Development of Land Valuation System using Machine Learning algorithms

Natalia Novosad

Ukrainian Catholic University, Lviv, Ukraine
nataliia.novosad@ucu.edu.ua

Abstract. Evaluating the real market value of land is a complicated and expensive process carried out by experts. The data we are considering is the land in Ukraine. Due to the land moratorium, the farmland in Ukraine is usually underestimated. The objective is to find its fair value. This paper aims to compare the different Machine Learning models and classical econometric models in three approaches: the income method, the comparative method, and the real options method.

Keywords: Machine Learning · Land Valuation · Pricing · Economic Factors · Descriptive Analytics.

1 Introduction

The land is one of the most crucial assets for the country. The policy of land management impacts the country's growth and development. Here we consider farmland as it is the key factor for production and economy. The existence of a farmland market ensures the optimal use of the asset. The land belongs to the farmers and investors that are interested in the productive usage of the asset. It causes business growth and, therefore, the country's growth.

Since 1992, the farmland market in Ukraine is closed. The moratorium was first signed in 2001 and has been continued ten times. It limits the development of the country and generates significant losses. In 2001 the agriculture formed 16% of GDP, and in 2018, only 10%. Even though the quality of land in Ukraine is high, the production (and therefore the prices) is low (Fig. 1,2). Due to the new law signed in March 2020, the moratorium is lifted from June 2021. We believe that it will open new opportunities for farmers.

The main challenge that will come up is the land valuation. As was said before, the farmland now has low prices compared to other countries. So we need to apply advanced methods to find its fair value. Developing a fair land evaluation system will help landowners earn more and manage their assets consciously and wisely.

This paper aims to compare different land valuation approaches (classic ones and ML) to estimate the land value fairly. All the methods are known for a long time, but we propose to use them specifically for Ukraine, which was not done before. Our task is to find the most appropriate method in Ukrainian reality,

where not all the data is fully available. We expect to use the nearest counties' data, the dynamics of weather, soil characteristics, rent prices, raw material costs, and others.

The remaining part of the paper includes the overview of existing methods in Section 2. Section 3 outlines the proposed approach. Finally, Section 4 concludes the paper.

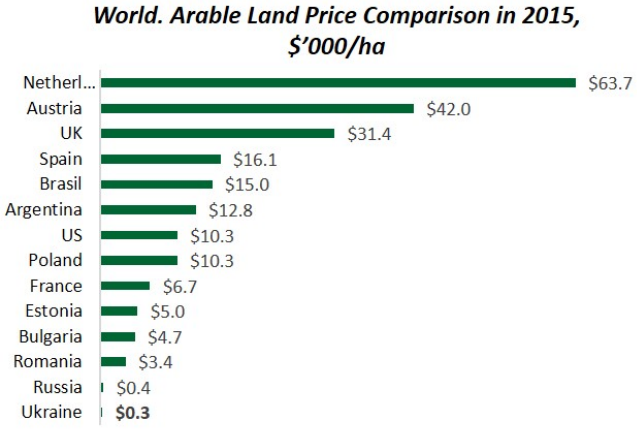


Fig. 1: The prices of arable land in different countries in 2015
Source: Eurostat

2 Related work and background

2.1 Normative farmland valuation methodology in Ukraine

Since the farmland in Ukraine has no market price, the government set a normative farmland valuation methodology. It is used to determine land tax, state duty, inheritance and donation of land, and others. The actual methodology was set in 2016 and is the following[1]:

$$NV = \sum(PA \times NVA) + P \times R \tag{1}$$

where
NV - normative farmland valuation (UAH)
PA - area of agricultural soils (hectare)
P - area of non-agricultural soils (hectare)
R - standardized rental income of non-agricultural land (UAH/hectare)

