Szent István University
Doctoral School of Environmental Sciences

Digital mapping of categorical soil properties

Thesis of Ph.D. dissertation
Laborczi Annamária

Gödöllő
2018

Ph.D. School: Doctoral School of Environmental Sciences

Discipline: Soil Science (Digital Soil Mapping)

Head of School: Csákiné Dr. Michéli Erika, D.Sc.
Professor
Szent István University
Faculty of Agricultural and Environmental Sciences
Institute of Environmental Sciences
Department of Soil Science and Agrochemistry

Scientific supervisor: Dr. Pásztor László
senior research fellow,
head of department
Institute for Soil Sciences and Agricultural Chemistry,
Centre for Agricultural Research,
Hungarian Academy of Sciences
Department of Soil Mapping and Environmental Informatics

…………………………	………………………
Approval of Head of School	Approval of Scientific Supervisor

1. INTRODUCTION AND OBJECTIVES

Demands on soil related information have been significant and are still increasing worldwide as well as in Hungary. In most cases, problems cannot be solved directly by the available data, only by some conversions and inferences. Traditional field survey and data collection is too expensive to be a realistic possibility. Accordingly, it is indispensable to rely on the available soil related and other environmental information in order to fulfil the demands. Thanks to the technical development, possibilities of soil mapping have been increasing. Spatial information related to the soil forming processes are widely accessible for lower and lower costs, in higher and higher spatial resolution. Tools of digital soil mapping (DSM) provide solutions for refining, inferring and involving environmental co-variables in order to supply proper spatial soil information.

Digital Soil Mapping can be defined as the creation and population of spatial soil information systems by numerical models inferring the spatial and temporal variations of soil types and soil properties from soil observation and knowledge and from related environmental variables (Lagacherie & McBratney 2007). The basis of DSM, the so-called SCORPAN model (**s**oil; **c**limate; **o**rganisms, vegetation, land cover; **r**, relief; **p**arent material; **a**ge; spatial positio**n**) formalizes the relation between the soil characteristic and the environmental co-variables (McBratney et al. 2003). Researchers are examining relations between spatial soil information and environmental factors in order to exploit more effectively the technical possibilities of mapping (Lagacherie & McBratney 2007; Hartemink et al. 2008; Boettinger et al. 2010; Minasny et al. 2012; Arrouays et al. 2014; Zhang et al. 2016).

Based on the above mentioned background, the main objectives of my research were the followings:

1. Mapping of representatively chosen, categorical soil variables by goal-oriented digital soil mapping methods. I intended to create such categorical maps, which did not exist before, or not in detailed spatial resolution. If possible, the maps should reflect the national and international requirements, accordingly they can provide further environmental models and support researches with spatial information.

2. Validation of the result maps, evaluation of their reliability on local and global level.

3. Goal-oriented mapping by different approaches; comparison, and evaluation of the result maps:
- direct prediction of the target variable vs. aggregated one from partial results;
- result maps originated from different digital soil mapping methods;
- predictions by the help of different environmental co-variables.

4. Goal-specific development of digital soil mapping methods, in order to refine prediction of categorical spatial variables.

5. In some interdisciplinary examinations and tasks, elaboration of thematic and spatial correlation between ecological soil-related categories and available spatial soil information.

2. MATERIALS AND METHODS

In my research, I involved the following spatial soil information as **reference data**, as **soil factor** of the SCORPAN model, and **validation** data:

- Digital Kreybig Soil Information System (DKSIS, Pásztor et al., 2012);
- AGROTOPO database (Várallyay et al. 1979; Várallyay et al. 1980);
- MÉM NAK genetic soil type map (Kocsis et al. 2015);
- Hungarian Soil Information and Monitoring System (SIMS, TIM 1995);
- Hungarian Detailed Soil Hydrophysical Database (MARTHA, Makó et al., 2010).

As **environmental co-variables**, in the form of predictor variables of the SCORPAN model, I used the following datasets:

Climate: layers of the Hungarian Meteorological Service, namely average annual precipitation, average annual temperature, average annual evaporation and evapotranspiration (Szentimrey & Bihari 2007);

Land cover, vegetation: MODIS satellite images RED, NIR (near infrared) bands, NDVI (Normalized Difference Vegetation Index) from two dates (16.03.2012. and 07.09.2013.), furthermore NDVI images from 16 day periods of two dates (03.2012. and 09.2013.) (NASA LP DAAC 2015).
Corine Land Cover (CLC50) database (Büttner et al. 2004).

Topography: In the first period of my investigations I applied ASTER GDEM (2000) digital elevation model and its derivatives (marked by * in the list below, and Profile Curvature). In later mapping processes EU-DEM (2015) played the role with the following derivatives: Slope*, Aspect*, General Curvature*, Vertical Distance to Channel Network, SAGA Wetness Index, Diurnal Anisotropic Heating, Real Surface Area, Channel Network Base Level, Multiresolution Index of Valley Bottom Flatness (MRVBF), Multiresolution Index of Ridge Top Flatness (MRRTF), LS factor (slope length and slope factor of the Universal Soil Loss Equation, USLE), Mass Balance Index, Stream Power Index, Topographic Position Index, Topographic Wetness Index*.

Lithology: Geological Map of Hungary 1:100.000 (FDT100; Gyalog & Síkhegyi 2005), Groundwater-level Map of Hungary 1:100.000 (Pentelényi & Scharek 2006).

Nationwide particle size fraction- and texture class mapping

Direct prediction of texture class maps, as well as derivation of them from particle size fraction maps

A possible method to create texture class map is based on using reference data in categorized form. In this case, spatial inference targets directly the texture classes. The other solution is making prediction of clay-, silt-, and sand content, and the texture class map comes unambiguously from the combination of them. According to the two approaches, I made nationwide predictions for the internationally most widespread USDA (United States Department of Agriculture) texture classes for the 0-5 cm layer.

1. Reference data (SIMS particle size distribution) were categorized into the texture classes. Spatial inference was carried out by Classification and Regression Trees (CART) in 150 m spatial resolution. I took into consideration the taxonomic distances of the categories, which were computed as the distance between polygons within the USDA textural triangle (Figure 1).
2. In same spatial resolution, I made independent predictions for clay-, silt-, and sand content by regression kriging (RK). I used two different groups of environmental co-variables. I made corrections in the values of the result maps because their sum have to be 100 %. The derived texture class maps were compared to each other, and to the map directly predicted by CART method (Figure 2).

