

THE IMPACT OF THE FULL-SCALE WAR ON AGRICULTURAL
PERFORMANCE IN UKRAINE

by

Oleksandr Petrenko

A thesis submitted in partial fulfilment of the
requirements for the degree of

BA in Business Economics, Social Sciences Department

Kyiv School of Economics

|

2025

Thesis Supervisor:

Professor Oleg Nivievskyi

Kyiv School of Economics

Abstract

THE IMPACT OF THE FULL-SCALE WAR ON AGRICULTURAL PERFORMANCE IN
UKRAINE

by

Petrenko Oleksandr

Thesis Supervisor:

Professor Oleg Nivievskyi

This paper assesses the impact of Russia's full-scale invasion on agricultural production in Ukraine by comparing production in the absence of war with actual production. Previous literature on losses in agricultural production due to the war provides a fairly comprehensive overview of the main components of losses and how the country's agricultural sector has changed and adapted. Financial losses in the agricultural sector are divided into damage to assets and losses related to lost income, which primarily concerns production cuts due to the occupation of territories. The literature also examines the impact of war on the structure of production, the impact on soil, and the forced closure of farms in the frontline zone due to military operations on their agricultural land. This study is an important step in assessing the impact of the war on Ukraine's agricultural production, because under the current conditions of agricultural business, it is necessary to clearly understand what losses production has suffered precisely because of the war, and not because of weather conditions. Forecasting is used to construct an alternative scenario for the development of agricultural production in the absence of war, including regular changes in natural factors in forecasts. The difference between actual and predicted yields showed the impact of the war on production performance. The yield assessment covers major crops such as sunflower, rapeseed, wheat, corn, soybeans, and barley. The forecast was based on data obtained from FAOSTAT and the World Bank for the years 1992 to 2023, using time series analysis with endogenous variables (ARIMAX) and

machine learning (Random Forest Regression) methods. The results showed that in 2022, the actual yield differs from the forecast in the absence of war by almost 0.2 tons per ha on average according to two analysis methods. The study also found that rapeseed was almost unaffected by the war, while corn performance was most significantly affected.

TABLE OF CONTENTS

LIST OF TABLES	4
LIST OF FIGURES	5
Chapter 1. Introduction	6
Chapter 2. Literature review	9
2.1 Ukraine's agricultural losses are related to the war	9
2.1 Methods for working with agricultural data	13
Chapter 3. Data	14
Chapter 4. Methodology	19
4.1 Random Forest Regression.....	19
4.2 ARIMA (Autoregressive Integrated Moving Average)	21
Chapter 5. Results	23
5.1 ARIMA Results.....	23
5.2 RFR Results.....	29
Chapter 6 Discussion.....	32
Chapter 7 Conclusion	34
Works Cited	36
Appendix 1	39

LIST OF TABLES

Number	Page
Table 1. Descriptive Statistics of crop yield.....	14
Table 2. Descriptive statistics of weather indicators	15
Table 3. Descriptive statistics of land area.....	17
Table 4. Descriptive statistics of variables	18
Table 5. Results of checking historical yield data for 1992-2023 to identify stationarity	24
Table 6. Structure of ARIMAX models for each crop separately and RMSE (Root Mean Square Error) with cross-validation of these models.....	26
Table 7. Results of testing endogenous variables for statistical significance in models for each crop separately. Statistical significance at the 5% level. If the significant p-value is less than 0.05, then the predictor is considered statistically significant (green), otherwise red....	26
Table 8. Forecasting results using ARIMA/ARIMAX models in centners per hectare. Loss shows the percentage difference between the forecasted and actual values. The difference is equal to the actual value minus the forecast.	28
Table 9. Mean square error (MSE) for each crop in RFR	30
Table 10. Estimated RFR yield for 2022 compared to actual yield in centners per hectare	31
Table 11. Agricultural crops and variables used in RFR models.....	39

LIST OF FIGURES

Figure 1. Structure of losses in Ukraine's agricultural sector as of 2024	10
Figure 2. Historical crop yields 1992-2023 in centners per hectare. The dotted line represents the year 2022(the start of full-scale war).....	15
Figure 3. Average temperature during the growing season across Ukraine from 1992 to 2022 and average annual precipitation across all regions from 1992 to 2023	16
Figure 4. The black line represents the actual yield in centners per hectare, while the orange line shows the projected yield. The dotted line marks 2022 as the start of full-scale war...	28

Chapter 1. Introduction

One of the leading economic activities in Ukraine has always been the agricultural sector. The agricultural sector is the third largest in the country's economy, accounting for 10.9 percent of the country's total GDP as of 2021 (the last year before Russia's full-scale invasion in 2022)(Albaladejo Román 3). During the full-scale invasion, the agricultural sector suffered from massive shelling, occupation of fertile agricultural land, loss of human resources, loss of established supply channels, etc. As of 2023, the share of value added by the agricultural sector in gross domestic product (GDP) had fallen to 7.4 per cent (World Bank). However, despite the difficult conditions in Ukraine, the country remains the world's leading exporter of sunflower oil, accounting for approximately half of global exports. Ukraine is also the world's third largest exporter of corn and sixth largest exporter of wheat (before the war, it was fifth) (Martyshev et al. 4). By the end of 2024, the country's total exports of agricultural products amounted to about \$24.5 billion, with the share of agricultural products in total exports amounting to 59%. This figure is the second highest in the record line after 2021, when Ukraine's agricultural exports totalled about \$27.7 billion (Vitaliy Koval).

Ukraine's agricultural land is almost two-thirds covered by fertile soil, namely black soil, which contains 9 per cent humus and covers about 28 million hectares, making our country super fertile in terms of crops (Albaladejo Román 2). The climate allows for sufficient rainfall, sunshine and warmth, and a large number of wide, steep rivers, lakes and reservoirs make the country rich in water resources. Approximately 41 million hectares were involved in agricultural activities in Ukraine, which is 68.5 percent of all Ukrainian territory before the full-scale invasion. In 2021, Ukraine's main agricultural crops were cereals, mainly wheat, corn and barley, which accounted for about 56% of the total sown area. Another 32 per cent of the acreage was allocated to major oil crops, such as sunflower, rapeseed, and soybeans (Albaladejo Román 3). In the western regions of

Ukraine, where the land and climate are less suitable for growing staple grains and oilseeds, large quantities of potatoes are common.

For more than two decades, Ukraine's agricultural sector has developed rapidly, demonstrating growth in crop yields, share of GDP, and exports. However, on February 24, 2022, Russia launched a full-scale invasion of Ukraine, causing significant damage to the agricultural sector and slowing down its potential development. The war has caused very serious losses and problems for the Ukrainian economy as a whole, and in three years, the country has more or less adapted to life under fire. The consequences were felt not only by the economy and manufacturing, but also by millions of Ukrainians who died or were forced to leave their homes. As of 2025, nearly 6.9 million Ukrainians live abroad, approximately 3 million of whom left before the outbreak of full-scale war (Kruglenko). By the beginning of 2024, there were about 80,000 dead Ukrainian soldiers and about 400,000 wounded, but these are only officially confirmed figures, and a large number of Ukrainian soldiers are currently in captivity or missing (Vira Khmelnytska). Among this number of people were also farmers who fed both the Ukrainian people and people from other countries. It is estimated that by 2022, approximately 400 million people worldwide were dependent on Ukrainian grain (Martyshev et al. 5).

As a result of Russia's full-scale war in Ukraine in 2022, the area under crops decreased by almost 21 per cent, but between 2022 and 2023, thanks to the resilience and endurance of the Ukrainian military, about 5 percent was recovered, and as of 2023, about 15 percent of the land under crops was still under occupation (Bogonos et al.). Due to the loss of land for crops, the agricultural industry lost up to US\$70 billion, but the loss of profit is not only due to the loss of land, also to the fact that the domestic price of products has been severely reduced due to weak sales and export opportunities. Due to the war, most key inputs increased in prices, driven by both inflation and higher logistics costs. These two effects combined, namely the extreme drop in commodity prices and the rise in prices for key inputs, led to a drop in agricultural income.

