LAND MARKETS UNDER STRESS: DETERMINANTS OF AGRICULTURAL LAND VALUE IN WARTIME UKRAINE

by

Yevheniia Mitriakhina

A thesis submitted in partial fulfillment of

the requirements for the degree of

MA in Economic Analysis

Kyiv School of Economics

2025

Thesis Supervisor: ______________________ Professor Oleh Nivievskyi

Approved by ____

____

____

____

____

Date ______________________________

Kyiv School of Economics

Abstract

LAND MARKETS UNDER STRESS: DETERMINANTS OF AGRICULTURAL LAND VALUE IN WARTIME UKRAINE

by Yevheniia Mitriakhina

Thesis Supervisor: Professor Oleh Nivievskyi

Ukraine's market for the sale of agricultural land opened in July 2021, following the end of a long-standing moratorium. Less than a year later, the country faced a full-scale Russian invasion, introducing extraordinary uncertainty into the newly liberalized market. This thesis examines the determinants of agricultural land value in Ukraine during this period of transition and wartime disruption, including the anticipated second phase of land reform extending access to the market to legal entities in 2024.

The analysis draws on a comprehensive dataset of verified transactions, with economic, spatial, and geographic variables constructed and merged at the plot level. A spatial hedonic pricing model is estimated, incorporating a spatial lag component to capture price interdependencies across neighboring areas.

Results show that official administrative valuation is the strongest predictor of land price, reflecting the anchoring role of the price floor – one of the key market restrictions. Wartime transactions are associated with a 29% discount, while the 2024 reform allowing legal entity participation in the market corresponds to a 4.3% price premium. Other key determinants include proximity to urban centers, infrastructure access, land use classification, and distance from occupied territories. The spatial lag term confirms significant spillover effects across the land market.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor, Prof. Dr. Oleg Nivievskyi, for his help in choosing the topic and for the invaluable feedback throughout the research process. I am also deeply grateful to the KSE faculty, whose academic rigor and high expectations defined the entire experience of this degree. I also wish to acknowledge the UCU, where I completed my undergraduate studies, for laying the academic foundation that enabled this research.

I extend heartfelt thanks to my family for their patience and support throughout this challenging year. I am beyond words grateful to the Armed Forces of Ukraine for making it possible to study and work under circumstances that are anything but ordinary.

And finally, thank you to my husband, whose genuine enthusiasm for my maps almost made up for the rest of the thesis talk.

LIST OF ABBREVIATIONS

CPI. Consumer Price Index.

DFRR. State Regional Development Fund.

EDRPOU. Єдиний державний реєстр підприємств та організацій України (Unified State Register of Enterprises and Organizations of Ukraine).

GMM. Generalized Method of Moments.

k-NN, kNN. k-Nearest Neighbors.

KOATUU. Класифікатор об'єктів адміністративно-територіального устрою України (Classifier of Administrative-Territorial Units of Ukraine).

LISA. Local Indicators of Spatial Association.

NMV. Normative Monetary Value.

OLS. Ordinary Least Squares.

OSM. OpenStreetMaps.

SEM. Spatial Error Model.

SLM. Spatial Lag Model.

SSSU. State Statistics Service of Ukraine.

Chapter 1

INTRODUCTION

Farmland is a critical asset in agricultural economies. It represents the majority of farm wealth and often serves as collateral for loans (Burns et al. 2018). As a result, understanding the determinants of farmland value is important for both market participants and policymakers. In other countries, farmland prices are shaped by economic conditions, land use regulations, and environmental factors, often following market trends.

In Ukraine, however, the farmland market has been defined by institutional and political changes. The legacy of the Soviet land system, followed by the post-Soviet redistribution of land rights and a moratorium on land sales from 2001 to 2021, meant that private owners were unable to sell most of agricultural land for two decades. Although the moratorium was introduced as a temporary safeguard, it ended up limiting landowners' ability to transfer property and discouraged long-term investment (Deininger and Nivievskyi 2019). The 2021 reform, which allowed land sales to private individuals, and the second phase in 2024 that opened the market to legal entities (Deininger et al. 2024), marked the first real steps toward liberalization. But this process was interrupted by the full-scale Russian invasion in 2022, adding major uncertainty and physical disruption. As a result, land values in Ukraine reflect not just economic or environmental factors, but a combination of institutional transition and wartime conditions.

This study develops a large-scale model for evaluating agricultural land in Ukraine, incorporating economic, geographic, and policy-related factors to provide a national-level analysis of land values. It focuses on the effects of spatial features, economic conditions, and regulatory interventions such as the legal price floor. In addition, the model captures how land valuation has been influenced by the war and the second phase of land market liberalization, which opened the market to legal entities. By identifying the main determinants of land value, the study contributes to a broader understanding of farmland markets in transition economies and conflict-affected regions. The empirical analysis is conducted using spatial econometric techniques.

This thesis is structured as follows: The introduction presents the research context, objectives, and hypotheses. Chapter 2 provides background information on Ukraine's land market, including its history and current state. Chapter 3 provides a review of the relevant literature. Chapter 4 outlines the methodology, focusing on the spatial econometric models employed to estimate the determinants of land value. Chapter 5 describes the data used in the analysis, detailing key variables, their transformations and their sources. Chapter 6 presents the estimation results, interpreting the effects of the various factors on land prices. Finally, Chapter 7 concludes with a summary of findings and policy implications.

Chapter 2

BACKGROUND INFORMATION

Since gaining independence in 1991, land market reform was a crucial stepping stone in transition from planned to market economy for Ukraine (Minich et al. 2021). Following the dissolution of the Soviet Union, the Ukrainian government initiated a comprehensive transfer of property rights, redistributing approximately 31.5 million hectares of land to private individuals and local governments (Wengle, Dankevych, and Mamonova 2024). By 2017, over half of Ukraine's land area—about 52.2%—had been privatized (Nizalov 2019). However, the sales of agricultural land became restricted under a moratorium that was meant to be a temporary solution, but lasted for 2 decades.

The moratorium on agricultural land sales, introduced in 2001, became a significant source of inefficiency in Ukraine's economy (Deininger et al. 2024). While its original intent was to safeguard rural populations and avoid concentration of the land in few hands, the prolonged ban prevented landowners' from fully utilizing their property rights. By restricting the sale of land, the moratorium discouraged investments, limited access to finance, and slowed productivity growth, undermining the potential of the agricultural sector to contribute to economic development of Ukraine. Halytsia, Nivievskyi, and Deininger (2022) estimate that the farmland sales moratorium reduced agricultural productivity growth in Ukraine by 54-57%, placing the sector on a lower long-term development trajectory. Its negative impact outweighed that of other policy constraints such as trade liberalization and agricultural tax benefits.

Due to the nature of the sales ban, quantifying its impact from the empirical evidence is challenging. However, estimates from a study by Deininger and Nivyevskyi (2019) suggest that lifting the moratorium could trigger substantial economic growth. In their analysis, under more liberal market conditions, where foreign and legal entity participation is allowed, Ukraine's GDP could increase by up to $10.57 billion over a 3-5 year period. In scenarios with more restrictions, such as caps on land supply and limited foreign investment, the GDP increase would range between $5 and $8 billion.

Lifting the moratorium would also trigger redistribution of the land. For many of the individuals who received or inherited agricultural land, farming was not a viable or desirable option. Instead, they leased their plots to agricultural producers, often for a fraction of the land's potential value Ukraine. Payments for these leases were frequently made in kind—such as bags of grain or bottles of sunflower oil—rather than in monetary terms, further reducing the financial benefits for landowners (Minich et al. 2021). In 2018, the European Court of Human Rights ruled that the moratorium violated property rights, emphasizing that no other Council of Europe state imposed such a blanket ban. The Court states that less restrictive measures, such as land ownership caps or targeted taxation, could have achieved similar goals while respecting the property rights of landowners (European Court of Human Rights 2018).

Following a series of reforms, including the digitalization of the land cadastre, introduction of data exchange between the Cadastre and the Registry of Rights, granting notaries the authority to register land rights alongside state registrars, efforts to correct errors in the cadastre, and the transfer of state land to amalgamated communities, the land market for private individuals officially opened in June 2021. The reform was set to progress in two stages: in June 2021, the market opened for private individuals with a cap of 100 hectares per person. Starting in January 2024, legal entities are allowed to purchase up to 10,000 hectares of agricultural land. Foreign individuals and firms with foreign capital are excluded from the market. Additionally, there is a price floor for buying land, set to the Normative Money Value calculated by the State Land Cadastre

(Wengle, Dankevych, and Mamonova 2024). In 2024, average appraised monetary value of arable land is 28,924 UAH per hectare (Center for Food and Land Use Research 2024). Current renters have a primary right to purchase the land they lease. Despite the remaining restrictions, the opening of the land market is a major milestone for Ukraine's land reform.

Less than a year passed between the opening of the land market and Russia's full-scale invasion of Ukraine. The opening of the market had been eagerly anticipated, with a survey conducted in May 2021 revealing that 7.4% of landowners planned to sell all or part of their land once the market was open (Foundation "Democratic Initiatives" and Razumkov Center 2021). Supply of land was anticipated to be between 2.3 and 3.1 million hectares, while demand from individuals involved in agricultural production was expected to be around 1.1 million hectares (Deininger, Nizalov, and Nivyevskyi 2017). However, many potential sellers adopted a "wait and see" approach, anticipating rising land prices. In the first seven months of the market's opening before the invasion, only 0.421% of agricultural land was traded. The Kharkiv region led the way, with over 1% of its farmland sold, while other regions such as Rivne, Lviv, and Ivano-Frankivsk saw minimal activity, with Ivano-Frankivsk registering just 0.083% of its farmland in circulation (KSE Agrocenter 2022).

The full-scale invasion by Russia in February 2022 caused significant damage to Ukraine's agricultural sector. Key agricultural regions, such as Kherson and Zaporizhzhia, faced occupation, destruction of infrastructure, and disruption to planting and harvesting cycles. Direct damages amounted to $8.7 billion, while indirect losses—including reduced crop and livestock production, logistics disruptions, and higher production costs—are estimated at $31.5 billion during the first year of invasion (KSE Agrocenter 2023).

The full-scale invasion of Ukraine disrupted the emerging agricultural land market. During March and April 2022, the market was closed as the government restricted access to the land ownership database. Operations resumed in May 2022 under wartime regulations, but trade volume fell from approximately 10,000 deals per month in 2021 to half that number due to the war. Regional patterns shifted, with trade volume collapsing in frontline regions while central Ukraine witnessed both relative price stability and, in some areas, significant price increases. The legal price floor prevented prices from dropping further in many regions (Matvieiev 2023).

In 2024, Ukraine followed through with the second phase of the land reform, allowing legal entities to purchase agricultural land. In Q3 2024, the market recorded 28,400 sales and purchase transactions covering 62,700 hectares—a 29% increase in the number of transactions and a 39% rise in the area of land in circulation compared to the same period in 2023 (KSE Agrocenter 2024). While transaction volumes remain below pre-invasion levels, the market has demonstrated steady growth during the 2022–2024 period.

Chapter 3

LITERATURE REVIEW

The study of land value determinants has evolved with the increasing availability of data, allowing for more sophisticated analytical approaches beyond traditional rent-based theories. Contemporary research integrates agricultural and non-agricultural factors, employing advanced econometric spatial modelling techniques to provide a more comprehensive understanding of land pricing dynamics. These advancements enable researchers to account for spatial dependencies and market heterogeneity, resulting in more accurate models.

Ricardo's theory of rent, remains foundational in understanding the determinants of land value. The theory suggests that land value is heavily influenced by its fertility, which in turn affects agricultural productivity. According to Ricardo, the more fertile land yields higher returns to its owner because it allows for greater agricultural output with fewer inputs. As a result, land with superior fertility tends to be more valuable in the market. Conversely, land that is less fertile or has poorer agricultural potential will yield lower returns, which is reflected in a lower market value. This difference in productivity across land types creates a rent differential, with more fertile land commanding higher prices (Clark 1973, 1–39). For the purposes of this study, the key takeaway from Ricardo's rent theory is the importance of land fertility and location as determinants of land value. In the context of modern land markets, this principle suggests that the inherent characteristics of land—particularly its soil fertility—should be closely considered when developing predictive models for land price estimation.

Rosen's (1974) hedonic pricing model extends the theory of land value by recognizing that land and other goods are valued for their specific attributes, rather than a single inherent value. In this model, land prices are determined by the implicit prices of various characteristics, such as soil quality, proximity to infrastructure, and environmental factors. These implicit prices are estimated through regression analysis, where the price of land is regressed on its attributes, allowing for the identification of how much consumers or investors are willing to pay for these characteristics. The hedonic model is commonly used in land valuation to assess the impacts of various factors such as accessibility, amenities, and land use. However, identifying the key determinants that reflect consumer preferences and market dynamics can be challenging.

In modern empirical studies of agricultural land value, multiple agricultural and non-agricultural determinants are used. The study by Borchers, Ifft, and Kuethe (2014) concludes that a range of factors beyond traditional agricultural productivity influences agricultural land values. While the discounted stream of expected returns remains a central determinant, non-agricultural attributes such as development potential, proximity to amenities like hospitals and college campuses, and higher median household incomes in the area also significantly contribute to farmland market values. The study finds that development potential is particularly important in explaining land value. However, the authors also note a difference between agricultural and market values, suggesting that non-agricultural factors might drive up land prices. If the ability to change land purpose from agricultural to suit for development is limited, like in Ukraine, these factors' impact may be diminished.