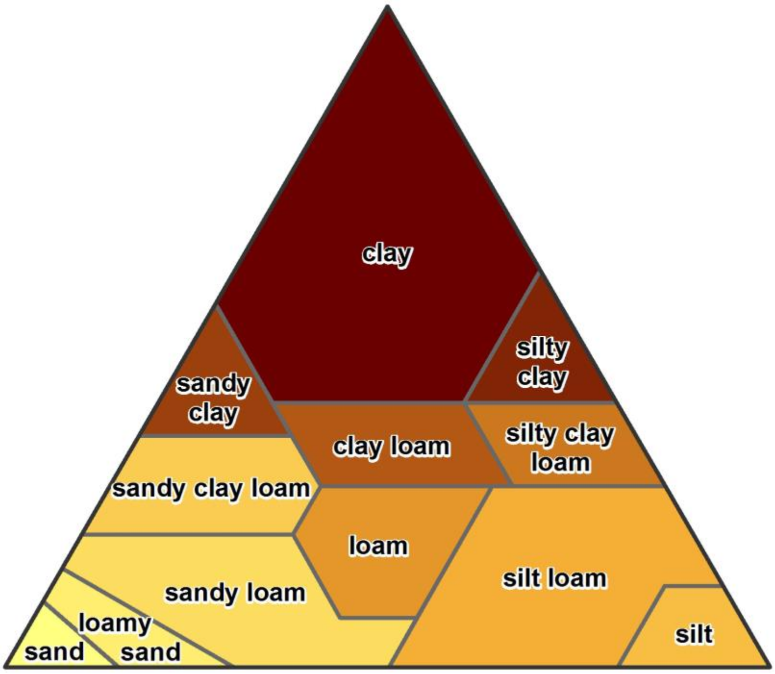

Figure 1. USDA textural triangle

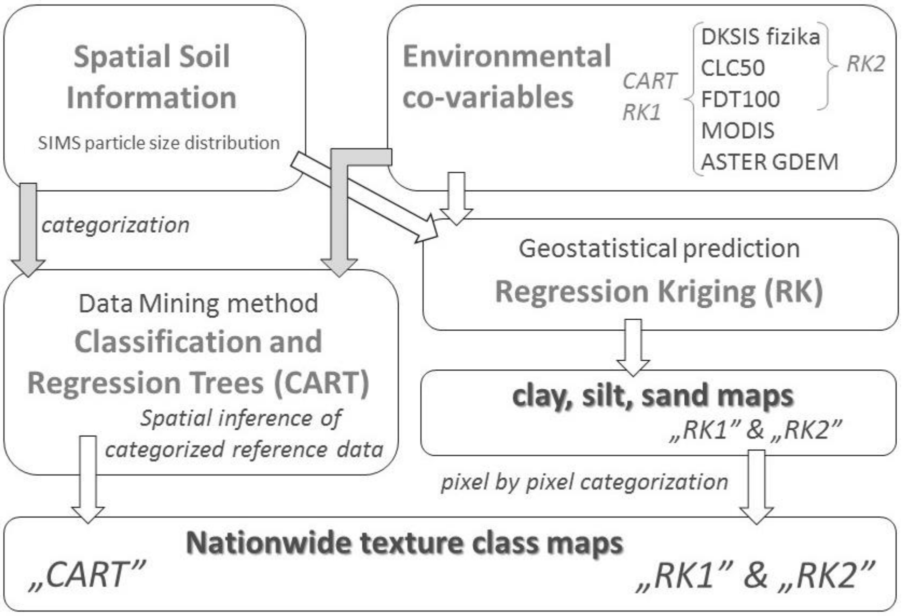

Figure 2. Process of direct prediction of texture class maps (by CART method), and deriving them from the results of regression kriging (RK)

Particle size fraction maps predicted by composite kriging, and texture class maps derived from them

Based on the previously described experiences, I made predictions by composite kriging for six standard depths of texture fraction maps, in 100 m spatial resolution. The depths (0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm, 100-200 cm) and the resolution was defined according to the GlobalSoilMap specifications. The goal of the international initiative is to compile a new, global soil map database. Mapping of clay-, silt-, and sand content are also part of the requirements (Arrouays et al. 2014, 2015). The specifications define only the parameters of the result maps and their reliability. The mapper always have to determine the proper methods, accommodating the local conditions of the country in order to reach the goals appropriately.

By composite kriging, target variables keep their complementarity. I made a so-called ALR transformation on the reference data, hence two variables contained the information of the three complementary variables. Regression kriging was carried out on this two variables, therefore, sum of the three fractions remains 100 % in the results, after back-transformation.

I derived texture class maps from the three fraction map results predicted by composite kriging.

Comparison of soil texture maps synthetized from standard depth layers with directly compiled products

Some agro-meteorological and hydrological models require spatial information inputs representing topsoil (0-30 cm) features. The optimal solution would be a goal-specifically compiled topsoil map. However, if the original source data are not available, the modelling expert can use the freely available databases with their specified depth layers, such as the internationally acknowledged GlobalSoilMap database. In these databases, there is not 0-30 cm soil layer, it is represented by three standard layers (0-5 cm, 5-15 cm, and 15-30 cm). In this case, the 0-30 cm map can be compiled by synthetizing the information featuring the three standard layers. Directly compiled and synthetized maps might differ, their comparison can establish their proper usage.

In order to investigate the above mentioned problem, I compiled the weighted average of the 0-5 cm, 5-15 cm, 15-30 cm clay, silt, sand maps predicted by composite regression kriging. I compared them with directly compiled maps of 0-30 cm layer.

By classification, a texture class map was also derived from the synthetized maps. It was compared to the texture class map directly compiled for 0-30 cm layer depth.

I compared the differences between the directly compiled and synthetized maps by the following tools. Histograms and scatter plots unfold the relationships of the clay and sand separate map-pairs. Spatial distribution of the differences was investigated by difference maps and their histograms.

Considering the pixel values of the directly compiled and the synthetized clay and sand content map as statistical population, I applied a non-parametric test for matched samples. Two tailed Wilcoxon-signed-rank test was used to evaluate significance of paired variables.

Disaggregation of thematic soil map

In order to increase the spatial resolution of AGROTOPO genetic soil type map by reclassification and formalization of the built-in soil-landscape model, I elaborated the following disaggregation method. Virtual sampling points were generated, and soil type was assigned to each site, according to the categories of the archive map. This data provided the dependent variable of the classification. Co-variables - assigned to the sites - constituted the set of independent variables. For spatial inference, random forest model was built by the set of training points. Multiple sampling leads more robust prediction of dominant patterns, therefore random forest models were run on 100 different sets of virtual sampling points. This resulted 100 different spatial predictions. The final result map came from the most frequently predicted class in each pixel, according to maximum likelihood rule. The applied method can be considered as a modification of random forest, because it originated from not only one, but 100 different training data sets.

Elaboration of correlation between objects of spatial soil information and other interdisciplinary objects and categories

Application of the results of MÉTA programme (Landscape Ecological Vegetation Database & Map of Hungary) generated some demands on soil-related information. I intended to provide solutions of the problems through two approaches by correlation between the ecological objects as well as concepts, and spatial soil information.

Integrating spatial soil information into ecological mapping units

I integrated spatial soil information into the 35 ha hexagonal net of MÉTA. Integration was carried out on three levels, according to the spatial and thematic resolution of soil maps. From AGROTOPO, based on the occupied space within the objects, dominant soil properties were assigned to each hexagon. Give more details within the objects is not worthy because of the smaller spatial resolution of the database. However, in the case of DKSIS, the heterogeneity was also taken into consideration. Beside dominant (physical-, chemical characteristics, and landscape management soil types) properties, I computed the proportion of them within the objects. Number of categories within hexagons was also given as a further indicator of their heterogeneity. On the third level, I integrated texture maps into the hexagonal net. Minimum, maximum, and average clay-, silt-, and sand content were calculated within each hexagon. Furthermore, the dominant USDA textural class was determined.

Correlation of categories differently defined in soil science and ecology

The other approach emerge from the demand of providing certain sites with soil type information. The soil types were defined by ecologists, and were not identical to the classical definitions of soil science. Simple conversion was not enough, therefore mutual discussions were needed to elaborate a correlation method. I made predictions of most (and second most) likely occurring soil categories at centroid of selected MÉTA hexagons. The categories were sand, loess, (weakly, medium, strongly) salt affected soils, peaty soils, local peaty characteristics, forest soils, alluvial soils. The correlation to existing spatial soil information was based on DKSIS. I took into consideration the point and polygon component of the database. Direct correlation was possible in the case of soil patches. In the cases of soil observation data, I had to make spatial inference (indicator kriging). The investigations were made within the 50 m circles around the centroids. The dominant type within the circle was defined as the most probable category (maximum likelihood rule).

Data, methods, and measures for validation

Independent validation of soil texture class maps was carried out by 692 points of MARTHA database. Measured values in control points were compared to the predicted values of the result maps at the same location. Comparison was based on internationally recommended measures: ME (Mean Error), MAE (Mean Absolute Error), RMSE (Root Mean Square Error). RMSE is the most commonly used value in comparison of maps (Hengl 2009).

Comparison of categorical type predicted and measured data was characterized by Overall Accuracy (OA), Overall Kappa, Producers' Accuracy (PA), Users' Accuracy (UA) values (Cohen 1960; Rossiter 2014). I compiled a similarity matrix based on the taxonomical distance of textural classes. I calculated the weighted version of the indicators by the help of the similarity matrix.

To compare the directly compiled and the synthetized clay and sand content map, considering the pixel values as statistical population, I applied a non-parametric test for matched samples (two tailed Wilcoxon-signed-rank test). The indicators were applied not only in relation with independent data, but we also investigated the agreement of the two (directly compiled and synthetized) texture class maps.

Disaggregated AGROTOPO map was validated by genetic soil type of SIMS. The categories of the two databases are not identical, therefore I merged or eliminated some types.

I validated the results of spatial soil information correlated with ecological soil types by different soil maps: Loess and sand categories were compared to the 0-30 cm texture class map compiled by composite kriging. Salt affected categories were validated by the electrical conductivity (EC) map compiled in the frame of delineation of 'Areas with Natural Constraints' project (Takács et al. 2016). Validation of the forest-, alluvial-, and peaty soils was carried out applying the disaggregated AGROTOPO genetic soil type map.

