UDC 631.4: 551.4: 004.942

Effect of DEM sources on quality indicators of predictive maps of soil cover

V. R. Cherlinka[1*], Y. M. Dmytruk[1], Y. H. Bodyan[2]

[1]**Yuriy Fedkovych Chernivtsi National University,** *Chernivtsi, Ukraine*
[2]**Chernivtsi branch of the State Institution "Soils Protection Institute of Ukraine",** *Chernivtsi, Ukraine*

ARTICLE INFO

Received 01.07.2020
Received in revised form 08.08.2020
Accepted 15.09.2020
Available online 05.11.2020

Keywords:

cartogram of agro-production soil groups; DEM; digital elevation model; modelling; morphometric parameters, predicative algorithms; soil map.

ABSTRACT

The aim of the study was to identify the impact of digital elevation models of different origins on the qualitative characteristics of forecast maps of soil cover or cartograms of agro-production soil groups using predictive modeling technologies. The current situation with large-scale soil cartographic data in Ukraine is analyzed and it is shown that the fastest and most cost-effective way to fill gaps in creating a continuous cartographic coverage for unexplored areas, which make up 33% of Ukraine, is mathematical simulation. The latter is based on morphometric analysis of digital elevation models, which distinguishes a number of predictors, which are further analyzed for links with existing cartographic soil materials by creating a mathematical predictive model using landscape reference points and associated soil type. The identified difference in the quality of predictive materials using the Cohen's kappa coefficient allows us to recommend individual sources of DEM as a basis for such tasks. A demonstration of a closed production cycle of creating predicative soil cartographic materials based on free software (GRASS and Quantum geographic information systems, language and environment for statistical computing and graphics R and shareware - Easy Trace vectorizer) was conducted.

Corresponding author: v.cherlinka@chnu.edu.ua

Citing: Cherlinka V.R., Dmytruk Y.M., Bodyan Y.H. 2020. Effect of DEM sources on quality indicators of predictive maps of soil cover. *Agrochemistry and Soil Science*. Collected papers. No. 90. Kharkiv: NSC ISSAR, P. 36-46. DOI: https://doi.org/10.31073/acss90-04.

1. Introduction

The complexity of the political and economic situation of the last 30 years in Ukraine and its current forecast lead to the conclusion that the current situation with large-scale soil cartographic materials should not be expected to improve. The existing problems of map relevance, quality and coverage [1-5] require the latest approaches to solve them. Filling unexplored areas with model data is a method widely used in the modern world. The number of studies on modeling the spatial distribution of soils is now growing rapidly [6-15]. The arsenal of mathematical methods used in this case is extremely wide: from multifactor regression analysis and neural networks to various classification trees [16]. The application of such methods is based on the use of reference points of landscapes and finding the dependence of soil units associated with them [17]. The main source of initial parameters is the digital elevation model (DEM), the analysis of which allows to identify a number of geomorphological and related indicators. Since soils data are a categorical type of data, and the indicators obtained from the DEM are numerical, only the use of modern mathematical methods allows to establish relationships between these parameters and build on this basis a model of soil cover [12, 18-22].

The general modeling procedure includes three main points: the analysis of the DEM (1), the construction of a training dataset for machine learning (2) and the actual modeling (3). This approach with minor variations is used by almost all researchers. Depending on the goal, the spatial coverage of the models can be local [23], national [24] or global [25, 26].

Because these models use DEM as the main source of input data, their availability and quality characteristics, which are significant for soil modeling, are important. A great role is played by such a parameter as spatial resolution. Its influence on the results of various modeling is widely discussed in the literature. Resolution affects, for example, the quality of modeling of landslide processes [27], morphometric and hydrological parameters [28], the quality of predictive simulation of soil cover [29].

The main sources of DEM now are the following [16]: topographic or kinematic GPS surveying, analog and digital photogrammetric approaches, radar technologies (SRTM, Aster GDEM and others), laser scanning of LiDAR and digitization of contours topographic map. From the point of view of accessibility in the realities of Ukraine, the most suitable still remain global DEMs based on radar technologies, digitization of contours topographic maps, and, more recently, digital photogrammetry according to UAV survey data.

In this paper, we will focus on the least expensive ways to obtain DEMs, which are suitable for obtaining data for altitude of topographic surface both nationally (global DEMs) and local coverage (global DEMs + digitization of topographic maps).

Global DEMs are freely available on the Internet and have different spatial resolutions, which allows you to select sources according to requirements. However, not every global DEM is suitable for large-scale predictive soil modeling: for example, GMTED2010 (Global Multi-resolution Terrain Elevation Data) [30] with some resolution levels (30-arc-second (1 kilometer), 15-arc-second (450 meters) and 7.5-arc-second (225 meters)) does not meet the needs of spatial resolution.

From this point of view, the new release of ASTER GDEM (Advanced Satellite Thermal Emission and Reflection Radiometer) global digital terrain model version 3 (GDEM 003 or ASTGTM V003) of the Ministry of Economy, Trade and Industry of Japan (METI) and of the US National Aeronautics and Space Administration (NASA) with a resolution of 1 arc-second (~30 m) seems to be one of the preferred options for use as a source of DEM [31].

SRTM (Shuttle radar topographic mission) SRTMGL1v003 data from the NASA and the US National Geospatial Intelligence Agency (NGA) have a similar to previous Aster GDEM resolution suitable for large-scale modeling [32].

Noteworthy is the global digital surface model with a resolution of 30 meters ALOS World 3D version 3.1 (AW3D30), distributed by the Japan Aerospace Exploration Agency (JAXA) [33]. Although it is declared a digital model of the surface, not a relief, its global coverage encourages verification as a source for soil modeling. Unfortunately, the digital surface model EU-DEM v1.1 with a resolution of 25 meters, which is freely available for the EU countries, does not extend to the territory of Ukraine [34]. It is a hybrid product based on SRTM and ASTER GDEM data combined by the weighted average method. Therefore, the assessment of its suitability for modeling the soil cover within Ukraine can be performed using the method of analogies.

The analysis shows [35] that now there are no available sources of complete geospatial data of domestic origin in Ukraine, even within the framework of the implemented pilot project of the prototype of the National Geospatial Data Infrastructure (NGDI). Moreover, the process of building the NGDI has virtually come to a halt, and crucial issues related to legislative, institutional and financial support have not been resolved. As a result, the current state of creation of geoinformation resources and provision of geoinformation services is characterized by systemic problems [36] and unpredictable vector of development. Accordingly, obtaining low-cost high-precision DEM of local scale in the absence of state support is possible only with the digitization of large-scale topographic maps, with the described reservations at [37].

The situation is similar with soil maps: the only way to obtain them in digital form - self-scanning (in the presence of the original source), digitization and attribution. This realizes the possibility of constructing a training dataset in the above-mentioned three-stage modeling procedure.

There are two clear distinctions to the construction of a set of training data [7, 9, 10]: data of soil pits in the field survey and a sample of clearly defined contours of soil map landfills. The first approach has good prospects, but requires a large database on verified soil profiles throughout the country, work on which is intensive [38, 39, 40]. Therefore, for the local scale, we use a different approach, as more relevant in the near future and easier to implement in the current modeling environment.