Distance to urban centres is also often considered in the literature as a determinant for land value as cities offer both economic opportunities and access to resources, which can drive up demand for surrounding agricultural land. The study by Guiling, Brorsen, and Doye (2009) finds that urban proximity significantly impacts agricultural land values in Oklahoma. As expected, land prices decrease with greater distance from major cities. The influence of urban proximity has slightly expanded over time due to increased population and real income. The key takeaway is that urban proximity is a valuable factor in determining land value, with real income and population size being the most influential elements.

Factors included in the hedonic price model vary widely based on the country, type of agriculture, and data availability. For example, Huang et al (2006) in a study on Illinois farmland values uses county-level cross-section time-series data to incorporate variables such as land productivity, parcel size, improvements, urban-rural index, proximity to cities like Chicago, livestock production, population density, income, and inflation. Spatial and serial correlation components significantly improve the model, revealing that farmland values increase with soil productivity, population density, and income, while they decrease with parcel size, ruralness, and distance to urban centres.

In addition to land characteristics, government agricultural policies play a significant role in determining farmland values. Weersink et al. (1999) find that in Ontario, government payments are capitalized into land prices differently than market-based returns. Using a present value model, they show that government subsidies are viewed as more stable and are discounted less than farm income. This suggests that policy stability plays a key role in farmland valuation.

Table 1 summarizes the most frequently used determinants in land valuation studies, offering an overview of the attributes that influence land prices.

Table 1. Land value determinants commonly used in the literature.

Category	Variable	Expected Effect	Studies
Intrinsic characteristics	Soil quality	Positive impact of better soils.	Breustedt and Habermann, 2008
	Topography	Less challenging topography is more valuable.	Kunwar and Bohara, 2017
	Parcel size	Larger parcels are cheaper on average due to economies of scale. Some other studies also include fragmentation of neighboring plots as s complementing factor.	Maddison, 2009
	Irrigation/ drainage	Well-established irrigation\draining is more valuable.	Kostov, 2009
	Share of arable land	Other types of land (pastures, forests) are typically less valuable so share of arable land is positive.	Borchers, Ifft, and Kuethe 2014

Table 1 – Continued.

Category	**Variable**	**Expected Effect**	**Studies**
	Value of amenities on the property	Natural (e.g., rivers) amenities and farm structures are more valuable as they can generate additional revenue.	Nilsson and Johansson, 2013
	Climate	Depending on specification, as a rule - plots in harsh climates are less valuable.	Kunwar and Bohara, 2017
Accessibility	Distance to major roads	Proximity to roads and their quality have positive impact on land value.	Marques and Telles, 2023
	Distance to amenities	Proximity to amenities (hospitals, colleges etc) has positive impact on land value.	Borchers, Ifft, and Kuethe 2014
	Distance to markets	Positive impact on land value.	Borchers, Ifft, and Kuethe 2014
Urban pressure	Distance to major cities	Positive impact on land value.	Maddison, 2009

Table 1 – Continued.

Category	**Variable**	**Expected Effect**	**Studies**
	Population in the area	Positive impact on land value.	Huang et al., 2006
	Degree of urbanization	Positive impact on land value.	Marques and Telles, 2023
	Development potential	Positive impact on land value.	Nilsson and Johansson, 2013
Agricultural activity	Share of different crops	Depends on specification, location etc.	Marques and Telles, 2023
	Crop rotation	Positive impact of more sustainable farming practices.	Maddison, 2009
	Livestock density	Depends on specification.	Huang et al., 2006
	Indicator of rented or owner-cultivated	Owner-cultivated plots are more valuable.	Choumert, and Phélinas, 2015

Table 1 – Continued.

Category	**Variable**	**Expected Effect**	**Studies**
Policy based	Zoning indicators	If land is desirable for alternative uses (development), it is typically more expensive.	Cotteleer et al., 2011
	Direct payments and other subsidies	Positive impact on land value.	Breustedt and Habermann, 2011
Buyer/seller characterstics	Indicators for forced sale/family sale	Price depends on type of sale.	Tsoodle, Golden, and Featherstone 2003
Spatial/location	Neighbouring rental and sales prices	There's positive spatial autocorrelation in prices	Burnett, Lacombe, and Wallander, 2024

Table 1 – Continued.

Category	**Variable**	**Expected Effect**	**Studies**
	Average income in the area	Positive impact on land value.	Borchers, Ifft, and Kuethe 2014

Research on land valuation in Ukraine is still in its early stages, as the market for agricultural land sales only opened in 2021 after a long-standing moratorium. Prior to this, most studies focused on rental rights transactions, examining lease prices as a proxy for land values. However, the availability of transaction data since market liberalization has sparked growing interest among researchers, leading to an increasing number of studies analysing land price determinants, valuation methods, and market efficiency.

One of the most recent contributions in this field is the work by Deininger et al. (2024), which proposes a framework for land valuation and taxation in Ukraine. Their model estimates parcel-level land prices using fixed effects at the most granular administrative level with sufficient sales data. Key determinants include land use, soil quality (e.g., quantity of nitrogen, organic matter, pH), infrastructure access (distance to roads, railways, ports, and cities), and compliance with the minimum price floor. Results indicate that market prices are, on average, 33% higher than the official "normative monetary value", with greater disparities in western regions. War-time transactions see a 7–16% price drop, with an additional 2.5% decline in 2024. Cross-validation shows the model explains about 50% of price variation, outperforming simpler averaging methods. The study highlights the feasibility of market-based mass valuation of agricultural land.

While the fixed-effects model helps control for unobserved heterogeneity across regions, increasing the granularity of fixed effects—such as moving from oblast-level to hromada- or rayon-level may absorb much of the spatial variation in land prices, making it difficult to identify broader spatial patterns. Moreover, they do not explicitly model spillover effects, meaning that dependence between neighbouring observations remains unaccounted for. If land values in one location are influenced by those in adjacent areas, failing to model these spatial interactions can lead to omitted variable bias and inefficient estimates.

Traditional econometric models often assume observations are independent, yet spatial dependence in land prices is well-documented (Anselin 2002). Spatial econometrics addresses this issue by explicitly modelling spatial interactions, recognizing that land values in one location influence, and are influenced by, those in nearby areas. Empirical studies confirm the presence of spatial autocorrelation in land price data, as well as spatial autocorrelation in error terms (Breustedt and Habermann, 2008). The omission of spatial dependence can lead to biased estimates and misinterpretation of key determinants.

The literature offers several spatial econometric approaches to account for spatial dependence. The Spatial Lag Model (SLM) includes a spatially lagged dependent variable to capture direct spillover effects, while the Spatial Error Model (SEM) accounts for spatial dependence in unobserved factors (Anselin 2002). More advanced specifications, such as the Spatial Durbin Model, allow for flexible modelling of spatial interactions in independent variables. The choice of model depends on the nature of spatial dependencies present in the data, with empirical studies employing diagnostic tests such as Moran's I to guide selection.

A critical component of spatial models is the spatial weight matrix, which defines the structure of spatial relationships. The literature commonly employs contiguity-based matrices (e.g., simple adjacency or Queen and Rook) and distance-based matrices (e.g., k-nearest neighbours, inverse distance, neighbours within given radius) (Cotteleer, Stobbe, and van Kooten 2010). Maddison (2009) introduces spatio-temporal approach arguing that the price of a land plot is influenced by the neighboring plots, but only by those that were sold prior to it. Making the right choice is important as it influences both spatial lag specification and efficiency of a maximum likelihood estimator, however the procedure for choosing one is not established and case-specific (Anselin, 2002) (Griffith and Lagona, 1998).

KSE Agrocenter plays a key role in monitoring Ukraine's emerging agricultural land market. The KSE Agrocenter’s "Invincible Land" project provides monthly and quarterly reports on market trends, including transaction volumes, pricing, and regional patterns. It offers valuable data and analysis, addressing gaps in market monitoring and informing research and policymaking.

To summarize, classical economic theories emphasize land fertility and agricultural productivity as primary determinants of land value. However, modern empirical studies highlight a broader set of factors, including non-agricultural influences like urban proximity, development potential, and policy interventions. More recent studies incorporate spatial econometric techniques to address spatial dependencies in land markets, improving the accuracy of land value estimation. In the context of Ukraine, research on agricultural land value remains an emerging field, with ongoing studies examining market dynamics, policy effects, and external shocks, yet significant gaps remain in understanding spatial dependencies and price formation mechanisms.

Chapter 4

METHODOLOGY

This study primarily follows the methodology outlined in Cotteleer, Stobbe, and van Kooten (2010), which applies spatial econometric techniques to hedonic pricing models of farmland values. Their approach is particularly relevant because it systematically examines the effects of alternative spatial weighting matrices and model specifications (e.g., inclusion of spatial components). While their analysis employs Bayesian model averaging to address model uncertainty, the present study focuses on the frequentist estimation of multiple spatial model types without incorporating the Bayesian component. The comparative evaluation of spatial weights and functional forms in Cotteleer, Stobbe, and van Kooten (2010) strongly informs the structure and diagnostics adopted here.

In terms of the choice of determinants, the study draws on Deininger et al. (2024), whose analysis of land valuation and taxation in Ukraine uses land use structure and crop type distribution. Furthermore, the literature overview by Tavares, Tavares, and Santos (2022) offers a comprehensive summary of farmland value determinants and modeling practices, helping to ground this analysis within the broader body of land valuation research. The spatial error modeling approach of Borchers, Ifft, and Kuethe (2014) also helped inform the choice of explanatory variables, as their study includes several useful determinants of land prices.

4.1. Notation

Throughout the text, P_i denotes the price of land plot i, $\mathbf{X}$ is the matrix of explanatory variables, X_k is the k-th column of $\mathbf{X}$ representing the k-th feature, β_k is the estimated coefficient for X_k, α is the intercept, and W represents the spatial weights matrix. The model specifications, including the notation and formulation, follow the framework from Rey, Arribas-Bel, and Wolf (2020).

4.2. Spatial Error Model and Spatial Lag Model

Spatial error model is employed to model spatial autocorrelation in the error terms, which accounts for unobserved factors that may affect land prices in neighbouring areas.

In the general form it looks like

$$logP_i = \alpha + \sum_k \beta_k X_k + u_i \tag{1}$$

$$u_i = \lambda u_{lag-i} + \epsilon_i \tag{2}$$

Where

$$u_{lag-i} = \sum_j w_{i,j} u_j \tag{3}$$

λ as the spatial autoregressive parameter, and the significant and positive value of it indicates positive spatial autocorrelation in the error terms, suggesting that unobserved factors influencing land prices are spatially correlated. This specification corrects for bias arising from omitted spatially correlated variables, improving model accuracy.

Meanwhile, the spatial lag model (Kelejian and Prucha 1998) accounts for spatial dependence directly in the dependent variable by incorporating the spatial lag of the dependent variable, enabling us to capture the influence of land prices in adjacent locations.

In SLM:

$$logP_i = \alpha + \rho \, log \, P_{lag-i} + \sum_k \beta_k X_k i + \epsilon_i \tag{4}$$

The parameter ρ measures the strength of spatial dependence in land prices. A positive value of it indicates that higher land prices in neighbouring plots positively influence the price of land plot i, suggesting that spatial interactions, such as proximity to high-value land, play a significant role in determining land values.

Combined spatial error and spatial lag model incorporates both treatments for spatial autocorrelations:

$$logP_i = \alpha + \rho \, log \, P_{lag-i} + \sum_k \beta_k X_k i + \lambda \sum_j w_{i,j} u_j + \epsilon_i \tag{5}$$

SEM and SLM are used to account for spatial autocorrelation in the error terms and dependent variable, respectively. In the case of OLS, spatial dependence between observations can lead to biased and inconsistent coefficient estimates. This bias arises because OLS assumes that the error terms are independent and identically distributed, but in the presence of spatial autocorrelation, errors in one observation may be correlated with errors in neighbouring observations.

In this study, all three specifications—SEM, SLM, and SEM+SLM—will be estimated and compared. The objective is to assess which model most effectively captures spatial dependencies while retaining interpretability and predictive accuracy. The primary specification will be selected based on both statistical performance and how clearly it attributes variation in land prices to observable explanatory variables.

4.3. Cross-Validation and Model Stability

To assess the generalizability and stability of the model, a five-fold cross-validation procedure will be conducted. In each fold, the dataset will be split into training and test sets. A new k-nearest neighbors spatial weights matrix will be constructed using the coordinates of both training and test observations. The spatial lag of the dependent variable will be computed using this matrix, but only training observations will be used in the calculation to prevent label leakage. The resulting spatial lag term will be included as a covariate in the regression model. Model performance will be evaluated using pseudo R-squared, mean absolute error (MAE), and root mean squared error (RMSE), while coefficient stability across folds will be tracked to confirm consistency in parameter estimates.

4.4. Weights Matrix Specification

The spatial weights matrix W is constructed using k-nearest neighbors with k=10 for computational efficiency. This approach connects each observation to its 10 nearest neighbors, and the matrix is row-standardized to ensure each row sums to one. The choice of k=10 for the k-nearest neighbors spatial weights matrix was made based on is based on preliminary experimentation; alternative values of k yielded similar model performance. k=10 provided a good balance between capturing spatial structure and maintaining computational efficiency. This is consistent with Cotteleer et al. (2010), who found that varying k between 5 and 10 had little impact on results.

4.5. Software and Tools

The analysis was conducted using Python 3.11, the spreg package (version 1.8.2) from PySAL (version 25.1), and GeoPandas for working with geometries. Python ensures efficient handling of large datasets, while spreg supports the GMM for estimating spatial models, including both spatial error and spatial lag components. GeoPandas was used to handle spatial data and perform spatial joins of different datasets and geometric operations.