3. RESULTS

Nationwide particle size fraction- and texture class mapping

Direct prediction of texture class maps, as well as derivation of them from particle size fraction maps

Result map of direct prediction of USDA texture classes, compiled by Classification and Regression Trees (CART), can be seen in Figure 3.

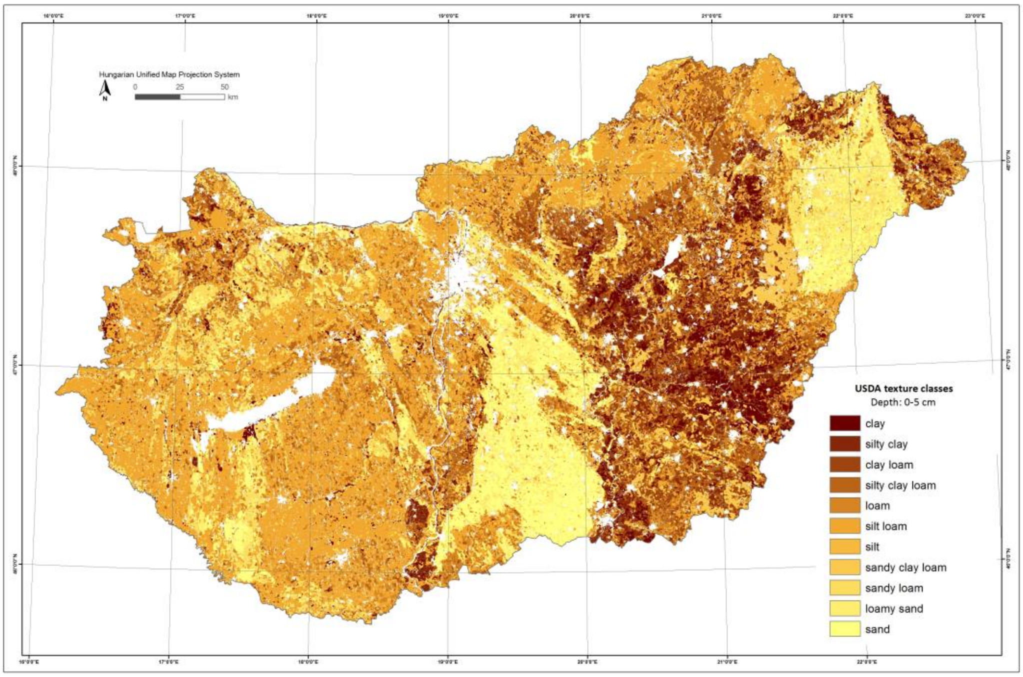

Figure 3. Result map of direct prediction of USDA texture classes of 0-5 cm layer depth, compiled by Classification and Regression Trees (CART); spatial resolution: 150 m

Maps of independent predictions for clay-, silt-, and sand content by regression kriging can be seen in Figure 4. ’RK1’ and ’RK2’ refer to the applied two sets of environmental co-variables. Spatial distribution information of reliability of the prediction is attached to the result maps.

From the two types of soil separate maps, by classification of the pixels into USDA categories, I derived two texture class maps (RK1, RK2).

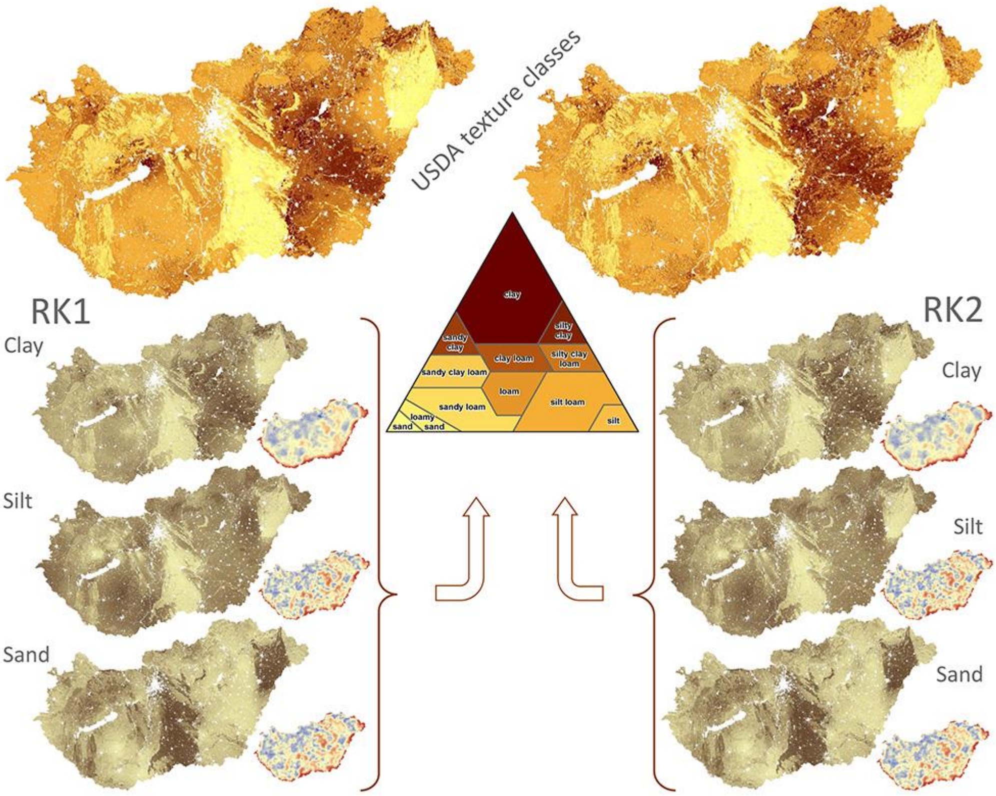

Figure 4. Result maps of independent predictions for clay-, silt-, and sand content by regression kriging (applying two sets of environmental co-variables - RK1, RK2) and the derived two texture class maps

The characteristics of the Hungarian landscapes are fairly well reflected in the result maps. In general, structures of CART map are coarser, while in RK maps the transitions are smoother, there are not any abrupt ’border lines’ between large areas with different texture class. This can be attributed to that regression kriging is a kind of spatial interpolation method, while CART is a classification method during which spatial inference is carried out disregard to spatial relations.

I compared RK1 and RK2 by the spatial distribution of the differences of their pixel values. Minor differences appear at Danube-Tisza Interfluve, and Nyírség. Greater differences can be found sporadically e.g. in Közép-Dráva-Valley, Mosoni-Plain, some areas of Sárrét, and River Tisza is outlined. However, taking the values into consideration, map pairs show only slight differences.

Comparison of texture class result maps show that the RK1 and RK2 texture class maps are identical in 79 % of the areas. Both RK1 and RK2 have 32 % identical area with the CART map (not in the same pixels). 27 % of the area have total agreement among the three result maps.

Based on independent validation by MARTHA database, User’s and Producer’s Accuracy values of RK maps show greater accuracy than the CART map for almost all of the texture classes. The two RK maps did not show significant differences. According to the Overall Accuracy, which is calculated from the correctly predicted- and the total number of validation points, 26 (CART), 33 (RK1), and 34 (RK2) percent of the points predicted the texture class accurately. Values of Overall Kappa, and weighted Overall Kappa (corrected by taxonomical distances) refer to slight and fair agreement, and in the case of RK result maps the values are greater than that of CART map. By MAE and RMSE values, calculated from the differences of particle size fraction maps and validation points, there is no significant differences between RK1 and RK2.

Based on accuracy assessment, the texture class maps derived from clay-, silt-, and sand content maps compiled by regression kriging, over performed the directly predicted texture class map created by Classification and Regression Trees. There were no significant differences between the results of regression kriging by the help of two sets of environmental co-variables.

Particle size fraction maps predicted by composite kriging, and texture class maps derived from them

Result maps for the six standard GlobalSoilMap layers, compiled by composite regression kriging can be seen in Figure 5. Their spatial resolution is 100 m. The characteristics of the Hungarian landscapes are fairly well reflected in the result maps. Based on the results of validation, the upper (0-5 cm), and the two deepest (60-100 cm and 100-200 cm) layers show greater error values. In the cases of deeper layers, the phenomenon can be explained by the decreasing predictive capacity of surface co-variables. Moreover, spline interpolation extrapolates in marginal layers.

