AGRICULTURE LAND APPRAISAL WITH USE OF REMOTE SENSING AND INFRASTRUCTURE DATA

Nataliia Kussul[1,2], Andrii Shelestov[1,2], Hanna Yailymova[1,2], Leonid Shumilo[1,3], Sophia Drozd[1]

[1] National Technical University of Ukraine "Igor Sikorsky Kiev Polytechnic Institute", Kyiv, Ukraine
[2]Space Research Institute NASU-SSAU, Kyiv, Ukraine
[3] Department of Geographical Sciences, University of Maryland, College Park, MD, USA

ABSTRACT

1st July 2021 the law on the creation of land market start effect in Ukraine. As a result, land appraisal became cornerstone task in Ukrainian agriculture sector. The official methodology on land appraisal includes use of soil fertility characteristics combined with coefficients related to the distance to the infrastructure objects or settlements and placing of field in specific functional areas, like recreational, or areas with high level of radiation pollution. In this study we collected open source infostructure geospatial information and characteristics of fields obtained from remote sensing data – crop types and Normalized Difference Vegetation Index to build land price predictive model trained on the official land market information. This work designed to investigate potential of geo-informational technologies and remote sensing in the land appraisal use. We separated all available ground truth land price data into three groups by fields size – very small, small, medium and big. We found different relationships between field characteristics and prices. For very small fields the most important features are area, altitude, slope, bonitet and distances to elevators, villages and roads. For small fields the most important are bonitet, altitude, area and distances to cities and roads. For medium and big field's area, slope, distance to cities, roads and historical NDVI.

***Index Terms*—** deep learning, Generative Adversarial Networks, GAN, super-resolution, Sentinel-2.

1. INTRODUCTION

March 31, 2020, Verkhovna Rada of Ukraine (main legislature of Ukraine) passed a low №2178-10 about creation and opening of land market in Ukraine. This low take effect from 1 July 2021 and give possibility to small and big business or common folks to buy or sell agricultural land. The official land appraisal methodology that in use for agricultural land is area-based and consider big set of coefficients related to the geographical position of specific field in regions, relatively to natural-agricultural zones, recreational zones, zones with radiation pollution, administrative units (cities, villages), infrastructure and value of soil fertility quality, called bonitet. However, conditions of open land market create some level of uncertainty. The average price per hectare for fields with similar characteristics can have high variation, due to unforeseen reasons. As an example, if landowner already have land in some specific area, he can pay more for land in the same area, because in this case he would have more economical benefits on use of equipment. And usually if farmer needs to buy few small fields surrounded by his fields to join them in one big field, he will overpay. Or sometime land can be bought for changing of land use type from agriculture to housing and as a result, the price will rise, especially if this land location is favorable for recreational purpose. However, some reasons can be explained by the economic benefits that can be found in the infrastructure or soil quality and can be estimated through the use of GIS technologies and remote sensing data.

Satellite data are widely used for the agriculture monitoring purpose. Right now, it is one of the most valuable sources of historical information. Landsat mission is the longest in history satellite mission for land monitoring started in the 1960s [1]. After adoption of free and open Landsat data policy in 2008 we can see high growth rate of smart agriculture technologies and commercial satellite data-based services development. Today, remote sensing agriculture monitoring systems are very common in the world. USA uses CropScape [2] that provide country-level crop classification maps produced based on the Landsat data by machine learning since 2008. In Ukraine, Space Research Institute NASU-SSAU produce crop classification maps with 10 m. spatial resolution since 2016 with use of Sentinel mission data and deep learning neural networks approach [3]. These data are used on the state level and published on official cadaster web portal within World Bank project "Supporting Transparent Land Governance in Ukraine" [4,5]. In Europe farmers are obliged to provide their field information in the Land Parcel Identification System [6], but European union is still developing and implement agriculture monitoring system based on the satellite data and machine learning crop classification approaches to checking and clarification of submitted by farmers information. This system has name Sen-4-CAP [7] and it is a good example of satellite data usage

for common agriculture policy support. In 2021 d'Andrimont et al [8] study demonstrated big potential of continental level crop classification mapping based on Sentinel mission. The main goal of crop and land cover type maps production in the world is performance of land use audit and accounting, essential for the governmental economy state assessment. However, for the agriculture business it gave additional benefits – now it is possible to defined history of land use for land evaluation and appraisal. A study of relationships between soil bonitet and potential yield, show that bonitet is not reflect actual economical potential of land [9] and it is better to change it to some more relatable characteristics – crop rotation history, agrometeorology or land productivity sub-indicator. Similar conclusion was made in the Zhichkin et al [10] research, where crop rotation history also was proposed as the alternative for bonitet.