C h a p t e r 5

DATA DESCRIPTION

The primary dataset used in this study is obtained from the State Service of Geodesy, Cartography, and Cadaster of Ukraine[1]. It contains comprehensive information on all transactions involving agricultural land, including sales, gifts, inheritance transfers, exchanges, and corporate contributions. The dataset has been collected since July 1, 2021, and is updated bi-weekly. It includes key details such as land area (in hectares), agreement type, transaction price, ownership type, registration date, and valuation of the land plot. The data is provided in tabular format and is compiled based on interactions between the State Land Cadaster and the State Register of Property Rights.

The dataset covers only agricultural land (category code "100") and includes transactions recorded through various legal agreements, such as sale and purchase contracts, gift agreements, and inheritance transfers. Each land parcel is linked to its respective administrative unit using KOATUU codes, allowing for geographic classification. For the purposes of further analysis, only sale agreements are considered, and all other types of transactions are filtered out.

Price reporting is not mandatory, leading to some missing values in the transaction price variable. The total percentage of missing prices across the dataset is 18.78%. However, nearly all missing values are concentrated within the first six months of market operation. The percentage of missing prices is generally consistent across regions, with the exception of Luhansk (64.7%) and Kherson (49.7%), which exhibit significantly higher proportions of missing data. In the

[1] Available at: https://land.gov.ua/monitorynh-zemelnykh-vidnosyn/

remaining regions, the percentage of missing prices tends to range between 10% and 30%.

Figure 1 shows the distribution of missing values over time. As the missing prices are heavily concentrated in the early months, these entries are omitted from the analysis for the sake of clarity. Additionally, valuations are missing for certain transactions, and these are also excluded, as they represent a crucial determinant for the model. By excluding transactions with missing prices and losing about 1,000 transactions per month due to missing valuations, the dataset goes from 332 to 235 thousand records. The remaining dataset is still large enough to ensure the analysis is reliable and robust.

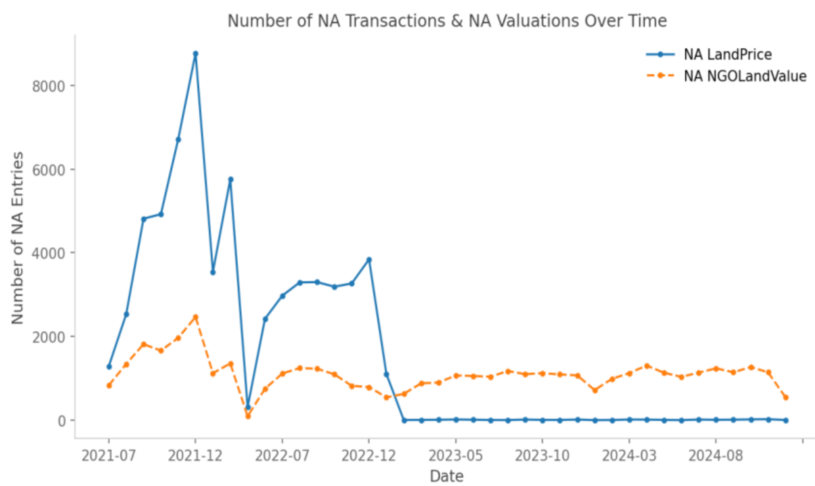


Figure 1. Distribution of missing prices and valuations over time aggregated by month.

The majority of land plots in the dataset are quite small, with the largest concentration of plots under 5 ha, as shown in the histogram below. The boxplot, shown in Figure 2, zoomed in on land areas under 120 ha, reveals presence of larger plots. The largest plot size is around 227 ha. This indicates a presence of larger transactions alongside the majority of smaller plots.

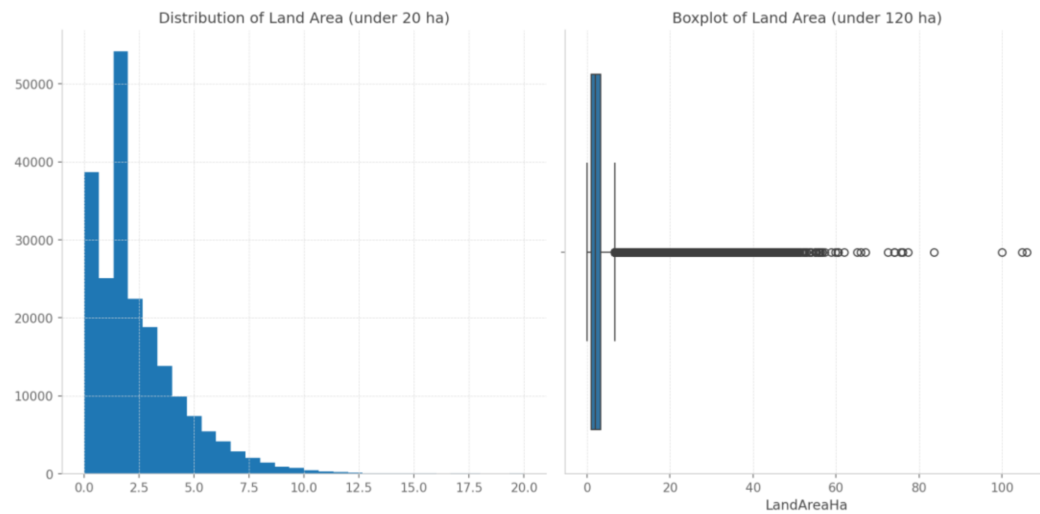


Figure 2. Distribution of land plot area

For a small subset of land plots, reported area values were corrected using the area derived from cadastral polygon geometry. Corrections were applied under specific conditions: when discrepancies between sources were extreme, or when one source appeared implausible (e.g., cadastral area above 370 hectares or polygon area below 0.1 ha). When the two sources closely matched, the original cadastral value was retained.

All prices and valuations that are originally provided in UAH have been adjusted for inflation using the monthly Consumer Price Index (CPI) data from SSSU[2] to ensure consistency over time. Additionally, prices have been normalized by dividing them by the corresponding corrected land area, allowing for a more accurate comparison across plots of different sizes. Figure 3 illustrates the median monthly price per hectare of agricultural land, shown in both nominal terms (blue

[2] Consumer Prices Index https://www.ukrstat.gov.ua/imf/arhiv/isc_e.htm

line) and inflation-adjusted terms (orange line). While the nominal price exhibits an upward trend with fluctuations, the inflation-adjusted price remains relatively stable after an initial decline. This suggests that the apparent price growth over time is largely driven by inflation rather than an actual increase in land value.

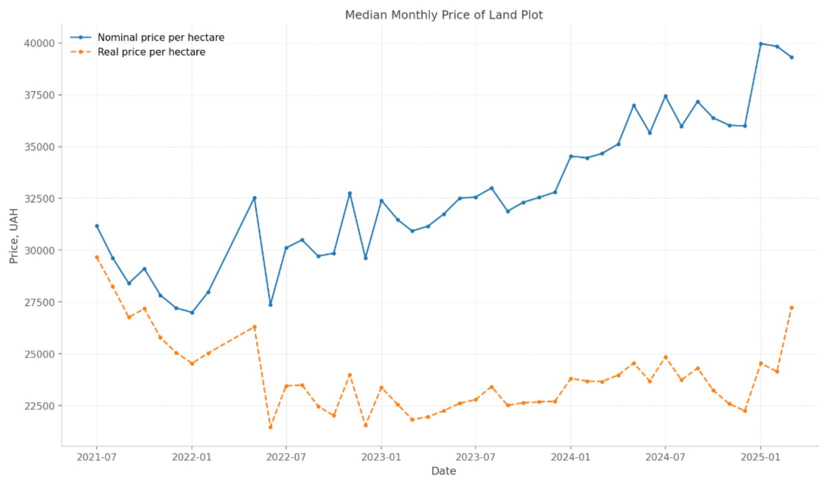


Figure 3. Median price of land plot per hectare aggregated by month.

While NMV can be interpreted as a binding price floor for agricultural land transactions, this does not apply uniformly across the dataset. In practice, the floor is set only for land plots allocated in kind to owners of land shares (*pai*), and does not extend to all ownership types or transaction contexts. Exceptions include state or communal land, as well as land sold at auction, where prices are based on expert valuation and competitive bidding. As a result, some observed prices fall below the corresponding NMV—not due to data errors, but due to the distinctions in land use purpose and ownership form.[3]

[3] Written response from the State Service of Geodesy, Cartography, and Cadaster to author's inquiry, April 2025.

The dataset includes fields for 'Region', 'District', and 'Settlement'. However, due to inconsistencies in the records, we use the KOATUU classification to construct the full address of the administrative unit each land plot belongs to. We then extract the geographic coordinates of these addresses using the HERE API[4], resulting in 5,187 unique addresses. Administrative unit may be a village or a city. It is not fully synonymous with hromada - multiple administrative units may belong to the same hromada.

Figure 4 displays the geographic distribution of all land plots, with markers indicating each administrative unit where transactions occurred. Marker size reflects the number of transactions associated with each unit. As shown, land plots are heavily concentrated in central Ukraine, highlighting a higher volume of transactions in this region relative to others.

Whenever possible, the location of the land plot is refined to be more accurate using cadaster resources[5]. It contains geographic polygons with the exact locations of land plots. However, it is important to note that it does not contain information on all land plots, about 6.6% of land plots are missing. If the exact coordinates are missing, the coordinates of the administrative centre of the respective KOATUU are applied instead.

[4] HERE API Geocoding https://geocode.search.hereapi.com/v1/geocode

[5]https://kadastr.live/api/parcels/

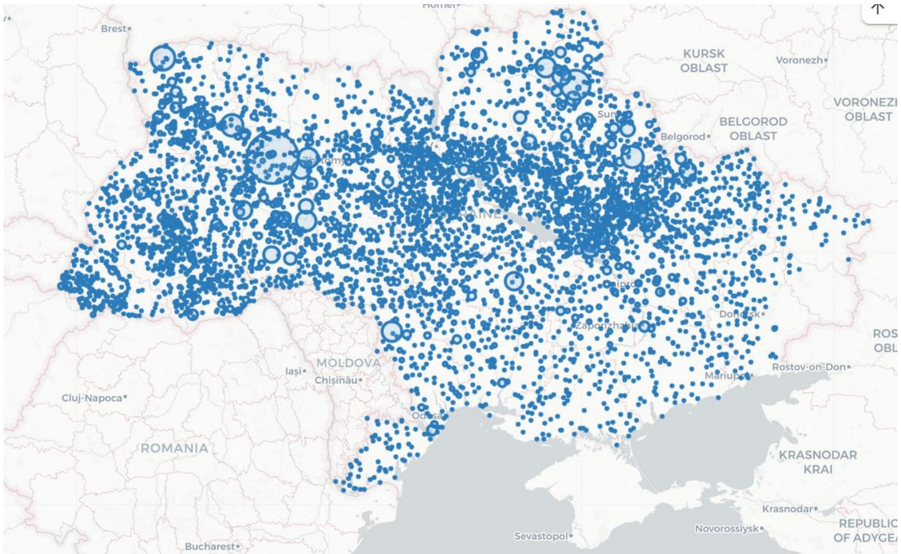


Figure 4. Distribution of plot locations with transactions count.

To improve the performance of k-nearest neighbours computations, jittering is applied to the coordinates assigned from administrative centres. Jittering introduces small, random perturbations to these coordinates, preventing multiple land plots from being assigned identical locations. This adjustment helps avoid artificial clustering effects that could distort distance-based spatial analysis and ensures that k-NN methods operate more effectively.

To account for the impact of hostilities on land transactions, areas of conflict were identified using data from the DeepState[6]. Each area is represented by a collection of polygons and categorized into three types: areas occupied since 2014, areas occupied between 2022 and 2025, and areas liberated since 2022. Based on these polygons and the locations of land transactions, two sets of spatial features were computed for each conflict category: (1) the closest distance from each land plot to the nearest polygon of the given type and (2) the average

[6]DeepState Map of hostilities - https://deepstatemap.live/#6/49.4383200/32.0526800

distance from each land plot to all polygons of that type. All distances are measured in kilometers. These features help capture the potential influence of proximity to conflict zones on land values and transaction dynamics.

To incorporate community-level characteristics, data from the KSE-Loc-Data-Hub[7] project were joined based on a spatial join (point of land plot is within hromada polygon). This dataset provides detailed geographic and administrative information at the hromada (amalgamated community) level, including boundaries, governance details, population statistics, infrastructure availability, and socio-economic indicators. Key features extracted include community type, population size and distribution, access to transportation, administrative classifications, education performance (e.g., Ukrainian External Independent Evaluation scores), youth and entrepreneurship support structures (e.g., number of youth centers, councils, and entrepreneurial support centers), healthcare coverage (e.g., number of health facility declarations). The dataset also includes information on tax revenues, financial capacity assessments, project costs financed by the State Regional Development Fund, and the status of war zones.

The Harmonized World Soil Database[8] was used to determine the predominant soil type at each location. The soil type was identified using the WRB2 classification and the HWSD1_SMU_ID, referred to further as soil name and soil type, respectively. Each soil mapping unit (HWSD1_SMU_ID) may represent a mix of soil types, so the classification with the highest share (WRB2) was selected as the dominant soil. While this data has potential to be useful, the nominal valuation already includes the soil bonitet coefficient, so its contribution will be evaluated at a later stage.

[7] https://github.com/kse-ua/KSE-Loc-Data-Hub

[8] https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v20/en/

Additional spatial features are determined using the OpenStreetMaps[9], including count of water features, distance to the nearest primary and secondary roads, density of all roads and paved roads in particular within 10 kilometres of land plot's centroid. River count reflects the proximity to water sources, while distance to road measures accessibility. Detailed description of these features extraction is provided in Appendix B.