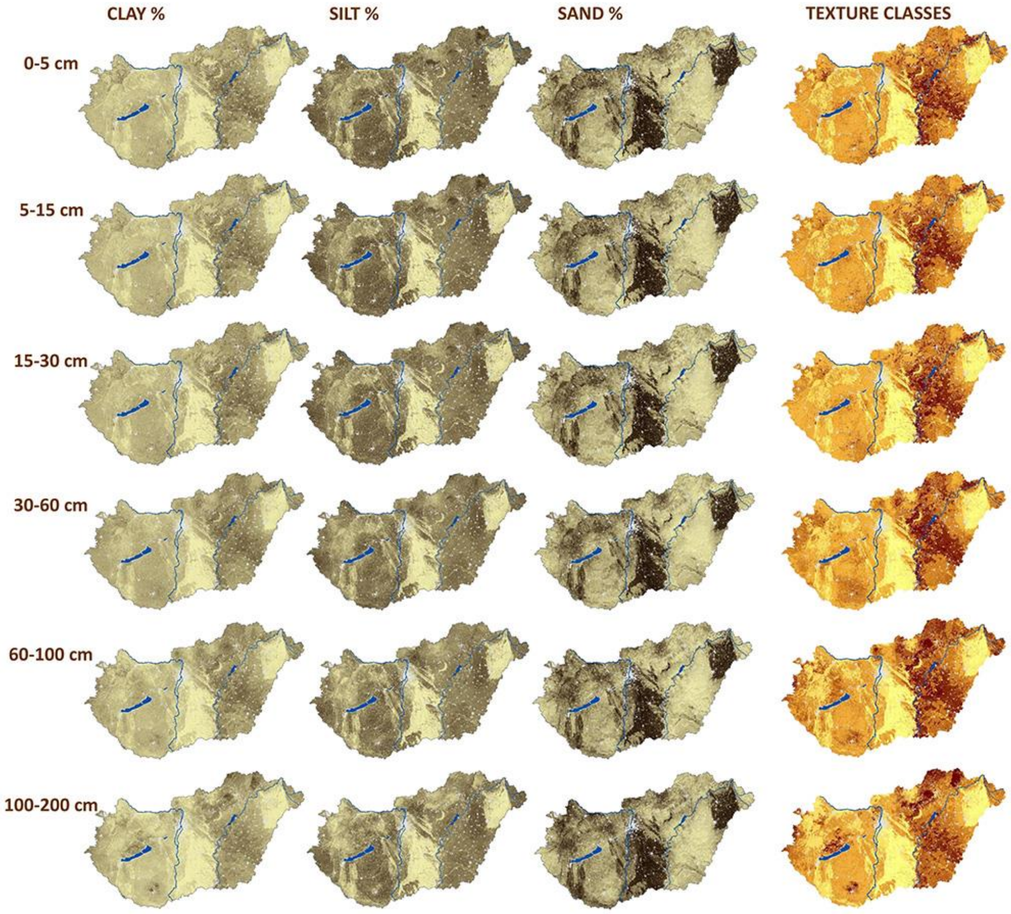

Figure 5. Nationwide clay-, silt-, and sand content maps compiled by composite regression kriging for standard layer depths, and derived USDA texture class maps

In order to compare the maps, I compiled difference maps of the consecutive layers. Toward deeper layers, the maps show greater differences.

According to the independent validation (Overall Accuracy, Overall Kappa, weighted Overall Accuracy, and weigthed Overall Kappa), texture class maps derived from particle size fraction maps show the tendency that the middle layers are characterized with greater accuracy.

Comparison of soil texture maps synthetized from standard depth layers with directly compiled products

I compiled synthetized maps of particle size distribution of 0-30 cm layer depth by weighted average of 0-5 cm, 5-15 cm, and 15-30 cm clay-, silt-, and sand maps (Figure 6).

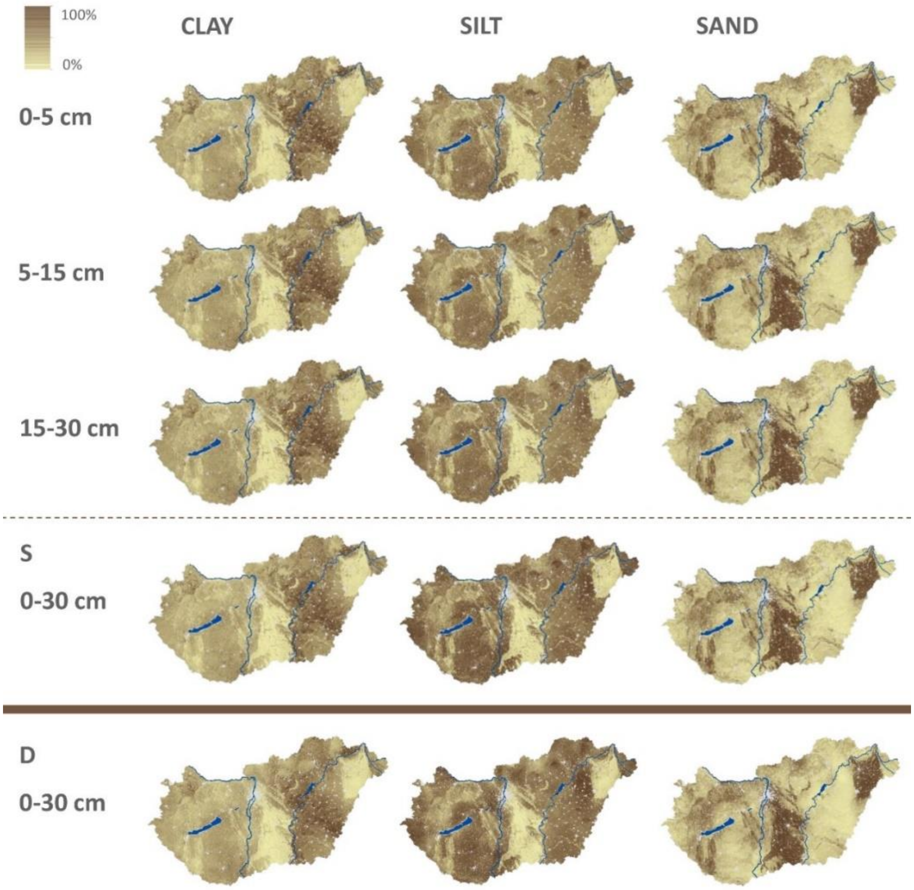

Figure 6. Soil separate maps synthetized from standard depths (S) and directly compiled maps (D)

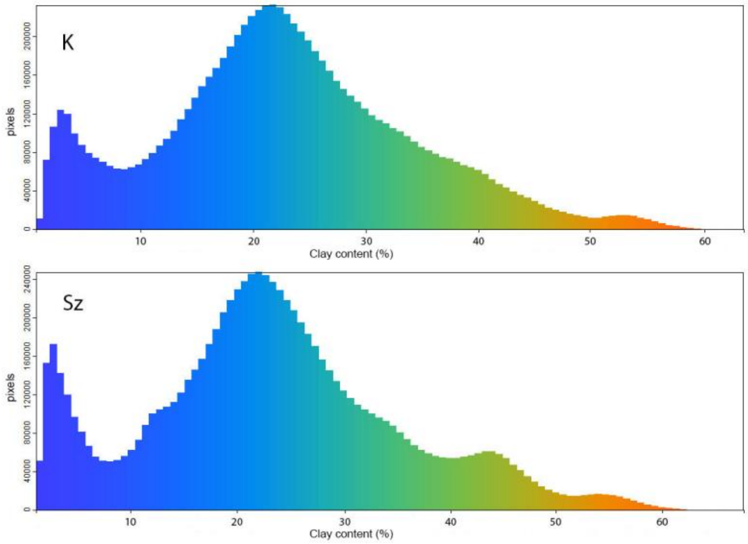

Figure 7. Histograms of the directly compiled (D) and the synthetized (S) clay content map

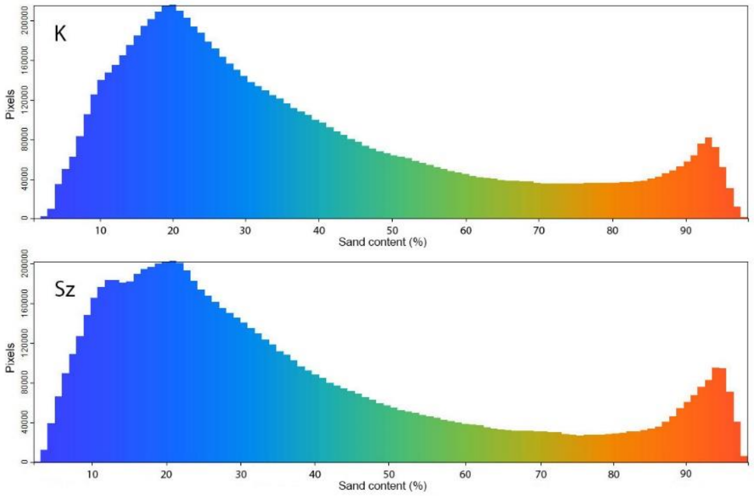

Figure 8. Histograms of the directly compiled (D) and the synthetized (S) sand content map

The histograms of the clay and sand maps show the differences of the map-pairs. In the cases of the sythetized maps, there is a steeper increase at the lower values of clay-, and sand content. There are additional peaks in the synthetized diagrams. The proportion of the areas with small and large clay-, and sand content was increased by synthetization (Figure 7 and 8).