[1] https://zakon.rada.gov.ua/laws/show/831-2016

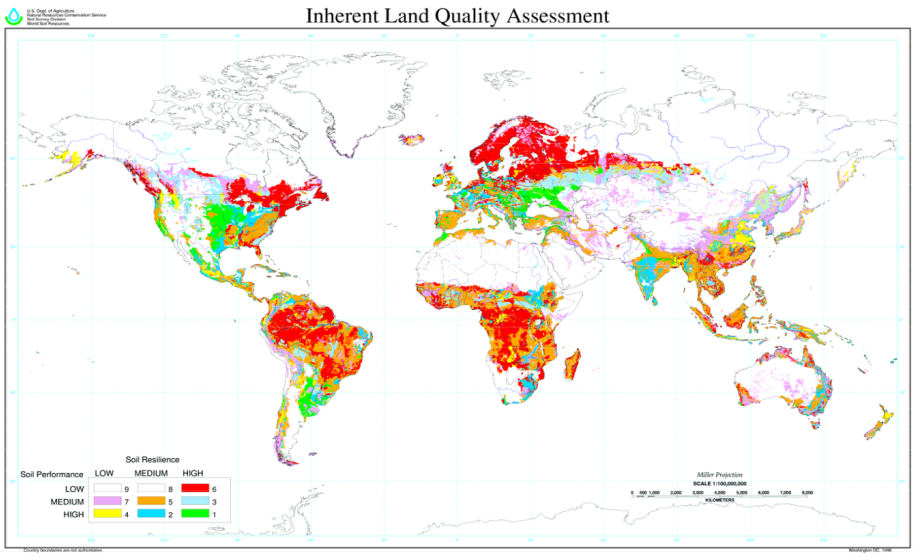


Fig. 2: World soil performance
Source: Natural Resource Conservation Service, USDA

NVA - normative valuation of agricultural land (UAH/hectare)
and the NVA is:

$$NVA = RA \times Q : QS \tag{2}$$

where
RA - standardized rental income of agricultural land (UAH/hectare)
Q - quality score of soil in the region
QS - average quality score of soil in Autonomous Republic of Crimea region, Kyiv city and Sevastopol city

This valuation is very rough and depends on 'standardized rental income' fixed and generalized for whole regions. The new project is developed, and the proposed methodology of normative valuation of any land is[2]:

$$NV = PA \times RA \times K1 \times K2 \times K3 \times K4 \times Kp \times Ku \times Ki \tag{3}$$

where
NV - normative farmland valuation (UAH)
PA - area of agricultural soils (hectare)
RA - standardized rental income of agricultural land (UAH/hectare)
$K1$ - coefficient that takes into account distance to big cities
$K2$ - coefficient that takes into account the resort and recreational value of settlements
$K3$ - coefficient that takes into account distance to zones of radiation pollution

[2] https://land.gov.ua/info

$K4$ - coefficient that takes into account the location
Kp - coefficient that takes into account the purpose of the land
Ku - coefficient that takes into account the specifics of land usage
Ki - indexation coefficient

We see that the new methodology includes a lot of pre-calculated indexes. Even if the indexes are calculated correctly, they contain only the baseline scenario and data at the moment of calculation. Therefore, this methodology is not adaptive to the situation. Furthermore, it is as generalized as the first methodology and does not consider the specifics of land.

2.2 Estimation methods of farmland value

As an alternative to normative valuation, the expert valuation exists. It is expensive and requires certain skills and experience. Usually, their valuation is based on classical methods. Each of them has advantages and disadvantages. The choice of the method depends on the available data and the situation.

Comparative method. The idea of this method is to evaluate the land based on the value of similar land. The similarity is measured with different properties: the area and shape, the quality of soils, the location, socio-demographic factors, climate, and others. So the straight-forward way here is to use the mean value of the similar lands:

$$VC = \frac{1}{n} \sum_{i=1}^{n} VC_i \tag{4}$$

where
VC - monetary valuation of the land by comparative method (UAH)
VC_i - monetary valuation of the similar land plots (for example those within a radius of 10 km)