To characterize land use composition at the plot level, high-resolution crop classification maps developed by Kussul, Lavreniuk, Skakun, and Shelestov (2017) were used. These maps were generated using multitemporal satellite imagery, combined with a deep learning framework that integrates both spectral and spatial features for pixel-level classification. The dataset enables detailed identification of major land cover types and crops, and is used here to quantify the share of cultivated land, built-up areas, water, and uncultivated base land for each plot. Technical details are provided in Appendix C.

Time-based features include year, month, week of the year, and day of the month for each transaction, along with their corresponding sin-cos transformations to capture cyclical patterns. All time-based features are calculated based on the RegistrationDate column. Additionally, two binary variables are included: is_war, which indicates whether the transaction occurred after February 24, 2022, marking the start of the full-scale war, and is_legal_market, which indicates whether the transaction date is after January 1, 2024, when legal entities were allowed to access the market.

For outlier treatment, 1 and 99 percentiles by price per hectare and valuation per hectare are filtered out. Table 2 presents descriptive statistics of the variables used in analysis.

[9] https://export.hotosm.org/v3/exports

Table 2. Descriptive statistics of selected variables

	mean	median	std	min	max
distance_to_kyiv_log	5.580	5.673	0.540	2.393	6.543
distance_to_obl_center_log	4.063	4.177	0.673	0.636	5.442
LandAreaHa_log	1.097	1.099	0.582	0.000	5.432
ValuationPerHectar_log	9.646	9.877	0.672	7.234	10.797
is_war	0.818	1.000	0.385	0.000	1.000
is_legal_market	0.481	0.000	0.500	0.000	1.000
distance_to_eu_log	5.436	5.659	1.031	0.343	6.800
income_log	11.022	11.052	0.859	8.710	15.721
mountain_hromada	0.015	0.000	0.123	0.000	1.000
near_seas	0.014	0.000	0.118	0.000	1.000
urban_pct	0.302	0.310	0.267	0.000	1.000
population_2022_log	9.613	9.620	0.816	7.492	14.167
inter_area_mount	0.007	0.000	0.068	0.000	2.734
inter_area_sea	0.018	0.000	0.177	0.000	3.866
percent_crops	0.780	1.000	0.369	0.000	1.000
percent_forest	0.020	0.000	0.107	0.000	1.000

Table 2 – Continued.

	mean	median	std	min	max
percent_built_up	0.014	0.000	0.105	0.000	1.000
percent_grassland	0.121	0.000	0.285	0.000	1.000
percent_water	0.008	0.000	0.066	0.000	1.000
is_not_cultivated	0.033	0.000	0.178	0.000	1.000
log_count_water_features	3.956	3.951	0.928	0.000	7.088
log_dist_to_primary_km	2.524	2.674	0.933	0.000	4.446
log_dist_to_secondary_km	2.113	2.222	0.808	0.012	4.239
edprou_indicator	0.105	0.000	0.306	0.000	1.000
paved_road_density	0.134	0.096	0.169	0.000	2.994
road_density	0.716	0.571	0.590	0.000	7.598
log_closest_dist_to _occupied_km	4.565	5.304	2.312	0.000	6.962
is_urban_hromada	0.324	0.000	0.468	0.000	1.000
has_water	0.038	0.000	0.192	0.000	1.000
no_primary_road_within_10k	0.013	0.000	0.115	0.000	1.000
distance_to_hromada _center_log	2.303	2.368	0.573	0.026	3.858

Table 2 – Continued.

	mean	median	std	min	max
hromada_area_log	6.102	6.158	0.674	2.625	7.823
purpose_commercial_farm	0.011	0.000	0.106	0.000	1.000
purpose_commodity_farm	0.437	0.000	0.496	0.000	1.000
purpose_gardening	0.012	0.000	0.107	0.000	1.000
purpose_other	0.301	0.000	0.459	0.000	1.000
Quarter_2	0.184	0.000	0.388	0.000	1.000
Quarter_3	0.231	0.000	0.422	0.000	1.000
Quarter_4	0.305	0.000	0.461	0.000	1.000
PricePerHectar_adjusted	34572.2	23676.5	44793.4	1514.9	742735.3
ValuationPerHectar_adjusted	18412.7	19484.1	8921.8	1384.6	48876.6

Full list of features used in the models, with source and description is provided in Appendix A.

In the final analytical sample, only transactions with cadastral (plot-level) coordinates were retained. This decision reflects the fact that certain spatial variables—particularly those derived from high-resolution crop classification maps and OpenStreetMap—are not meaningful when applied to jittered administrative coordinates. While the model could in principle be re-estimated using the full sample with KOATUU-based locations, doing so would require excluding these spatially granular features. Restricting the analysis to plots with known cadastral coordinates ensures greater spatial precision and validity in distance-based computations and spatial econometric modeling.

C h a p t e r 6

ESTIMATION RESULTS

6.1. Diagnostic tests

Global spatial autocorrelation in land prices is confirmed by Moran's I, which yields a value of 0.496 ($p = 0.001$). This indicates a strong and statistically significant tendency for similar land prices to cluster geographically. The presence of positive spatial dependence justifies the use of spatial econometric models over conventional OLS. To further explore the spatial structure of the data, local indicators of spatial association (LISA) are computed. A map of resulting clusters is presented in Figure 5.

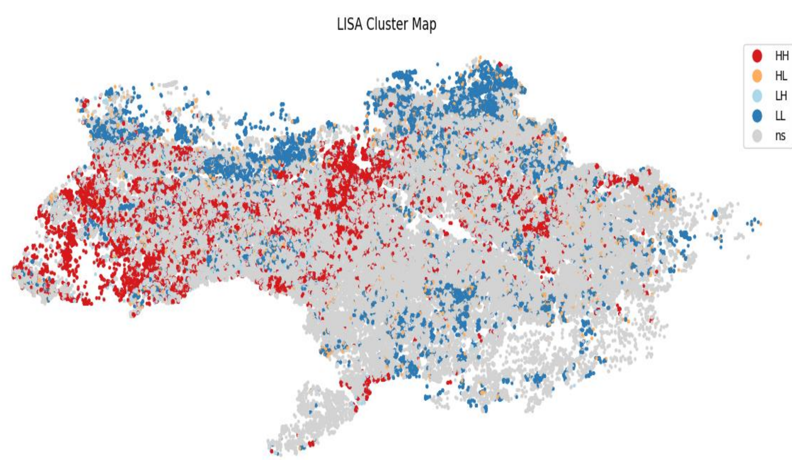


Figure 5. Local Indicators of Spatial Association.

Local Indicators of Spatial Association reveal substantial regional clustering in land prices. High-price clusters (HH) are prominently concentrated in the western and central regions, while low-price clusters (LL) dominate parts of the north and northeast. Scattered outliers also appear, such as high-price parcels surrounded by low-price areas (HL) and vice versa (LH), although these are relatively infrequent. A large portion of the country shows no statistically significant clustering (ns), especially in the south and parts of the east, likely reflecting lower transaction density or more heterogeneous pricing. Overall, 15.2% of observations fall into HH clusters, 17.2% into LL, 2.6% into LH, 3.0% into HL, while the majority (62.1%) show no statistically significant local spatial autocorrelation. While Moran's I confirms strong and statistically significant global spatial dependence in land prices, the LISA results reveal that this dependence is concentrated in specific regions, indicating that local clustering is present but spatially uneven and limited in overall scope.

A spatial weights matrix was constructed using the k-nearest neighbours algorithm with k=10, assigning each observation its ten geographically closest neighbours. The resulting graph contains no isolated observations (islands), but is not fully connected: it consists of 24 weakly connected components. The largest component includes over 210,000 observations, while the remaining 503 are distributed across smaller disconnected clusters. This structure is a result of spatial fragmentation in areas with low transaction density, and it does not affect model estimation, as the main component dominates the dataset.

As shown in Figure 6, the disconnected observations fall into two broad categories. Some are located in mountainous or sparsely populated areas in the northwest, where limited local transactions naturally restrict spatial connectivity. Others are situated in the eastern and southern regions, where transactions effectively ceased following the start of the full-scale invasion in February 2022. These disconnected clusters reflect either long-standing geographic isolation or the impact of war on market activity, rather than data quality issues.

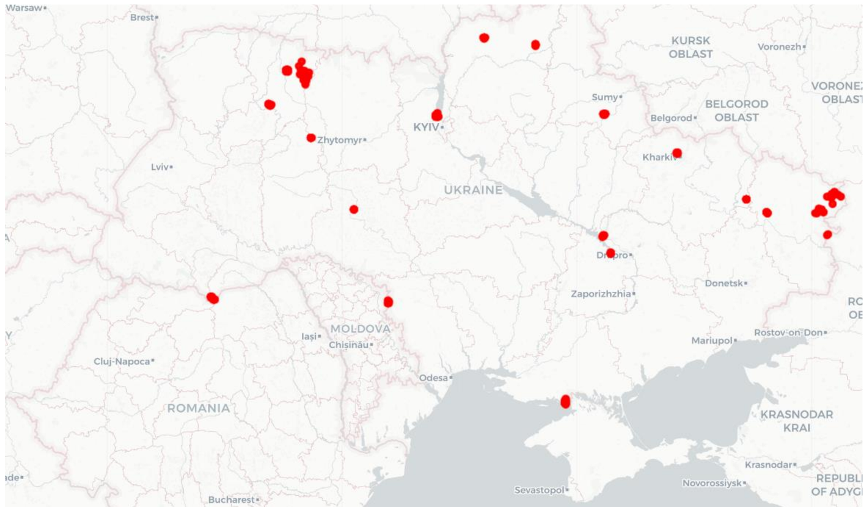


Figure 6. Disconnected observations in the k-NN spatial weights matrix.

6.2. Estimates for SEM, SLM and Combo models

Table 3 reports coefficient estimates for the spatial lag, spatial error, and combined models. All three specifications are estimated on the same dataset using an identical set of explanatory variables and an identical k-nearest neighbors spatial weights matrix. Coefficients are reported alongside z-statistics, with significance levels denoted by asterisks ((***) $p < 0.01$, (**) $p < 0.05$, p (*) < 0.1).

Table 3. Estimates of SLM, SEM and combined models

	Coefficient Combo	z-Statistic Combo	Coefficient Lag	z-Statistic Lag	Coefficient SEM	z-Statistic SEM
CONSTANT	4.4975 (***)	49.277	4.3804 (***)	54.483	5.7526 (***)	107.679
distance_to_eu_log	-0.0438 (***)	-17.299	-0.0427 (***)	-17.609	-0.0459 (***)	-17.820
distance_to hromada_center log	-0.0306 (***)	-9.549	-0.0304 (***)	-9.542	-0.0307 (***)	-9.595
distance_to_kyiv_log	-0.1591 (***)	-44.372	-0.152 (***)	-44.760	-0.1656 (***)	-45.136
distance_to_obl _center_log	-0.032 (***)	-11.415	-0.0298 (***)	-10.771	-0.0331 (***)	-11.756
edprou_indicator	0.3686 (***)	70.468	0.3666 (***)	71.288	0.3698 (***)	70.235
has_water	-0.1051 (***)	-10.625	-0.1083 (***)	-10.860	-0.1047 (***)	-10.628
hromada_area_log	-0.0914 (***)	-24.750	-0.0905 (***)	-24.761	-0.0932 (***)	-25.141
income_log	0.1135 (***)	20.017	0.1258 (***)	22.613	0.1119 (***)	19.612
inter_area_mount	-1.2703 (***)	-37.287	-1.2674 (***)	-37.150	-1.2697 (***)	-37.330
inter_area_sea	-0.2148 (***)	-12.425	-0.2175 (***)	-12.547	-0.214 (***)	-12.404
is_legal_market	0.0436 (***)	9.431	0.0425 (***)	11.394	0.0578 (***)	11.141
is_not_cultivated	-0.4223 (***)	-32.521	-0.4376 (***)	-33.560	-0.4161 (***)	-32.131
is_urban_ hromada	-0.0135 (***)	-2.877	-0.0074	-1.589	-0.0158 (***)	-3.357
is_war	-0.3533 (***)	-22.461	-0.341 (***)	-23.130	-0.3701 (***)	-22.702
lambda	0.2293				0.3388	
LandAreaHa_log	-0.0986 (***)	-29.449	-0.0973 (***)	-29.265	-0.0976 (***)	-29.078
log_closest_dist to_occupied_km	0.0658 (***)	24.191	0.0639 (***)	24.547	0.0682 (***)	24.597

Table 3 – Continued.

	Coefficient Combo	z-Statistic Combo	Coefficient Lag	z-Statistic Lag	Coefficient SEM	z-Statistic SEM
log_count_water_features	0.045 (***)	23.166	0.0436 (***)	22.781	0.0461 (***)	23.609
log_dist_to_primary_km	0.0131 (***)	7.490	0.0143 (***)	8.212	0.0128 (***)	7.264
log_dist_to_secondary_km	0.0144 (***)	7.196	0.0169 (***)	8.479	0.014 (***)	6.978
mountain_hromada	1.4939 (***)	75.760	1.4889 (***)	75.534	1.4995 (***)	76.090
near_seas	0.3824 (***)	14.434	0.3864 (***)	14.661	0.3804 (***)	14.341
no_primary_road_within_10k	-0.0687 (***)	-5.255	-0.0722 (***)	-5.609	-0.0687 (***)	-5.225
Ownership Cooperative	1.7478 (***)	25.106	1.703 (***)	25.396	1.7973 (***)	25.442
Ownership Private	-0.124 (***)	-3.164	-0.1296 (***)	-3.316	-0.1256 (***)	-3.205
Ownership StateCommunal	0.019 (***)	4.677	0.021 (***)	5.318	0.018 (***)	4.369
paved_road_density	0.0204	1.335	0.0496 (***)	3.285	0.0089	0.577
percent_built_up	0.0445 (**)	2.432	0.0287	1.558	0.0468 (**)	2.564
percent_crops	-0.5901 (***)	-46.961	-0.6064 (***)	-48.091	-0.5841 (***)	-46.587
percent_forest	-0.3788 (***)	-20.599	-0.3955 (***)	-21.441	-0.3735 (***)	-20.359
percent_grassland	-0.5294 (***)	-40.070	-0.551 (***)	-41.530	-0.5228 (***)	-39.670
percent_water	-0.4038 (***)	-13.057	-0.4193 (***)	-13.474	-0.3984 (***)	-12.930
population_2022_log	-0.0191 (***)	-2.674	-0.0324 (***)	-4.610	-0.0154 (**)	-2.138
Purpose Commercial farm	0.0545 (***)	3.653	0.0571 (***)	3.836	0.0539 (***)	3.610
Purpose Commodity farm	0.1418 (***)	34.744	0.1447 (***)	35.860	0.1411 (***)	34.435
Purpose gardening	0.9302 (***)	61.335	0.9199 (***)	60.774	0.9319 (***)	61.454
Purpose other	0.0548 (***)	12.359	0.0545 (***)	12.708	0.0559 (***)	12.457

Table 3 – Continued.