Additional characteristic that have high potential for land appraisal is vegetation quality analysis. Normalized difference vegetation index (NDVI) is widely used for the agricultural monitoring and can indicate the economy potential of land. Franch B. et al. [11] research on wheat yield forecasting for 7 main cereals producers in the world shows that using MODIS NDVI data with other remote sensing products it is possible to build model with 5-15% error at national level and 7-20% error at the sub-national level. At the same time, despite the small size of available time-series, high spatial resolution data are also useful in the yield forecasting tasks [12]. In addition, United Nations methodology for sustainable development goals indicator 15.3.1 "Proportion of land that is degraded over total land area" [13] is consider use of NDVI as sub-indicator that showing lad productivity trends and indicate land degradation. UN methodology is based on the MODIS data; however it is possible to use high spatial resolution data [14] for this task and estimate maps that are showing level of degradation and desertification for any field.

2. MATERIALS

In this study we are using information about 15.126 cadastral plots in Ukraine that were sold in 2021. For each field we have his official cadastral number, area, geographical coordinates, and price. Bonitet value for each field estimated with use of official soil type map [9]. To estimate distances to cities we used map of official administrative units of Ukraine that include cities and villages and oblasts. Roads and elevators for the same purpose were obtained from Open Street Map. To estimate land use history, we are using land cover and crop type maps produced by Space Research Institute NASU-SSAU with use of Sentinel-1 and 2 data, ground-truth data collected through ground surveys along the roads in Ukraine and deep neural network model runed in the Amazon cloud environment [15]. We also used year maximum NDVI values for fields calculated based on the Sentinel-2 mission data for 2019-2021.

3. METHODOLOGY

To conduct experiment, we are using correlation analysis and regression analysis approach for statistical analysis of dependences between agriculture field prices and characteristics. For correlation analysis we are using Pearson correlation coefficient that shows strength of linear dependences between x and y vales:

$$r = \frac{(\sum xy) - \sum x \sum y}{\sqrt{(\sum x^2 - (\sum x)^2)(\sum y^2 - (\sum y)^2}},$$

For regression analysis we are using linear regression function to build multivariable model for land appraisal and estimate coefficient of determination R^2:

$$y = a_0 + \sum a_i * x_i,$$

$$R^2 = 1 - \sum \frac{(y_i - y_f)^2}{(y_f - \Delta y_{fi})^2},$$

The country level data are very heterogeneous, so to reduce variativity we split data set into 3 main groups – very small fields (<0.2 ha), small fields (0.2-2 ha) and medium and big fields (>2 ha). As a result, we are building one model for each group. In each group data also split for train and validation with sizes 40% and 60%. Result of correlation analysis in the regression model is used for definition of informative regressors for the model to avoid overfitting.

4. RESULTS

Firstly, we conduct analysis of potential regressors informativity for further model fitting. Table 1 shows result of this experiment. As we can see, groups are differed by dependences between variables. Also, interesting fact that can be seen – for small and very small fields there is no correlation between NDVI values, meanwhile for medium and large this correlation can be observed. Some of visible trends can be explained by the land use purpose and agricultural technologies. As a example – small and very small fields usually owned by small busines and it is important to them have short distance to elevators and primary roads. Meanwhile big busines have no problems with logistic and main interest is in the areas, economic benefits, and possible yield of fields. It also explains, why NDVI is informative only for big fields. In addition, we can see that as for other markets, landowners by buying land with bigger prices are getting better prices. However, additional factor should be further investigated – some of this agriculture lands are owned to change land use purpose. It gives a big bias in the data and produce high rate of error for fields with small and very small sizes, because land purchased for housing have higher price then land for agriculture.

Table 1. Pearson's correlation coefficient between regressors and fields prices

Regressor \ Class	Very small	Small	Medium and big
Area	0,96	-0,18	-0,54
Bonitet	0,13	0,34	0,12
NDVI_2019	0,059	0,033	0,4
NDVI_2020	-0,0097	0,064	-0,12
NDVI_2021	-0,006	-0,037	-0,28
Altitude	0,25	0,19	0,1
Slope	-0,16	-0,054	0,45
Distance to cities	-0,044	0,11	0,038
Distance to towns	-0,083	0,057	-0,023
Distance to villages	-0,12	-0,072	-0,55
Distance to elevators	-0,13	-0,047	-0,027
Distance to primary roads	-0,12	0,052	0,38
Distance to secondary roads	-0,083	-0,11	0,061
Distance to country roads	-0,098	-0,12	-0,047
Land cover type class	-0,01	-0,034	-
Natural-agricultural zones	0,28	0,53	0,389