However, this evaluation has disadvantages: choosing the number of land plots to take or what factors take into account for similarity measure. Therefore, another approach is to use regression analysis. The main idea is to find such a function that will predict the land's value based on its characteristics (Eq. 5). Such models have to use the most recent data to make an up-to-date estimation. What is more, we need a lot of data to cover the space of features. Also, we cannot estimate the value of land with properties that were not seen before. For example, if we build the model calibrated on land located in a sunny climate, we cannot use the same model for rainy weather. So the availability of the most recent data and its amount plays a crucial role in the comparative method. An example of using this method for real estate is [2].

$$VC = f(x_1, x_2, x_3, \ldots, x_n) \tag{5}$$

where
VC - monetary valuation of the land by comparative method (UAH)
$x_1, x_2, x_3, \ldots, x_n$ - land characteristics
f - any function from characteristics space to monetary valuation

Income method. In this case, we estimate the possible revenue that the land can produce for a buyer. The revenue is usually estimated with discounted cash flows (DCF analysis). The evaluation formula for the income method is shown in Eq. 6. The main task here is to predict the future income for some period taking into account the risks and opportunities. The source of income can be different: the rent payments, the selling of harvest, building the infrastructure.

$$VI = I/r \tag{6}$$

where
VI - monetary valuation of the land by income method (UAH)
I - net income (UAH)
r - capitalization rate that include risk-free interest rate, industry risk, liquidity risk and others

In [1] the authors used the rental price of land in Ukraine and estimated the growth rate based on past progress. They compared the resulting farmland value with the normative one and concluded that the model overestimates the value (or is more likely the normative approach underestimates it). They simplify the real-life a lot and used average region values, but it was enough to show the land market's importance.

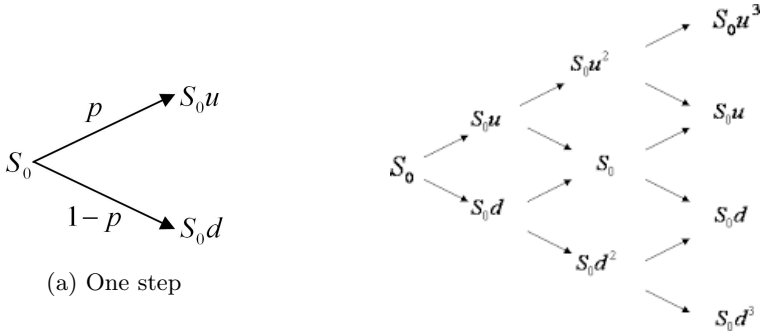
Cost method. It is also called the development method. The method is based on the idea that land price now is equal to the price of the same land with the potential development minus the costs for this development. In other words, it is equal to the residual after some development was build and sold. In this case, the challenge is to evaluate the costs and future demand and supply on made improvement.

Real option method. Real options are used to estimate the value of future opportunities. An example of options is the development method discussed before: the option to wait when to build the development also has a price. This method has an advantage over the income method (DCF analysis) because it considers the active management [8]. Buying the land, the investor has at least two options: when to use the land and how to use the land. In [3] the option of what crop should be seeded on the farmland is discussed.

The difficulty in option valuation is not to figure out expected cash flows but to discount them according to uncertainty. Therefore the valuation of options needs a different framework than DCF analysis. The two most common option valuation methods are the Binomial Method and the Black-Scholes formula. They are based on one-step valuation: the asset price can go up or down at maturity time (for example, one year) with unknown probabilities (Fig. 3a). Two approaches can be taken further that will end up with the same value [8]:

1. constructing the replication portfolio (based on the principle that payoffs from the levered investment in the stock are equal to the payments from the call option);

2. risk-neutral valuation (finding unknown probability such that there is no arbitrage on the market assuming risk-neutral investors)



(a) One step

(b) Multiple step

Fig. 3: The Binomial Method for option valuation.

But the options change faster than once per year. Therefore we need to take more steps. The Binomial method constructs the binomial tree with n steps (Fig. 3b). The n can go to infinity as we assume smaller and smaller periods. In 1973 Black and Scholes derived the formula for $n \to \infty$ [9]. To find the option value by the Black-Scholes formula, we need the following parameters: stock price, exercise price, a standard deviation of annual returns, years to maturity, interest rate.