Hence, the task of this study was to cover the options of predictive modeling when using as input data available maps of soils and sources of DEM and characterization of their qualitative parameters, which give the analysis to choose the best of them as a result of modeling. Accordingly, the aim of our work was to study the influence of variants of DEM sources on the qualitative characteristics of predictive soil maps. Large-scale soil and topographic maps (M 1:10000), freely available global elevation/surface models ASTGTM V003 (hereinafter ASTER), SRTMGL1v003 (hereinafter SRTM), AW3D30 (hereinafter ALOS) and free software were used for this purpose (GRASS [41] and Quantum GIS [42], the Easy Trace vectorizer [43] and R - a language and environment for statistical computing [44]).

2. Materials and methods

In accordance with the set goal, we identified a range of tasks that needed to be solved: 1) digitization and attribution of cartographic materials; 2) construction of a series of reference DEMs based on digitized topographic maps with a spatial resolution of 5, 15 and 25 m; 3) downloading and redesign data to the coordinate system of the project (Pulkovo 1942/CS63 zone X2, code epsg 7826) a global digital models Aster, Srtm and Alos; 4) analysis of digital elevation/terrain models and excretion from them in GIS GRASS of a set of maps of morphometric and other derived characteristics; 5) creation of training dataset for machine learning; 6) modeling of soil cover in R-statistics using the predictive algorithm Random Forest [45, 46, 47]; 7) analysis of the obtained results and conclusions about the optimal source DEM for predictive modeling.

A fragment of the territory of Ukraine (Fig. 1a) within the Chernivtsi region (Fig. 1b), dated to the Prut-Dniester interfluve (Northern Bukovyna) with contrasting geomorphological conditions and administratively belonging to the Kitsman district (Fig. 1c), was chosen as an object. We mentioning, that he has been involved in some previous studies [48]. This fragment has different administrative subordination and differences in the economic use of individual parts, and when choosing him, typical problems were solved [35, 37, 48, 49]. The Pulkovo 1942/CS63 zone X2 coordinate system was adopted as the base, in which scanned sheets of topographic maps M 1:10000 (M-35-124-Vg- {1,2,3,4}, M-35-124- Vb -3 and M-35-124-Vv-2) are contained (Fig. 1d). Free software tools were used for data processing: georectification of cartographic data - GIS Quantum and digitization - Easy Trace.

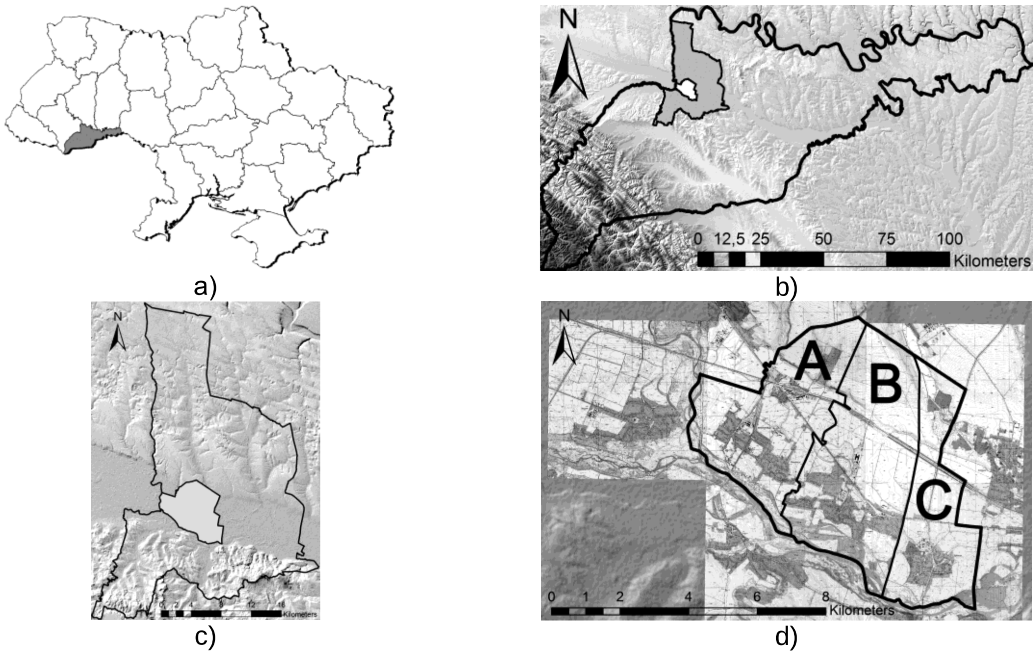

Fig. 1. Geographical location of the study region within Ukraine (a), of Chernivtsi region (b), Kitsman district (c) and test site scheme (d) *for the background used SRTM data

Topographic maps were georectified using the created vector mathematical basis and using 40-45 pinpoints per sheet, and the corresponding contours were digitized and attributed (Fig. 2a).

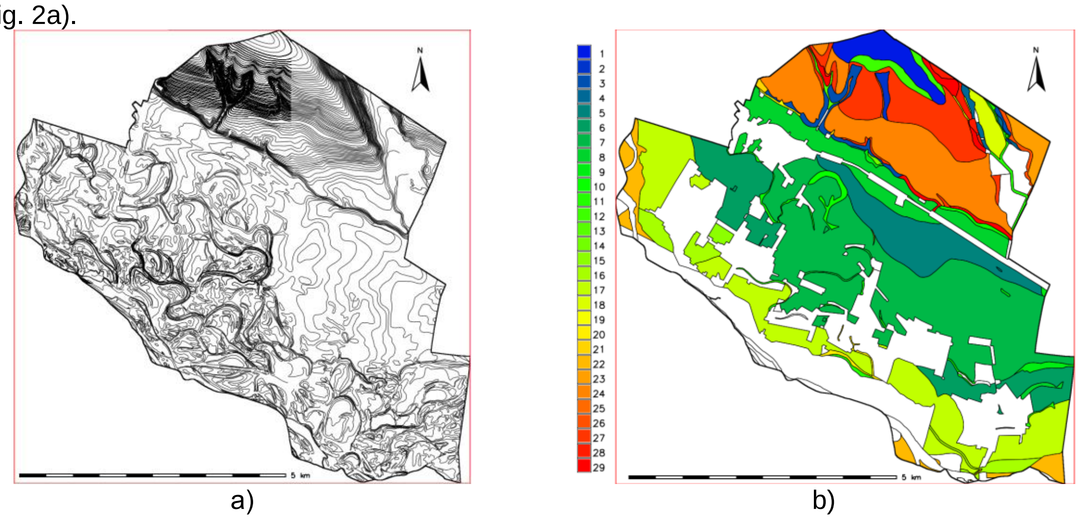

Fig. 2. *Digitized contours (a) and boundaries of the soil units (b)*

Georectification of soil maps was carried out to the characteristic points of the area and administrative boundaries of the currently existing village councils: Nepolokivtsi (village Nepolokivtsi) - "A", Beregomet (villages Beregomet and Revakivtsi) - "B", and Dubivtsy (village Dubivtsi) - "C" of Kitsman district of Chernivtsi region. Soil materials were based on a series of legacy soil maps of the collective farm "Soviet Ukraine" (survey of 1957 and correction of 1974). After generalization of the nomenclature list of soils and coordination of their contours, preliminary data on the percentage of soil coverage by soil surveys were obtained: from 4424.32 ha of total area for 24.86%, data are completely absent (Fig. 2b, Table 1).