According to the scatter plots of the synthetized and the directly compiled soil separate maps, they are strongly correlated in both cases of clay and sand. However, in certain locations, the values can differ in a wide range (Figure 9).

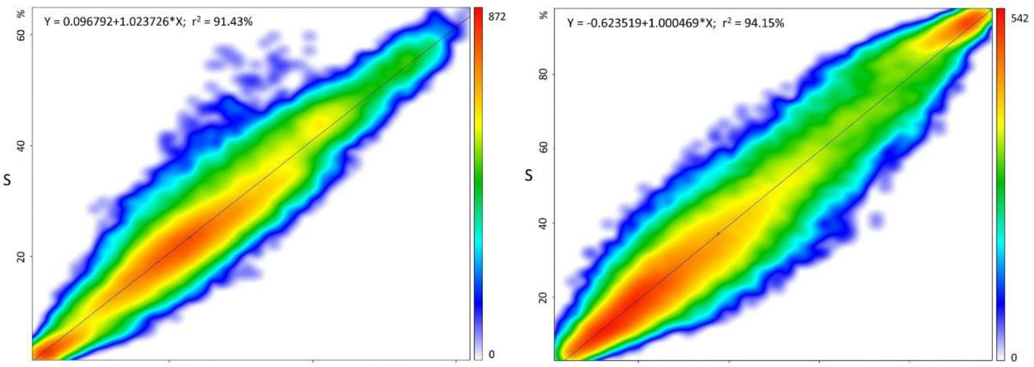

Figure 9. Scatter plot of the directly compiled (D) and the synthetized (S) clay- (left) and sand (right) content map

I investigated the difference maps of directly compiled and synthetized clay and sand maps. As for spatial pattern, they seem to be complementary, where the synthetized map overestimates in the clay map, it underestimates in the sand map at the same location. The differences show symmetrical distribution in the case of sand (~ ± 35 %). In the clay content map however, the differences are not symmetrical: the synthetized map overestimates up to 33 %, while underestimation occurs down to only 18 %. This asymmetry can be seen in the histogram of the clay difference map. However, statistically the values (in the cases of both clay, and sand) do not diverge from normal distribution.

According to the two tailed Wilcoxon-signed-rank test, there is significant difference between the directly compiled and the synthetized clay- as well as sand maps.

The validation by independent data, provided by MARTHA database, resulted that the error of the synthetized maps is slightly higher both for clay and sand.

The pixel by pixel combination of the three synthetized and directly compiled particle size fraction maps according to the USDA categories can be seen in Figure 10. The results of their comparison are the followings.

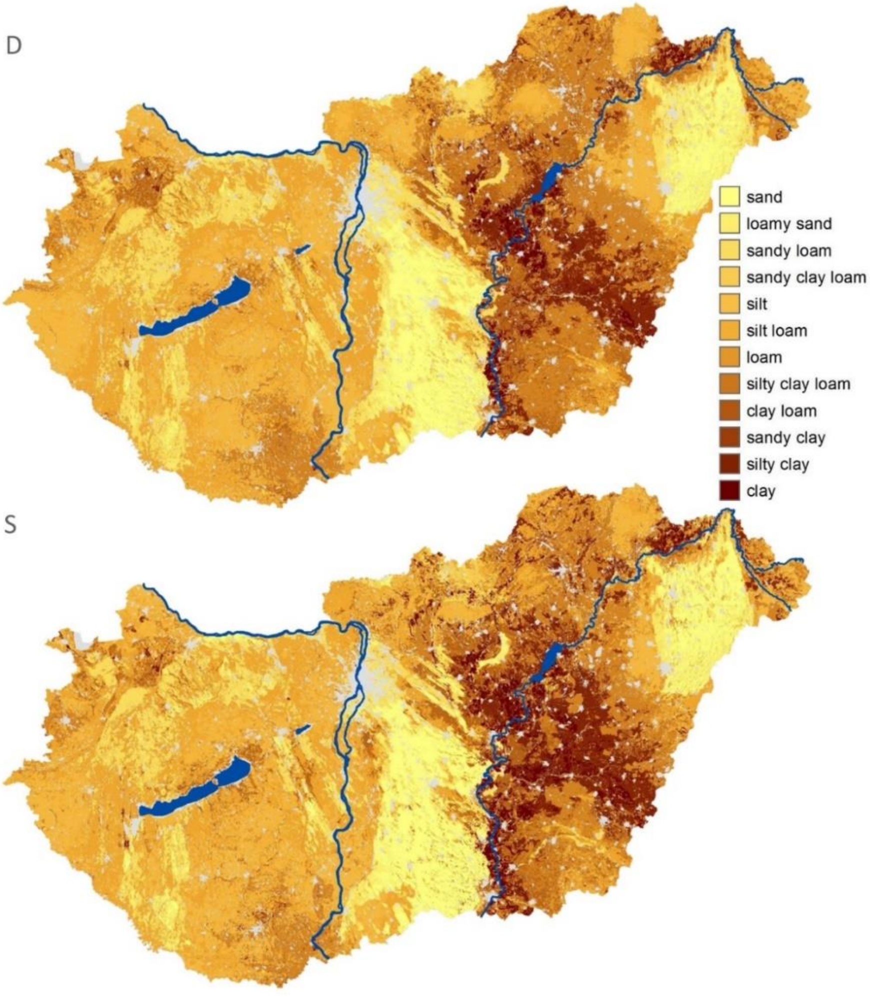

Figure 10. Directly compiled (D) and synthetized (S) USDA texture class maps for 0-30 cm depth

Spatial extension of the texture classes predicted by the two sorts of textural class maps can be seen in Figure 11. I concluded that remarkable differences occur typically in the more extreme categories. However, according to the two tailed Wilcoxon-signed-rank test, there is no significant difference between the two texture class maps globally.

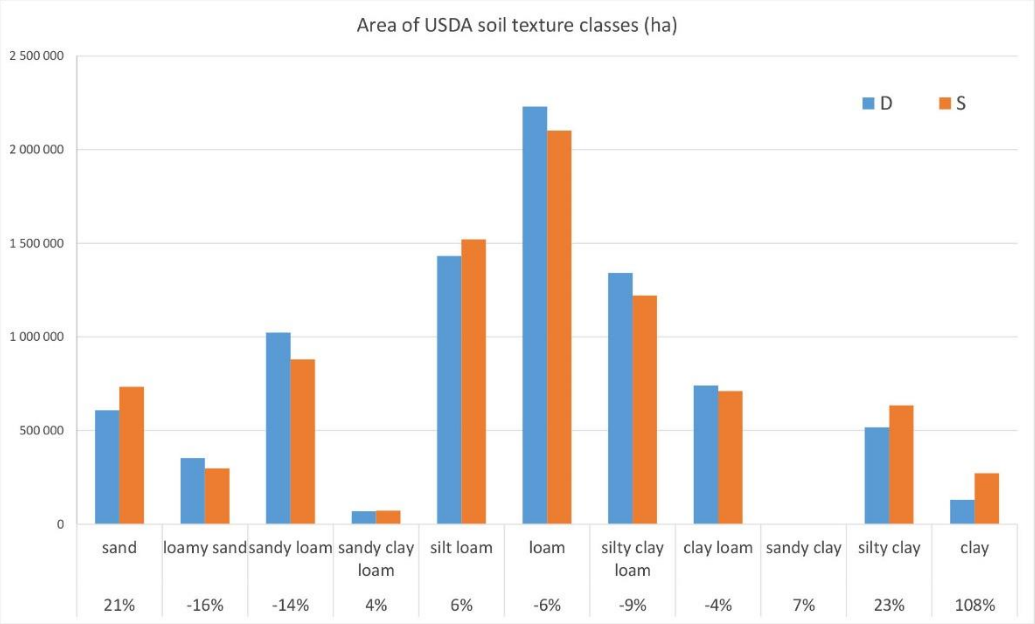

Figure 11. Area of USDA soil texture classes in the directly compiled (D) and in the synthetized (S) texture class map, and the differences between the two values in percentage

According to the differences between the directly compiled and synthetized texture class maps – taking into account taxonomical distances –, on 68 % of the area, the two maps predict the same class. The validation by independent data, provided by MARTHA database, resulted that the error of the synthetized maps is slightly higher both for clay and sand. The two texture class maps compared to each other show ‘substantial agreement’ (Kappa value: 0,62, weighted Kappa: 0,76), however, it does not reach the ‘almost perfect agreement’ category (Landis & Koch 1977).