Although the described models were constructed for stock options, the Black-Scholes model can be used for real options as well. In this context, the parameters will be the following: discounted future cash flows as a stock price, future investments as an exercise price, standard deviation of the industry, years to maturity, and interest rate [8]. The example of using the Black-Scholes formula and the Binomial model to find the real estate value is [7].

2.3 Machine learning approaches

The ML algorithms work very well on classification and regression. It can solve the tasks that have defined frames. It requires a lot of labeled data, and it is hard to explain. Never-the-less the ML proved to be the most accurate approach because of its ability to capture the nonlinearities.

Comparative method. This is the same method discussed above, but applying the ML algorithms as a function f (Eq. 5): Neural Network, Random Forest and Support Vector Machine and others. Using ML, we solve the task of regression. ML was used to evaluate the housing market, real estate, and land price [4,5]. It was proved that the ML approach could give better performance when the task is easy enough (like regression) and has many data.

Forecasting. ML can be used to forecast time series, for example, inflation or corn prices (Fig. 4). Forecasting is not the final estimate of land value, but it can be helpful in classical methods.

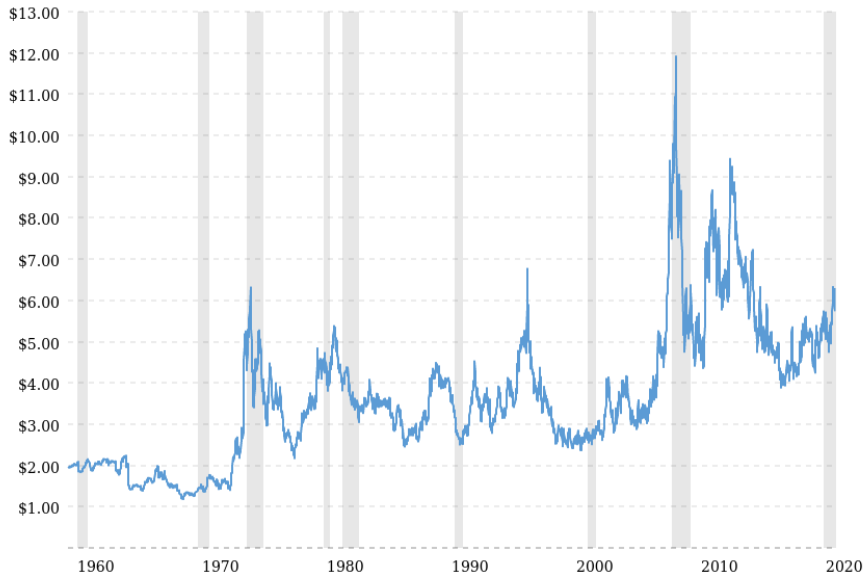


Fig. 4: Wheat Prices - 40 Year Historical Chart
Source: https://www.macrotrends.net/2534/wheat-prices-historical-chart-data

Alternative data. Using alternative data (texts and articles) makes investment decisions better, but this topic is not thoroughly researched. Also, the research on predicting poverty with satellite imagery [6] showed promising results. Alternatively, satellite images can be used to evaluate the quality of the land.

3 Research proposal

3.1 Problem setting

The thesis paper aims to test and compare the comparative method (classical methods and ML), real options method, income method, and normative valuation of farmland in Ukraine. Based on the background and motives, we assume that the land's actual prices are underestimated, and we want to find a reasonable approach for fair valuation.

3.2 Data

The data collection is still in process. It is a big challenge because not all of the data in Ukraine is easily accessible. The primary resource is The State Service of Ukraine for Geodesy, Cartography and Cadastre[3]. Already found data:

- the rent prices for land (appendix A)
- normative valuation [4]
- soil characteristics [5]
- historical data about harvest of different crops [6]

For the land valuation in Europe, to compare with Ukraine, we will use the Eurostat[7]. Also we will use weather data and historical crop prices.