After this we defined regressors for the land appraisal models. For very small fields we are using area, bonitet, altitude, slope, distance to village, elevator, primary roads, and natural agricultural zones. For small fields we are using area, bonitet, altitude, distance to cities, distance to secondary roads and country roads and natural-agricultural zones. For medium and big fields we are suing area, bonitet, NDVI for 2019-2021, altitude, slope, distance to village, primary roads, and natural-agricultural zones. We see that the price of very small fields is very hard to predict. R^2 for small fields is 0.17 – more than 2 times lower in comparison with small sizes (0.36) and 3 times in comparison with medium and big fields (0.53). It can be explained by the fact that very small fields are include the biggest percent of fields purchased for lad use change. At the same time, medium and big fields have high determination coefficient (figure 1) and low RMSE (figure 2). It means that very accurate appraisal of agricultural fields with size bigger then 2 ha is possible.

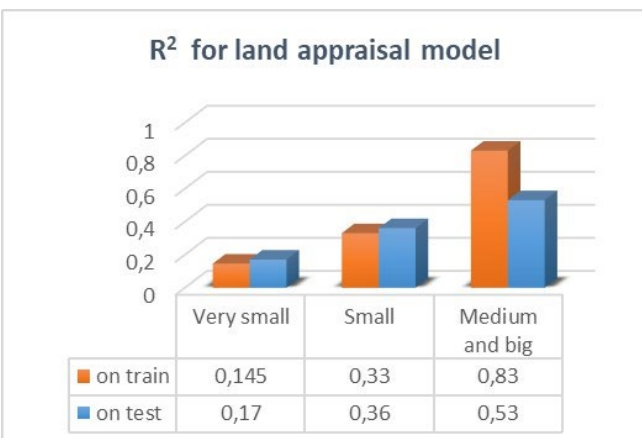

Figure 1. R^2 for country level models for land appraisal

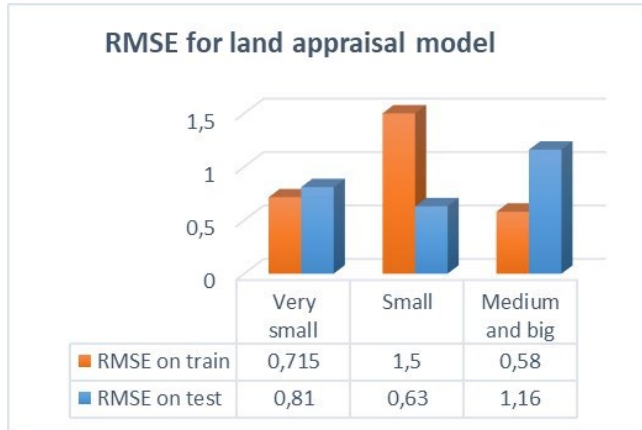

Figure 2. RMSE for country level models for land appraisal

However, land price is very dependent on the region of country. Table 2 shows correlation coefficients between used for land appraisal characteristics and prices of all fields in different regions of Ukraine. As we can see variation and informativity of coefficients is strongly dependent on the geographical location in administrative borders of country. If area in Luhansk region have 92% correlation, it can have -73% in Chernivtsi or if distance to cities is -98% it can be 81 in Ivano-Frankivsk. So, to improve quality of models it is essential to also consider this geographical socio-economical aspect. But such model require big number of recorded data for each region and each field size group.

Table 2. Pearson's correlation coefficient between regressors and price for the region of Ukraine