Table 1
List of soil types of the study area

Id soil	Code of soil	Type of soil (transliteration from Ukrainian)
0	No data	The survey not conducted
1	1 l	Temno-siri lisovi
2	10 l	Chornozemy opidzoleni serednozmyti z pliamamy 30-50% sylnozmytykh
3	11 l	Chornozemy opidzoleni sylnozmyti
4	12 ad	Chornozemno-luchni mocharysti
5	13 a	Luchni hlyboki vyluhovani
6	14 a	Luchni pyluvato-serednosuhlynkovi
7	15 a	Luchni pyluvato-vazhkosuhlynkovi
8	16 a	Luchni hleiovi
9	17 a	Luchno-bolotni osusheni
10	18 a	Bolotni pyluvato-vazhkosuhlynkovi na davnomu aliuviiu
11	19 al	Bolotni pyluvato-vazhkosuhlynkovi na suchasnomu aliuviiu
12	2 l	Temno-siri lisovi slabozmyti
13	20 d	Bolotni pyluvato-vazhkosuhlynkovi na suchasnomu deliuviiu
14	21 d	Bolotni mocharysti
15	22 a	Dernovi hlyboki karbonatni
16	23 al	Dernovi karbonatni supishchani
17	24 al	Dernovi karbonatni pishchano-lehkosuhlynkovi
18	25 al	Dernovi karbonatni hleiovi namyti
19	26 l	Slabozadernovani skhyly yariv ta krutykh ustupiv pyluvato-vazhkosuhlynkovi na lesopodibnykh suhlynkakh
20	27 a	Slabozadernovani skhyly yariv ta krutykh ustupiv pyluvato-vazhkosuhlynkovi na davnomu aliuviiu
21	28 l	Vykhody porid
22	29 al km	Ruslovi vidklady
23	3 l	Chornozemy opidzoleni pyluvato-lehkosuhlynkovi
24	4 l	Chornozemy opidzoleni pyluvato-serednosuhlynkovi
25	5 dl	Chornozemy opidzoleni hleiuvati namyti
26	6 l	Chornozemy opidzoleni slabozmyti pyluvato-lehkosuhlynkovi
27	7 l	Chornozemy opidzoleni slabozmyti pyluvato-serednosuhlynkovi
28	8 l	Chornozemy opidzoleni slabozmyti z pliamamy 10-30% serednozmytykh
29	9 l	Chornozemy opidzoleni serednozmyti

3. Results

For the spatial resolution of a series of reference DEMs based on digitized topographic maps, the following values were selected: 5, 15 and 25 m (Fig. 3a, b, c). DEM generation was performed in GRASS GIS using a thin plate spline interpolation with regularization and covariables in the module v.surf.tps [50, 51].

The downloaded global elevation models Aster, Srtm [52] and Alos [53] have the identical 1-arc-second resolution (~30 meters), but for the latitude and longitude of the study area when reprojection in coordinate system Pulkovo 1942/CS63 zone X2, horizontal resolution of data pixel is ~20.9 m and the vertical is ~30.6 m. Because the rectangular data pixel is difficult to use for modeling, the data of global elevation models were resampled in Quantum GIS to 25x25 m spatial resolution. The remote sensing data mentioned above usually have a high noise level. This does not allow them to be used for analysis without pre-treatment [54].

Therefore, we used the r.denoise tool implemented in GRASS GIS using the algorithm [55], which allowed to eliminate random noises, maintaining clear terrain and smoothing the original data with minimal changes (Fig. 4).

The prepared series of DEMs was used to obtain maps of morphometric characteristics of the terrain, which served as predictors in the modeling, in particular the slope and aspect, surface curvature (longitudinal and maximum), solar radiation, landforms. Additional maps of hydrological indicators were also generated: topographic wetness index, accumulation, direction and length of water flows and distance to them.

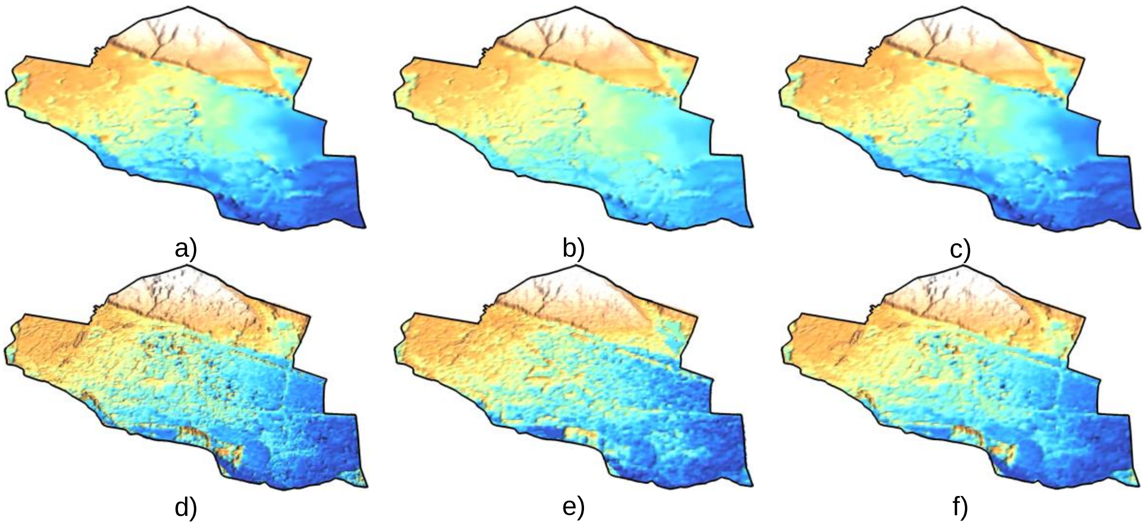

Fig. 3. *3D visualization of the generated DEMs (a - 5 m, b - 15 m and c - 25 m) and obtained on the basis of remote sensing data (d - Aster, e - Srtm, f - Alos) *view point from the south, camera height 3000 m, vertical scale 1:5*

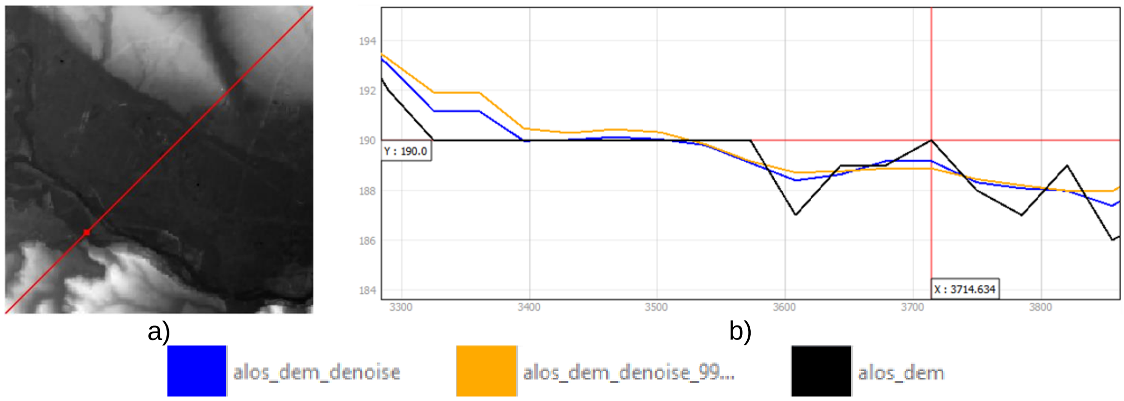

Fig. 4. *Transect of research area based on Alos data (a) and the difference between the original (black) and two variants (b) of denoised data (blue and yellow)*

For simulations of soil maps, we used the language and environment for statistical computing R-statistic [44], in particular the script, which included a number of adaptations to solve problems and implemented the predictive algorithm Random Forests [45-47]. Its latest implementation in the ranger package is very fast and can be used on large datasets [56].