3.3 Proposed Solution

The research workflow starts with data collection. After that, we plan to implement the comparative method, real options method, and income method. We will base the implementation itself on existing literature. However, we will test it on the completely new data.

3.4 Evaluation

The baseline for evaluation will be normative farmland valuation methodology. The results will be compared with the baseline: if methods underestimate or overestimate the baseline value and how much.

3.5 Timelines

The complete picture of the research pipeline is in Fig. 5. We plan to collect all the necessary data before February. The implementation starts in March and the Evaluation in the middle of April.

4 Conclusion

We consider the problem of underestimated land in Ukraine as the result of a twenty-year-long moratorium. We propose the methods to find the fair value of farmland and the roadmap for the research project. The paper's main objective is to use well-known methods in new Ukrainian realities.

[3] https://land.gov.ua
[4] https://zakon.rada.gov.ua/laws/show/831-2016
[5] https://data.gov.ua/dataset/23499cbc-326c-4917-8f93-2bff444bf918
[6] http://ukrstat.gov.ua/druk/publicat/Arhiv_u/07/Arch_sg_zb.htm
[7] https://ec.europa.eu/eurostat/web/agriculture/data/database

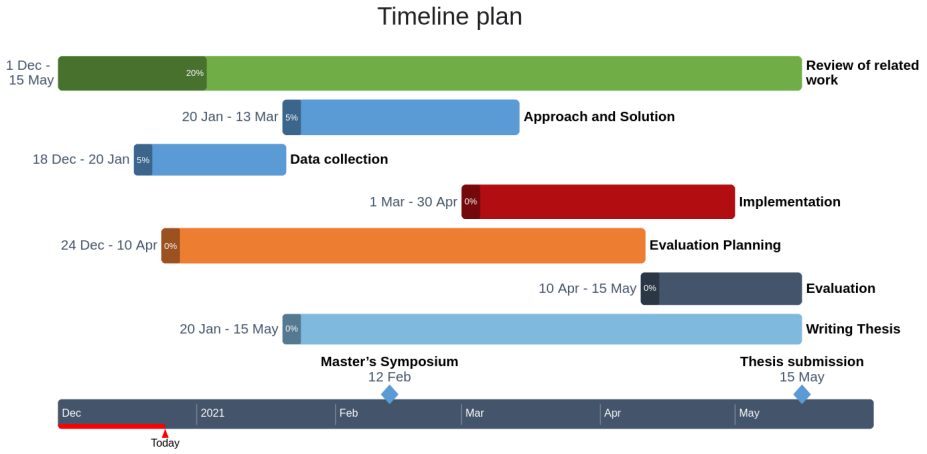


Fig. 5: The timeline plan of writing Master's Thesis

Acknowledgements

We want to thank Yevhen Pentsak, the Professor at Kyiv-Mohyla Business School, for consulting and bringing value to the research. We are grateful to CNEWA, the Vatican agency, for financial support during my Master's studies.

References

1. Nivievskyi, O., Kandul, S.: The Value of Farmland - Expected Farmland Prices in Ukraine after lifting the Moratorium on Farmland Sales.(2011) https://www.researchgate.net/publication/282119641
2. Minkyu, C., Taeha, Y., Meereh, K., Ji-Hyun L.: Land Price Prediction System using Case-Based Reasoning. (2019) https://www.semanticscholar.org/paper/Land-Price-Prediction-System-using-Case-Based-Choi-Yi
3. Moreno, M., Navas, J. F., Todeschini F.: Land valuation using a real option approach. (2009) https://www.researchgate.net/publication/28318918
4. Masías, V.H., Valle, M.A., Crespo, F., Crespo, R., Vargas, A., Laengle, S.: Property Valuation using Machine Learning Algorithms:A Study in a Metropolitan-Area of Chile. (2016) https://www.researchgate.net/publication/317137020
5. Sampathkumara, V., Santhib, M.H., Vanjinathanc, J.: Forecasting the land price using statistical and neural network software. (2015) https://www.researchgate.net/publication/283185838
6. Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B, Ermon, S.:Combining satellite imagery and machine learning to predict poverty. Science Vol. 353, Issue 6301, pp. 790-794 (2016)

7. Čirjevskisa, A., Tatevosjansb, E.:Empirical Testing of Real Option in the Real Estate Market. (2015)
https://www.sciencedirect.com/science/article/pii/S2212567115006115
8. Brealey, R., Myers, S.: Principles of Corporate Finance. (2019)
9. Black F., Scholes M.: The Pricing of Options and Corporate Liabilities. The Journal of Political Economy, Vol. 81, No. 3, pp. 637-654 (1973).