The question that will be answered in this study concerns crop yields, which measure the efficiency of agricultural production, namely the number of centners of crop harvested per hectare. Many studies devoted to the topic of agricultural losses due to the war have already answered the question of what financial losses Ukraine has suffered. In addition to previous studies, this research focuses specifically on crop yields and the losses that farmers have suffered in production as a result of the war. Since this study models crop yields using forecasting methods, we will be able to see the actual impact of the war on agricultural output and calculate losses per hectare in centners. As mentioned earlier in this paper, the rise in prices for key inputs and the decline in profitability logically entailed losses in the amount of inputs used. In this article, crop losses due to the war will be calculated by comparing them with the projected harvests for 2022–2023 in the absence of war. The difference between the projected and actual yields will reflect the impact of the war shock and allow us to calculate the production losses incurred by Ukrainian farmers per hectare of crops.

This section examined the importance of the agricultural sector for Ukraine's economy and food security, as well as the impact of the war on the country. In the second section, we review the existing literature on the impact of the war on the agricultural sector itself. The third section discusses the data and data sources used to build the models, which will also be discussed in the methodology section. Part of the results will consist of empirical studies that calculate the difference between actual and projected yields, as well as losses.

Chapter 2. Literature review

2.1 War related agricultural losses in Ukraine

The existing literature on agricultural losses has approached the issue both in terms of the value of assets that have been partially or completely destroyed, as well as in terms of lost profits or extra costs of fixing the problems that Russia has brought to Ukraine, and, last but not least, some studies have also focused on farms that have been forced to cease operations, either because of the hostilities on their land or because they are incapacitated. The total value of destroyed assets was estimated at \$10.3 billion, with machinery used in agricultural production accounting for \$5.8 billion of this figure as of December 2023. In total, about 19 per cent of all available agricultural machinery was either partially or completely damaged (Neyter et al. 8). As of December 2024, losses in agricultural assets have fortunately not changed very significantly compared to December 2023. In financial terms, losses in assets increased by about 9.2 per cent over the year, due to the fact that most of the assets were in the frontline and had already been destroyed in the first year of the war (World Bank et al. 117). Therefore, as of December 2024, losses in assets had already reached approximately \$11.2 billion (World Bank et al. 117).

As for financial losses, in December 2023, they amounted to approximately \$69.8 billion (Neyter et al. 12). These losses are associated with a decline in the profitability of agricultural activities, including reduced production volumes, falling prices, and higher production costs. As of December 2024, losses had already reached \$72.7 billion, with the structure of losses more closely linked to the decline in annual crop production, which accounts for 51% of total losses (see Fig. 1)(World Bank et al. 118). Losses related to reduced production are followed by a decline in prices for key agricultural products such as corn, wheat, barley, and oilseeds. The structure of losses is 34%

related to price declines (see Fig. 1) (World Bank et al. 118). Another 6% of losses are related to rising raw material prices (see Fig. 1). Other losses are demonstrated at Figure 1. Overall, according to the World Bank, Ukraine has adapted to the conditions of war and demonstrated good harvests in 2023, prompting the authors to revise the structure of losses for 2023. In 2024, however, yields were not very encouraging, largely due to weather conditions, but as the country had already adapted well to the conditions of war, the share of losses due to the war decreased (World Bank et al. 118).

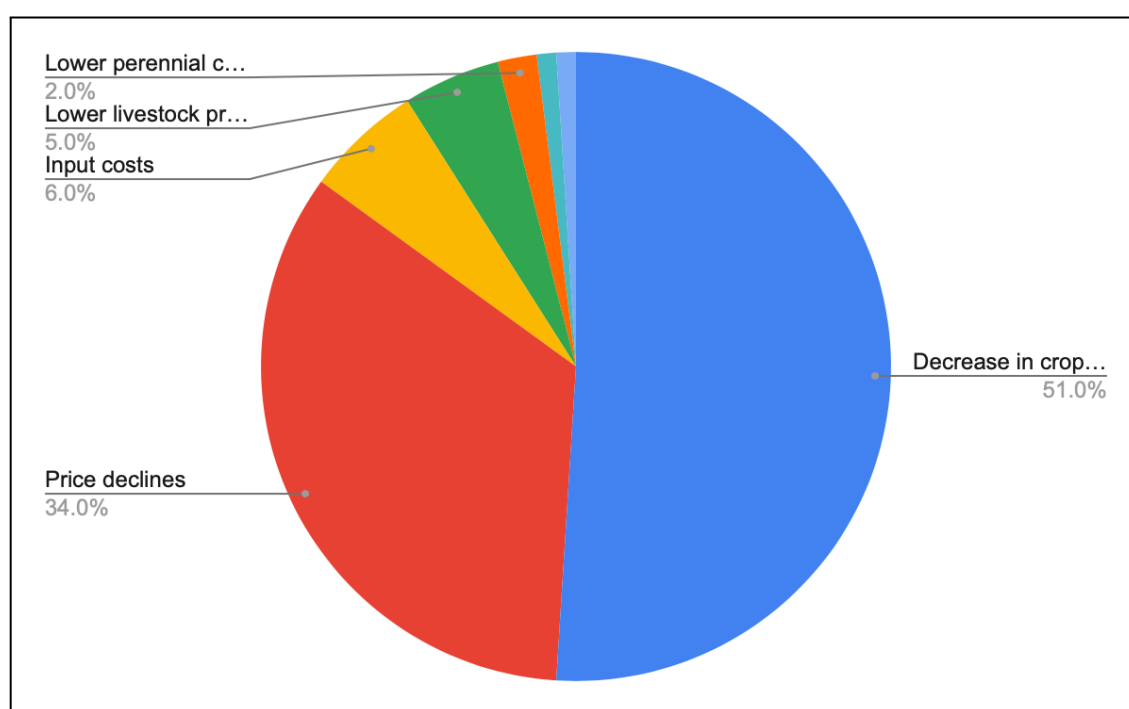


Figure 1. Structure of losses in Ukraine's agricultural sector as of 2024

Source: Based on World Bank RDNA 4

Roman Neyter et al. estimates losses based on a comparison of actual data on harvested volumes for 2022-2023 and the volumes that could have been produced by Ukraine based on the 2021 acreage and average yields for 2000-2021, thus using average yields, they try to smooth out the impact of weather conditions, which are a very significant factor in agriculture (20). The World Bank et al., in their collective work Fourth Rapid Damage and Needs Assessment (RDNA4), shared the limitations faced by the assessment of war damage, namely the limitations in the information content on the

territory of Ukraine, which are beyond the control of the Ukrainian government. This makes it impossible to conduct reliable optimization and research on reliable losses (119).

As noted in RDNA4, 6 % of losses are related to higher raw material prices (see Fig1). The rise in prices for key resources used in production was caused by high inflation amid serious disruptions and changes in logistics channels (Bogonos et al.). Among the main changes in the structure of costs and agricultural production are factors such as oil (fuel) prices, especially in crop production, where agricultural machinery is a key component of the labor force. The share of fertilizer costs remained almost unchanged for most crops, with the exception of sunflower, rapeseed, and oats, where it decreased (Bogonos et al.). Research on changes in the structure of costs is very important for understanding which factors of production have led to losses in crop production, so Bogonos et al. address this need with their own survey in 2023, which showed that the increase in costs led to a decline in the incomes of agricultural producers.

In addition to declining agricultural profitability, including reduced production volumes, falling prices, and rising production costs, the physical impact on the land also negatively affects agricultural performance. The question arises as to the impact of the war on the chemical condition of Ukrainian soils and crop yields. The soil cover is degrading due to the movement of heavy equipment, which leaves behind both fuel and oil residues, leading to the destruction of vegetation (Sushchuk). Artillery shelling, in addition to creating large craters from falls and explosions, scatters a significant amount of metal fragments and heavy metals across the fields. Numerous fires lead to the degradation of the Ukrainian ecosystem, and explosive substances severely pollute the soil, water, and plants. More than 90,000 square kilometers are partially or completely unsuitable for agricultural activities due to the above factors (Sushchuk). Therefore, economic factors are not the only ones that have a long-term impact on the performance of agricultural production. Taking as an example agricultural holdings of up to 250 hectares, which account for 65 percent of all

holdings, their number decreased by 7.7 percent with the start of the full-scale war, and 87 percent of the holdings that were forced to cease operations were located on the front line, making their continued existence impossible (FAO). As of 2022, 80 percent of the land on the front line was inaccessible for harvesting due to mines, unexploded ordnance, or other remnants of munitions, with an average of 14 percent of land inaccessible for work in the central regions (FAO).