Region	Area	Bonitet	NDVI_2019	NDVI_2020	NDVI_2021	Altitude	Slope	Distance to cities	Distance to towns	Distance to villages	Distance to elevators	Distance to primary roads	Distance tsecondaryroads	Distance to country roads	Land cover type class	Natural-agricultural zones
Kirovohrad	0,06	0,04	-0,08	0,07	0,09	0,24	-0,18	0,09	-0,07	-0,08	-0,21	-0,21	0,0059	-0,082	0,12	0,19
Luhansk	0,92	-0,6	-0,68	-0,8	-0,37	-0,01	-0,98	-0,19	-0,93	0,51	0,48	0,9	-0,24	-0,3	-	0,4
Lviv	0,02	0,24	0,29	-0,1	0,02	0,01	-0,03	0,09	0,06	-0,05	-0,04	-0,1	-0,35	-0,053	0,17	0,16
Mykolaiv	0,53	0,19	0,2	0,18	0,02	0,44	-0,32	-0,21	0,09	0,04	-0,35	0,06	0,31	0,21	-0,67	-0,087
Odessa	-0,03	-0,16	-0,69	-0,14	0,22	0,6	0,48	0,22	0,2	0,53	0,23	0,6	0,29	-0,13	-	-0,1
Poltava	0,05	0,07	0,02	-0,03	-0,11	0,24	0,03	0,07	0,01	-0,08	-0,1	0	0,0029	-0,12	-0,03	0,31
Rivne	-0,28	0,09	-0,13	-0,23	-0,29	-0,12	-0,21	-0,28	0,13	0,01	-0,49	-0,22	0,12	-0,35	-0,11	-0,13
Sumy	0,34	0,3	-0,3	-0,26	-0,27	0,46	-0,24	0,08	-0,3	-0,16	-0,15	-0,27	-0,12	-0,013	-0,056	0,12
Ternopil	0,11	0,17	0,16	-0,19	0,04	-0,51	-0,31	-0,24	0,14	-0,18	0,03	0,1	-0,15	-0,021	-	0,53
Kharkiv	0,41	0,06	-0,21	-0,16	-0,25	0,02	-0,41	0,32	0,13	-0,06	0,25	0,15	0,0049	0,37	-0,14	0,089
Kherson	-0,03	-0,03	0,04	0,03	0,15	0,42	-0,26	0,13	0,05	-0,06	0,1	0,31	-0,031	0,13	0,0011	0,3
Khmelnytsky	0,03	0,3	-0,07	-0,08	-0,1	-0,03	0	0,15	-0,06	-0,19	-0,08	-0,3	0,06	-0,17	-0,22	0,33
Cherkasy	0,12	0,39	-0,14	-0,25	-0,08	-0,01	-0,3	0,23	0,02	0,11	0,14	0,2	0,036	-0,053	0,063	0,2
Chernivtsi	-0,73	-0,33	0,29	0,39	0,12	0,3	0,3	-0,43	-0,29	0,14	-0,09	0	-0,067	-0,031	0,039	-0,031
Chernihiv	0,23	0,18	-0,24	-0,3	-0,29	0,2	-0,08	0,14	0,09	0,19	0,1	0,18	-0,029	-0,05	-0,43	0,081
Vinnytsia	-0,06	0,15	-0,28	-0,07	-0,16	-0,05	-0,09	-0,07	0,16	-0,18	0,28	0,41	-0,021	-0,11	0,2	0,17
Volyn	0,15	0,3	-0,04	0,05	-0,06	0,5	0,28	0,07	-0,43	-0,16	-0,13	-0,21	-0,076	-0,043	0,0096	0,54
Dnipropetrovsk	0,19	-0,83	0,12	0,07	0,19	0,1	-0,23	0,12	0,1	-0,13	0,09	0,12	3,20E-05	0,26	-0,24	0,44
Donetsk	0,34	-0,22	-0,17	-0,01	-0,44	0,1	-0,23	0,17	-0,2	0,26	0,27	-0,21	-0,084	-0,068	-0,16	-0,11
Zhytomyr	-0,19	0,32	0,19	0,13	0,19	-0,17	0,25	-0,15	0,02	-0,04	-0,2	-0,03	0,028	-0,028	-0,22	0,35
Transcarpathia	-0,5	0,16	-0,16	0,01	0,18	0,2	-0,13	-0,2	0,32	-0,52	0,22	-0,11	0,14	0,0097	-	-0,3
Zaporozhye	-0,34	0,01	0,03	0,18	0,17	-0,14	-0,35	-0,14	-0,12	-0,29	-0,13	-0,01	-0,24	-0,082	-0,28	0,3
Ivano-Frankivsk	-0,01	-0,21	0,13	-0,1	-0,07	0,88	0,81	0,78	-0,09	-0,42	0,87	0,43	0,084	0,064	0,54	-0,71
Kyiv	-0,16	0,03	-0,09	-0,08	-0,05	0,05	-0,26	0,03	-0,04	-0,16	-0,16	0,05	-0,19	-0,029	-0,3	0,15