A The average rent price for agricultural land plots of state property purchased for rent on a competitive basis (land auctions)

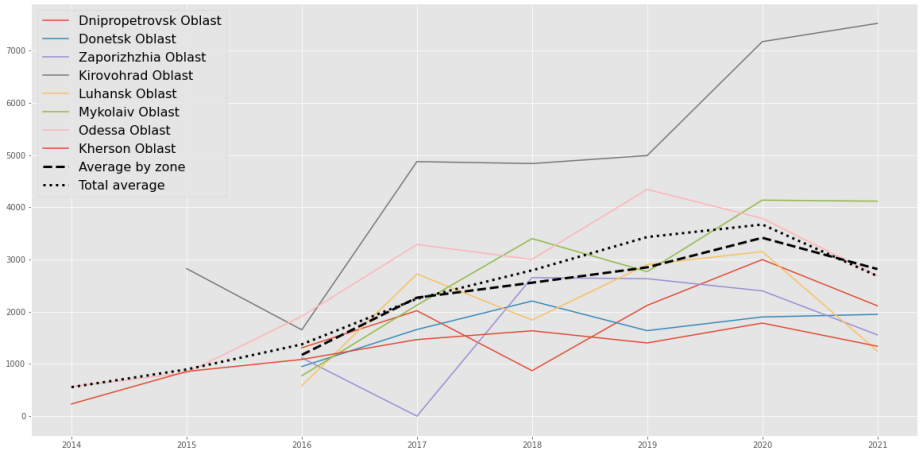


Fig. 6: Average rent price in Steppe climate zone

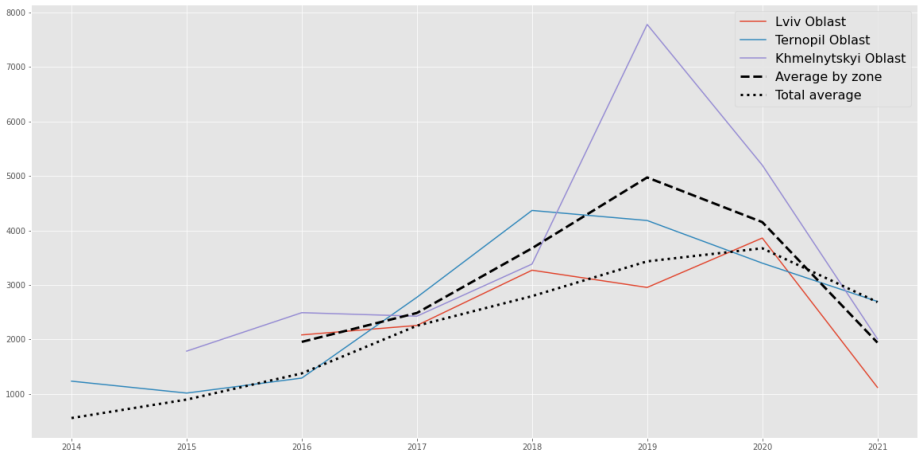


Fig. 7: Average rent price in Forest climate zone