Based on the wide range of losses in Ukraine's agricultural sector as a result of the war, as described in the literature, and the large number of assets destroyed by Russia, the needs of the agricultural sector for effective recovery can be estimated for at least the next 10 years. Overall, the needs of the agricultural sector are estimated at \$56 billion (Nivievskyi et al.). When calculating the financial needs, they were divided into restoration needs and recovery needs. The methodological approach proposed by Nivievskyi et al. is that restoration needs are calculated as compensation for what was destroyed plus a 20% bonus to restore everything better than it was. Recovery needs are calculated based on the principle of restoring production to pre-war levels and include such sub-groups as immediate needs, long-term needs, and needs for support to strengthen state institutions in the agricultural sector.

Previous studies on the impact of war on Ukraine's agricultural sector are extensive and cover almost the entire sector; the literature addresses both asset values and profit losses as well as future prospects. However, this study analyzes how the war affected actual crop yields. Ukraine has lost more than 15% of the land on which grain and oil crops were grown (Bogonos et al.), so crop volumes are significantly lower and, as mentioned above in the World Bank et al. report, 51% of all profit losses are due to reduced production. Instead, in this paper, we look at the losses in agricultural production profitability due to the difference between yields in peacetime and wartime. Yields will be forecast in the absence of war in order to assess the difference, which will allow us to isolate the actual impact of the war, taking into account other factors affecting yields.

2.2 Methods for working with agricultural data

Time series analysis is widely used to forecast long-term crop yields. Often, time series analysis allows predicting future values using only past values and their errors. However, for greater accuracy and confidence in forecasting in the agricultural sector, models with additional exogenous variables are used. Traditional time series methods, such as Autoregressive Integrated Moving Average (ARIMA), for example, are widely used in yield forecasting, but these models avoid external factors that influence agricultural yields, such as weather conditions (Kaushik et al.). In order to make the analysis more explanatory, it is worth using exogenous variables, and for this purpose, the Auto-Regressive Integrated Moving Average with Exogenous Variables (ARIMAX) model is suitable, which allows external factors such as climatic conditions to be included in the forecast (Kaushik et al.). Models with external predictors, such as ARIAMAX, tend to show stronger explanatory power than those that rely solely on historical values (Ghosh et al.). Ghosh et al. use rainfall data and various data formats in their work to study the impact of exogenous variables on time series analysis.

Machine learning techniques, in particular regression models, are actively used in work on yield with large data sets. Regression models allow us to identify dependencies between various factors, such as weather conditions or pesticide use, for processing historical data to build a yield forecast. Random Forest Regression (RFR) method was mentioned as the most successful in working with large agricultural data sets because it allows overcoming overfitting in noisy data sets (Jorvekar et al. 133). Also RFR has an advantage over other methods in working with agricultural data in terms of explanatory power and loading, because the method selects different combinations of data and identifies the most effective ones (Gupta et al.).

Chapter 3 Data

To work on yield, it is necessary to assess historical yield indicators in Ukraine. The study uses data on the yield of major crops such as barley, wheat, rapeseed, corn, sunflower, and soybeans. Time series data were obtained from FAOSTAT, starting in 1992 and limited to the last year available on the service, 2023. Yields are estimated in centners per hectare, where one centner equals 100 kg. The summary statistics are presented in Table 1, and Figure 2 shows the dynamics of yields during 1992–2023.

Table 1. Summary Statistics of crop yield in centners per hectare

Source: Own calculations based on FAOSTAT data of crop yields

Variable	N	Mean	SD	Min	Max
Barley	32	25.66	6.59	14.62	38.17
Corn	32	46.58	17.44	23.61	78.43
Rapeseed	32	17.39	7.67	6.63	29.25
Soy	32	15.95	5.71	7.2	26.4
Sunflower	32	15.87	5.26	8.9	25.59
Wheat	32	32.78	6.91	19.75	46.42

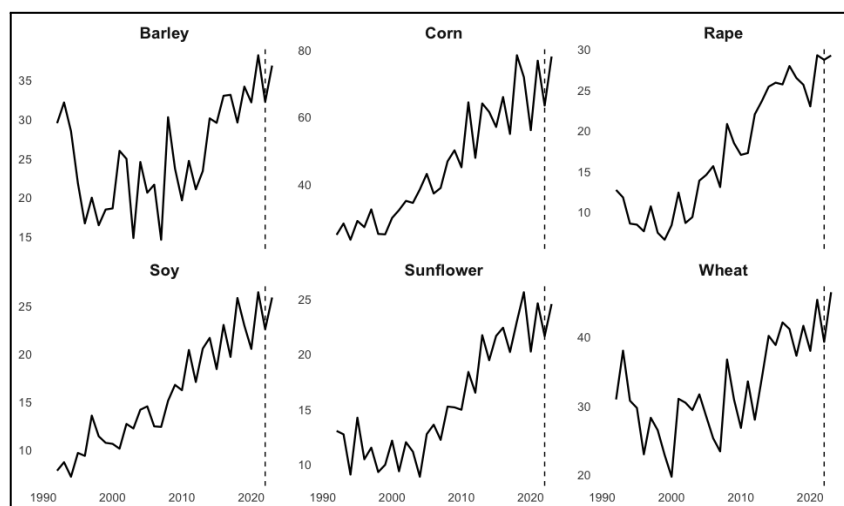


Figure 2. Historical crop yields 1992-2023 in centners per hectare. The dotted line represents the year 2022(the start of full-scale war).

Source: Calculations based on FAOSTAT data of crop yields

One of the key growth periods is spring, so the average temperature will be estimated during the growing season from April to August and measured in degrees Celsius. Precipitation, which is necessary for successful plant growth, is measured as the total amount in millimeters for the entire year, weighted by the average value for all regions of Ukraine. Data was obtained from the World Bank. Summary statistics on weather conditions are presented in Table 2. Figure 3 shows how the average temperature during cultivation changed with the total amount of precipitation per year during 1992-2023.

Table 2. Descriptive statistics of weather indicators

Source: Own calculations based on World Bank data

Variable	N	Mean	SD	Min	Max
Average temperature in degrees Celsius (growing season)	32	17.25	0.86	15.50	19.04
Millimeters of rain	33	559.73	53.88	455.26	669.38