5. DISCUSSION AND CONCLUSIONS

Correlation analysis for very small fields shows that the strongest dependence for price is area (96%). For land appraisal model also can be used altitude, slope, bonitet,

distance to towns, villages, elevators, primary roads that have absolute values between 12% and 28. For small fields, the strongest correlation is with Natural-agricultural zones (58%) and bonitet (34). Additional possible regressors in this case are area, altitude, slope and distances to cities, secondary and country roads (absolute values from 11% to 19%). Medium and big fields have the strongest correlation with distance to village (-55%) area (-54%), slope (45%), distance to primary road (38%) and natural agriculture zone (38%). Negative correlation with area can be explained by the fact that in Ukraine a lot of citizens are interested in the selling of their land and ready to lower the price for owner who ready to buy big areas. Also, we see that these regressors are very dependent on the region in which they are located. The number of samples is not giving possibility right now to build statistically significant models for each region, however correlation analysis shows that in addition to splitting by size it is better to split fields by regions to build accurate land appraisal models. The R^2 for land appraisal models that we obtained on country level are 0.17, 0.36 and 0.53 respectively to size, so we see that the larger the size of the field, the more stable is price.

6. ACKNOWLEDGMENT

The authors acknowledge the funding received by the National Research Foundation of Ukraine from the state budget 2020/01.0273 “Intelligent models and methods for determining land degradation indicators based on satellite data” (NRFU Competition “Science for human security and society”).

7. REFERENCES

[1] Loveland, T. R., & Dwyer, J. L. (2012). Landsat: Building a strong future. Remote Sensing of Environment, 122, 22-29.

[2] Han, W., Yang, Z., Di, L., & Mueller, R. (2012). CropScape: A Web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. Computers and Electronics in Agriculture, 84, 111-123.

[3] Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. (2017). Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14(5), 778-782.

[4] Supporting Transparent Land Governance in Ukraine - https://ukraine-landpolicy.com/

[5] Kussul, N., Shelestov, A., Lavreniuk, M., Kolotii, A., & Vasiliev, V. (2019, July). Transparent land governance in Ukraine within world bank program. In 2019 IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON) (pp. 1077-1080). IEEE.

[6] Kocur-Bera, K. (2019). Data compatibility between the Land and Building Cadaster (LBC) and the Land Parcel Identification System (LPIS) in the context of area-based payments: A case study in the Polish Region of Warmia and Mazury. Land Use Policy, 80, 370-379.

[7] De Vroey, M., Radoux, J., Zavagli, M., De Vendictis, L., Heymans, D., Bontemps, S., & Defourny, P. (2021, July). Performance Assessment of the Sen4CAP Mowing Detection Algorithm on a Large Reference Data Set of Managed Grasslands. In 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS (pp. 743-746). IEEE.

[8] d'Andrimont, R., Verhegghen, A., Lemoine, G., Kempeneers, P., Meroni, M., & van der Velde, M. (2021). From parcel to continental scale--A first European crop type map based on Sentinel-1 and LUCAS Copernicus in-situ observations. arXiv preprint arXiv:2105.09261.

[9] Shumilo, L., Lavreniuk, M., Skakun, S., & Kussul, N. (2021). Is Soil Bonitet an Adequate Indicator for Agricultural Land Appraisal in Ukraine?. Sustainability, 13(21), 12096.

[10] Zhichkin, K., Nosov, V., Zhichkina, L., Zhenzebir, V., & Sagina, O. (2020). Cadastral appraisal of lands: agricultural aspect. In IOP Conference Series: Earth and Environmental Science (Vol. 421, No. 2, p. 022066). IOP Publishing.

[11] Franch, B., Vermote, E., Skakun, S., Santamaria-Artigas, A., Kalecinski, N., Roger, J. C., ... & Sobrino, J. A. (2021). The ARYA crop yield forecasting algorithm: Application to the main wheat exporting countries. International Journal of Applied Earth Observation and Geoinformation, 104, 102552.

[12] Skakun, S., Kalecinski, N. I., Brown, M. G., Johnson, D. M., Vermote, E. F., Roger, J. C., & Franch, B. (2021). Assessing within-Field Corn and Soybean Yield Variability from WorldView-3, Planet, Sentinel-2, and Landsat 8 Satellite Imagery. Remote Sensing, 13(5), 872.

[13] UN stats SDG 15.3 metadata - https://unstats.un.org/sdgs/metadata/?Text=&Goal=15&Target=15.3

[14] Kussul, N., Lavreniuk, M., Kolotii, A., Skakun, S., Rakoid, O., & Shumilo, L. (2019). A workflow for Sustainable Development Goals indicators assessment based on high-resolution satellite data. International Journal of Digital Earth.

[15] Shelestov, A., Lavreniuk, M., Vasiliev, V., Shumilo, L., Kolotii, A., Yailymov, B., ... & Yailymova, H. (2019). Cloud approach to automated crop classification using Sentinel-1 imagery. IEEE Transactions on Big Data, 6(3), 572-582.