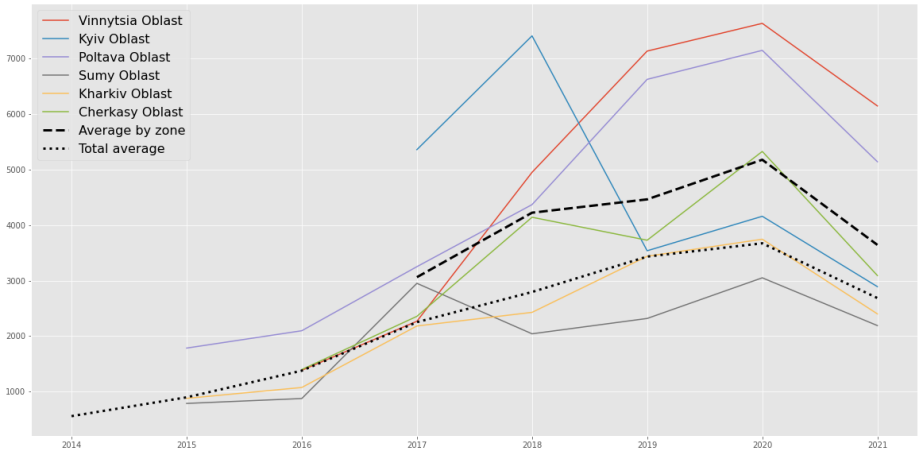


Fig. 8: Average rent price in Forest-steppe climate zone

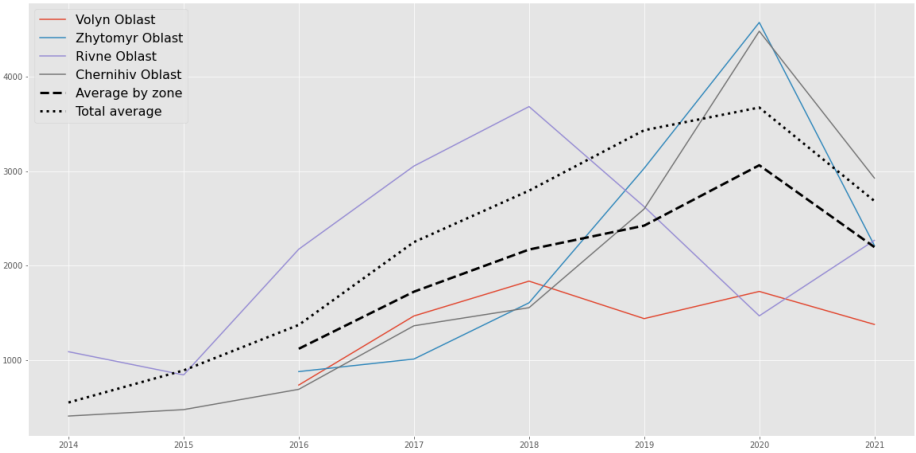


Fig. 9: Average rent price in Polissya climate zone

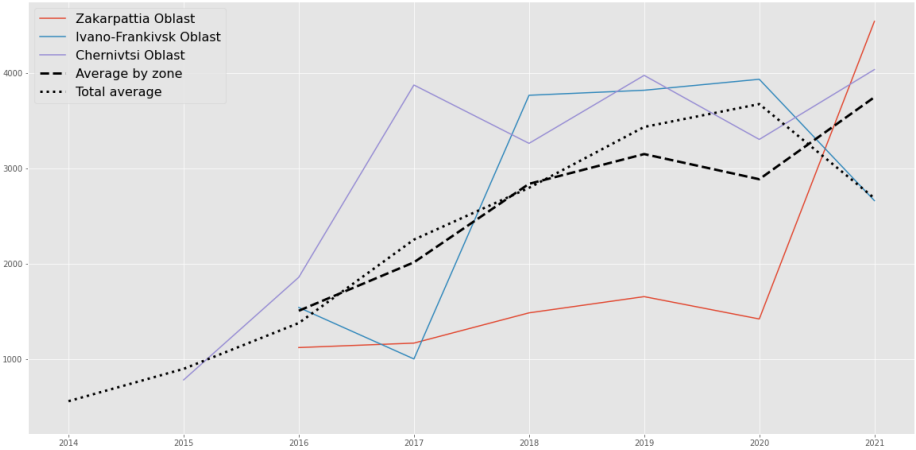


Fig. 10: Average rent price in Carpathians climate zone