The evaluation of the quality of the obtained models was performed on the basis of the index (quality functional) of Kappa Cohen [12, 21, 57-60]. In addition to the accuracy of the model solution, this indicator also takes into account the probability of accuracy, which could be obtained by chance. In general, the kappa shows the degree of correspondence between the original and simulated data.

As mentioned above, there are almost no materials available at the state level for soil modeling within Ukraine. Announced as a map of soils on the Public cadastral map [61] and cartograms of agricultural groups of soils on the site of the Regulatory Monetary Assessment [62] data are not available for download. In the first case, it is not really a map of soils, but a cartogram of agricultural groups of soils, and is presented in a scale of 1:200000, which allows it to be used purely for informational purposes. Given the whole set of problems that accompany these data [4, 35, 63], in practice to obtain reliable results should use self-scanned and digitized soil maps.

In this experiment, we used a method of randomized weighted training sample with 35% coverage of the area of the surveyed soils [22]. Modeling of soil cover on the basis of different sources of DEM and variations of resolution (only 6 variants) gave quite interesting results. The first thing to say is that the resolution above 5-15 m significantly reduces the compliance of resampled versions of the soil map to its scanned version. This is due to the "roughening" of small elements and contours when increasing the pixel size of the data. Therefore, for example, on the forecast map of soils with a resolution of 5 m and the corresponding Cohen's kappa 0.98 you can see almost complete coincidence of the contours of the predicate areas of soils to their scanned version (Fig. 5a, b). In the case of a similar DEM, but with a resolution of 25 m (Fig. 5c), the kappa is 0.84 and even in well-predicted areas, the contours do not match very well due to the significant pixelation of the data. Analysis of predictive map variants also reveals a significant number of single pixels of soils, which are scattered in larger areas of other soils. We believe that this can be regarded as a model "noise", which can be largely eliminated by generalizing the results. This is done by us using the GRASS r.neighbors module, which looks at each pixel in the input raster forecast map file and checks the values assigned to it in a certain "neighborhood" around it. As a result, we obtain a new layer of the raster map, in which each pixel is assigned a value that is a function of the values around this pixel. The experiment showed that the best results (increasing the value of the kappa) are obtained by generalizing the function "Moda" and the running window 3x3 cells, while the central pixel of this window is also included in the calculation. When applying this approach, the value of Cohen's kappa increases from 0.84 to 0.88 for the model version of the predict soil map based on interpolated DEM with 25 m resolution (Fig. 5d).

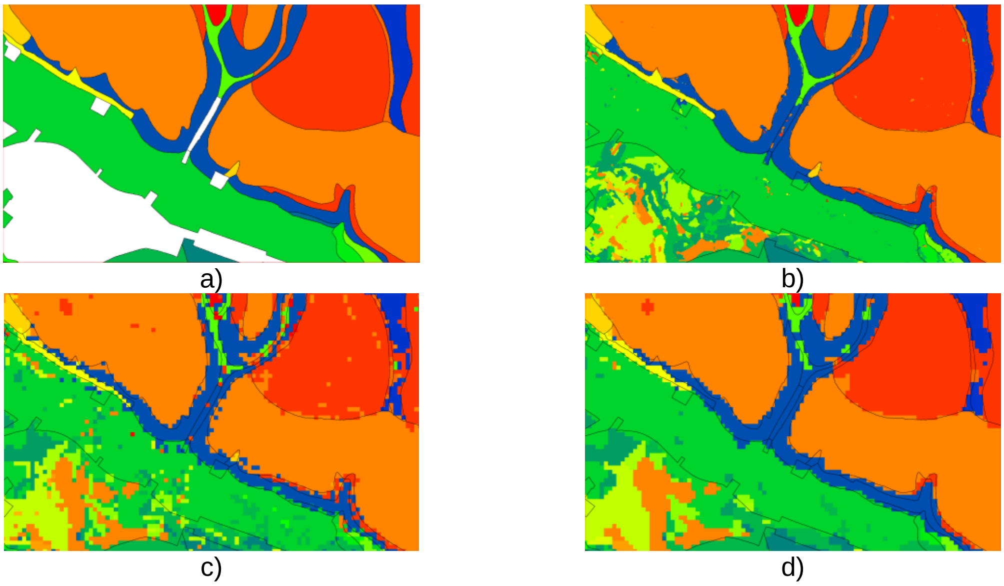

Fig. 5. *Comparison of the original (some fragment) soil map (a) with it predict variants based on interpolated DEM: resolution 5 m (b), resolution 25 m (c) and its generalized variant (d)*

4. Discussion

In general, the obtained array of simulations (Fig. 6) shows a clear pattern in reducing the qualitative characteristics (Cohen's kappa) of predicative maps depending on the source of DEM and its spatial resolution. They can be arranged in the following order, according to the Kappa decreasing trend: DEM 5 m → DEM 15 m → DEM 25 m → Alos → Srtm → Aster. This is a trend due to some features of modelled maps that occur during their generalization. Thus, it was found that during generalization kappa increases in the range from 0.018 to 0.068, i.e. in any case, generalization improves Cohen's kappa (Fig. 7, graph, right scale). Accordingly, this leads to the fact that the generalized kappa of one DEM source variant may be better than the non-generalized variant of another (e.g. Alos and generalized Aster), or be close to it (DEM 15 m and generalized DEM 25 m). We get a similar result when analyzing the coincidences of pixel values of specific soil units (Fig. 7, histogram, left scale). Summarizing the data on the size of the kappa and the number of matching pixels, we can say that Alos gives better simulation results than analogues Srtm and Aster, and the last two terrain models, although differing in detail, but in modeling the soil cover show very similar results.

It is also worth noting that in addition to the generally accepted method in the literature for estimating the quality parameters of kappa-based modeling, where the original map is used as a comparison with the original map in the same resolution, we analyzed the option when model maps for all data sources are compared with the original soil map at 5m resolution. This makes it possible, in addition to the overall accuracy of the prediction, to estimate its decrease at the boundaries of soil units when enlarging the data pixel. This parameter is called by us kappa reference, and it shows a smaller value and similar to the usual kappa trends, but with a slightly longer interval of variation - from 0.018 to 0.075.