Coefficient Combo	z-Statistic Combo	Coefficient Lag	z-Statistic Lag	Coefficient SEM	z-Statistic SEM	
Quarter_2	0.0041	0.724	0.004	0.896	0.0055	0.830
Quarter_3	0.033 (***)	6.196	0.0332 (***)	7.936	0.0393 (***)	6.421
Quarter_4	0.013 (***)	2.579	0.013 (***)	3.286	0.0133 (**)	2.305
road_density	0.0969 (***)	19.400	0.0952 (***)	19.197	0.0977 (***)	19.516
sale_order	0.0604 (***)	6.418	0.058 (***)	6.314	0.0589 (***)	6.210
urban_pct	-0.1122 (***)	-12.010	-0.1144 (***)	-12.380	-0.1145 (***)	-12.211
ValuationPerHectar_log	0.5469 (***)	209.180	0.5457 (***)	207.801	0.5486 (***)	210.055
W_Log Land Price	0.1228 (***)	16.253	0.1295 (***)	19.927		

Table 4 reports model fit statistics for the three model specifications. All three spatial models show reasonable overall fit on the dataset, with pseudo R^2 values ranging from 0.3421 to 0. 3563. Contrary to expectations, the Spatial Lag Model achieves the highest pseudo R^2 (0. 3563), along with the highest spatial pseudo R^2 (0. 343), indicating that it explains the most variance of the three. The Spatial Error Model follows with a pseudo R^2 of 0. 3421, while the combined GM Combo model performs similarly in this comparison, with a pseudo R^2 of 0. 3556 and a spatial pseudo R^2 of 0. 3428. These results suggest that, despite its additional complexity, the Combo model does not offer a meaningful gain in explanatory power.

Table 4. Model fit comparison.

Model	Pseudo R-squared	Spatial Pseudo R-squared
Spatial Lag	0.3563	0.343
Spatial Error	0.3421	-
GM Combo	0.3556	0.3428

A reassuring finding is that the vast majority of coefficients are stable and significant across SEM, SLM, and the combined model. 42 out of 45 predictors are statistically significant ($p < 0.05$) in each model specification, indicating a robust set of core drivers. Key variables like distance to Kyiv, land area, land valuation per hectare, war impact, income levels, indicator for mountain region, etc., have very similar coefficients and z-statistics across all three models. This consistency suggests that the relationships in the data hold regardless of how spatial autocorrelation is modelled.

That said, there are a few notable differences in coefficient significance that show the trade-offs between models. One such difference is the variable for paved road density: in the SLM it appears positive and significant, but in the SEM and combined model it drops out as insignificant. This suggests that in a pure lag model, road density was picking up some spatially correlated effect (perhaps regions with better roads also had higher prices), but once spatial error correlation is accounted for (SEM/Combo), that effect is no longer distinct.

In contrast, the "urban hromada" dummy shows the opposite pattern – it is insignificant in the SLM, but becomes significant (with a small negative coefficient) in the SEM and even in the Combo model. This implies that the SLM's spatial lag term may have been absorbing some of the effect of urban status (since neighboring areas often share urban/rural characteristics), whereas the SEM/Combo isolating spatial error allowed the urban influence to emerge. Another small difference is seen in the built-up area percentage: it is only borderline significant in the combined model and not in SLM, but reaches significance in SEM, hinting that controlling for spatial errors might reflect subtle effects of land use type.

Crucially, no major coefficient flips sign or drastically changes magnitude between the models. The few variables that switch significance are those with relatively marginal effects to begin with. The primary drivers remain solid across all specifications, showing that the core inferences are robust to model choice. This boosts confidence that whichever model is chosen, the main findings (e.g., the impact of distance, war, valuation, etc.) are reliable.

The Spatial Lag Model is selected as the preferred specification due to a combination of empirical performance and interpretability. While it achieves the highest pseudo R^2 and spatial pseudo R^2 among the three models, the decision is not based on fit alone. The estimated coefficients in the SLM are stable, significant, and broadly consistent with those from the SEM and combined models. Importantly, the SLM preserves interpretability for key structural variables without absorbing their explanatory power into the spatial error term, as observed in the Combo model. Moreover, the lag specification aligns with theoretical expectations of spatial spillover in land prices and is supported by diagnostic tests, which point to substantive spatial interaction rather than unobserved spatially correlated shocks. Taken together, these factors suggest that the SLM offers a robust and transparent representation of spatial dependencies in the data, making it the most suitable choice for the analysis.

6.3. K-fold cross-validation

To assess the robustness and generalizability of the model, a five-fold cross-validation procedure was conducted. The model demonstrates consistent out-of-sample performance across folds, with an average R^2 of 0.356 (SD = 0.0022), mean absolute error (MAE) of 0.487 (SD = 0.0016), and root mean squared error (RMSE) of 0.671 (SD = 0.0024), as illustrated in Table 5 Coefficient estimates remain highly stable in both sign and magnitude across folds.

Figure 7 illustrates the range of estimated coefficients for each variable, showing that no signs are reversed and variation remains minimal even for variables with weaker effects. This confirms that the SLM generalizes well and captures structural relationships in the data.

Table 5. Model errors across cross-validation folds.

Fold	Pseudo R^2	MAE	RMSE
1	0.3572	0.4865	0.6695
2	0.3581	0.4844	0.6666
3	0.3567	0.4892	0.6734
4	0.3518	0.4875	0.6718
5	0.3561	0.4881	0.6716
Average	0.3560	0.4871	0.6706

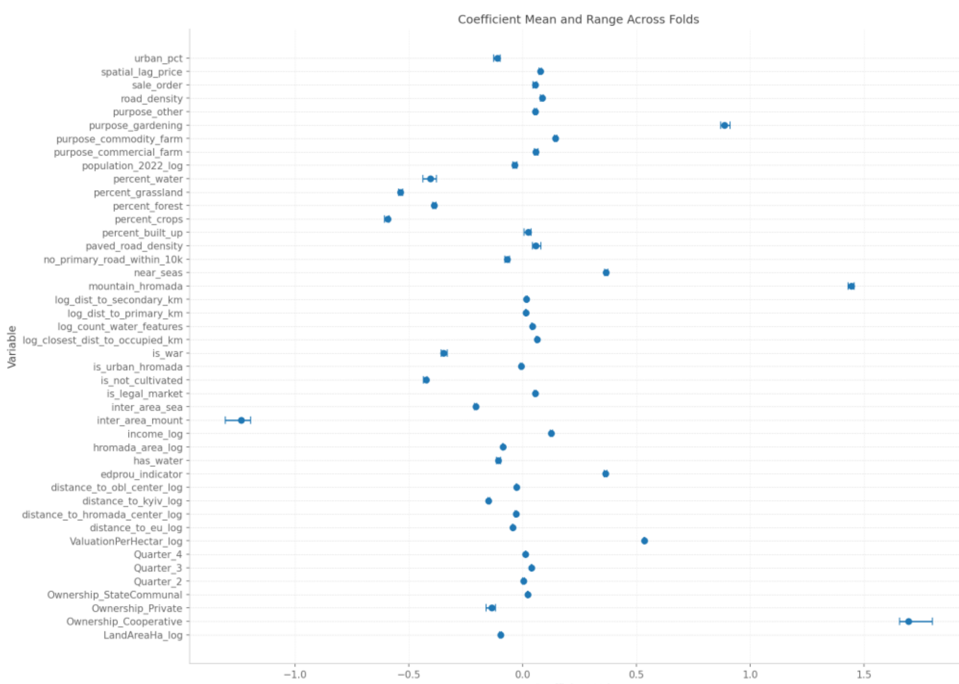


Figure 7. Coefficients mean and range across folds.

6.4. Discussion of determinants

The results from the spatial lag model reveal a rich structure of factors influencing agricultural land prices in Ukraine.

As expected, location is quite important. Prices are significantly higher in areas closer to Kyiv and oblast centers, consistent with greater economic integration, market activity, and infrastructure concentration. Among all geographic indicators, distance to Kyiv has the largest effect: a 1% increase in distance from the capital is associated with a 0.152% decrease in land price. Prices also decline with greater distance to oblast centers and hromada centers, with effects of −0.0298% and −0.0304% per 1% increase in distance. Land closer to the EU border is also more expensive; a 1% increase in distance from the border corresponds to a 0.043% decrease in price, possibly reflecting expectations of future integration or trade opportunities in the western regions.

One of the most consequential variables in the model is is_war, which captures whether the transaction took place after the start of the full-scale invasion on February 24, 2022. The estimated effect is substantial: land sold during the wartime period is priced nearly 29% lower than land sold before. This sharp discount reflects the profound uncertainty and disruption introduced by the war. Beyond immediate risks to property and personal safety, buyers face heightened concerns about future accessibility, and the long-term viability of agricultural activity both in affected regions and in the rest of the country.

In addition, land farther from occupied territories is more valuable. Each 1% increase in distance from the frontline raises price by 0.064%. This effect becomes substantial over Ukraine's geographic scale, as land plots located hundreds of kilometers from the frontline can be valued significantly higher than similar plots situated closer. The pattern likely reflects how buyers assess a broad set of war-related risks—from direct strikes and infrastructure damage to the threat of renewed hostilities. In this way, the war has reshaped the land market not only through physical destruction, but through the risk environment.

The model confirms that the official valuation per hectare is a strong predictor of land price. This likely reflects the dual role of formal valuation: not only does it anchor price expectations, but it also is a binding price floor for most transactions. The coefficient implies that a 1% increase in valuation is associated with approximately a 0.55% increase in actual land price, underscoring how closely market outcomes follow this administrative formula.

Two variables in the model capture the effect of legal entity participation in the land market. The first, is_legal_market, indicates whether the transaction took place after January 1, 2024, when legal entities were officially allowed to buy land. The second, edprou_indicator, flags whether the transaction actually involved a legal entity, based on the presence of a EDRPOU code. Both variables are positively associated with land price. Transactions that occurred after the reform are priced approximately 4.3% higher than those that took place earlier. This suggests that the policy change had an effect on the market, possibly by raising demand or altering expectations, even for plots that were not purchased by legal entities.

The effect is stronger when the buyer is a legal entity. The presence of an EDRPOU number is associated with a 44.3% increase in price compared to transactions involving individuals. This substantial difference likely reflects the types of plots that legal entities target—often larger, more investment-oriented, or located in more desirable areas. It may also signal a greater willingness to pay, especially for land intended for commercial or strategic use. Together, these results show that the entry of legal entities into the market has not only expanded the buyer base but has shifted the pricing structure of agricultural land.

Several variables describe characteristics of the hromada to which each plot belongs. Among them, average income has the strongest effect: a 1% increase in hromada income is associated with a 0.13% increase in land price, indicating that buyers are willing to pay more in areas with stronger economic conditions. In contrast, hromada population has a slight negative effect: a 1% increase in population corresponds to a 0.032% decrease in price. This may reflect land structure in more densely populated areas, where parcels tend to be smaller or more fragmented. Hromada area also matters. A 1% increase in total area is associated with a 0.09% decrease in price, possibly because larger hromadas are more rural and less economically concentrated. A 10 percentage point increase in the urban population share corresponds to a 1.14% decrease in land price, suggesting that more urbanized hromadas have less demand for agricultural land.

The model identifies strong location premiums for plots in designated mountain or coastal hromadas. Land in mountain areas is priced approximately 343% higher than elsewhere. This large premium likely reflects the unique features of mountain regions: recreational or tourism value, constrained land supply, or alternative uses beyond standard agriculture. A similar, though smaller, pattern holds for coastal areas. Plots near seas are priced approximately 47.2% higher than inland plots, possibly due to natural amenity value or better access to transportation infrastructure. Coastal proximity, especially near ports, may facilitate exports and thus attract buyers engaged in commercial agriculture.

The model shows that larger land parcels tend to be valued less per hectare. Specifically, a 1% increase in area is associated with a 0.1% decrease in price. This pattern is common in land markets and may reflect several factors, including reduced per-hectare demand for large holdings, limited pool of potential buyers, or economies of scale.

The model also includes interactions between plot size and location in mountain or coastal areas. These interactions are both negative, indicating that the price penalty for larger plots is even steeper in these regions. These results suggest that while small plots in scenic areas may have a premium, that premium does not scale with size. Larger parcels in such locations may be harder to sell or valued less because they exceed what typical buyers—especially private individuals or small investors—are looking for.