According to the results, the directly computed and the synthetized maps show various differences. The proportion of the areas with small and large clay-, and sand content was increased by synthetization. In the case of texture classes, differences also occurred typically in the more extreme categories. However, spatial extension of the categories globally do not show statistically significant difference. Based on the results, I conclude that further usage of the maps requires due foresight, and it is recommended to use directly compiled maps.

Disaggregated nationwide genetic soil type map

Result map compiled from 100 predictions using maximum likelihood rule can be seen in Figure 12. Thematically it is in good agreement with the original AGROTOPO map, however, its spatial resolution has been increased remarkably. The apparent similarity of the two maps show that the main structures and characteristics remained, however, the homogene mapping units has been divided. Differences are realized on local levels.

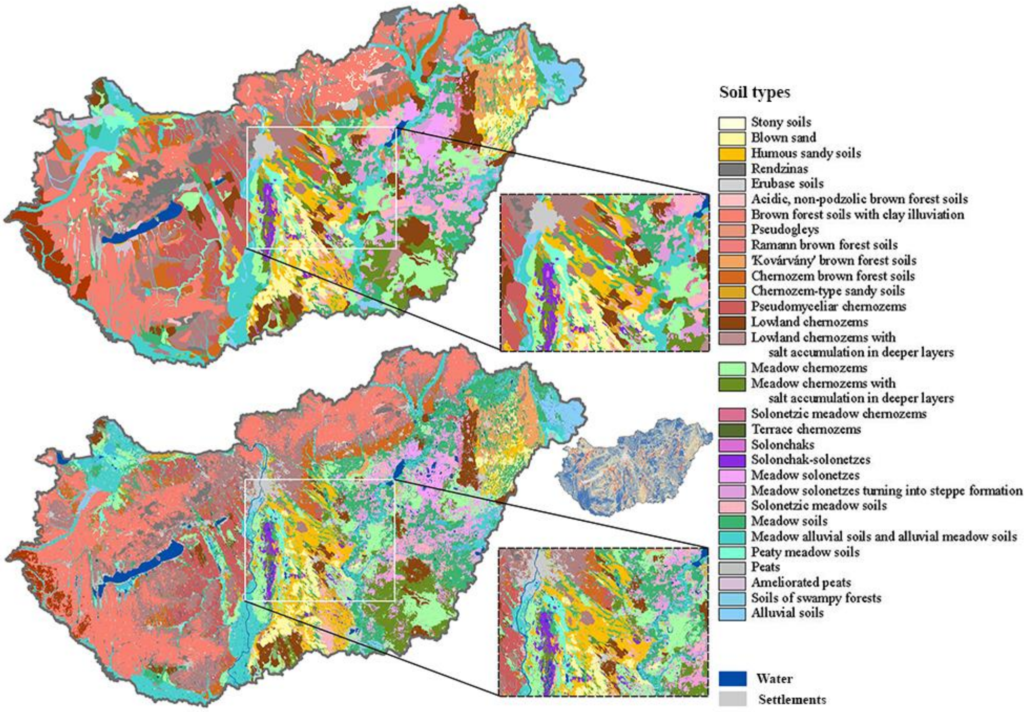

Figure 12. AGROTOPO (upper) and disaggregated 100 m resolution genetic soil type map (lower), with map of prediction model stability

Spatial information on the reliability of the prediction is also attached to the result map. It shows how many classes were predicted for the given pixel. Actually the map does not refer to the accuracy of information, but to the stability of the prediction model.

According to validation by SIMS points, Overall Accuracy is 43 %. Overall Kappa is 37 %, which refers to ’fair’ agreement. Since some categories of the two databases differ, the results of accuracy assessment can be considered a little vaguely.

Summing up the 100 model results, prediction stability was determined type by type. These results verify the assumption that the larger structures of the map remain, the disaggregated map can contribute to distinguish local heterogeneity. Furthermore, the results of stability examinations are suitable for not only the characterization of the model, but it can support further investigations in taxonomical distances of soil types.

The elaborated method can be considered as a modified random forest method, and it can be suggested for increasing spatial resolution of thematic soil maps. Experiences of disaggregation of AGROTOPO soil type map can support further soil mapping projects. The method can contribute to further disaggregation tasks related to other categorical soil maps.

Elaboration of correlation between objects of spatial soil information and other interdisciplinary objects and categories

Integrating spatial soil information into ecological mapping units

I integrated spatial soil information into the 35 ha hexagonal net of MÉTA on three levels. After intersection with AGROTOPO, dominant soil properties were assigned to each hexagon. Adaptation has been carried out to all of the nine soil properties of AGROTOPO database, for the whole territory of Hungary.

Integrating DKSIS into MÉTA net has been carried out to the territory of the Great Hungarian Plain. Not only the dominant soil properties, but also the proportion of them has been computed. Furthermore, the number of categories within hexagons was also given.

On the third level, minimum, maximum, and average clay-, silt-, and sand content were calculated in percentage, within each hexagon for 0-30 cm and 30-60 cm depth layers. Furthermore, the dominant USDA textural class was determined also for these two layers. Conversion has been carried out for the whole territory of Hungary.

Correlation of categories differently defined in soil science and ecology

I carried out correlation between spatial soil information and ecological soil categories (sand, loess, weakly-, medium-, strongly salt affected soils, peaty soils, local peaty characteristics, forest soils, alluvial soils) in the centroid of approximately 3500 MÉTA hexagons (Figure 13).

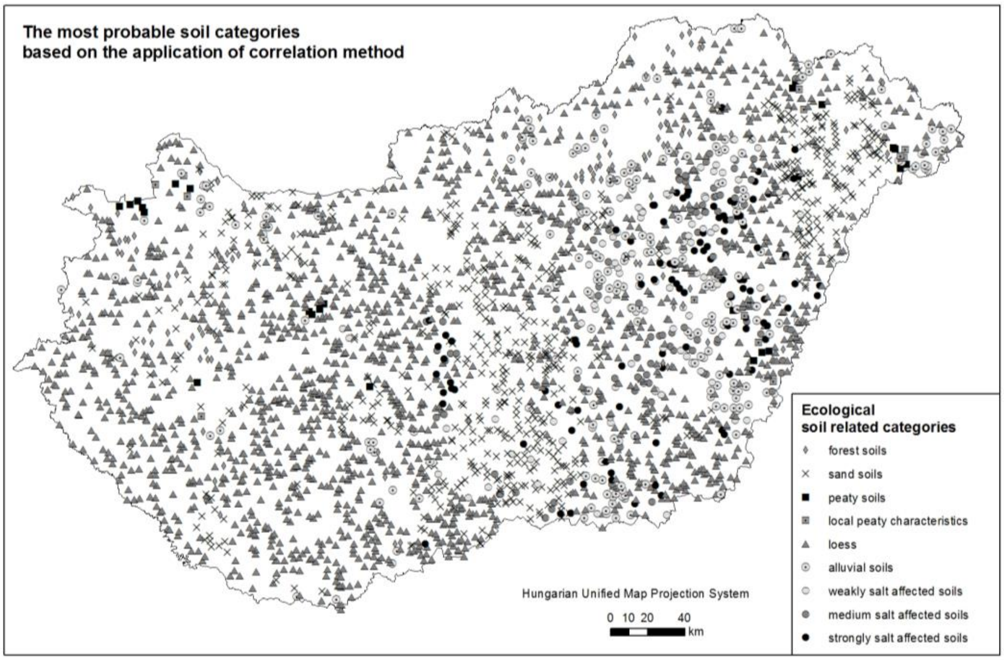

Figure 13. The most probable soil categories based on the application of correlation method at centroids of selected MÉTA hexagons

According to validation by independent data, sites classified as sand strongly overlap with sand and loamy sand areas of the texture class map (74 %). In the case of sand, loamy sand and sandy loam, the result is 91 %. The average EC values of the weakly- and strongly salt affected soils are in good agreement provided by the international categorization. Medium salt affected soils form a transitional category. The forest soil category resulted 63 % accuracy. In the case of alluvial soils, the correlation method was weak, accuracy is only 14 %. Peaty soils in large areas show 88 % accuracy, based on the ‘peats’, and ‘ameliorated peats’. The category of ‘local peaty characteristics’ resulted 26 % accuracy. In this case however, as comparing a local chararacteristics to a nationwide map, the relevancy of validation is rather questionable.