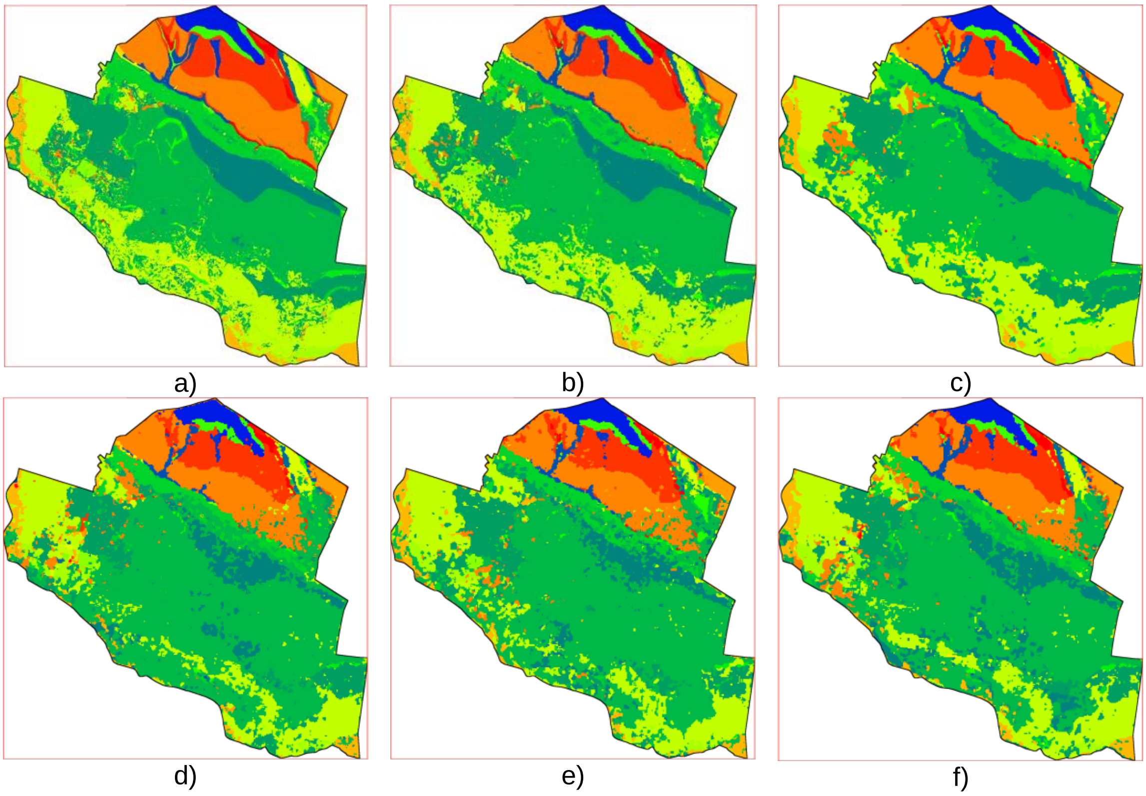

Fig. 6. *The results of the prediction of soil cover based on different DEM sources: interpolated from the topographic map (a - 5m, b - 15 m and c - 25 m) and global DEM (d - Aster, e - Srtm, f - Alos)*

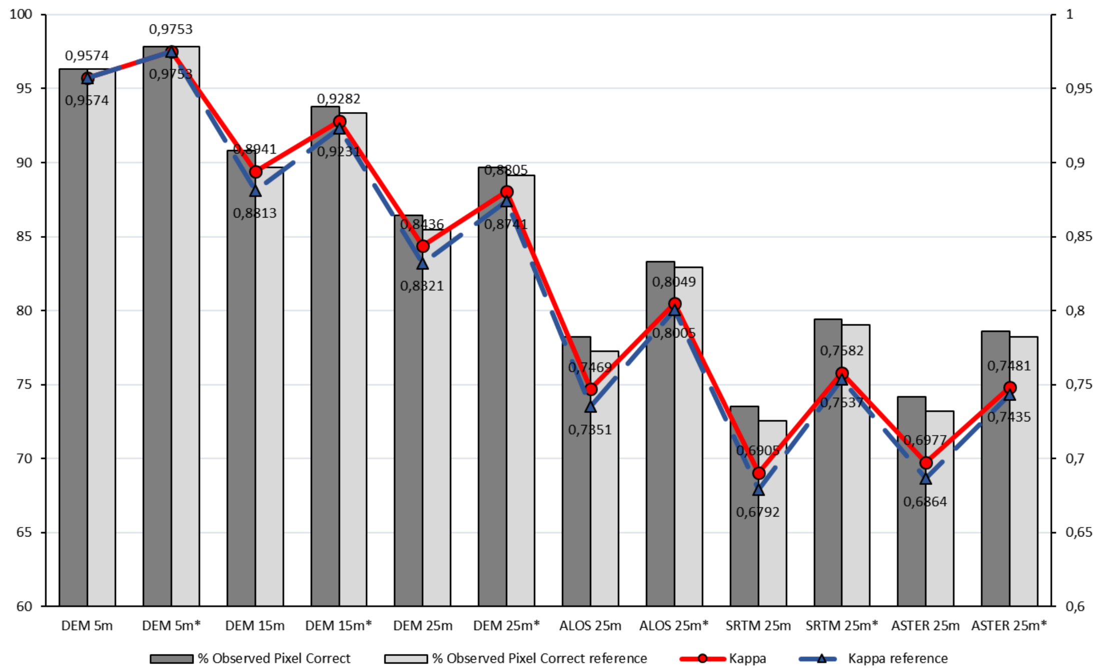

Fig. 7. *Differentiation of the quality of predicative soil maps depending on the source of DEM*
**generalized versions of maps*

The resulting array of model variants of the soil cover is interesting in terms of its compliance with the original maps, and, in particular, the prognostic ability in areas without soil information. The rather high coincidence of model and real data obtained by us indicates a certain level of statistical reliability of data in the areas of "white spots". Moreover, although these results can be called quite encouraging in this regard, visual analysis of predicate values in these white spots often shows a rather chaotic data set. We consider this not so much as the shortcomings of the applied forecasting algorithm, but most likely as the problems of input data, in particular the originals of soil maps. Very often, the resolution of the contours of the soil situation in large-scale soil research wanted to be better. Accordingly, the generated training sample based on such maps carries the same errors and inaccuracies that are inherent in the original maps and transfers them to the model in the area of "white" spots. Therefore, when planning future rounds of large-scale soil surveys, they should be based on the latest developments to avoid the mistakes of the past, which concerns, first, the quality recognizes of soil units and their spatial differentiation.

It should be noted that our proposed method allows obtaining a complete analogue of the map for areas of "white" spots, and because it is based on archival and, accordingly, outdated materials, the result of such modeling cannot exceed the quality of the input data. Nevertheless, when using the latest surveys as training data, devoid of the shortcomings of the old mapping technique, the model situation will be much more reliable. And since the modern wide-scale and at the same time large-scale surveys within Ukraine are not yet planned, we are forced to state the fact that in the near future even such data obtained by us will be significantly informative compared to the situation when they are absent. It is also worth mentioning that based on legacy data cartograms of agricultural groups of soils underlie the normative monetary valuation of agricultural land of Ukraine. Since the proposed method gives good results in the modeling of cartograms of agricultural production groups, we believe that it can be widely applied, despite the described problems. Note also that when modeling one mapping unit (usually one collective farm or state farm), all probable errors that were made during a large-scale survey are completely transferred to the model. However, when analyzing a larger data set, such as an administrative district or region, such errors are often offset by the prediction algorithms used.