The variable sale_order captures how many times a parcel has changed hands and reveals a clear upward trend in prices across repeated transactions. Each additional sale raises the price by about 5.8%, compounding over time—second sales are roughly 6.2% higher than first, third sales 12.7%, and so on. Due to transaction costs involved, it is likely that only a selected few land plots re-enter the market in the studied period and they have other favorable attributes, such as good location, legal clarity, or investment potential. This pattern might also point out to selection and learning effects in the land market. Prices may rise simply because more is known about the land. Early transactions often happen under uncertainty—about boundaries, ownership, or productivity. As these frictions resolve, subsequent buyers are willing to pay more. In some cases, speculative or institutional buyers purchase land specifically to resell it.

The significant spatial lag term ($\rho = 0.13$) shows that land prices are influenced by nearby values: plots tend to be more expensive if their neighbors are. In addition to direct effects, the spatial lag model shows meaningful indirect effects—spillovers transmitted through the land market via the prices of neighboring plots. Since spatial dependence in the model is defined using a 10-nearest-neighbors matrix, these effects do not follow administrative boundaries but are instead based on physical proximity between plots. That is, changes in one parcel's characteristics or context can influence land prices in the surrounding ten geographically closest parcels, regardless of whether they fall within the same hromada or oblast.

For example, both the normative valuation and the edprou indicator exhibit meaningful indirect effects. The positive indirect effect of the normative valuation (0.0812) suggests that high-valued plots tend to be located near other high-valued plots. Similarly, the edprou_indicator—which identifies transactions involving legal entities—has an indirect effect of 0.0545, implying that institutional purchases are associated with higher prices not only for the transacted plot but also for nearby parcels. Full table of indirect effects is provided in Appendix E.

Accessibility remains a clear determinant. Road density—both paved and total—is positively associated with land value. Conditional on overall road density, the model finds that greater distance to primary and secondary roads is linked to slightly higher land prices: a 1% increase in distance corresponds to price increases of approximately 0.014% and 0.016%, respectively. This counterintuitive result likely reflects the structure of the model, where road density already captures the general level of accessibility. What remains is the marginal effect of being near a major road, which may correlate with land fragmentation, pollution, or other externalities that can depress the value of agricultural land.

In this sense, while being in a well-connected area is clearly beneficial, direct adjacency to a primary or secondary road may not add value—and may even reduce agricultural suitability due to competing land uses or smaller parcel sizes. Moreover, the absence of a primary road within 10 km lowers land price by about 7% on average.

Official land use purpose has an effect on price, even after controlling for plot characteristics. The base category is personal farming, which includes subsistence or household-scale agricultural use. Compared to this group, plots designated for gardening are valued approximately 151% higher per hectare — a striking premium that likely reflects their location near settlements, smaller size, and the possibility that they already contain mature or planted gardens, making them immediately usable.

Land classified for commodity farming is priced approximately 15.6% higher than personal-use land. These parcels are likely oriented toward commercial-scale crop production, making them more attractive to larger or professional buyers. Land intended for commercial farming, a more general category, is valued about 5.9% higher.

The model also captures differences in land prices based on ownership type, using farm enterprises as the reference category. Land owned by private individuals is sold at prices approximately 12.2% lower, a difference that may reflect smaller average plot sizes, lower bargaining power, or differences in how these parcels enter the market. In contrast, state or communal land sells for about 2.1% more than land owned by farm enterprises. This modest premium may stem from auction-based procedures or stricter valuation protocols tied to public ownership.

The most striking result is the estimated 449% premium for land owned by cooperatives. While the magnitude is large, it must be interpreted with caution. This estimate is based on only 109 transactions and may reflect outliers or atypical cases—such as institutional restructuring, high-value land transfers, or transactions involving strategically located plots.

Seasonal effects are also present but modest: prices in the third quarter are about 3.4% higher, and in the fourth quarter about 1.3% higher, relative to the first quarter, possibly reflecting timing around harvests or annual land market activity cycles.

Water-related features show mixed effects on land prices, depending on their form and proximity. Land plots with water features directly on-site are priced approximately 10.3% lower, possibly reflecting marshy conditions, flood risk, or regulatory restrictions on land use. In contrast, a greater number of water features within a 10 km radius is associated with higher prices: a 1% increase in nearby water bodies raises land value by about 0.044%. This suggests that while excess water on a plot can be a liability, nearby access to water may enhance value through scenic amenities or improved conditions for irrigation.

The model includes land cover features based on satellite crop classification data from summer 2023. While these variables offer detailed spatial coverage, their interpretation should be approached with caution. Plots identified as not cultivated or bare land—are priced about 35% lower. Other land cover shares, including crops, forest, grassland, and water, are also linked to lower prices. For example, a 10 percentage point increase in cropland share is associated with a roughly 5.9% price decrease. Taken individually, these effects might reflect real market preferences or specific local constraints. But taken together, the pattern is puzzling: all coefficients are negative, suggesting that no combination of land covers leads to a price premium.

This raises the possibility that the model's linear structure is too rigid to capture how land composition affects value. In reality, small amounts of forest or water might add scenic or practical value, while larger shares could reduce usability. Similarly, a plot dominated by crops might be productive but also carry risks tied to monoculture or irrigation needs. The timing of the satellite snapshot is another limitation—it doesn't match the full transaction period (2021–2025), so some land use changes may be missed. Classification noise, especially on small plots, adds further uncertainty. In short, while these features bring valuable spatial detail, their combined interpretation is likely distorted by nonlinearity, timing mismatches, and measurement issues. These results should be treated with caution and may benefit from robustness checks using alternative specifications or better-aligned land use data.

The spatial lag model shows that land prices in Ukraine are driven less by any single factor and more by the interaction of institutional, spatial, and market forces. Rather than reflecting pure productivity or location, prices follow patterns shaped by policy constraints, geography, and uneven accessibility. Administrative valuation continues to act as a price anchor, while land reforms have raised the market price for everybody. Spatial patterns reflect both exposure to conflict and access to markets — but not all infrastructure delivers value equally. What matters is not just being close to roads, but being embedded in a broader network. Plot-level characteristics, too, are interpreted through context: size, land cover, and use designation affect price in ways that reflect not just utility but regulation, scarcity, and local demand. While land prices are still shaped by state-set benchmarks, differences in location, legal rules, and market expectations are starting to matter more — showing that the market is slowly becoming more responsive and dynamic.

6.5. Comparison with existing literature

The most closely related study is the land valuation model developed by Deininger et al. (2024) for the World Bank, which uses the same core dataset. While the modeling approach differs, many of the variables overlap, allowing for meaningful comparison. The World Bank study models log land price per hectare in USD, whereas this study uses inflation-adjusted UAH. Although differences in currency and deflation methods may affect coefficient magnitudes, the signs and statistical significance of many variables remain broadly comparable.

A notable discrepancy concerns the coefficient values associated with parcel area. The World Bank model includes both log area and its square, with both coefficients positive. This implies that prices increase with size at an increasing rate—an implausible result that contradicts standard economic intuition and empirical evidence. Moreover, the estimated coefficients are implausibly large: the log area terms range from approximately 11 to 15, and the squared terms exceed 30 across all specifications. By contrast, this study finds a negative elasticity of -0.1, consistent with the typical pattern of decreasing per-hectare prices for larger plots, reflecting economies of scale.

Both models identify a strong negative effect of war, though the magnitude differs: this study estimates a price drop of nearly 29% for post-invasion transactions, while the WB model reports a smaller effect (~6–17%). This discrepancy likely arises from different model controls. The WB study includes year dummies that may absorb part of the war effect, whereas this model isolates war onset more explicitly.

Some variable definitions also differ. In this model, road accessibility is measured using both road density and the presence of nearby roads, and distance to them. This may explain the positive association between road distance and price—an effect that could reflect disamenities such as fragmentation or pollution—whereas the World Bank model, which relies solely on distance-to-road measures, reports the expected negative relationship.

Another divergence concerns repeated transactions. The present model finds a strong and significant price premium for resales—approximately 5.8% per additional transaction—whereas the WB effect is weaker and significant only in the pooled specification.

While the World Bank model includes a dummy for the year 2024, which coincides with the start of the second stage of land reform, it does not isolate the reform effect. The estimated coefficient for 2024 in their model is negative relative to the base year 2021, though it is smaller in magnitude than the coefficients for 2022 and 2023. In contrast, this model includes two reform-specific variables: one indicating whether the transaction occurred after the market opening for legal entities, and another identifying legal entity buyers. The results suggest that the reform is associated with a 4.3% increase in prices, and that transactions involving legal entities are priced 44% higher than those involving individuals—evidence of a substantial structural shift in market behaviour not separately captured in the other specification.

Soil quality features also differ between the models. The WB study includes granular indicators like pH and acidity, while this study excludes soil type after robustness tests. These variables were found to be either insignificant or redundant, likely because bonitet—a composite measure of soil productivity—is already included in NMV. Moreover, soil variables likely exhibit non-linear effects that are poorly captured in linear models. Given the lack of expertise to interpret detailed soil measures metrics meaningfully, their inclusion in the model is not justified.

Finally, while both studies include land use and land cover controls, crop map-derived features that yield more coherent effects in the WB model. This study faces interpretability issues likely due to timing mismatches, measurement error, and lack of functional thresholds. All land cover variables (crops, grassland, forest, etc.) yield negative coefficients, which is difficult to reconcile with expected market valuations. These results are acknowledged as tentative and may require future refinement using more flexible model structures or more granular data.

Overall, while the two models differ in structure and scope, they converge on the importance of geography, and conflict exposure as key determinants of land price. This study also incorporates institutional and spatial dependencies not covered by the World Bank analysis, offering an alternative view of market dynamics during a period of reform and instability.

6.6. Directions for future research.

While the selected model captures spatial dependencies in land prices, several limitations in the current specification suggest directions for future research.

First, the spatial weights matrix used in this analysis is purely spatial and static. It does not incorporate the timing of transactions. In reality, land prices can only be influenced by past transactions, not future ones. A spatio-temporal weights matrix—where each observation is only influenced by earlier transactions within a spatial neighborhood—would better capture the directional nature of price spillovers. Using such a matrix would prevent the model from implicitly assuming symmetric influences over time. However, due to the size of the dataset, applying a spatio-temporal weights matrix would require custom algorithm implementation and optimized memory management beyond the scope of this study.

Second, the analysis relies on a single k-nearest neighbors specification. While k-NN offers computational efficiency, it may obscure relevant spatial structures in areas with variable settlement density. Alternative specifications, such as inverse distance matrices or contiguity-based weights, could reveal different spatial interaction patterns. Testing multiple matrix forms could improve the robustness of spatial effect estimates and help evaluate whether pricing spillovers are more localized or continuous.

In addition, the effects of land cover variables derived from crop classification maps remain inconclusive. This suggests that the current specification—based on linear transformations of surface cover shares—may not adequately capture the complex relationship between land use and price. Future work should consider refining the functional form, improving feature extraction methods or testing for nonlinear and interaction effects to better isolate the contribution of land cover to land valuation.

C h a p t e r 7

CONCLUSIONS

This study investigates the determinants of agricultural land prices in Ukraine using three spatial econometric models: a spatial lag model, a spatial error model, and a combined spatial autoregressive error model. All three confirm the presence of spatial dependence in land prices, but the spatial lag model is selected as the preferred specification due to its comparable performance and greater interpretability.

Among all factors, wartime conditions exert the most pronounced negative effect. Transactions occurring after the beginning of the full-scale invasion are associated with a 28.9% price discount, reflecting direct risks and market uncertainty. Furthermore, each percentage increase in distance from occupied territories is associated with a 0.064% increase in land price, showing how proximity to conflict zones shapes spatial variation in land value.

In contrast, the second stage of market reform in 2024 allowing legal entities to participate in the market is associated with a 4.3% increase in price, suggesting that demand-side pressure and increased competition are contributing to upward price movement in the post-reform period. When the buyer is a legal entity, the price premium rises sharply to 44.3%. These entities likely target more desirable, investment-oriented parcels, and their presence may increase competition or signal strategic value. Spatial lag effects further indicate that these premiums spill over to nearby plots, affecting local pricing dynamics.

Normative Monetary Valuation, which serves as a legally binding price floor for most transactions, remains closely associated with observed market prices. A 1% increase in NMV is associated with a 0.55% increase in transaction price, indicating that NMV continues to structure price formation. While the share of transactions occurring above the floor has increased—particularly following the 2024 reform—NMV still acts as a strong reference point, especially in less liquid segments of the market.

Sale order is positively associated with price: each additional transaction increases the price by approximately 5.8%, suggesting that parcels re-entering the market tend to have more favorable characteristics or benefit from reduced informational frictions over time. Beyond these effects, land prices are also systematically influenced by geographic and structural variables. Prices are lower for larger plots and for parcels located farther from Kyiv and oblast centers, consistent with scale effects and reduced accessibility. Additional factors such as land use designation, local income levels, and proximity to water and infrastructure also contribute meaningfully to price variation.

These findings demonstrate that land values in Ukraine are shaped by both policy reform and conflict exposure, with legal market liberalization and wartime conditions exerting large and opposing effects on price. While reform efforts have expanded participation and stimulated price growth, the war has introduced spatially uneven risks that continue to suppress land values and limit liquidity.

WORKS CITED

Anselin, Luc. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics* 27 (3): 247–267. https://doi.org/10.1016/s0169-5150(02)00077-4.

Borchers, Allison, Jennifer Ifft, and Todd Kuethe. 2014. "Linking the Price of Agricultural Land to Use Values and Amenities." *American Journal of Agricultural Economics* 96 (5): 1307–1320. https://doi.org/10.1093/ajae/aau041.