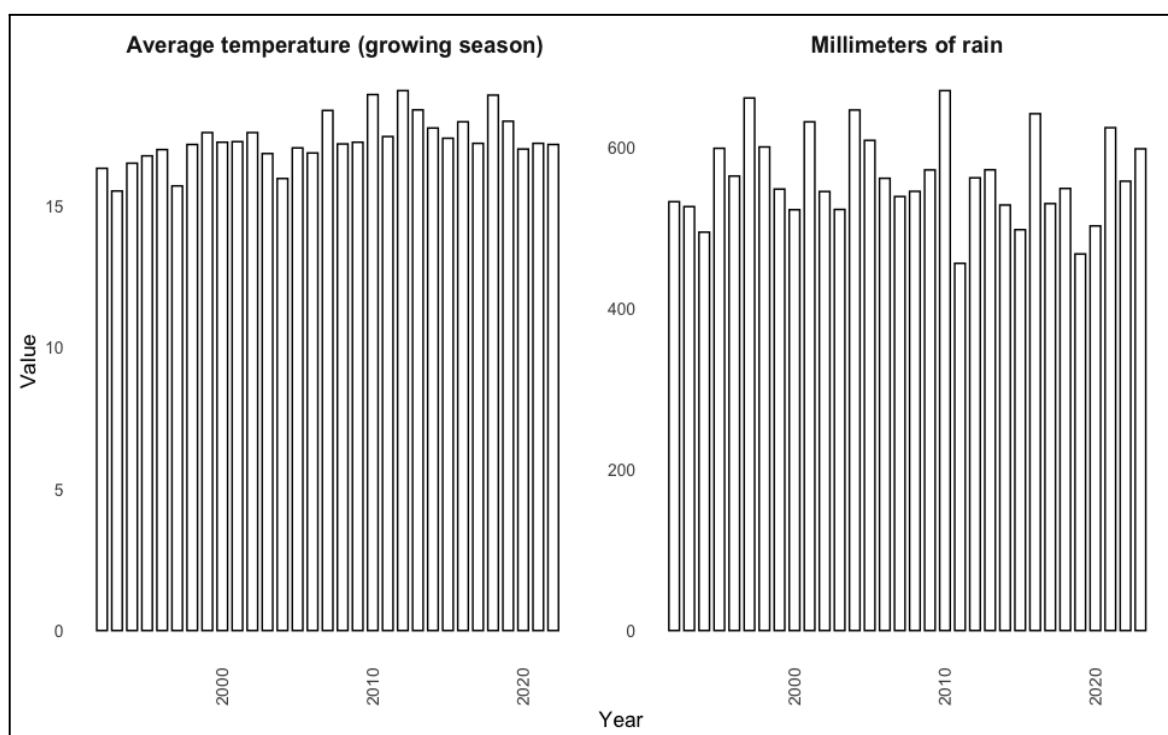


Figure 3. Average temperature during the growing season across Ukraine from 1992 to 2022 and average annual precipitation across all regions from 1992 to 2023.

Source: Own calculations based on World Bank data

The study also takes into account the number of thousands of hectares with a particular crop across the country. A hectare is a commonly used unit of measurement for land area, equal to 10,000 square meters. For the purposes of the study, in order to forecast yields for 2022 in the absence of war, it was assumed that the area of land used for growing specific crops would remain the same as in 2021, as the share of crops remains more or less unchanged over the years. Historical data on the number of hectares under individual crops were obtained from FAOSTAT for the period 1992 to 2023. The summary statistics are presented in Table 3 below.

Table 3. Descriptive statistics of land area (hectares)**Source: Own calculations based on FAOSTAT data**

Variable	N	Mean	SD	Min	Max
Rapeseed	32	557,851.4	475,865.9	21,250	1,431,600
Barley	32	3,559,258.9	945,024.7	1,494,300	5,236,200
Corn	32	2,726,070.6	1,611,695.4	651,900	5,481,800
Soy	32	811,399.8	739,728.4	13,520	2,135,600
Sunflower	32	4,018,115.5	1,586,184.9	1,629,000	6,665,100
Wheat	32	6,052,838.1	776,632.7	3,890,000	7,099,400

In order to assess the impact of the actual reduction of fertilisers and chemicals, the quality of the soil needs to be assessed, namely its content of essential chemical elements. From FAOSTAT, the data is received for Ukraine on the balance of nutrients in the soil (NPK), which is calculated as the sum of input resources minus output resources (at the output point in the nutrient balance, there is a yield reduction that depletes the soil.). N stands for the nitrogen content of the soil, P for the phosphorus content, and K for the potassium content of Ukrainian soils. The unit of measurement for the key nutrients in the soil is kilograms per hectare. Since FAOSTAT calculates the nutrient balance as the difference between the sum of inputs and outputs (harvesting), NPK for 2022 can not be used to estimate yields in the absence of war. The war has made serious adjustments to the content of inputs to the soil, so if NPK for 2022 will be used, the war factor will not be avoided. However, under peaceful conditions, crop cultivation technologies remain more or less unchanged over the years, so this paper uses the 2021 NPK data for the soil. In order to estimate the natural impact on soil organic matter in 2022, such natural factors as atmospheric deposition, biological

fixation, leaching, which are almost not affected by the war, but have an impact on soil nitrogen content are used.

Table 4. Descriptive statistics of variables

Source: Own calculations based on FAOSTAT data of Cropland Nutrient Balance

Variable	N	Mean	SD	Min	Max
AtmosphericDeposition (Kg per hectare)	31	7.63	1.178	6.69	12.37
BiologicalFixation (Kg per hectare)	31	4.32	2.32	1.405	8.72
Leaching (Kg per hectare)	31	-4.703	1.397	-9.18	-3.08
Cropland nitrogen per unit area (Kg per hectare)	31	14.55	12.084	1.63	56.55
Cropland phosphorus per unit area (Kg per hectare)	31	-0.297	3.152	-4.16	9.48
Cropland potassium per unit area (Kg per hectare)	31	7.73	12.15	-5.03	46.09

In this paper, the data generalised across Ukraine is used, as work on a regional basis is not available at the moment due to the lack of necessary regional data. The FAOSTAT service provides a portion of key and very informative data, but all data is aggregated for the whole country. FAOSTAT works with most countries in the world, collecting large data sets, they focus on a generalised view of the country's agricultural sector, not a regional breakdown.

Chapter 4 Methodology

This study is dedicated to examining the relationship between key agricultural factors and the yield of major crops in Ukraine, such as corn, sunflower, wheat, rapeseed, soybeans, and barley. To forecast yields in the absence of war, we use two methods with different methodological approaches. The first method, namely random forest regression, is a method based on key resources in the construction process and is used to investigate which factors influence yields and to what extent. The second method, namely time series analysis — ARIMAX (autoregressive integrated moving average with exogenous variables) — is used to forecast yields in the absence of war, which in this method is based not only on historical yields as a factor, as assumed in the usual ARIMA, but also on external weather factors. The RFR model simulates complex relationships between variables, overcoming multicollinearity and overfitting, while ARIMAX tracks trends and the impact of past deviations from the mean as factors influencing the future.

4.1 Random Forest Regression

In this paper, one of the regression analysis methods used to investigate the yields of major crops in Ukraine in the absence of war was Random Forest Regression, as in previous work on yield forecasting, this method is leading in terms of R^2 and in dealing with the sharpened agricultural data. This is a popular method used to solve regression problems in data processing, which avoids overfitting (Jorveka et al.). When working with Random Forest Regression as a model for forecasting yields for 2022, the model takes into account the complexity of the interaction between key factors that have an impact on the dependent variable, namely yield. Let's assume an interaction between the average temperature during the growing season and kilograms of nitrogen per hectare. The principle of this method is that each tree in the forest operates separately between the different

views of the data, providing different ways of processing the same data. All the different views of the data (trees) are added to the forest by calculating the average of the final dependent variable. Thus, with different sets of interactions between data, random forest regression provides insight into the importance of different features and avoids overfitting. The RFR model looks like this:

$$\hat{y}_t = (1/B) \times (f_1(X_t) + f_2(X_t) + \dots + f^B(X_t))$$

\hat{y}_t is the predicted value of the yield in time, let's assume in our case it will be 2022. B is responsible for the number of trees that will be used. $f_1(X_t) + f_2(X_t) + \dots + f^B(X_t)$ are responsible for each tree individually, and this is precisely what is predicted for each tree. Each X_t is the set of predictors at year t (such as average temperature during the growing season, precipitation, NPK balance in the soil). Each individual decision tree models a random sample of data differently, and the final result is the arithmetic mean of all individual trees.

The random forest method is well suited for processing very noisy agricultural data sets containing a large number of interactions between factors. This method is used in a large number of studies related to agricultural data because it handles the problem of nonlinear relationships well without excessive data transformations and is also suitable for working with high degrees of correlation, which is common in agricultural datasets. In this work, it is very important to identify the factors that most influence yield because the data sample is quite weak. RFR is suitable for this purpose since it automatically selects the most influential predictors and evaluates their impact on the dependent variable. If some variables have a negative impact on the explanatory power of the model, these variables can be excluded. Another argument in favor of using this method is that it allows you to overcome the problem of outliers and missing values by dividing all possible data sets into separate trees, so that individual missing values or outliers do not affect the overall

performance of the model, since all trees with different data sets are ultimately combined into an average forest.