When we analyze and model the data of archival domestic surveys, we observe a high correspondence between predict and source maps. This accuracy is not inferior to the level of similar studies in the literature, and often significantly exceeds their prediction accuracy. This directly indicates the drawing of the boundaries of soil units by relief or its derivatives (slope steepness map) according to the survey methods used in the past. However, when we analyze the freely available data of foreign datasets, in them, due to the selection of predictors, we obtained better kappa values than those published by the authors, although lower in absolute terms than in the Ukrainian data. The currently proposed set of predictors gives consistently high results. However, we continue to work in this direction, and plan to expand it in future. Often the set of predictors is limited by the unavailability of large-scale data, and low-scale in the analysis of a small area act simply as additional "coefficients", without affecting the qualitative parameters of the final model. Although up to 100 predictors are used in global SoilGrid250 models [26], their results are still very inaccurate, including due to the small scale of a number of predictors.

It should also be noted that the proposed technique, unfortunately, does not yet allow to involve in the modeling of genetic features of soils. For example, our forecasting approach does not allow the separation of, for example, soils such as chernozem podzolic and dark gray soil. Differences in the genesis of these soils are due to the influence of forest vegetation, and take into account its former presence place (in the case of agricultural land) is very problematic. Therefore, the current predictive algorithms currently divide the soil areas purely based on the use of reference points of landscapes and finding the dependence of the soil type units associated with them. In general, the above allows us to offer the use of such model approaches in the applied problems of soil science, agronomy, land management and land management, i.e. areas where the need for such data is most acute.

5. Conclusions

The study found that there is a significant influence of DEM sources on the qualitative characteristics of predictive maps of soil cover. It was found that depending on the source of DEM and its spatial resolution using the Random Forests algorithm; it is possible to obtain predictive cartographic materials with the Kappa Cohen's from 0.73 to 0.98. The main problems of availability and access to large-scale cartographic materials are highlighted and ways to solve them are shown.

An extended evaluation of predictive soil maps with ordinary and reference kappa was performed and it was shown that in the absence of highly detailed DEMs the most promising is the use global DEM of ALOS World 3D version 3.1 resampled to 25 m.

Acknowledgement

This work (programming and mathematical modelling) was partially funded by the National Scholarship Programme for the support of mobility of students, PhD students, university teachers, researchers and artists were established by the approval of the Government of the Slovak Republic [2016/2017:id17680]. NSP is funded by the Ministry of Education, Science, Research and Sport of the Slovak Republic. The programme is managed by SAIA, n.o. (Slovak Academic Information Agency). The authors gratefully thank the anonymous reviewer for their valuable comments and constructive recommendations.

References

1. Achasov A.B., Titenko G.V., Kurilov V.I. 2015. Data of remote sensing as a basis of soil mapping: economic aspect. *Bulletin of VN Karazin Kharkiv National University*. Series: Ecology (1104. Issue 10), 60–66. (Ukr.). URL http://journals.uran.ua/visnukkhnu_ecology/article/download/25458/33191.

2. Cherlinka V. 2017. Using Geostatistics, DEM and Remote Sensing to Clarify Soil Cover Maps of Ukraine. In: Dent D., Dmytruk Y. (Eds.), Soil Science Working for a Living: Applications of soil science to present-day problems. Springer-Verlag GmbH, Cham, Switzerland, Ch. 7. P. 89–100. URL: https://link.springer.com/chapter/10.1007/978-3-319-45417-7_7.

3. Cherlinka V., Dmytruk Y., Barabas D. 2019. Methods of verification of soils prediction maps: a case study from Chernivtsi region, Ukraine. *Geographia Cassoviensis*. Vol. 13, no. 2. P. 141-160. DOI: https://doi.org/10.33542/GC2019-2-04.

4. Cherlinka V.R., Lobova O.V. 2018. Methodical approaches to the coordination of soil cartographic materials on the borders of administrative-territorial units of Ukraine. Scientific reports of NULES of Ukraine. №6(76). P. 1-15. DOI: https://doi.org/10.31548/dopovidi2018.06.010 (Ukr.).

5. Polchyna S.M., Nikorych V.A., Danchu O.A. 2004. Application of the modern FAO / WRB soil classification system to the soil cover map of Chernivtsi region. *Soil Science*. Vol. 5(1–2). P. 27–33. URL: http://arr.chnu.edu.ua/jspui/bitstream/123456789/471/1/Nikorich.pdf.

6. Browning D.M., Duniway M.C. 2011. Digital soil mapping in the absence of field training data: A case study using terrain attributes and semiautomated soil signature derivation to distinguish ecological potential. Applied and Environmental Soil Science. DOI: https://doi.org/10.1155/2011/421904.

7. Brungard C.W., Boettinger J.L., Duniway M.C., Wills S.A., Edwards T.C. 2015. Machine learning for predicting soil classes in three semi-arid landscapes. *Geoderma*. Vol. 239. P. 68–83. DOI: https://doi.org/10.1016/j.geoderma.2014.09.019.

8. Caten A.T., Dalmolin R.S.D., Pedron F.D.A., Ruiz L.F.C., Silva C.A.D. 2013. An appropriate data set size for digital soil mapping in Erechim, Rio Grande do Sul, Brazil. *Revista Brasileira de Ciência do Solo*. 37(2). P. 359–366. DOI: https://doi.org/10.1590/s0100-06832013000200007.

9. Heung B., Ho H.C., Zhang J., Knudby A., Bulmer C.E., Schmidt M.G. 2016. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping. *Geoderma*. Vol. 265. P. 62–77. DOI: https://doi.org/10.1016/j.geoderma.2015.11.014.

10. Heung B., Hodúl M., Schmidt M.G. 2017. Comparing the use of training data derived from legacy soil pits and soil survey polygons for mapping soil classes. *Geoderma*. Vol. 290. P. 51–68. DOI: https://doi.org/10.1016/j.geoderma.2016.12.001.

11. MacMillan R. 2008. Experiences with Applied DSM: Protocol, Availability, Quality and Capacity Building. In: Hartemink A.E., McBratney A., Mendonça-Santos M. (eds) Digital Soil Mapping with Limited Data. Springer, Dordrecht. P. 113–135. DOI: https://doi.org/10.1007/978-1-4020-8592-5_10.

12. Malone B.P., Minasny B., McBratney A.B. 2016. Using R for Digital Soil Mapping. Progress in Soil Science. Springer International Publishing. DOI: https://doi.org/10.1007/978-3-319-44327-0.

13. McBratney A.B., Santos M.L.M., Minasny B. 2003. On digital soil mapping. *Geoderma.* Vol. 117(1-2). P. 3–52. DOI: https://doi.org/10.1016/s0016-7061(03)00223-4.

14. Scull P., Franklin J., Chadwick O. A., McArthur D. 2003. Predictive soil mapping: a review. *Progress in Physical Geography*. 27(2). P. 171–197. DOI: https://doi.org/10.1191/0309133303pp366ra.

15. Walter C., Lagacherie P., Follain S. 2006. Integrating pedological knowledge into digital soil mapping. In: Lagacherie P., McBratney A.B., Voltz M. (Eds.). Digital Soil Mapping: An Introductory Perspective. Vol. 31 of Developments in Soil Science. Elsevier, Amsterdam. Ch. 22. P. 281–301. DOI: https://doi.org/10.1016/s0166-2481(06)31022-7.

16. Florinsky I.V. 2016. Digital Terrain Analysis in Soil Science and Geology. 2 edition. Amsterdam : ACADEMIC PRESS / Elsevier. 506 p. https://doi.org/10.1016/c2015-0-02363-2.