4. CONCLUSIONS

During my research, I compiled numerous novel digital soil maps, which contribute to the national spatial soil information. Although nationwide clay, silt, and sand content estimates, as well as USDA texture class maps appeared as demands, but they did not exist earlier. Predictions for particle size fractions were compiled according to the GlobalSoilMap specifications, supporting the contribution of Hungary to the initiative. The disaggregation of the widely used AGROTOPO genetic soil type map resulted a thematically identical map with more detailed spatial resolution. The maps and their related elements correspond to the crucial requirements of digital soil mapping, that is, accuracy and reliability information are attached to the maps.

The result maps can be used for further research tasks in soil science and other disciplines. They can be input parameters of erosion-, biodiversity-, ecosystem service-, agro-meteorological-, as well as water management models. Fitting into the goals of GlobalSoilMap, they can help solving global problems related to food production, starvation, climate change, environmental contamination. They can help the work of decision makers in agri-environmental-, water management-, environment protection-, nature conservation-, and landscape management topics.

The results can support further digital soil mapping tasks and the proper application of the maps. I concluded that in the case of countrywide texture mapping, the application of regression kriging offers more accurate results than the use of decision trees. Using more auxiliary variables does not necessarily lead to more accurate results.

According to the results of comparison of soil texture maps synthetized from standard depth layers with directly compiled products, it is recommended to use directly compiled maps. Using synthetized maps in models might involve some risks, because synthetization can intensify extremities. In the cases of climate change models, which are elaborated for investigations of extreme circumstances, using synthetized map can bias scenarios in different ways.

Soil texture properties are well handleable soil characteristics, which is suitable for testing methods and co-variables applied in digital soil mapping. Further investigations can be worthy, referring to other soil properties or depth layers. Moreover, it would be interesting to parametrize models with different spatial soil information, and to compare their results.

Elaborating a method that is basically a modification of random forest, I performed the spatial disaggregation of the AGROTOPO genetic soil type map. The method can be suggested for increasing spatial resolution of thematic soil maps. Experiences of disaggregation of AGROTOPO soil type map can support further soil mapping projects. The method can contribute to further disaggregation tasks related to other categorical soil maps. Spatial distribution information on the reliability of the prediction belongs to the result map. The results of stability examinations can support further investigations in taxonomical distances of soil types.

Elaboration of correlation between spatial soil information and other interdisciplinary objects and categories can support ecological researches and models with input data. SoilMÉTA can complement the original goals of MÉTA, supporting nature conservation and landscape ecological analyses with proper information. It can be a basis of landscape history researches, climate change scenarios, trend assessment of priority natural habitats. It would help in interpretation of soil-plant interrelations, vegetation-, and landscape evaluation researches. It can contribute to predict probability distribution of potential natural vegetation, or investigations of edaphic plant associations.

It might be an important point of view, that the vegetation database concentrates primarily on the natural habitats, while soil databases lay more emphasis on agricultural areas. SoilMÉTA supplement MÉTA from this point of view as well, supporting nationwide agri-environmental and water management decisions.

5. NEW SCIENTIFIC RESULTS

1. I compiled unique soil texture maps with regard to their representation and spatial resolution, based on digital soil mapping methods. Nationwide USDA soil texture class maps, as well as clay-, silt-, and sand content maps were compiled in 100 m and 150 m spatial resolution. The result maps support Hungary's contribution to GlobalSoilMap initiative, which aims at establishing a new, detailed, freely available global soil map database.
2. I compared the predictive capacity of variously compiled soil texture maps. Directly compiled maps by decision trees has been compared to those derived from particle size fraction maps mapped by regression kriging. I also examined the performance of regression kriging with two combinations of co-variables. Spatial resolution of the maps is 150 m, validation has been carried out by independent data points. I concluded that regression kriging resulted globally more accurate maps. However, I demonstrated, that using more co-variables do not necessarily leads to more accurate results.
3. I carried out comparison investigations of soil texture maps synthetized from standard depth layers (0-5 cm, 5-15 cm, 15-30 cm) with directly compiled products (0-30 cm). According to the results, the directly computed and the synthetized particle size fraction maps show various differences. The proportion of the areas with small and large clay and sand content was increased by synthetization. In the case of texture classes, differences also occurred typically in the more extreme categories. However, spatial extension of the categories globally do not show statistically significant difference. I concluded that it is recommended to use directly compiled maps.
4. I compiled particle size fraction (clay-, silt-, and sand content) maps according to the GlobalSoilMap specifications, for the standard 0-5 cm, 5-15 cm, 15-30 cm, 30-60 cm, 60-100 cm, 100-200 cm depth layers, in 100 m spatial resolution. The mapping process was carried out by composite kriging. I derived

texture class maps from the predictions also for the six layers. Based on independent validation dataset, I concluded that the upper (0-5 cm), and the two lower layers (60-100 cm and 100-200 cm) show greater error. I explained the results by the decreasing predictive capacity of surface co-variables, and by the unavoidable extrapolation of spline interpolation in marginal layers.

5. I elaborated a method that is considered to be a modification of random forest. The method is suitable for increasing spatial resolution of thematic soil maps, by which I carried out disaggregation of AGROTOPO genetic soil type map. Spatial information on the reliability of the prediction is also attached to the result map, whose spatial resolution is 100 m. The results of stability examinations can support further investigations in taxonomical distances of soil types.

6. I elaborated a spatial correlation method, by which probability prediction can be provided on soil related categories defined in ecology. The correlation method is based on the Digital Kreybig Soil Information System (DKSIS).

6. REFERENCES

Arrouays, D., McBratney, A.B., Minasny, B., Hempel, J.W., Heuvelink, G.B.M., MacMillan, R.A., Hartemink, A.E., Lagacherie, P., McKenzie, N.J., 2015. Specifications Tiered GlobalSoilMap products, Release 2.4.

Arrouays, D., McKenzie, N., Hempel, J., Richer de Forges, A., McBratney, A.B., 2014. GlobalSoilMap: Basis of the global spatial soil information system. CRC Press.

ASTER GDEM, 2000. Advanced Spaceborne Thermal Emission and Reflection Radiometer, Global Digital Elevation Model.

Boettinger, J.L., Howell, D.W., Moore, A.C., Hartemink, A.S., Kienast-Brown, S., 2010. Digital soil mapping; bridging research, environmental application and operation, Progress in Soil Science. Springer Netherlands, Dordrecht.

Büttner, G., Maucha, G., Bíró, M., Kosztra, B., Pataki, R., Petrik, O., 2004. National land cover database at scale 1:50000 in Hungary. EARSeL eProceedings 3, 323–330.

Cohen, J., 1960. A Coefficient of Agreement for Nominal Scales. Educ. Psychol. Meas. 20, 37–46.

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

Gyalog, L., Síkhegyi, F., 2005. Magyarország földtani térképe 1:100.000. Magyar Állami Földtani Intézet, https//map.mfgi.hu/fdt100/.

Hartemink, A.E., McBratney, A.B., Mendonça-Santos, M. de L., 2008. Digital Soil Mapping with Limited Data. Springer, Dordrecht.

Hengl, T., 2009. A Practical guide to Geostatistical Mapping, Scientific and Technical Research series.

Kocsis, M., Berényi Üveges, J., Várszegi, G., Sisák, I., 2015. A MÉM NAK genetikus talajtérkép bemutatása és talajosztályozási kategóriáinak elemzése. Agrokémia és Talajtan 64, 53–72.

Lagacherie, P., McBratney, A.B., 2007. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping, in: Lagacherie, P., McBratney, A.B., Voltz, M. (Eds.), Digital Soil Mapping - an Introductory Perspective. Elsevier, Amsterdam, pp. 3–22.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159–174.