4.2 ARIMA (Autoregressive Integrated Moving Average)/ ARIMAX (Autoregressive Integrated Moving Average with Exogenous Variables)

As a supplement to the regression forecasting model, in this work, where the main objective is to forecast crop yields in Ukraine in the absence of war, time series analysis is also used, namely the ARIMAX (autoregressive integrated moving average with exogenous variables) model to forecast crop yields in the absence of war for 2022–2023. The ARIMA model is often used in long-term forecasting because it works with trends in the sector and structural models, differentiating data. On the other hand, ARIMAX has the same meaning as ARIMA, but is supplemented by additional predictive factors that are used to increase the explanatory power and flexibility of the forecast depending on the values of the exogenous variables that will be used.

Yield data are predominantly time series affected by a number of factors. The ARIMA/ARIMAX model is suitable for yield forecasting because it uses past values and past errors as actual predictors. Technically, the ARIMA (p,d,q) model can be described using the following model:

$$(1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p)(1 - L)^d Y_t = (1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q) \varepsilon_t$$

Y_t is the value that is predicted in time t and it is forecasted based on its historical values where: L is the lag which means $LY_t = Y_{t-1}$, d is the number of differentiations that must be made for the data to become stationary, $\phi_1, \phi_2, \dots, \phi_p$ are coefficients showing how strongly past lags affect future values, $\theta_1, \theta_2, \dots, \theta_q$ coefficients showing how past random shocks will affect future values, ε_t

stands for the noise. The model also uses p , d , and q , which determine how much past data used for prediction, how many times the data is differentiated, and how many past errors are used.

ARIMAX (p,d,q) does not differ significantly from the model presented above, the only difference being that at the end the following term is added:

$$ARIMA + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt}$$

$X_{1t}, X_{2t}, \dots, X_{kt}$ are responsible for the values of factors at time t . $\beta_1, \beta_2, \dots, \beta_k$ coefficients that correspond to the influence of factors on our dependent variable, in this case yield. Coefficients can have both a positive effect, increasing yield with their own growth, and a negative effect, i.e., decreasing yield with their own growth.

The model can be broken down into components such as the autoregressive (AR) part, which is responsible for the influence of past values on future ones. Next comes the differential part (I), which is responsible for ensuring data stationarity by eliminating trends, and the moving average part, which uses coefficients (θ) to estimate the impact of past unexpected shocks on future values. In the ARIMAX model, X is added, which is responsible for exogenous variables. In the case of the Ukrainian agricultural sector, the ARIMA model for forecasting crop yields would be easy to use, as it does not require a large array of complex data, but a simple analysis of time series without the use of independent variables will not allow us to form a clear picture of agricultural yields in Ukraine in the absence of war, as it is based only on past yield data without variable weather conditions. To build a clear model for forecasting crop yields in the absence of war, we must include significant factors that were not affected by the war; climate values did not have a direct impact from the war and will be used to adjust the analysis for specific years, namely 2022 and 2023.

Chapter 5 Results

This section of the results will answer the main question of this work, namely: crop losses, particularly grain crops, in Ukraine as a result of the war. To calculate the losses, we used methods for forecasting yields for 2022–2023, assuming that there will be no war, which will allow us to compare the forecast yield with the actual yield. To forecast yields and compare them with actual yields, two methods with different methodological approaches were used, which were described in the previous section, namely ARIMAX and the random forest regression method. First, the results of the time series analysis for each crop are presented, followed by the RFR results.

5.1 ARIMAX Results

In order to make yield forecasts in the absence of war for the six main crops of Ukrainian agricultural production, namely wheat, barley, corn, rapeseed, soybeans, and sunflowers, the ARIMAX model used historical data for these crops from 1992 to 2021 as training data which was used to create forecast values. The training data, which reflects the state of relative peace, is used to continue the trend of relative peace in the forecast for 2022-2023. Since ARIMAX is a model that also relies on predictors in forecasting, various factors independent of the war that could be significant were used, and a test for statistical significance will be demonstrated in the results.

First, the data is checked for stationarity. Stationarity means that the data has a stable mean value and a stable variance over time. Raw data cannot be used in time series analysis, especially Ukrainian crop yield data, which increases over time, because time series analysis requires a constant mean value and cannot assess the relationships between points when there is a growing trend. In order to make the data stationary and verify this, differentiation is first performed.

Stationarity is tested using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test, which allows us to assess whether the first differentiation helped to overcome non-stationarity, remove the trend, and check whether the time series has become stationary around a stable level. The test assumes that at a statistical significance level of 5%, the critical value will be 0.463, i.e., KPSS test values higher than this critical value will require differentiation.

The first differentiation allowed us to overcome non-stationarity for all crops, namely corn, wheat, sunflower, soybean, rapeseed, and barley, and the data can be further used to build time series analysis, as evidenced by the fact that the KPSS value after differentiation for each crop became less than the critical value, i.e., for each subsequent ARIMAX model, $d = 1$ (see Table 5).

Table 5. Results of checking historical yield data for 1992-2023 to identify stationarity.

Source: Own calculations based on FAOSTAT data of crop yields

Crop	KPSS_Level_Stat	KPSS_Level_5pct	KPSS_Diff_Stat
wheat	0.6569	0.463	0.1516
corn	0.8854	0.463	0.0972
sunflower	0.8007	0.463	0.2486
barley	0.5248	0.463	0.2221
rapeseed	0.8384	0.463	0.2018
soy	0.888	0.463	0.0652

The next step, once the data is differentiated and we are confident that the first level of differentiation ($d=1$) is sufficient for all cultures, is to move on to training and testing the different AR(p) and MA(q) parts. Testing the most suitable p and q values for the model will be done using cross-validation, which will allow us to evaluate how the model gradually works with the prediction of each subsequent value, gradually comparing it with the actual value. This method is more acceptable for evaluating the effectiveness of model forecasting, as it evaluates not the error and

residuals of unpredicted values over the entire period, but specifically the effectiveness of the model in working with new data.

In this method, the data is divided into training and testing data, with the mandatory condition that the training data cannot be further in the time series than the testing data, which is the essence of the chronological order of testing. Thus, the effectiveness of the model is evaluated through its prediction of each subsequent value after updating the training set in the same way as each subsequent value. In the case of agricultural data, this can be described as follows: start with the initial training period from 1992 to 2005 and test it on the 2006 forecast, then add one year to the training period and test the next one, let's say 2007 based on 1992-2006. This allows us to evaluate the effectiveness of the model based on data it has not seen and avoid overfitting. The effectiveness of the model is evaluated using the RMSE (Root Mean Square Error) indicator, which is the average of the errors between the predicted test values and the actual values squared, which is taken to the root. This allows us to evaluate the effectiveness of the model: the lower the RMSE value, the more accurate the model's predictions.

In order to select the most effective model, namely the indicators (p, q), since the indicator d, as was found in the stationarity test, is suitable at level 1, the `auto.arima()` function can be used, which is available from the forecast library in R. This function allows us to estimate the most suitable model parameters through automatic selection combined with cross-validation, where the model with the lowest RMSE is considered the best for further work. The model for each agricultural crop is tested separately on historical yield data from 1992 to 2021 (see Table 6). Corn shows the highest RMSE value (9.1), indicating the complexity of predicting the yield of this crop, so it can be assumed that it is the least predictable.

Table 6. Structure of ARIMAX models for each crop separately and RMSE (Root Mean Square Error) with cross-validation of these models

Source: Own calculations in R based on FAOSTAT data of crop yields

Crop	ARIMAX (p,d,q)	RMSE
wheat	ARIMAX (1,1,1)	3.86
corn	ARIMAX (2,1,0)	9.1
sunflower	ARIMAX (2,1,0)	2.46
barley	ARIMAX (1,1,1)	4.38
rapeseed	ARIMAX (1,1,1)	3.03
soy	ARIMAX (1,1,0)	2.9

For each model, the most appropriate ARIMA structure (p, d, q) was selected to balance the explanatory power of the model and the accuracy of the forecast (see Table 6). P shows how many past members (historical yield data) are used in the analysis, or, in other words, how many past yield values influence the forecast value; d shows how many times the data has been differentiated to become stationary, or how many times the yield value needs to be differentiated to eliminate the trend, and the last one is the q part, which is part of the moving average, showing us the relationship between past forecast errors and their impact on future yield values. The following factors were candidates for endogenous variables: average temperature from April to August in degrees Celsius (growing_season_avg(t)), biological nitrogen fixation (BiologicalFixation), and annual rainfall in millimeters (Millimeters of rain).