17. Lagacherie P., Robbez-Masson J.M., Nguyen-The N., Barthès J.P. 2001. Mapping of reference area representativity using a mathematical soils cape distance. *Geoderma*. Vol. 101(3-4). P. 105–118. DOI: https://doi.org/10.1016/S0016-7061(00)00101-4.

18. Giasson E., Figueiredo S.R., Tornquist C.G., Clarke R.T. 2008. Digital soil mapping using logistic regression on terrain parameters for several ecological regions in Southern Brazil. In: Hartemink A.E., McBratney A.B., de Lourdes Mendonça-Santos M. (Eds.), Digital Soil Mapping with Limited Data. Springer Netherlands, Amsterdam, Ch. 19. P. 225–232. DOI: https://doi.org/10.1007/978-1-4020-8592-5_19.

19. Kempen B., Brus D.J., Heuvelink G.B.M., Stoorvogel J.J. 2009. Updating the 1:50,000 Dutch soil map using legacy soil data: A multinomial logistic regression approach. *Geoderma.* Vol. 151(3). P. 311–326. DOI: https://doi.org/10.1016/j.geoderma.2009.04.023.

20. Debella-Gilo M., Etzelmüller B., 2009. Spatial prediction of soil classes using digital terrain analysis and multinomial logistic regression modeling integrated in GIS: Examples from Vestfold County, Norway. *Catena.* Vol. 77(1). P. 8–18. DOI: https://doi.org/10.1016/j.catena.2008.12.001.

21. Hengl, T., 2009. A practical guide to geostatistical mapping, 2nd Edition. Office for Official Publications of the European Communities, Luxembourg. URL http://spatial-analyst.net/book/system/files/Hengl_2009_GEOSTATe2c1w.pdf.

22. Cherlinka V.R. 2017. Variations of predictive efficiency of soil maps depending on the methods of constructing educational samples of predicative algorithms. *Ecology and noospherology*. T. 28(3-4). P. 55-71. (Ukr.) http://erae.dp.ua/index.php/erae/article/view/20.

23. Esfandiarpour-Boroujeni I., Shahini-Shamsabadi M., Shirani H., Mosleh Z., Bagheri-Bodaghabadi M., Salehi M.H. 2020. Assessment of different digital soil mapping methods for prediction of soil classes in the Shahrekord plain, Central Iran. *Catena*, Vol. 193. Art.104648. DOI: https://doi.org/10.1016/j.catena.2020.104648 .

24. Kidd D., Searle R., Grundy M., McBratney A., Robinson N., O'Brien L., Jones E. 2020. Operationalising digital soil mapping – Lessons from Australia. *Geoderma Regional*. DOI: https://doi.org/10.1016/j.geodrs.2020.e00335.

25. Batjes N.H., Ribeiro E., van Oostrum A. 2020. Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019). *Earth Syst. Sci. Data*. Vol. 12. P. 299–320. DOI: https://doi.org/10.5194/essd-12-299-2020.

26. Hengl T., Mendes de Jesus J., Heuvelink G.B.M., Ruiperez Gonzalez M., Kilibarda M., Blagotić A. [et al.]. 2017. SoilGrids250m: Global gridded soil information based on machine learning. *PLoS ONE*. Vol. 12(2): DOI: https://doi.org/10.1371/journal.pone.0169748.

27. Chen Z., Ye F., Fu W., Ke Y., Hong H. 2020. The influence of DEM spatial resolution on landslide susceptibility mapping in the Baxie River basin, NW China. *Natural Hazards*, 101. P.853-877. DOI: https://doi.org/10.1007/s11069-020-03899-9.

28. Munoth P., Goyal R. 2019. Effects of DEM source, spatial resolution and drainage area threshold values on hydrological modeling. *Water Resources Management*. Vol. 33(9). P. 3303-3319. DOI: https://doi.org/10.1007/s11269-019-02303-x.

29. Cherlinka V.R. 2017. Influence of resolution of digital relief models on the quality of predicative simulation of soil cover. *Soil Science*. T. 18. №1-2. P. 79-95. URL: http://www.dnu.dp.ua/docs/zbirniki/fbem/program_5cf026b4bbf65.pdf. (Ukr.).

30. Danielson J.J., Gesch D.B. 2011. Global multi-resolution terrain elevation data 2010 (GMTED2010) (p. 26). US Department of the Interior, US Geological Survey. DOI: https://doi.org/10.5066/F7J38R2N.

31. NASA/METI/AIST/Japan Spacesystems, and U.S./Japan ASTER Science Team. ASTER Global Digital Elevation Model V003. 2019, distributed by NASA EOSDIS Land Processes DAAC. DOI: https://doi.org/10.5067/ASTER/ASTGTM.003. Accessed 2020-09-29.

32. NASA JPL. NASA Shuttle Radar Topography Mission Global 1 arc second. 2013, distributed by NASA EOSDIS Land Processes DAAC. DOI: https://doi.org/10.5067/MEaSUREs/SRTM/SRTMGL1.003. Accessed 2020-09-09.

33. Takaku J., Tadono T., Tsutsui K., Ichikawa M. 2018. Quality Improvements of 'AW3D'Global Dsm Derived from Alos Prism. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium. P. 1612-1615. IEEE. URL: https://ieeexplore.ieee.org/document/8518360.

34. EU-DEM Upgrade Documentation EEA User Manual. Available online: https://land.copernicus.eu/user-corner/technical-library/eu-dem-v1-1-user-guide (accessed on 5 September 2020).

35. Cherlinka V.R. Digital elevation models in soil science: theoretical and methodological bases and practical application: Extended Abstract of Dr. Biol. Sciences: 03.00. Dissertation, Chernivtsi, Yuriy Fedkovych Chernivtsi National University, 2019. 538 p. (Ukr.). https://drive.google.com/open?id=1TZubbaD3fNIk7FQUkSyZOdK_dPpozqpN.

36. Kameneva T. 2020. Formation of the National Infrastructure of Geospatial Data in Ukraine and its legal regulation. *Public opinion on lawmaking*. № 9 (194). P. 21–25. (Ukr.) URL: http://nbuviap.gov.ua/images/dumka/2020/9.pdf.

37. Cherlinka V.R. Adaptation of large-scale soil maps to their practical use in GIS. 2015. *Agrochemistry and Soil Science*. Collected papers. No. 84. ISSAR. Kharkiv. P. 20-28. (Ukr.). URL: http://agrochemsoilsci.org/84/84-03.html.

38. Resolution of the Presidium of the National Academy of Agrarian Sciences of Ukraine No. 09/05 of 26.06.2019 (Minutes №9) on the establishment of the Ukrainian Soil Information Center on the basis of the National Scientific Center "Institute for Soil Science and Agrochemistry Research named after O.N. Sokolovsky" (in Ukrainian). URL: http://www.issar.com.ua/downloads/postanova-1.pdf.

39. Laktionova T.M. 2018. The experience of creating and using seven soil databases in the Soil-Geoecophysics Laboratory. *Agrochemistry and Soil Science*. Collected papers. No. 87. ISSAR. Kharkiv. P. 63-71. (Ukr.). DOI: https://doi.org/10.31073/acss87-10

40. Laktionova T.N., Bigun O.N., Nakisko S.G., Uvarenko K.Yu. 2020. Design and functional features of the world's leading soil databases. Analytical review. *Agrochemistry and Soil Science.* Collected papers. No. 89. Kharkiv: NSC ISSAR, P. 4-17. (Ukr.). DOI: https://doi.org/10.31073/acss89-01.