Breustedt, Gunnar, and Hendrik Habermann. 2008. "Determinants of Agricultural Cash Rents in Germany: A Spatial Econometric Analysis for Farm-Level Data." Paper presented at the 2008 International Congress of the European Association of Agricultural Economists, Ghent, Belgium, August 26–29. https://doi.org/10.22004/ag.econ.44462.

Breustedt, Gunnar, and Hendrik Habermann. 2011. "The Incidence of EU Per-Hectare Payments on Farmland Rental Rates: A Spatial Econometric Analysis of German Farm-Level Data." *Journal of Agricultural Economics* 62 (1): 225–243. https://doi.org/10.1111/j.1477-9552.2010.00286.x.

Burnett, J. Wesley, Donald J. Lacombe, and Steven Wallander. 2024. "Bayesian Spatial Analysis of US Agricultural Land Values." *Spatial Economic Analysis*, May, 1–16. doi:10.1080/17421772.2024.2342350.

Burns, Christopher, Nigel Key, Sarah Tulman, Allison Borchers, and Jeremy Weber. 2018. *Farmland Values, Land Ownership, and Returns to Farmland, 2000–2016*. Economic Research Report No. 245. Washington, DC: U.S. Department of Agriculture, Economic Research Service.

Choumert, Johanna, and Pascale Phélinas. 2015. "Determinants of Agricultural Land Values in Argentina." *Ecological Economics* 110: 134–140. https://doi.org/10.1016/j.ecolecon.2014.12.024.

Clark, Colin. 1973. *The Value of Agricultural Land*. Revised edition. Amsterdam: Elsevier.

Cotteleer, Geerte, Tracy Stobbe, and G. Cornelis van Kooten. 2010. "Bayesian Model Averaging in the Context of Spatial Hedonic Pricing: An Application to Farmland Values." *Journal of Regional Science* 51 (3): 540–557. https://doi.org/10.1111/j.1467-9787.2010.00699.x.

Deininger, Klaus, and Oleh Nivyevskyi. 2019. "Economic and Distributional Impact from Lifting the Farmland Sales Moratorium. Full Version." *Vox*

Ukraine, November 2019. https://voxukraine.org/en/economic-and-distributional-impact-from-lifting-the-farmland-sales-moratorium-in-ukraine-2.

Deininger, Klaus, Daniel Ayalew Ali, Eduard Bukin, and Andrii Martyn. 2024. *Reforming Land Valuation and Taxation in Ukraine: A Path towards Greater Sustainability, Fairness, and Transparency*. Policy Research Working Paper 10998. Washington, DC: World Bank.

Deininger, Klaus, Denys Nizalov, and Oleh Nivyevskyi. 2017. "Lifting the Moratorium on Farmland Sales: Why Supply and Demand Must Balance." *Vox Ukraine*, October 2017. https://voxukraine.org/en/lifting-the-moratorium-on-farmland-sales-why-supply-and-demand-must-balance.

European Court of Human Rights. 2018. *Zelenchuk and Tsytsyura v. Ukraine*, Nos. 846/16 and 1075/16. Accessed December 23, 2024. https://hudoc.echr.coe.int/eng-press#{"itemid":["003-6089956-7847640"]}.

Foundation "Democratic Initiatives" named after Ilko Kucheriv, and Razumkov Center. 2021. "Do Ukrainians Support the Opening of the Land Market, Under What Conditions, and What Protection of Their Interests Do They Expect?" National Survey Conducted from May 14 to May 19, 2021. https://dif.org.ua/article/chi-pidtrimuyut-ukraintsi-vidkrittya-rinku-zemli-na-yakikh-umovakh-ta-na-yakiy-zakhist-svoikh-interesiv-ochikuyut.

Griffith, Daniel A., and Francesco Lagona. 1998. "On the Quality of Likelihood-Based Estimators in Spatial Autoregressive Models When the Data Dependence Structure Is Misspecified." *Journal of Statistical Planning and Inference* 69: 153–174.

Guiling, Pam, B. Wade Brorsen, and Damona Doye. 2009. "Effect of Urban Proximity on Agricultural Land Values." *Land Economics* 85 (2): 252–264. https://doi.org/10.3368/le.85.2.252.

Huang, Haixiao, Gary Y. Miller, Bruce J. Sherrick, and Miguel I. Gómez. 2006. "Factors Influencing Illinois Farmland Values." *American Journal of Agricultural Economics* 88 (2): 458–470. https://doi.org/10.1111/j.1467-8276.2006.00871.x.

Kelejian, Harry H., and Ingmar R. Prucha. 1998. "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances." *The Journal of Real Estate Finance and Economics* 17 (1): 99–121. https://doi.org/10.1023/A:1007707430416.

Kostov, Philip. 2009. "A Spatial Quantile Regression Hedonic Model of Agricultural Land Prices." *Spatial Economic Analysis* 4 (1): 53–72. doi:10.1080/17421770802625957.

KSE Agrocenter. 2022. *Land Governance Review Ukraine*. January 2022. https://kse.ua/wp-content/uploads/2022/06/eng_jan2022_land_monitoring.pdf.

KSE Agrocenter. 2023. *Agricultural War Damages, Losses, and Needs Review*, Issue 3. April 24, 2023. https://kse.ua/wp-content/uploads/2023/05/RDNA2.pdf.

KSE Agrocenter. 2024. *Land Market Review Ukraine Q3 2024*. The Invincible Land Project. https://kse.ua/wp-content/uploads/2024/11/Land-Market-review-Q3-.pdf.

Kunwar, Samrat B., and Alok Bohara. 2017. "Climate Sensitivities and Farmland Values in Nepal: A Spatial Panel Ricardian Approach." *Journal of Development and Agricultural Economics* 9 (6): 145–161. https://doi.org/10.5897/JDAE2017.0822.

Marques, Felipe César, and Tiago S. Telles. 2022. "Spatial Effects Are Determinants of Agricultural Land Prices in Brazil." *Revista de Economia e Sociologia Rural 61*: e262420. https://doi.org/10.1590/1806-9479.2022.262420en.

Matvieiev, Mykhailo. 2023. "Ukraine's Agriculture and Farmland Market: The Impact of War." *Vox Ukraine*, May 2023. https://voxukraine.org/en/ukraines-agriculture-and-farmland-market-the-impact-of-war.

Minich, Ruslan, Orestis Panayiotou, Christoforos Pissarides, and Ilke Dagli-Hustings. 2021. *Land Reform in Ukraine: Attitudes & Expectations Based on SCORE 2021 Ukraine*. Kyiv: International Republican Institute and UNDP. https://www.scoreforpeace.org/files/land-reform/land-reform-en-final.pdf

Nilsson, Pia, and Sara Johansson. 2013. "Location Determinants of Agricultural Land Prices." *Jahrbuch für Regionalwissenschaft* 33 (1): 1–21. https://doi.org/10.1007/s10037-012-0071-4.

Nivievskyi, Oleg, Olha Halytsia, and Klaus Deininger. 2022. *The Impact of Farmland Sales Ban on Agricultural Productivity*. SSRN. https://ssrn.com/abstract=5144750.

Nizalov, Denys. 2019. "Land Reform Strategy in Ukraine Until 2024." Vox Ukraine. https://voxukraine.org/en/land-reform-strategy-in-ukraine-until-2020.

Rey, Sergio J., Dani Arribas-Bel, and Levi J. Wolf. 2020. *Geographic Data Science with Python*. "Spatial Regressions." Accessed February 22, 2025. https://geographicdata.science/book/intro.html.

Tavares, Vasco Capela, Fernando Tavares, and Eulália Santos. 2022. "The Value of Farmland and Its Determinants—The Current State of the Art" *Land* 11, no. 11: 1908. https://doi.org/10.3390/land11111908

Tsoodle, Leah, Bill Golden, and Allen Featherstone. 2003. "Determinants of Kansas Agricultural Land Values." Paper presented at the Southern Agricultural Economics Association Annual Meeting, Mobile, AL, February 1–5.

Wang, Haoying. 2018. "The Spatial Structure of Farmland Values: A Semiparametric Approach." *Agricultural and Resource Economics Review* 47 (3): 568–591. https://doi.org/10.1017/age.2017.35.

Weersink, Alfons, Steve Clark, Calum G. Turvey, and Rakhal Sarker. "The Effect of Agricultural Policy on Farmland Values." *Land Economics* 75, no. 3 (1999): 425–39. https://doi.org/10.2307/3147188.

Wengle, Susanne, Vitalii Dankevych, and Natalia Mamonova. 2024. *Farming the Black Earth: The History and Politics of Ukraine's Agricultural Land Markets Before, During and After the War.* Land Tenure Journal No. 1–24. Rome: Food and Agriculture Organization of the United Nations.

Appendix A

Table 6. Description of variables used in the model.

Variable	Source	Description
distance_to_hromada_ center_log	Cadaster Open Data	Log of distance (in km) from land plot centroid to the administrative center of the hromada.
distance_to_kyiv_log	Cadaster Open Data	Log of distance (in km) from land plot centroid to Kyiv.
distance_to_obl_ center_log	Cadaster Open Data	Log of distance (in km) from land plot centroid to the administrative center of the oblast.
inter_area_mount	Calculated	Interaction calculated as log land area multiplied by mountain hromada indicator.
inter_area_sea	Calculated	Interaction calculated as log land area multiplied by near seas indicator.
has_water	Cropmaps	Indicator that water or wetland is present on the plot.
is_not_cultivated	Cropmaps	Dummy variable equal to 1 if more than 50% of the plot is classified as bare or uncultivated.
percent_built_up	Cropmaps	Share of the plot area classified as artificial structures.
percent_crops	Cropmaps	Share of the plot area classified as crops.
percent_forest	Cropmaps	Share of the plot area classified as forest.
percent_grassland	Cropmaps	Share of the plot area classified as grassland.
percent_water	Cropmaps	Share of the plot area classified as water or wetland.

Table 6 – continued.

Variable	Source	Description
log_closest_dist_to _occupied_km	Deepstate, calculated	Log of the distance (in km) from the plot to the closest geo polygon identified is occupied.
edprou_indicator	Land monitoring relations	Presence of EDPROU code for a transaction in the land monitoring dataset. Indicator of whether buyer is a legal entity.
is_legal_market	Land monitoring relations	Dummy variable indicating if the transaction occurred after legal entities were allowed to purchase land (post-2024).
is_war	Land monitoring relations	Dummy variable indicating whether the transaction occurred after the start of the full-scale invasion (February 24, 2022).
LandAreaHa_log	Land monitoring relations	Log of the land area of the plot in hectares.
OwnershipType	Land monitoring relations	Dummy variables for ownership form: Cooperative, Private, State/Communal. Base category: Farm Enterprise.
LandPurpose	Land monitoring relations	Dummy variables for land use purpose. Categories: Commercial Farming, Commodity Farming, Gardening, Other. Base category: Personal Farming.
Quarter	Land monitoring relations	Number of the quarter when the transaction takes place.
sale_order	Land monitoring relations	Order of sale transaction for plots that are traded repeatedly. Range: 1-4
ValuationPerHectar _log	Land monitoring relations	Log of the official normative monetary valuation per hectare of the plot adjusted for inflation by CPI.

Table 6 – continued.

Variable	Source	Description
log_dist_to_ primary_km	OSM	Log of distance (in km) from land plot centroid to the closest primary road.
log_dist_to_ secondary_km	OSM	Log of distance (in km) from land plot centroid to the closest secondary road.
no_primary_road _within_10k	OSM	Indicator that there is not primary road within 10 km of land plot centroid.
paved_road _density	OSM	Index of paved road density within 10 km radius.
road_density	OSM	Index of all road density within 10 km radius.
distance_to _eu_log	Repository of Hromada-Level Data in Ukraine	Log of distance (in km) from the hromada center to the nearest point on the EU border.
hromada_area_log	Repository of Hromada-Level Data in Ukraine	Log of total land area of the hromada in square kilometers.
income_log	Repository of Hromada-Level Data in Ukraine	Log of hromada-level tax revenue per capita; based on 2021 values for transactions before 2022 and 2022 values otherwise.
is_urban_hromada	Repository of Hromada-Level Data in Ukraine	Dummy variable equal to 1 if the hromada is officially classified as urban.
mountain_hromada	Repository of Hromada-Level Data in Ukraine	Dummy variable equal to 1 if the hromada is officially classified as mountanous.

Table 6 – continued.

Variable	Source	Description
population_ 2022_log	Repository of Hromada-Level Data in Ukraine	Log of population of hromada in 2022.
urban_pct	Repository of Hromada-Level Data in Ukraine	Percent of urban dwellers in hromada.

Appendix B

OpenStreetMap (OSM) data were used to extract spatial features describing the physical and infrastructural context surrounding each cadastral land plot. The data were downloaded using the HOT Export Tool and manually configured to include relevant geographical feature types.

Unlike the cropmaps, which were processed using a high-resolution grid, the OSM data were exported as 17 manually defined large tiles. Each tile was sized to ensure complete coverage of Ukraine with adequate buffer zones and to guarantee the inclusion of all relevant nodes, ways, and relations necessary for constructing full road and waterway geometries.

Three geometry layers were included in each .gpkg tile:

- planet_osm_line: for roads, railways, and waterways
- planet_osm_polygon: for water bodies and terminal buildings
- planet_osm_point: for ferry terminal point features

The export was filtered using SQL-like WHERE clauses. Key selections included:

- Roads: highway IS NOT NULL
- Railways: railway IS NOT NULL
- Waterways: waterway IN ('river', 'canal', 'stream')
- Water bodies: natural='water'
- Transport nodes: amenity='ferry_terminal', building='ferry_terminal', building='train_station'

For each tile line geometries were filtered to extract road segments, classified by highway=*, and paved/unpaved status determined via surface=*. Polygon and line geometries were also used to identify water bodies, rivers, and streams. Points and polygons representing transport terminals were retained for potential accessibility analysis. Geometries were converted to EPSG:4326, validated, and merged into a single feature set. Duplicate osm_ids and overlapping geometries across tiles were removed during pre-processing.