Table 7. Results of testing endogenous variables for statistical significance in models for each crop separately. Statistical significance at the 5% level. If the significant p-value is less than 0.05, then the predictor is considered statistically significant (green), otherwise red.

Source: Own calculations in R based on FAOSTAT data of crop yields

Endogenous variable	wheat	corn	sunflower	barley	rapeseed	soy
Biological Fixation	0.001	0.0951	0.0093	0.0569	0.0269	0.3941
Growing season avg(t)	8e-04	0.993	0.1405	0.0016	0.0793	0.6786
Millimeters of rain	0.3375	0.9854	0.1258	0.7382	0.342	0.1032

Table 7 shows the historical statistical significance of variables for our dependent variable in the form of yield for each crop separately. Table 7 shows that the yield of crops such as soybeans and corn is not statistically influenced by the variables, so the variables will not be used in the forecast. However, wheat yield is simultaneously influenced by both average temperature and biological nitrogen fixation, so both factors were used in the model.

The results presented in Table 8 show that, overall, the actual yield in 2022 was lower than predicted based on data for 1992–2021. The largest difference for 2022 is more than 4 centners for corn, which is the difference between the forecast of 67.61 centners per hectare in the absence of war and the actual figure of 63.49, but it should be noted that corn relies solely on historical values in forecasting (ARIMA is used). If to take into account the ARIMAX results with the endogenous variable, the largest difference is in barley, namely 3.46 for 2022, where the forecast value is higher. Next, there is a difference of 1.36 centners per hectare for wheat and 2.46 for sunflower. Soybeans, like corn, rely only on historical values due to the lack of statistical significance of factors, and the difference is 0.41 centners per hectare, where the forecast value is higher. The actual rapeseed yield results are slightly higher than predicted. As noted in the literature, 2023 saw very favorable weather conditions for agricultural production and adaptation to production conditions during the war, so Ukrainians got super high actual yields that beat the predicted results, except for barley. Despite this, the forecast for 2023 showed higher yields than in 2022, so it can be concluded that the forecast is correct and logically constructed.

Table 8. Forecasting results using ARIMA/ARIMAX models in centners per hectare. Loss shows the percentage difference between the forecasted and actual values. The difference is equal to the actual value minus the forecast.

Source: Own calculations using the ARIMAX models based on FAOSTAT data

Crop	Year	Actual	Forecast	Loss (%)	Difference
wheat	2022	39.25	40.61	-3.35	-1.36
	2023	46.42	42.06	10.37	4.36
corn	2022	63.49	67.61	-6.09	-4.12
	2023	78.06	68.72	13.59	9.34
sunflower	2022	21.63	24.09	-10.21	-2.46
	2023	24.53	23.31	5.23	1.22
rapeseed	2022	28.7	27.56	4.14	1.14
	2023	29.22	28.7	1.81	0.52
soy	2022	22.55	22.96	-1.79	-0.41
	2023	25.86	24.97	3.56	0.89
barley	2022	32.23	35.69	-9.69	-3.46
	2023	36.85	37.12	-0.73	-0.27

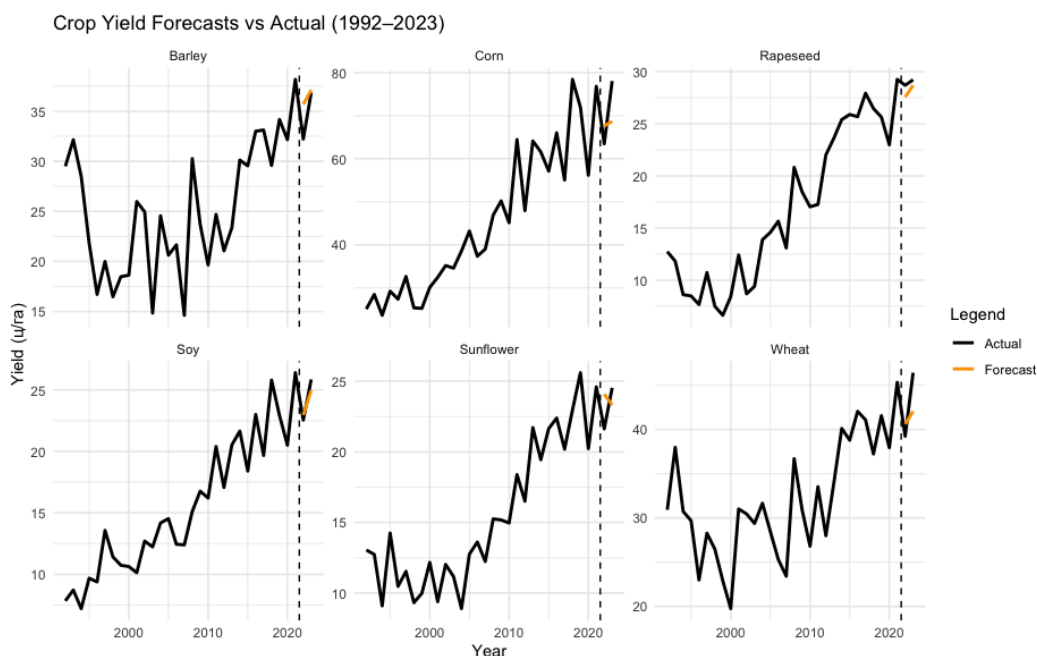


Figure 4. The black line represents the actual yield in centners per hectare, while the orange line shows the projected yield. The dotted line marks 2022 as the start of full-scale war.

Source: Own calculations using the ARIMAX models based on FAOSTAT data

5.2 RFR Results

As an addition to the ARIMAX model in the previous section, this section presents the results of Random Forest Regression (RFR). As mentioned earlier, their main difference lies in the fact that in RFR, the focus is on predictors, their interactions, and nonlinear dependencies, which are quite common in agricultural data.

To assess the impact of full-scale war in Ukraine on agricultural production, namely wheat, barley, soybean, corn, sunflower, and rapeseed yields, each RFR model was trained on data from a relatively peaceful period representing the period from 1992 to 2021, and used to forecast yields for 2022 in the absence of war. The difference between the forecast yield and the actual yield for 2022 reflects the impact of Russia's military intervention in Ukraine. The yield forecast, or hypothetical assumption, was constructed in such a way as to combine both factors that could be affected and those that do not depend on the war and the restrictions caused by it. To avoid the impact of the war on the forecast, factors that could be affected by the war, such as nitrogen per hectare, phosphorus per hectare, and potassium per hectare, were taken from 2021.

The quality of the model for each crop is assessed by the Mean Square Error (MSE), which represents the quadratic error of the test data. The smallest error was obtained for sunflower, equal to 1.54, while the largest error was for corn (38.79); all other crops showed an acceptable level of quadratic error (see Table 9). As in ARIMAX, it can be seen that sunflower showed the most accurate prediction, and corn showed the worst (see Table 6), which may be related to both the low informativeness of the data and the low predictability of the crop due to its dependence on a number of other factors that are not used in the model (see Table 9)

Table 9. Mean square error (MSE) for each crop in RFR

Source: Own calculations on R based on FAOSTAT data of crop yields and Cropland Nutrient Balance

Crop	MSE
Barley	12.20
Rapeseed	2.95
Soy	4.04
Sunflower	1.54
Wheat	7.47
Corn	38.79

Table 11 in Appendix 1, shows a list of variables and their values for the model for each crop separately. The significance of a variable in the model is assessed by metrics such as the percentage increase in the mean square error (%IncMSE) and IncNodePurity, which means the increase in node purity. The first value, %IncMSE, shows us how much the error will increase if to remove this value, while the second value shows us how well the variable contributes to branching in decision trees. The higher these indicators are, the better the variable is evaluated in terms of its impact on the explanatory power of the model.