41. GRASS Development Team. 2020. Geographic Resources Analysis Support System (GRASS GIS) Software. Version 7.6. URL: http://grass.osgeo.org.

42. QGIS Development Team, 2020. QGIS Geographic Information System. URL http://qgis.osgeo.org

43. EasyTrace group, 2015. Easy Trace 7.99. Digitizing software. URL http://www.easytrace.com.

44. R Development Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: http://www.r-project.org.

45. Breiman L. 2001. Random forests. *Machine learning* 45(1), 5–32. DOI: https://doi.org/10.1023/A:1010933404324.

46. Cutler A., Cutler D.R., Stevens J.R. 2012. Random Forests. Springer US, Boston, MA, P. 157–175. DOI: https://doi.org/10.1007/978-1-4419-9326-7_5

47. Xu R., Nettleton D., Nordman D.J. 2016. Case-specific random forests. *Journal of Computational and Graphical Statistics*. 25(1). P. 49-65. DOI: https://doi.org/10.1080/10618600.2014.983641.

48. Cherlinka V.R., Dmytruk Y.M., Zaharovskyy V.S. 2017. Comparative estimation of the accuracy of simulation modeling of soil cover and forecast of cartograms of agro-industrial groups. *Biological systems*. Vol. 9. Issue. 2. P. 298-306. http://nbuv.gov.ua/UJRN/biolsist_2017_9_2_23.

49. Cherlinka V.R, Dmytruk Y.M. 2014. Problems of creation, georectification and use of large-scale digital elevation models. *Geopolitics and ecogeodynamics of regions*. Vol. 10(1). P. 239–244. (Ukr.). https://www.researchgate.net/publication/296596582_Problemi_stvorenna_georektifikacii_ta_vikoristanna_krupnomasstabnih_cifrovih_modelej_relefu.

50. Hutchinson M.F. 1995. Interpolating mean rainfall using thin plate smoothing splines. *International journal of geographical information systems*. Vol. 9(4). P. 385-403. DOI: https://doi.org/10.1080/02693799508902045.

51. Wahba G. 1990. Spline models for observational data. Society for industrial and applied mathematics. DOI: https://doi.org/10.1137/1.9781611970128.

52. USGS Earth Explorer. URL: https://earthexplorer.usgs.gov/.

53. ALOS Global Digital Surface Model "ALOS World 3D - 30m (AW3D30)" URL: https://www.eorc.jaxa.jp/ALOS/en/aw3d30/

54. Stevenson J.A, Sun X., Mitchell N.C. 2009. Despeckling SRTM and other topographic data with a denoising algorithm. *Geomorphology*. 144. P. 238-252. DOI: https://doi.org/10.1016/j.geomorph.2009.07.006.

55. Sun X., Rosin P.L., Martin R.R., Langbein F.C. 2007. Fast and Effective Feature-Preserving Mesh Denoising. IEEE. *Transactions on Visualisation and Computer Graphics*. 13(5). P. 925-938. DOI: https://doi.org/10.1109/TVCG.2007.1065.

56. Wright M.N., Ziegler A. 2017. ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software*. 77(1). P. 1–17. DOI: https://doi.org/10.18637/jss.v077.i01.

57. Grinand C., Arrouays D., Laroche B., Martin M.P. 2008. Extrapolating regional soil landscapes from an existing soil map: Sampling intensity, validation procedures, and integration of spatial context. *Geoderma*. Vol. 143(1). P. 180–190. DOI: https://doi.org/10.1016/j.geoderma.2007.11.004

58. Kuhn M. 2008. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software.* Vol. 28(5). P. 1–26. DOI: https://doi.org/10.18637/jss.v028.i05.

59. Landis. J.R., Koch G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*. Vol. 33, No. 1. P. 159–174. URL https://doi.org/10.2307/2529310

60. Li W., Zhang C. 2007. A Random-Path Markov Chain Algorithm for Simulating Categorical Soil Variables from Random Point Samples. *Soil Science Society of America Journal*. Vol. 71(3). P. 656–668. DOI: https://doi.org/10.2136/sssaj2006.0173.

61. Public cadastral map of Ukraine. URL: https://map.land.gov.ua/ (accessed on 5 September 2020) (Ukr.).

62. National (all-Ukrainian) normative monetary valuation of agricultural lands https://ngo.land.gov.ua/uk/map/ (accessed on 5 September 2020) (Ukr.)

63. Cherlinka V.R., Dmytruk Y.M. 2018. Solving existing problems with soil maps in Ukraine. *Biological systems*. Vol. 10(1) P. 298-308. DOI: https://doi.org/10.31861/biosystems2018.01.094.

УДК 631.4:551.4:004.942

Вплив джерел ЦМР на якісні показники предикативних карт ґрунтового покриву

В. Р. Черлінка[1*], Ю. М. Дмитрук[1], Ю. Г. Бодян[2]

[1]**Чернівецький національний університет ім. Ю. Федьковича,** *Чернівці, Україна*
[2]**Чернівецька філія ДУ «Інститут охорони ґрунтів України»,** *Чернівці, Україна*
**E-mail: v.cherlinka@chnu.edu.ua*

Метою дослідження було виявлення впливу цифрових моделей різного походження на якісні характеристики прогнозних карт ґрунтового покриву чи картограм агровиробничих груп ґрунтів за використання технологій предикативного моделювання. Проаналізовано поточну ситуацію з великомасштабними ґрунтовими картографічними даними в Україні та показано, що найбільш швидким та економічно ефективним способом заповнення прогалин при створенні суцільного картографічного покриття для необстежених територій, які складають до 33 % площі України, є математична симуляція. В основі останньої лежить морфометричний аналіз цифрових моделей рельєфу, на основі якого виділяють ряд предикторів, які в подальшому аналізують на предмет наявності зв'язків з існуючими картографічними ґрунтовими матеріалами шляхом створення математичної предикативної моделі з використанням опорних точок ландшафтів та приурочених до них ґрунтових таксонів. Виявлена різниця в якості предикативних матеріалів з використанням індексу kappa Коена (Cohen's *kappa* coefficient) дозволяє рекомендувати окремі джерела ЦМР як базові для такого роду завдань. Проведено демонстрацію замкнутого виробничого циклу створення предикативних грунтових картографічних матеріалів на базі безкоштовного програмного забезпечення – геоінформаційних систем *GRASS* та *Quantum*, мови статистичних розрахунків *R-Statistic* та умовно-безкоштовного – векторизатора *Easy Trace*.

Ключові слова: карта ґрунтів; картограма агровиробничих груп ґрунтів; морфометричні параметри; цифрова модель рельєфу (ЦМР), предикативні алгоритми, моделювання

Форма цитування: Cherlinka V.R., Dmytruk Y.M., Bodyan Y.H. Effect of DEM sources on quality indicators of predictive maps of soil cover. *Agrochemistry and Soil Science*. Collected papers. No. 90. Kharkiv: NSC ISSAR. 2020. P. 36-46. DOI: https://doi.org/10.31073/acss90-04.