All spatial metrics were calculated using the centroids of cadastral plots as reference points. The following metrics were derived:

- Count of water features: For each centroid, a 10 km buffer was created. All OSM features tagged as water-related (e.g., natural=water, waterway=river, canal, stream) that intersected the buffer were counted. Only distinct geometries were counted per centroid.

- Distance to primary road: The minimum Euclidean distance from the centroid to the nearest road with highway tag equal to 'primary' or 'primary_link'. Only roads with valid geometries were considered.

- Distance to secondary road: As above, but using highway tags 'secondary' and 'secondary_link'.

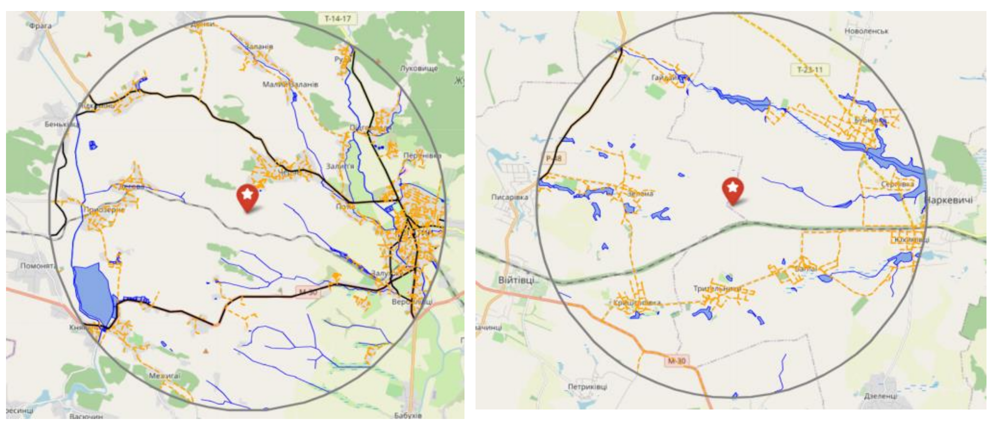
- Road density: The total length of all roads (from the list of known highway types) intersecting the 10 km buffer was calculated. These lengths were summed and divided by the area of the buffer to yield a road density in kilometers per square kilometer (km/km^2)

- Paved road density: A subset of the roads identified above was selected based on the surface tag. Roads whose surface matched a predefined list of paved types (e.g., asphalt, concrete, paving_stones) were included. Their total length was divided by the buffer area, as above.

Some limitations apply. Tag completeness and geometric detail in OSM vary across regions. In rural or sparsely mapped areas, surface=* may be missing, which can lead to underestimation of paved road access. The analysis does not account for road condition, width, or seasonal accessibility.

To verify the spatial accuracy and interpretability of the derived metrics visual verification was conducted for a subset of land plots. Two random examples are provided below (Figure 9). Each example overlays the following elements on a map:

- The centroid of the land plot (red marker)
- A 10 km buffer around the centroid (grey outline)
- All intersecting road segments (coloured by surface type: black for paved, orange for unpaved)
- All intersecting water features (blue lines or polygons)



Plot A has denser paved road network and more water features.

Plot B has some secondary roads but only 1 primary in the buffer zone.

Figure 8. Example land plots with 10 km buffer and extracted spatial features.

The maps are accompanied by the corresponding metric outputs for each plot (Table 7), including the cadastral number, geographic coordinates, water feature count, distances to primary and secondary roads, and computed road density.

These examples serve to confirm that the spatial joins and buffer operations were executed correctly and illustrate how different landscape contexts result in varying metric values, and provide qualitative validation of the paved/unpaved classification logic applied to road surfaces.

Table 7. OSM features for random plots

Feature name	**Value for Plot A**	**Value for Plot B**
Cadastral Number	2624487200:02:004:0061	3223386800:03:004:0007
Longitude	24.535021	31.743323
Latitude	49.414926	50.143198
count_water_features	211	47
dist_to_primary (meters)	2907.392838	30529.35463
dist_to_secondary (meters)	7181.236311	13556.2649
road_density	1.078643	0.436514
paved_road_density	0.21623	0.074923

Appendix C

This appendix describes the procedure used to extract crop classification data from the publicly available "2023-summer" layer hosted at ukraine-cropmaps.com via a WMS service.

The spatial extent of the cadastral dataset was used to define the bounding box for downloading cropmap tiles. A 0.1° buffer was added to each side of the bounding box to ensure coverage at the edges. The territory was divided into tiles of 0.25° × 0.25°, resulting in approximately 2,600 tiles.

Each tile was requested using the WMS GetMap endpoint with the following settings: CRS:84, Format: image/png, Image size: 1024 × 1024 pixels.

This corresponds to a ground resolution of approximately 22–25 meters per pixel, depending on latitude. Each pixel represents ~498 m². Each cadastral plot was matched to the tile it intersects. The tile image was loaded and the plot rasterized onto the same grid. Pixels within the masked region were extracted and classified into crop types based on predefined RGB-to-category mappings. The output was a per-plot percentage breakdown of dominant land cover types.

The selected resolution provides sufficient detail for identifying dominant land use types in plots ≥1 hectare. Smaller plots may be underrepresented, especially in sparsely covered or mixed-use areas. Some plots near borders may not have the most accurate breakdown as it is averaged between the breakdowns from multiple tiles without regard for proportion of the plot that belongs to each tile. The classification includes a range of agricultural crops, natural land covers, and built-up areas. The full list of categories available in the dataset is as follows:

Wheat, Rapeseed, Buckwheat, Maize, Sugar beet, Sunflower, Soybeans, Barley, Peas, Alfalfa, Potato, Grape, Other crops, Not cultivated, Grassland, Forest, Damaged forest, Wetland, Water, Gardens and parks, Bare land, and Artificial surfaces.

For about 2,400 plots (1.1% of all plots), no matching crop categories were identified in the rasterized cropmap layer. Inspection suggests that many of these plots are small (<0.5 ha), and their footprint intersects only a few pixels in the cropmap raster. At the chosen resolution (~498 m^2 per pixel), such plots may fall between classified areas or be dominated by edge effects, leading to missing or unmatched crop assignments.

To simplify analysis, the detailed cropmap categories were grouped into broader land use indicators. First, all crop-related percentages (such as wheat, sunflower, and maize) were summed to calculate the total share of cultivated land in each plot, stored as percent_crops. Forested areas, grasslands, and artificial surfaces were kept as individual features and renamed to percent_forest, percent_grassland, and percent_built_up. Water-related land types (Water and Wetland) were combined into a single variable, percent_water, and a binary indicator has_water was added to flag plots where any water was present. Similarly, a variable called is_not_cultivated was defined to identify plots where more than 50% of the area was labeled as either Not cultivated or Bare land.

Summary statistics of features derived from cropmaps are shown in Table 8.

Figure 9 illustrates tile #904. The full tile is shown on the right; the left panel displays a cropped version highlighting a single land plot with its rasterized mask overlaid (gray fill with red border). The plot is approximately 2 hectares in size and corresponds to 43 pixels at the current resolution. It is classified as 100% sunflower crops.

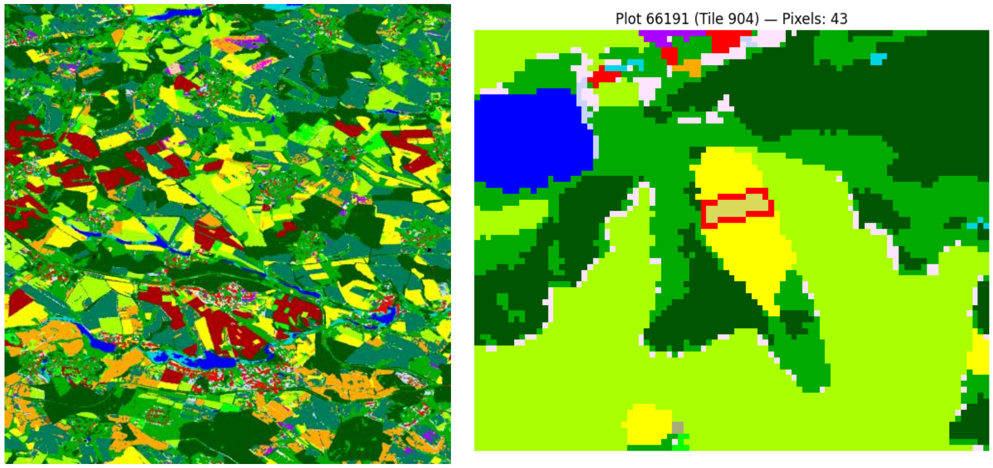


Figure 9. Example of a cropmaps tile.

Table 8. Summary statistics of crop features at plot level.

	Percent crops	Percent forest	Percent grassland	Percent water	Has water	Percent built up	Is not cultivated
mean	0.766	0.021	0.128	0.009	0.04	0.017	0.033
std	0.379	0.112	0.293	0.07	0.196	0.117	0.177

Appendix D

Table 9. Data Sources

Name	Link	Description
Land monitoring relations	https://land.gov.ua/monitorynh-zemelnykh-vidnosyn/	Main dataset used in the analysis. Contains detailed records of land transactions, including sale prices, plot characteristics, ownership types, and land use purposes. Regularly updated as part of Ukraine's national land monitoring system.
DeepState Occupied Territories	https://deepstatemap.live/	Live conflict tracker used to extract polygon boundaries of occupied territories. Provides geospatial data on frontline dynamics, which was used to compute distances from land plots to the nearest occupied area.
Repository of Hromada-Level Data in Ukraine	https://github.com/kse-ua/KSE-Loc-Data-Hub	KSE-Loc-Data-Hub is an open-access repository developed by the Kyiv School of Economics that provides comprehensive hromada-level data in Ukraine. It includes administrative, demographic, economic, and geospatial datasets, along with analytical scripts and visualizations. The repository supports research on decentralization reforms (2014–2022) and community resilience during the Russian invasion.

Table 9 – Continued.

Name	Link	Description
Crop Maps	https://ukraine-cropmaps.com/	Used to identify land cover and dominant agricultural use on each plot. Provides satellite-derived classifications for cropland, forest, grassland, water, and bare land, used to construct plot-level land use features.
OpenStreetMap	https://export.hotosm.org/v3/	Used to extract infrastructure and accessibility features, including road networks and nearby amenities. Data was used to calculate road density, distance to roads, and presence of key infrastructure within defined radii around each plot.
Harmonized World Soil Database	https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v20/en/	Used to assign dominant soil type and classification to each plot. Provides standardized global data on soil properties, used here primarily for categorical soil indicators such as fertility class and WRB classification.
Cadaster Open Data	https://kadastr.live/	Data on exact coordinates and polygons of land plots.

Appendix E

Table 10. Indirect effects in SLM model.

Variable	Direct	Indirect	Total
distance_to_kyiv_log	-0.152	-0.0226	-0.1746
distance_to_obl_center_log	-0.0298	-0.0044	-0.0342
LandAreaHa_log	-0.0973	-0.0145	-0.1117
ValuationPerHectar_log	0.5457	0.0812	0.6269
is_war	-0.341	-0.0507	-0.3917
is_legal_market	0.0425	0.0063	0.0488
distance_to_eu_log	-0.0427	-0.0063	-0.049
income_log	0.1258	0.0187	0.1445
mountain_hromada	1.4889	0.2214	1.7103
near_seas	0.3864	0.0575	0.4439
urban_pct	-0.1144	-0.017	-0.1314
population_2022_log	-0.0324	-0.0048	-0.0372
inter_area_mount	-1.2674	-0.1885	-1.4559
inter_area_sea	-0.2175	-0.0323	-0.2498
percent_crops	-0.6064	-0.0902	-0.6966
percent_forest	-0.3955	-0.0588	-0.4544
percent_built_up	0.0287	0.0043	0.033
percent_grassland	-0.551	-0.0819	-0.6329

Table 10 - Contimued.

Variable	Direct	Indirect	Total
percent_water	-0.4193	-0.0624	-0.4816
is_not_cultivated	-0.4376	-0.0651	-0.5027
log_count_water_features	0.0436	0.0065	0.05
log_dist_to_primary_km	0.0143	0.0021	0.0164
log_dist_to_secondary_km	0.0169	0.0025	0.0194
edprou_indicator	0.3666	0.0545	0.4211
paved_road_density	0.0496	0.0074	0.057
road_density	0.0952	0.0142	0.1093
log_closest_dist_to_occupied_km	0.0639	0.0095	0.0734
is_urban_hromada	-0.0074	-0.0011	-0.0085
has_water	-0.1083	-0.0161	-0.1244
no_primary_road_within_10k	-0.0722	-0.0107	-0.083
distance_to_hromada_center_log	-0.0304	-0.0045	-0.0349
hromada_area_log	-0.0905	-0.0135	-0.1039
sale_order	0.058	0.0086	0.0667
purpose_commercial_farm	0.0571	0.0085	0.0656
purpose_commodity_farm	0.1447	0.0215	0.1662
purpose_gardening	0.9199	0.1368	1.0567
purpose_other	0.0545	0.0081	0.0626
Ownership_Private	-0.1296	-0.0193	-0.1489

Table 10 – Continued.

Variable	Direct	Indirect	Total
Ownership_Cooperative	1.703	0.2532	1.9562
Ownership_StateCommunal	0.021	0.0031	0.0241
Quarter_2	0.004	0.0006	0.0046
Quarter_3	0.0332	0.0049	0.0382
Quarter_4	0.013	0.0019	0.0149