Table 10 shows the results of Random Forest Regression forecasting for six major crops. For most crops, except sunflower, rapeseed, and soybeans, the forecast showed higher yields than the actual yields. In the case of rapeseed, the actual yield is taken. However, if we also refer to the ARIMAX results (see Table 8), it can be seen that the predicted result almost matches with actual result, so it can be concluded that the war had almost no impact on rapeseed yields. Sunflower yield showed the smallest error in the RFR forecast (1.54) (see Table 10), with a predicted yield of 22.30 centners per

hectare compared to the actual yield of 21.63. Soybeans are also similar to sunflowers, with a low error and a relatively small difference between the forecast and actual values: 22.93 centners per hectare compared to 22.55.

Table 10. Estimated RFR yield for 2022 compared to actual yield in centners per hectare.

Source: Own calculations on R based on FAOSTAT data of crop yields and Cropland Nutrient Balance

Crop	Predicted_Yield	Actual_2022_Yield	Loss
Barley	34.74	32.23	2.51
Rapeseed	27.48	28.70	-1.22
Soy	22.93	22.55	0.38
Sunflower	22.30	21.63	0.67
Wheat	41.10	39.25	1.85
Corn	66.26	63.49	2.77

Chapter 6 Discussion

The main objective of this study was to examine the impact of full-scale war in Ukraine on production indicators, i.e., the yield of major agricultural crops that are part of the country's exports and rank high even in global ratings. The aim of the study was to demonstrate higher yields in the absence of war and, based on projected yields, to assess the impact of the war in centners per hectare. Based on the results obtained, it can be concluded that in 2022, almost all crops, except rapeseed, experienced an average yield reduction of approximately 2 centners per hectare. In addition, it can also be concluded that corn was the most affected crop, as the difference between the forecast and actual values in one of the assumptions reached more than 4 centners (see Table 8). However, corn did not have statistically significant endogenous variables for forecasting in the time series analysis, so it was only forecast based on historical values. However, if to take only models with endogenous ARIMAX variables, as well as RFR results, the war in 2022 had the greatest impact on barley yields (3.46 centners per hectare ARIMAX, 2.51 RFR) (see Table 8, Table 10).

Yields are generally influenced by a number of factors, and in 2022, agricultural production was particularly negatively affected by the war, which disrupted normal production processes during a period of relative peace. As noted in the literature review, the share of fertilizer costs remained almost unchanged for most crops, with the exception of sunflower, rapeseed, and oats, where it decreased (Bogonos et al.). Oats were not included in this study, but the example of sunflowers demonstrated the difference (1.56 on average) between the yield assumed in peacetime and the actual yield during the war (see Table 8, Table 10). Rapeseed showed almost the same yield in both the forecast models and in reality. However, it is interesting to note that with the start of the full-scale invasion in 2022, more rapeseed was harvested than even in 2021. This is primarily due to the fact that more was sown and that yields remained almost at pre-war levels. However, the

number of hectares of rapeseed harvested in 2022 was primarily sown in the fall of 2021, which is also due to the fact that in 2021, it had a fairly high price.

As mentioned in RDNA4 from the World Bank, in 2023, Ukraine demonstrated very good harvests in agricultural production thanks to favorable weather conditions and partial adaptation of production to wartime conditions. The yield forecasts in this study showed lower yields in 2023 than the actual ones, except for barley. However, the forecast results for 2023, which are lower than the actual results, are not taken into account, and very effective management of Ukrainian agriculture can be observed, thanks to which, in the second year of the war, some crops yielded even higher yields than in peacetime (see Table 8). For example, wheat and corn yielded higher yields in 2023 than in 2021: 76,818 versus 78,060 centners per hectare for corn and 45,332 versus 46,422 for wheat (see Figure 4). Crops such as rapeseed and sunflower yielded almost the same harvests in 2023 during the war as in 2021. However, this does not cover the fact that in 2023, approximately 15 percent of the land used for agricultural production was still occupied or inaccessible for work (Bogonos et al.).

Overall, predictive analysis provides a deeper understanding of the fact that crop yields in agricultural production are difficult to predict, as it is impossible to account for the volatility of most factors, such as unpredictable weather conditions or price volatility for production factors that affect the internal management of each individual producer. This study found that Ukrainian agricultural production is showing a trend toward higher yields for major crops thanks to the development of the agricultural sector and closer interaction with EU markets (see Fig. 4) (Bogonos et al.). However, this study identified several crops with the highest level of error in explanatory power for building predictive models, namely corn. The fact that it had a higher level of error than other crops is confirmed by two methods, so it can be assumed that this crop is the most unpredictable in terms of cultivation and forecasting in general.

Chapter 7 Conclusion

As already mentioned in this article, Ukrainian agriculture is very rich and competitive in the world, but the full-scale war launched by Russia in 2022 has brought serious changes to Ukraine's agricultural production. The aim of this study was to examine the impact of the war on agricultural production results, measured by yield, specifically in centners per hectare in the context of this work. To assess the impact on yield, specifically to measure it in numerical terms, forecasting methods were used to estimate yield in the absence of war, which allowed for a comparison between predicted and actual yield. In this study, two methods were used to build yield prediction models: ARIMAX time series analysis and the Random Forest Regression machine learning method, which differ radically in their methodological approach. The first method built forecasts based on historical yield data using an endogenous variable for greater accuracy, while the second selected the most influential independent factors affecting yield. The forecast was made based on generalized historical data for Ukraine, so the yield was also generalized.

As expected, the results of both methods in this study showed that yields in the absence of war were higher than the actual results for 2022, while 2023 was a very successful year for agricultural production in Ukraine, so the predicted values were even lower, except barley (see Table 8). Both methods showed an average difference between the predicted and actual yields of 2 centners per hectare. Based on these results, the answer to the question about the impact of the war on yields in 2022 will be an average of 2 centners per hectare for major crops such as sunflower, wheat, corn, soybeans, and barley. Rapeseed was not included in this list because the predicted yield results coincided with the actual results (see Table 8, Table 10).

The results of the RFR machine learning model suggest that the independent variable of yield was most influenced by soil nutrient balance variables, namely phosphorus, potassium, and nitrogen (see

Table 11). Other variables related to atmospheric influences on yield, such as precipitation, average temperature during the main growing season, and biological fixation in the soil, had a less significant impact on average. However, the main limitation of this study is the data, which was either limited or unavailable. This study focused on generalized yield across the country due to a lack of data on soil conditions and weather conditions at the regional level. In the future, for a more in-depth study of the impact on yield, it is necessary to focus on regional or local levels, but this requires filling data registers.

Even during the war, Ukraine's agricultural production is showing very good results both locally and in terms of exports. A large part of the world's total exports depends on Ukraine's agricultural production, while the war continues and brings cumulative losses. In order to restore agricultural production to its pre-war level, huge investments are needed, as well as time and labor. Financial resources are needed to restore damaged or completely destroyed assets, while time is needed to restore processes, management, and return territories to Ukraine. These territories need to be restored to their ability to engage in agricultural activities, as many of them require prior demining. Financial investments concern not only external partners, but also internal ones, as the development of agricultural production depends primarily on internal conditions favorable for this. This also applies to a favorable political regime and the development of internal institutions, which require additional changes that will benefit agricultural production in the future. Ukrainian politicians must address this issue in order to bring the country closer to the EU.