Makó, A., Tóth, B., Hernádi, H., Farkas, Cs., Marth, P., 2010. Introduction of the Hungarian Detailed Soil Hydrophysical Database (MARTHA) and its use to test external pedotransfer functions. Agrokémia és Talajtan 59, 29–38.

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52.

Minasny, B., Malone, B., McBratney, A.B., 2012. Digital Soil Assessments and Beyond. Taylor and Francis Group, London.

NASA LP DAAC, 2015. MODIS 09 (16.03.2012, 07.09.2013) and MODIS13Q (03.2012, 09.2013) products. USGS/Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. http://modis.gsfc.nasa.gov/data.

Pásztor, L., Szabó, J., Bakacsi, Zs., Matus, J., Laborczi, A., 2012. Compilation of 1:50,000 scale digital soil maps for Hungary based on the digital Kreybig soil information system. Journal of Maps 8: 215-219

Pentelényi, A., Scharek, P., 2006. Magyarország talajvízszint mélység térképe (0-20m) 1:100.000. https://map.mbfsz.gov.hu/tvz100_251020/.

Rossiter, D.G., 2014. Technical Note: Statistical methods for accuracy assesment of classified thematic maps. Tech. Rep. ITC.

Szentimrey, T., Bihari, Z., 2007. Mathematical background of the spatial interpolation methods and the software MISH (Meteorological Interpolation based on Surface Homogenized Data Basis), in: Proceedings from the Conference on Spatial Interpolation in Climatology and Meteorology. Budapest, pp. 17–27.

Takács, K., Szatmári, G., Bakacsi, Zs., Laborczi, A., Szabó, J., Tóth, T., Pásztor, L., 2016. Target-specific digital soil mapping supporting spatial planning in Hungary, in: Sarjakoski, T., Santos, M.Y., Sarjakoski, L.T. (Eds.), The 19th AGILE International Conference on Geographic Information Science. pp. 1-7.

TIM, 1995. Talajvédelmi Információs és Monitoring Rendszer. Módszertan.

Várallyay, Gy., Szűcs, L., Murányi, A., Rajkai, K., Zilahy, P., 1979. Magyarország termőhelyi adottságait meghatározó talajtani tényezők 1:100.000 méretarányú térképe I. Agrokémia és Talajtan 28, 363–384.

Várallyay, Gy., Szűcs, L., Murányi, A., Rajkai, K., Zilahy, P., 1980. Magyarország termőhelyi adottságait meghatározó talajtani tényezők 1:100.000 méretarányú térképe II. Agrokémia és Talajtan 29, 35–76.

Zhang, G.-L., Brus, D.J., Liu, F., Song, X.-D., Lagacherie, P., 2016. Digital Soil Mapping Across Paradigms, Scales and Boundaries. Springer Science+Business Media, Singapore.

7. PUBLICATIONS RELATED TO THE DISSERTATION

Journal articles with impact factor

Laborczi A, Szatmári G, Kaposi A D, Pásztor L. 2018. Comparison of soil texture maps synthetized from standard depth layers with directly compiled products. *Geoderma,* (accepted for publication).

Pásztor L, Laborczi A, Bakacsi Zs, Szabó J, Illés G. 2018. Compilation of a national soil-type map for Hungary by sequential classification methods. *Geoderma* 311: 93-108.

Somodi I, Molnár Zs, Czúcz B, Bede-Fazekas Á, Bölöni J, Pásztor L, Laborczi A, Zimmermann N E. 2017. Implementation and application of Multiple Potential Natural Vegetation models – a case study of Hungary. *Journal of Vegetation Science* 28, 1260–1269.

Laborczi A, Szatmári G, Takács K, Pásztor L. 2016. Mapping of topsoil texture in Hungary using classification trees. *Journal of Maps* 12(5): 999-1009.

Pásztor L, Szabó J, Bakacsi Zs, Laborczi A. 2013. Elaboration and applications of spatial soil information systems and digital soil mapping at Research Institute for Soil Science and Agricultural Chemistry of the Hungarian Academy of Sciences. *Geocarto International* 28(1): 13-27.

Pásztor L, Szabó J, Bakacsi Zs, Matus J, Laborczi A. 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): 215-219.

Journal articles without impact factor, in English

Pásztor L, Laborczi A, Takács K, Szatmári G, Dobos E, Illés G, Bakacsi Zs, Szabó J. 2015. Compilation of novel and renewed, goal oriented digital soil maps using geostatistical and data mining tools. *Hungarian Geographical Bulletin* 2015;64/1: 49–64.

Pásztor L, Dobos E, Szatmári G, Laborczi A, Takács K, Bakacsi Zs, Szabó J. 2014. Application of legacy soil data in digital soil mapping for the elaboration of novel, countrywide maps of soil conditions. *Agrokémia és Talajtan* 63(1): 79-88.

Journal articles without impact factor, in Hungarian

Illés G, Fonyó T, Pásztor L, Bakacsi Zs, Laborczi A, Szatmári G, Szabó J. 2016. Az Agrárklíma 2 projekt eredményei: Magyarország digitális talajtípus térképének előállítása. *Erdészettudományi Közlemények* 6(1):17-24.

Bakacsi Zs, Laborczi A, Szabó J, Takács K, Pásztor L. 2014. Az 1:100 000-es földtani térkép jelkulcsának és a FAO rendszer talajképző kőzet kódrendszerének javasolt megfeleltetése. *Agrokémia és Talajtan* 63(2): 189-202.

Pásztor L, Laborczi A, Bakacsi Zs, Szabó J. 2013. Kategória típusú talajtérképek térbeli felbontásának javítása kiegészítő talajtani adatok és adatbányászati módszerek segítségével. *Agrokémia és Talajtan* 62(2): 205-218.

Laborczi A, Bakacsi Zs, Szabó J, Pásztor L. 2013. Térbeli talajinformációk konverziója és regionalizálása élőhely osztályok becsléséhez. *Tájökológiai Lapok* 11(2): 233-244.

Book chapter, in English

Pásztor L, Laborczi A, Takács K, Bakacsi Zs, Szabó J. 2016. Variations for the implementation of SCORPAN's "S". In: Zhang G-L et al. (eds.) *Digital Soil Mapping Across Paradigms, Scales and Boundaries*. Springer Science+Business Media Singapore, pp. 331-342.

Pásztor L, Szabó J, Bakacsi Zs, Laborczi A, Dobos E, Illés G, Szatmári G. 2014. Elaboration of novel, countrywide maps for the statisfaction of recent demands on spatial, soil related information in Hungary. In: Arrouays D et al. (eds.) *Global Soil Map: Basis of the global spatial soil information system*. London: Taylor & Francis Group, pp. 207-212.

Conference proceedings

Laborczi A, Takács K, Bakacsi Zs, Szabó J, Pásztor L. 2015. A Duna-Tisza köze genetikai talajtípus térképének előállítása adatbányászati eszközökkel. In: Boda J (szerk.) *Az elmélet és gyakorlat találkozása a térinformatikában. VI. Térinformatikai Konferencia és Szakkiállítás,* Debreceni Egyetemi Kiadó, Debrecen, 269–275.

Pásztor L, Bakacsi Zs, Szabó J, Takács K, Laborczi A. 2014. Tematikus talajtérképek dezaggregálási lehetőségei digitális talajtérképezési eljárásokkal. In: Balázs B (szerk.) *Az elmélet és a gyakorlat találkozása a térinformatikában V.: Térinformatikai konferencia és szakkiállítás*. Debreceni Egyetemi Kiadó, Debrecen, 253–261.

Pásztor L, Bakacsi Zs, Szabó J, Laborczi A. 2013. Célspecifikus digitális talajtérképek előállítása adatbányászati eszközök felhasználásával. In: Lóki J (szerk.) *Az elmélet és a gyakorlat találkozása a térinformatikában IV.*, Debreceni Egyetemi Kiadó, Debrecen, 315-324.

Laborczi A, Pásztor L, Szabó J, Bakacsi Zs, Dombos M. Pedological Support of the Landscape-Ecological Mapping of Hungary. *Cereal Research Communications* (Suppl.) 2008; 36: 503-506.

Laborczi A, Szabó J, Bakacsi Zs, Dombos M. 2008. Az élőhely-térképezés talajtani támogatása (TalajMÉTA). *Talajvédelem* (különszám), Talajtani Vándorgyűlés, 2008. május 28-29., Nyíregyháza, 627-632.