőgazda Kiadó.

Szabolcs, I., 1966. *A genetikus üzemi talajtérképezés módszerkönyve* I. Szabolcs, ed., Budapest: OMMI.

Várallyay, G. et al., 1979. Map of Soil Factors Determining the Agro-Ecological Potential of Hungary (1:100000). I. [In Hungarian]. *Agrokémia és Talajtan*, 28(3–4), pp. 363–384.