Works Cited

- Albaladejo Román, Antonio. *Ukrainian Agriculture: From Russian Invasion to EU Integration*. European Parliamentary Research Service, Apr. 2024, [https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/760432/EPRS_BRI\(2024\)760432_EN.p](https://www.europarl.europa.eu/RegData/etudes/BRIE/2024/760432/EPRS_BRI(2024)760432_EN.p)
- Bogonos, Mariia, et al. *AGRICULTURE IN UKRAINE: PRE-WAR, STATUS QUO AND A LOOK AHEAD*. KSE Agrocenter, Kyiv School of Economics, July 2023, <https://kse.ua/wp-content/uploads/2024/03/Market-analysis-and-Outlook-of-Ukraine-2023.pdf>
- Bogonos, Mariia, et al. *Ahrovyhliad Ukrainy 2024–2033 [Agri-Outlook of Ukraine 2024–2033]*. KSEA, PNU, and NUBiP, 2024, <https://kse.ua/wp-content/uploads/2024/04/Agroviglyad-2024-2033.pdf>
- FAO. *Ukraine: Impact of the War on Agricultural Enterprises*. FAO, 28 Apr. 2023, <https://openknowledge.fao.org/items/7010613d-950c-4b7b-b184-0f73622f8159>
- Ghosh, Soham, et al. *Impact of Rainfall Risk on Rice Production: Realized Volatility in Mean Model*. arXiv:2504.10121, arXiv, 14 Apr. 2025. *arXiv.org*, <https://arxiv.org/pdf/2504.10121>
- Gupta, Ishaan, et al. *Innovations in Agricultural Forecasting: A Multivariate Regression Study on Global Crop Yield Prediction*. arXiv:2312.02254, arXiv, 14 Mar. 2024. *arXiv.org*, <https://arxiv.org/pdf/2312.02254>

- Jorvekar, Priti Prakash, et al. 'Predictive Modeling of Crop Yields: A Comparative Analysis of Regression Techniques for Agricultural Yield Prediction'. *Agricultural Engineering International: CIGR Journal*, vol. 26, no. 2, June 2024, pp. 125–40.
- Kaushik, Surbhi, et al. 'Forecasting Mustard Yield Using Weather Parameters : An Application of the ARIMAX Model'. *Journal of Computational Analysis and Applications (JoCAAA)*, vol. 33, no. 08, 08, Dec. 2024, pp. 2431–37.
[tps://www.eudoxuspress.com/index.php/pub/article/view/2121/1406](https://www.eudoxuspress.com/index.php/pub/article/view/2121/1406)
- Kruglenko, Alexandra. 'Vyiizd ukrainsiv do Yevropy u 2025 rotsi zmenshyvsia, ale viddik liudei tryvaie'. *Зеркало недели | Дзеркало тижня | Mirror Weekly*,
<https://zn.ua/ukr/UKRAINE/vijizd-ukrajintsiv-do-jevropi-u-2025-rotsi-zmenshivsja-ale-vidtik-ljudej-prodovzhujetsja.html>. Accessed 22 May 2025.
- Martyshev, Pavlo, et al. *AgroDigest Ukraine January 2025*. Kyiv School of Economics, Jan. 2025.
 Zotero, https://kse.ua/AgroDigest_Ukraine_January_2025.pdf.
- Neyter, Roman, et al. *Zbytky, Vtraty Ta Potreby Silskoho Hospodarstva Cherez Povnomasshtabne Vtorhnennia [Damages, Losses and Needs of Agriculture Due to the Full-Scale Invasion]*. Kyiv School of Economics, World Bank, 2024,
https://kse.ua/wp-content/uploads/2024/02/RDNA3_ukr.pdf
- Nivievskyi, Oleg, et al. 'Impact of Russia's War on Ukrainian Agriculture'. *Ukrainian Analytical Digest*, application/pdf, no. 005, Mar. 2024, pp. 2–7. *DOI.org (Datacite)*,
<https://doi.org/10.3929/ETHZ-B-000665476>.
- Sushchuk, Maksym. 'The Impact Of War In Ukraine On Global Food Security'. *EOS Blog*, Aug. 2024, <https://eos.com/blog/the-impact-of-war-in-ukraine-on-global-food-security/>.

Vira Khmelnytska. 'The Wall Street Journal підрахував, скільки людей загинуло з обох сторін за період війни'. (TSN.ua), 17 Sept. 2024,

<https://tsn.ua/ato/the-wall-street-journal-pidrahuvav-skilki-lyudey-zaginulo-z-oboh-storin-za-period-viyni-2661153.html>

Vitaliy Koval. 'U 2024 rotsi ahrarna produktsiia sklala 59% v zahal'nomu eksporti'. *МІНІСТЕРСТВО АГРАРНОЇ ПОЛІТИКИ ТА ПРОДОВОЛЬСТВА УКРАЇНИ*, 3 Jan. 2025,

<https://minagro.gov.ua/news/u-2024-rotsi-ahrarna-produktsiia-sklala-59-v-zahalnomu-eksporti>

World Bank. 'Agriculture, Value Added (% of GDP) – Ukraine'. *World Bank Open Data*, 3 May 2025, <https://data.worldbank.org>

World Bank,. *Ukraine: Fourth Rapid Damage and Needs Assessment (RDNA4), February 2022 – December 2024.*, Feb. 2025,

<https://openknowledge.worldbank.org/server/api/core/bitstreams/c178c51a-47ea-47c0-a424-e6b38c700dce/content>

Appendix 1

Table 11. Agricultural crops and variables used in RFR models.

Source: Own calculations in R based on FAOSTAT data of crop yields and Cropland Nutrient Balance

Crop	Variable	%IncMSE	IncNodePurity
Barley	AtmosphericDeposition	-1.12459056	55.317146
Barley	BiologicalFixation	5.32795569	256.376879
Barley	growing_season_avg	0.04132143	20.741337
Barley	Leaching	2.55545972	73.534294
Barley	Cropland_phosphorus_per_unit_area	3.78051820	109.860668
Barley	Cropland_potassium_per_unit_area	3.29279781	94.487248
Barley	Cropland_nitrogen_per_unit_area	2.42151430	50.604514
Barley	winter_season_avg	2.02548384	55.859128
Corn	AtmosphericDeposition	3.92552915	739.800672
Corn	BiologicalFixation	3.74296657	691.932586
Corn	growing_season_avg	2.55311694	340.881786
Corn	Leaching	2.32187434	189.645521
Corn	Cropland_phosphorus_per_unit_area	4.67682577	945.338872
Corn	Cropland_potassium_per_unit_area	6.07966363	1,371.603399
Corn	Cropland_nitrogen_per_unit_area	3.47562895	531.332999
Rapeseed	AtmosphericDeposition	1.91887384	62.224415
Rapeseed	BiologicalFixation	3.68639744	142.982089
Rapeseed	growing_season_avg	1.95258727	30.535399
Rapeseed	Leaching	3.86554115	61.751817
Rapeseed	ММ_дош	0.91764172	9.466325
Rapeseed	Cropland_phosphorus_per_unit_area	4.56111592	210.389345
Rapeseed	Cropland_potassium_per_unit_area	5.51946480	331.803987
Rapeseed	Cropland_nitrogen_per_unit_area	2.05464490	16.778977
Soy	AtmosphericDeposition	2.78886307	52.520787

Soy	BiologicalFixation	3.92844423	64.402726
Soy	growing_season_avg	1.64096939	19.536635
Soy	Leaching	4.53952416	25.097805
Soy	ММ_дош	0.67009534	4.238521
Soy	Cropland_phosphorus_per_unit_area	4.68462488	104.994215
Soy	Cropland_potassium_per_unit_area	6.32138195	211.343190
Soy	Cropland_nitrogen_per_unit_area	2.35617678	48.593967
Soy	autumn_season_avg	0.76833636	7.440419
Sunflower	AtmosphericDeposition	1.47083916	31.313083
Sunflower	BiologicalFixation	3.92579680	86.563837
Sunflower	Leaching	1.21554035	33.296475
Sunflower	ММ_дош	0.42884059	17.467777
Sunflower	Cropland_phosphorus_per_unit_area	4.19285228	111.989318
Sunflower	Cropland_potassium_per_unit_area	5.48775777	125.343496
Sunflower	Cropland_nitrogen_per_unit_area	0.53974631	29.253486
Wheat	AtmosphericDeposition	2.48733508	57.159250
Wheat	BiologicalFixation	5.87434353	240.748299
Wheat	growing_season_avg	1.72403394	56.521374
Wheat	Leaching	3.02993910	101.970181
Wheat	Cropland_phosphorus_per_unit_area	5.91257475	193.632108
Wheat	Cropland_potassium_per_unit_area	4.53598960	138.687194
Wheat	Cropland_nitrogen_per_unit_area	3.56948181	74.967472
Wheat	winter_season_avg	0.85148026	84.255270